Cornelius Lassen Hjort
Hans Erik Bjørkum Heum

Fermentation Prediction through Machine Learning and its Potential Use in Production Planning and Control

Master's thesis in Ingeniørvitenskap og IKT
Supervisor: Anita Romsdal
June 2023

Master's thesis

NTNU
Norwegian University of Science and Technology
Faculty of Engineering
Department of Mechanical and Industrial Engineering

NTNU
Norwegian University of
Science and Technology

Cornelius Lassen Hjort
Hans Erik Bjørkum Heum

Fermentation Prediction through Machine Learning and its Potential Use in Production Planning and Control

Master’s thesis in Ingeniørvitenskap og IKT
Supervisor: Anita Romsdal
June 2023

Norwegian University of Science and Technology
Faculty of Engineering
Department of Mechanical and Industrial Engineering

NTNU
Norwegian University of
Science and Technology

Preface

This Master's thesis is the final report in the Engineering and ICT master program with a specialization in production management at the Norwegian University of Science and Technology (NTNU).

We would like to thank our supervisor Anita Romsdal for the feedback, comments, and support during the semester.

We would also like to thank Plaato for access to data and cooperation during the project. In particular we would like to thank software engineers Erik Olseng and Kristian Pedersen for sharing valuable insights, good suggestions and support.

Cornelius Hjort & Hans Erik Heum

Trondheim, June 2023

Abstract

Craft breweries face a multitude of challenges when it comes to production planning and control (PPC), largely due to the complex and unpredictable nature of their production processes combined with variable customer demand and the perishability of their products. Fermentation is the most time consuming production process and arguably the most critical. It adds significant uncertainty due to its variable duration, which poses significant challenges for PPC in craft breweries. Despite inconsistent results from previous attempts to enhance fermentation predictability, emerging technologies allow real-time tracking of the process combined with more data, presenting a potential for machine learning applications.

The purpose of this study is to investigate the feasibility of accurately determining the completion of a beer fermentation process and how this knowledge could be employed to mitigate production planning and control issues related to fermentation process time uncertainty. The research delves deeper into whether machine learning methods can be used to predict how long beer fermentation takes. It also explores the challenges craft breweries face due to the uncertainty, and how fermentation forecasts can be used to reduce them.

A theoretical background is conducted to identify key elements of production planning and control in general, along with relevant aspects of machine learning. Furthermore, essential aspects of production planning and control in the beer production industry are further identified via an empirical background. This also includes findings from a multiple case study focusing on production planning and control in craft breweries.

This thesis proposes four predictive models, using a number of input parameters to estimate the completion time of a fermentation process. A neural network, two gradient boosted forests and an automatic ML algorithm were applied to datasets of 40, 60 and 80 hours of information in the fermentation. The best models were more accurate than a baseline model that predicted the average. However, the models are struggling to predict accurately on batches that deviate from normal fermentation activity. Additionally, the study revealed that access to a substantial amount of data with high quality is an important factor when using machine learning combined with IoT.

A single case study involving one participating craft brewery is conducted to understand the current state of its production planning and control activities. Followed by an analysis of the information collected by the case company. As a result, we identified that fermentation process time uncertainty contributes to several PPC challenges in craft breweries, such as managing the flow of material, effective inventory management and accurate production planning and scheduling. These challenges impact capacity utilization and the actual production capacity by prolonging each production cycle, thus limiting the amount a craft brewery can produce. Furthermore, it results in idle time in production and increased production throughput time, impacting the amount of inventory held.

Overall, we identified that it is possible to retrieve insights about the fermentation process by using machine learning combined with industrial sensors. Secondly, we identified that these insights could cause substantial value in a craft brewery. Numerous topics for further development are identified, most importantly to actually implement and monitor the value gain of a machine learning model in a real craft brewery, but also the retrieval of more data.

Sammendrag

Håndverksbryggerier står overfor en rekke utfordringer når det kommer til produksjonsplanlegging og -kontroll (PPC), hovedsakelig på grunn av komplekse og uforutsigbare produksjonsprosesser, kombinert med varierende etterspørsel fra kunder og produktets begrensede holdbarhet. Fermentering er den mest tidkrevende produksjonsprosessen og den mest kritiske. På grunn av fermenteringens variable varighet fører den til PPC utfordringer i bryggerier. Tidligere studier har utforsket hvordan fermenteringsprosessen kan bli mer forutsigbar ved hjelp av tradisjonelle metoder. Utviklingen av nye teknologier og økt antall mengde data tilgjengelig fører til et stort potensiale for bruk av maskinlæring.

Formålet med denne oppgaven er å undersøke hvor nøyaktig og hvor tidlig man i forkant kan predikere når en fermenteringsprosess i øl er ferdig, og hvordan denne kunnskapen kan brukes til å redusere utfordringene som kommer av usikkerheten i fermentering. Oppgaven utforsker hvilke utfordringer fermenteringsusikkerheten fører til, og hvordan fermenteringsprognoser kan brukes til å redusere dem.

Gjennom en teoretisk bakgrunn identifiserer vi de viktigste temaene i PPC og maskinlæring. Videre identifiseres viktige PPC aspekter i ølbryggerier gjennom en empirisk bakgrunn. Denne inkluderer også funn fra et case-studie som fokuserer på PPC i ølbryggerier. Relatert forskning blir også presentert.

Denne oppgaven utvikler fire maskinlæringsmodeller for å predikere når fermenteringsprosessen er ferdig. Et nevralt nettverk, to gradient boosted rammeverk og en automatisk ML-algoritme ble anvendt på datasett med 40, 60 og 80 timers informasjon om fermenteringsprosessen. De beste modellene predikerte mer nøyaktig enn en utviklet baseline-modell, som kun predikerte gjennomsnittet av dataen den trente på. Modellene har dog problemer med å predikere nøyaktig når batchen den predikerer på avviker fra normal fermenteringsaktivitet. Studien viser at en avgjørende faktor for gode prediksjoner når man bruker maskinlæring med IoT er et stort volum av data med høy kvalitet.

Vi gjennomførte et casestudie med et håndverksbryggeri for å få innsikt i deres næverende PPC aktiviteter. Som et resultat ble det identifisert at usikkerheten i fermenteringsprosessen bidrar til utfordringer knyttet til PPC. Som for eksempel å håndtere materialflyten, ineffektiv lagerstyring og unøyaktig produksjonsplanlegging. Disse utfordringene påvirker kapasitetutnyttelsen og den faktiske produksjonskapasiteten, siden hver produskjonssyklus blir forlenget. Det resulterer også til uvirksom tid i produksjonen og at produksjone tar lenger tid, som gjør at lagernivået øker.

Oppgaven viser at det er mulig og et stort potensiale til å få innsikt i fermenteringsprosesser ved å kombinere maskinlæring med industrielle sensorer. Denne innsikten kan skape betydelig verdi for ølbryggerier. Vårt studie legger grunnlaget for å faktisk implementere en maskinlæringsmodell i et bryggeri, og måle og evaluere verdiene det skaper. Studiet viser også at mer data med høy kvalitet er nødvendig.

Contents

List of Figures

List of Tables

Abbreviations

ABV - Alcohol by volume

AI - Artificial Intelligence

ANN - Artificial neural network

APP - Aggregate production planning

ARIMA - Autoregressive integrated moving average

ATO - Assemble-to-order

BBL - Beer barrels

BBL - Barrels of Beer = 117 Liters

BOM - Bill of material

CODP - Customer order decoupling point

CPU - Central processing unit

CPS - Cyber physical systems

EDA - Exploratory data analysis

ERP - Enterprise resource planning

ETO - Engineer-to-order

FG - Final gravity

GPU - Graphics processing unit

HORECA - Hotel, restaurant and café

ICT - Information and communication technology

IoT - Internet of things

IT - Information technology

KNN - K-nearest neighbour

ML - Machine learning

MPS - Master production schedule

MRP - Material requirements planning

MROs - Maintenance, repair, and operational supplies

MTS - Make-to-stock

MTO - Make-to-order

OG - Original gravity

PAC - Production activity control

PLC - Product life cycle

PPC - Production planning and control

RNN - Recurrent neural network

RQ - Research question

SaaS - Software as a service

SCIM - Supply chain inventory management

SCP - Supply chain planning

SCM - Supply chain canagement

S&OP - Sales and operation planning

SVM - Support vector machine

WIP - Work-in-process

Chapter 1

Introduction

This chapter presents the background and motivation of our thesis, followed by a formulation of the problem and research questions and objectives. Lastly, we present the scope and an overview of the thesis.

1.1 Background

Beer is a beverage that has been part of human culture for thousands of years and is today the third most consumed beverage in the world and the most consumed alcoholic beverage (Salanță et al. 2020). Its popularity is reflected in the fact that there are more than 25.000 breweries globally, which produce approximately 1,8 billion hectoliters of beer every year (Conway 2021). This production is valued at nearly 800 billion dollars (Conway 2021).

Beer production is one of the most well-studied processes in the food sector (Pöllänen et al. 2001), which in general is known to be challenging to control due to its variability and unpredictability (Pöllänen et al. 2001). The production of beer consists of three interlinked, but still distinct, stages (Stewart et al. 2016). These are wort preparation, fermentation and post-fermentation processing (Stewart et al. 2016). Among these, fermentation is arguably the most important process step in the production of beer (Pöllänen et al. 2001). It is during this stage that yeast metabolizes the sugar to produce alcohol, carbon dioxide and other compounds that contribute to beer's flavor and aroma (Dequin 2001). However, the fermentation process is also known to be a significant production planning and control (PPC) challenge (Pöllänen et al. 2001). This is because the fermentation times in seemingly equivalent settings can vary considerably (Pöllänen et al. 2001).

Being able to predict the duration of beer fermentations would be useful with regard to production management (Gopal et al. 1993; Johnson et al. 1998). Through a multiple case study, we have earlier interviewed, analysed and identified PPC challenges occurring by the uncertainty in fermentation (Heum and Hjort 2022). Examples of such effects are difficulties in effective scheduling and low capacity utilization. Tools capable of mitigating the uncertainty of the durations of beer fermentations can thus be of high value to all breweries.

There currently exist research that explores forecasting techniques on fermentation. For instance, Montague et al. (2008) demonstrates how case based reasoning can be used to forecast future bioprocess conditions. Syu et al. (1994) made an artificial neural network to predict the effects of varying input conditions on hypothetical fermentations, and Speers et al. (2003) developed a non-linear regression model of brewing fermentations. However, these studies focus on predicting a target variable at a certain time step, by using other features at the same time step. In addition, previous studies does not research how increased predictability affects PPC.

A common weakness in all previous studies is the lack of data. Some studies had access to data monitored from sensors, but they express the inconsistent quality of the measures, which caused challenges when analysing and using the data. In this research, we have access to advanced monitoring sensors that frequently take measures of parameters in the brewery. We want to use this substantial amount of data, combined with the current development of machine learning models to develop forecasting techniques. We investigate how insights about the beer fermentation process can improve PPC activities in breweries.

1.2 Problem Formulation

Technological advancements have made real time monitoring of wort density available in industrial-scale brewing. This development has drastically increased the amount of data available to base predictive models on, opening up new possibilities for improving the predictability of beer fermentation times.

This paper will describe machine learning models for the prediction of beer fermentation duration and will, as opposed to earlier research, be based on both parameters available in industrial-scale brewing and complete fermentation curves from real time density monitoring. This paper will further investigate what the exact value of increased predictability in beer fermentation potentially can be, considering different levels of the model's accuracy, for the craft brewing industry.

This paper aims to contribute to the ongoing efforts to improve the efficiency and effectiveness of beer production, with a particular focus on the critical and challenging process of fermentation. The study also aims to fill the research gap regarding improved fermentation insights combined with production planning and control. This work will provide valuable insights and tools for breweries that want to start using machine learning. The research aims to motivate further research of the implementation of machine learning models in breweries, and monitoring its effect on PPC.

1.3 Research Questions and Objectives

Based on the problem formulation the following research questions have been formulated and serve as guidance through the thesis:

RQ1: *How can ML be used in order to retrieve insights about the fermentation process in breweries?*

The purpose of this research question is to address to what degree it is possible to accurately determine when a beer fermentation process will be finished. "Accurately" means how many days beforehand we can assure with an error deviation that the process will be completed. Included in this RQ is research about finding and making the most consistent and precise machine learning models that are suitable and can be reused in another context.

In order to answer this research question it has been broken down into the following two key objectives to complete:

- Identify forecasting approaches suited for forecasting the fermentation process in breweries.
- Develop machine learning models and evaluate their performance.

RQ2: *How can increased predictability regarding the duration of beer fermentations mitigate production planning and control challenges in breweries?*

The purpose of this research question is to determine the positive ripple effects on PPC in a brewery with more insights regarding when a fermentation process is finished.

In order to answer this research question has been broken down into the following three key objectives to complete:

- Identify key production planning and control challenges in craft breweries caused by the uncertainty of the fermentation process time.
- Identify how challenges caused by the uncertainty of the fermentation process time can be mitigated.
- Identify the benefits mitigating the challenges caused by the fermentation process time can have for craft breweries.

1.4 Research Scope

The scope of this project is to investigate how to increase the predictability of the beer fermentation process and how increased predictability can affect PPC activities. In order to understand how the machine learning problem can be formalized and solved, the theoretical background consists of fundamental theory regarding machine learning. The scope is limited to supervised learning on time series, as other machine learning problems are irrelevant to our thesis. Fundamental theory regarding PPC is also presented, with a focus on important activities in planning and control, supply chain management and digitalization which is relevant to our thesis. Even though lean manufacturing is often used in the context of smart PPC and digitalization, the area is deemed too broad, and could be a thesis on its own. Sustainability is another interesting topic that could be affected by machine learning. Nevertheless, this topic is deemed too broad and therefore excluded from this thesis.

The data used in the development of machine learning models are limited to the data available through the collaboration with Plaato. This data primarily consists of information regarding density, temperature, yeast and time stamps. Previous study has shown great success by using additional information, such as amino concentration, oxygen concentration and biomass. Even though combining this information with measurements from sensors could give interesting results, the study is limited to Plaato's data. We apply machine learning models that are highly regarded in the literature. The aim of the paper is not to develop new machine learning models specific to this problem. Our aim is to use current state-of-the-art time series forecasting models in order to investigate how accurately current models can predict the completion of the fermentation process.

We do not have access to fermentation monitoring data from macrobreweries, which are large, multinational corporations with significant market share. The database we have access to rather consists of data from several craft breweries, whereas the majority are located within the United States. As a result, the case study in this thesis is conducted with the American craft brewery Lock 27 Brewing Company. All gathered data was provided by them based on our scope for the case study, which is production planning and control. However, in order to thoroughly conduct an analysis of production planning and control we considered it relevant to investigate supply chain management as well because of its impact on production planning and control activities. The following aspects were thus investigated: Procurement and order handling, inventory management, forecasting and demand management, production planning and scheduling and supply chain management.

The case study has a broad focus on production planning and control, with the reason being that a thorough understanding of the current production planning and control situation is needed to analyze the connection to fermentation uncertainty and thus investigate how increased predictability of beer fermentations can affect PPC. However, aspects of operations management, such as social, environmental and economic sustainability is not within the scope. It is arguably relevant to the subject of this thesis, but deemed to broad to include. Furthermore, Lock 27 Brewing Company has implemented sensors for fermentation monitoring. However, the case study does not implement increased predictability in practice, but rather utilizes a theoretical approach.

1.5 Thesis Structure

Chapter	Description
Introduction	Introduces the background for the thesis and describes the research objective, research problem, and scope of the thesis.
Research methodology	Describes how the literature used in the theoretical background was retrieved, how the models were developed and how the interviews were gathered.
Theoretical background	Presents relevant literature needed in order to answer the RQs. Includes theory regarding PPC, machine learning and presents related research.
Empirical background	Investigates existing literature, research, the project thesis other sources such as trade association reports to describe the beer production industry.
Model development	Develops machine learning algorithms, evaluate their results and elaborates on the their performance.
Case study	Investigates the challenges and possible improvements of PPC in a craft brewery. Applications and improvements by a machine learning model is also discussed.
Discussion	Combines findings from the theoretical background, empirical background, model development and the case study to answer and discuss the research questions.
Conclusion	Summarizes key takeaways and conlcudes the research.

Table 1.1: Overview of thesis structure

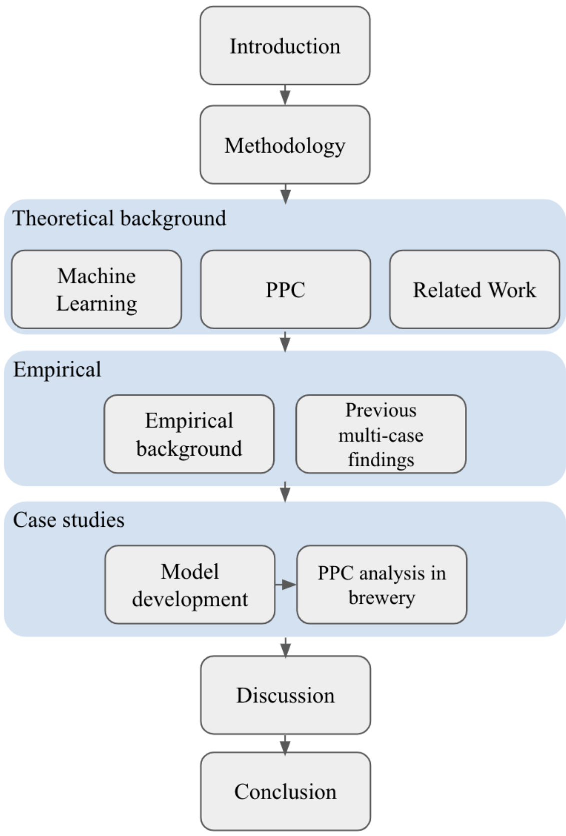


Figure 1.1: Explaination of each chapter in the thesis.

Chapter 2

Research Methodology

In this thesis, the research methodology consists of both a literature study, the development of models and a case study. This chapter presents and describes the research process, the research design, the data collection and the analyses performed.

Introduction

This chapter outlines the research methodology undertaken for this thesis. As stated by Kothari (2004), research methodology represents the systematic way in which a research problem is addressed, and the research methods are the tools used to perform the research. An important component of the research methodology involves describing the reasoning behind the selection and use of certain methods to investigate the research topic (Kothari 2004). Two distinct methodologies, namely qualitative and quantitative research, can be differentiated based on the attributes of the methods employed. Quantitative methods prioritize the analysis of numerical data via mathematical and statistical tools (Moser and Korstjens 2018), while qualitative methods aim to interpret meaning from non-numerical data such as literature reviews, observations and interviews (Moser and Korstjens 2018).

In this thesis, we adopted a quantitative approach focusing on the machine learning model, but also qualitative approaches such as a literature study and in-depth interviews. The methodological choices were taken considering the given research questions and the scope of the thesis. One of the objectives was to identify the challenges breweries are facing in terms of PPC. In order to do this, we conducted in-depth interviews with a representative brewery. Another objective was to create and implement a model that could capture the complex patterns present in the fermentation process of beer brewing and predict the time to completion. This required a systematic approach, involving the examination of the theoretical framework and comparison with related work. In order to understand and interpret the insights given from the interviews and in order to apply machine learning to PPC, we conducted a literature study. The combination of a literature study, the development of a machine learning model, and the analysis of PPC in breweries we aim to answer and discuss the research questions. An overview of how the thesis is structured regarding the objectives can be seen in figure 2.1.

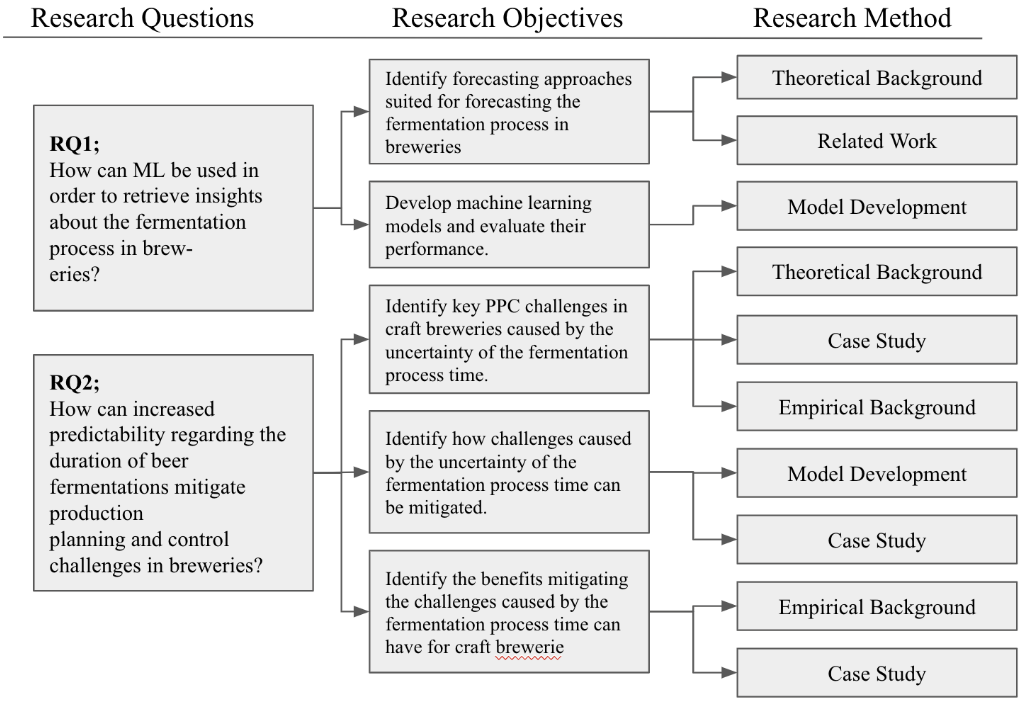


Figure 2.1: The link between RQ's, objectives and research method.

2.1 Literature Study

We chose to conduct a literature review as part of our research methods. Moser and Korstjens (2018) explains that a literature review provides a broad view of existing research on relevant topics and gives a clear picture of state-of-the-art solutions. It's an essential step to ensure that the research contributes differently from other research (Snyder 2019). Moreover, a literature study is essential to narrow the scope of the thesis and ensure the focus is on the research gap (Karlsson 2010). In order to gain a fundamental academic understanding of the research scope, and in order to interpret the insights from the interviews and how to apply a machine learning model, the literature study is divided into three parts. The theoretical background focuses on establishing a common ground knowledge base, by explaining the principles of production planning and control and machine learning and its domain. The empirical background aims to understand the context of brewing and PPC in brewing in particular. We also review prior studies in the field, with the objective of identifying the most suitable methodologies, as well as those that may not be as effective.

2.1.1 Theoretical background

The theoretical background primarily involves a theoretical investigation, which includes analyzing existing literature to establish the credibility and relevance of the research. By doing so, it will also help to identify research gaps that limit the research's scope (Moser and Korstjens 2018). The theoretical background was executed to gain insights into relevant topics and identify research opportunities. In particular, the study was required for this research in order to acquire and understand insights into production planning and control in breweries. The literature study was carried out to get knowledge on production planning and control in order to understand the case company and the challenges they face. As PPC is a broad topic, we have focused on the most relevant topics in PPC regarding our thesis. Production characteristics, production planning and Industry 4.0 in PPC are examples of such topics. Further, the theoretical background presents important topics within machine learning. This part is important in order to explore possible solutions in the machine learning domain and to be able to implement the best machine learning suitable to our problem statement. The theoretical background is used as support to answer the research questions.

Production planning and control cover a wide range of concepts. To get a better overview of the topic and in order to understand what principles in PPC that is relevant to our thesis we initiated the literature study by reading the books Fundamentals of Production Planning & Control (Chapman 2006) and Operations Management (William J Stevenson et al. 2014). These books gave us a comprehensible overview of the topic and helped us describe topics in the theoretical background section. However, these books may contain outdated information and are insufficient on some particular topics.

Machine learning is a broad topic, and we, therefore, establish an academic foundation with the most important topics relevant to our research scope through the theoretical background. Most of the conducted literature is from machine learning books such as *Artificial Intelligence - a Modern Approach* (Russell 2010), *The elements of statistical learning: data mining, inference, and prediction* (Hastie et al. 2009) and *Introduction to time series and forecasting* (Brockwell and Davis 2002). This theoretical part consists of an introduction to machine learning, what problems machine learning may solve, time-series and current state-of-the-art approaches and models. The scope of this thesis is not to deeply understand how each model functions mathematically. However, it is important to understand their scope and where they are applicable. The theoretical part, therefore briefly explains the functionality of different machine learning models.

Because some of the books contain outdated information and does not cover desired topics sufficiently, we systematically gathered research papers and articles relevant to our research. Initially, we gathered more information regarding PPC and machine learning through Google searches. This was done to get a better overview of the topics before retrieving articles about specific topics. PPC and machine learning in breweries specifically, is an area with few papers. We therefore also looked at PPC and machine learning in food processes as well, because they share several of the same characteristics as breweries.

After gaining a broader understanding of the research topic, we implemented a building block search as a strategy to retrieve further knowledge. We first made a list of primary search words, that aimed to to cover important topics in PPC and machine learning. Further, we made a list of search words belonging to each of the topics in the primary list, seen in table 2.1. These words were then joined with the main words and were used to search academic databases. We mainly used Google Scholar, Science Direct and Research Gate to collect literature. This search resulted in a significant amount of research papers and articles. In order to exclude irrelevant articles and duplicates, we applied a filtering approach. First of all, we reviewed the title and keywords written about the paper. Further, we reviewed the abstract and conclusion to ensure the relevancy and quality of the paper. While searching through the literature, we also checked the quality and reliability of the paper. This was done by examining the authors, publisher, amount of citations and when it was published. The citation count was given less importance if the paper was recently published, considering that it could still maintain considerable quality (Snyder 2019). If an article was deemed relevant and passed the preliminary quality assessment, we read the article in its entirety. Afterward, we classified the articles regarding what topic the research contributes to, and a short review was written. After the initial set of papers was identified, the snowball sampling technique was employed to expand the search. This involved using the references cited within the previously identified papers to uncover additional relevant papers (Goodman 1961).

Primary search words	Additional search words
Production planning and control	fundamentals, operations management, brewing, food production, sensors, automated, smart, challenges, Industry 4.0
Inventory management	brewery, smart, challenges, machine learning, improvement
Flow of material	brewery, bottleneck, smart, challenges, machine learning, improvement
Machine learning	fermentation, implementation, production planning and control, sensors, IoT, exploratory data analysis, missing data, forecating
Neural networks	fermentation, brewery, production planning and control, hyperparameter tuning, complexity, optimization, LSTM, RNN, forecasting
Decision tree	ensemble learning, gradient boosted forest, XGBoost, CatBoost, hyperparameter tuning, brewery, fermentation, forecasting
Time series	forecasting, multiple, multivariate, fermentation, supervised, machine learning, production planning and control

Table 2.1: Overview of used search words

2.1.2 Existing literature

Related work is a vital component of a thesis as it situates your research within the academic landscape and shows the understanding of the existing body of knowledge (Karlsson 2010). It is important to analyze current literature that either aligns with or contradicts the concepts and theories of our research. The aim of reading previous work is to identify gaps or inconsistencies in existing literature, and in doing so, demonstrate the necessity of our study. Additionally, it is useful for getting inspiration from techniques and approaches that have shown great success earlier. It may also provide information regarding which techniques to avoid. This extensive review of literature can enrich our analysis, enabling us to answer the research questions more comprehensively and with support from earlier research.

The majority of the papers used in this section were found through the systematic approach mentioned in section 2.1.1. However, we had to expand our research further to find literature more similar to our scope. We divided the related work into three separate areas to gain research about. The first topic consisted of research focusing on PPC in breweries and food production. The next topic aimed at articles about forecasting and prediction techniques of the different fermentation processes. The last topic combines the latter topics and focuses on how information regarding fermentation can affect PPC in a brewery or in food production.

In addition to gathering research papers through the literature search, many relevant and similar studies was found in the *Technical Quarterly* which is a journal published by Master Brewers Association of the Americas. The Technical Quarterly is an online journal. It features papers covering wide technical aspects of brewing ingredients, the brewing process, brewing by-products, environmental considerations in breweries, beer packaging, and beer flavor and physical stability (*Master Brewers Association of the Americas: Publications* 2023). Brewing professionals in production, engineering, quality control, research, packaging and material handling, maintenance, and administration in brewing, malting, and associated industries utilize this essential resource (*Mas-*

ter Brewers Association of the Americas: Publications 2023). Membership in the Master Brewers is not a requirement for publication. However, in order to read the journal, one must pay for a subscription.

2.1.3 Empirical background

To gain insight into how production planning and control is executed in the beer production industry, and particularly craft breweries, we supplement our theoretical background with empirical evidence. The empirical foundation shares numerous features with the theoretical one but focuses specifically on food, beverage, and beer production. An empirical investigation proved to be suitable for this project, incorporating resources such as annual reports and industry-specific journals from independent associations, alongside scientific papers.

The empirical background also consists of a summary of a case study we conducted in the project thesis (Heum and Hjort 2022). The study focuses on production planning and control in multiple lager breweries around the world. This research was gathered through several in-depth interviews. The goal of including information from the multiple case study is to inform general knowledge regarding how PPC actually is managed in breweries. The study is interesting in terms of how much improvement a machine learning model could cause when applied.

2.2 Model Development

In order to gather up-to-date information on machine learning, we adopted a more modern approach in addition to the literature review. We recognized that the field of machine learning evolves rapidly, and therefore, modern and relevant research is found in public notebooks, online articles, technical blogs and other digital platforms where professionals and enthusiasts share their findings, insights, and breakthroughs. This approach provided us with a broader perspective and more contemporary understanding, ensuring our thesis was anchored in the most recent developments within the field. Therefore, some of the research regarding machine learning is from sources such as Machine Learning Mastery (Brownlee 2016), Medium (Medium 2023), Towards Data Science (*Towards Data Science: Your home for data science. A Medium publication sharing concepts, ideas, and codes* 2023) and Kaggle (*Kaggle: Your Machine Learning and Data Science Community* 2023).

Our methodology for the development of a machine learning model is primarily focused on the exploration of various machine learning algorithms, aiming to predict the fermentation process. Rather than choosing a single model, we decided to implement multiple models. This approach allowed us to examine the predictive capabilities of different algorithms and identify the one that provides the most accurate prediction. A goal of this thesis is to determine to what extent it is possible to predict the fermentation process, rather than how well it optimally can predict. Therefore, several models will give relevant insights regarding the scope of the thesis.

While traditional approaches such as logistic models and optimization models are widely used in research, we intentionally had a different approach. The popular field of machine learning presents substantial opportunities for enhanced predictive accuracy, particularly given the substantial volume of data at our disposal. The case company records information about temperature, density and recipe with an interval of 30 minutes from the Internet of Things sensors installed at several breweries. This rich, high-frequency dataset provides a valuable opportunity for implementing advanced machine learning algorithms, capable of analyzing large volumes of data more efficiently and accurately than traditional methodologies.

The available data consists of information about fermentation processes in batches from several different breweries. Even though our case study consists of one brewery in particular, the model development approach uses data from several breweries. Machine learning models are heavily influenced by the amount of data available, and by utilizing the large amount of data available, the models will perform better. A machine learning model that learns patterns from all the available batches will still be able to perform forecasting on a desired single brewery. Furthermore, the scope of our thesis is to investigate if it is possible to predict when a batch is finished fermenting, and that applies to all breweries and not a single one.

Keeping our research objective in mind, we selected a few machine learning models that had demonstrated great performance in previous studies. To maintain consistency and facilitate accurate comparison, we cleaned the training data identically for all models, thereby ensuring that they operated on the same input data. The models underwent hyperparameter tuning on a validation set, which consisted of 10% of the training data. Subsequently, the models were trained on a separate dataset that represented 80% of all the available data. Evaluation of the models was carried out on a test set, constituting the remaining 20% of the data.

Model performance was evaluated based on their ability to predict the test data. In particular, we used the root mean squared error (RMSE) and the coefficient of determination (R2) as metrics. RMSE quantifies the difference between predicted and observed values, thus serving as a measure of prediction error. A lower RMSE indicates a better fit of the model (Bishop and Nasrabadi 2006). RMSE is also used as a metric because the error will be given in the same unit as the target value, meaning it is easy to interpret the results (Russell 2010). R2 represents the proportion of variance in the target variable that can be predicted from the training features. Higher R2 values, nearing 1, denote that a larger portion of variance is accounted for by the model (Tayo 2022). By using these two metrics in combination, we can assess both the accuracy and the goodness of fit of our models, providing a more comprehensive evaluation of their performance.

To systematically document the impacts of different data and hyperparameter modifications on the predictions, we maintained a log in an external spreadsheet. This practice was beneficial for multiple reasons. Firstly, it created a transparent record of the model development process, enabling us to trace and justify each decision and action. Additionally, it facilitated a clear overview of the performance of different models under varying conditions. In this way, the log served as an important tool for effective model development and evaluation.

2.3 Case Study

The case study methodology was chosen due to its suitability for in-depth, context-specific exploration of complex issues (Voss et al. 2002). Specifically, a single case study was chosen because it allows for a thorough, in-depth analysis of the case (Voss et al. 2002). However, in order to successfully conduct a case study for this thesis, it was necessary to focus on a company willing to dedicate a significant amount of time to participate in interviews. Consequently in order to not be dependant on several case companies for the progress of this thesis, we decided to proceed with a single case study.

The case study focused on Lock 27 Brewing Company, a craft brewery from Ohio, US. Lock 27 Brewing Company is a certified member of The Brewers Association and have been operating for 11 years. They also have a close relationship to the company that provided data for the model development in this thesis. As a result, we consider the data retrieved from this case study to be of high quality. Furthermore, our research in Heum and Hjort 2022 identified that the degree in which craft breweries have implemented production planning and control activities varies. Not every craft brewery have designated roles for production planning and some craft breweries do not have a structured approach where activities are documented. Lock 27 Brewing Company however, have implmented production planning and control activities which are the responsibility of their operations manager. Consequently, Lock 27 Brewing COmpany was considered a good fit for this research.

Both qualitative and quantitative data was collected in the case study. Several unstructured interviews were conducted with the production manager from the case company in order to collect qualitative data. The information collected was then sent to the case company for control and confirmation. To further ensure its viability each interview was also recorded and transcribed.

Interviews as an approach for data collection was chosen in order to achieve sufficient depth of observation. Production planning and control is also a subject area where specific subject terms can occur, so in order to ensure understanding between us and the case company a need for interviews were considered necessary. The alternative of sending out surveys might open up misunderstandings and remove the possibility of follow up questions. Consequently, it could have reduced the quality and accuracy of the data collected.

Our flexible interview guideline, which can be found in the appendix, wasn't strictly followed, but it naturally guided our discussions, covering most prepared questions. Initially, the interviews touched on the interviewee's role and the brewery's general overview, before delving deeper into their production planning and control activities, including the responsibilities and challenges involved. We then investigated specific challenges and opportunities within their specific brewery and the industry. Towards the end, the conversation focused on available data, the kind of insights that could improve their production planning, and particularly, their perspective on the benefits of enhanced predictability in the fermentation process. Additional quantitative data, such as production plans and schedules, were provided via email between interview sessions.

Our data analysis was primarily guided by thematic analysis, a method that allowed us to identify, analyze, and report patterns within the data. We started by thoroughly reviewing the interview transcripts and notes to get a comprehensive understanding of the content and to begin spotting potential patterns. Next, we marked sections of the data that seemed important to our research question. These marked sections helped us identify key aspects of the data. We then looked for patterns among these marked sections, grouping related ones together.

After identifying these potential patterns, we checked them to ensure they accurately represented the data. This involved adjusting the patterns, splitting or merging them as necessary, and discarding any that lacked sufficient support. We then finalized and named the patterns, making sure they accurately represented the key aspects of the data.

Furthermore, suggestions were provided with the goal of mitigating identified challenges from the analysis and exploiting the identified potential for improvement with regard to production planning and control at Lock 27 Brewing Company. Lastly, the potential value of these suggestions was explored and presented.

Chapter 3

Theoretical background

This chapter presents the theoretical foundation needed in order to follow our key takeaways in the study. The chapter starts by introducing production planning and its most important operations and concepts. Further, concepts of machine learning and approaches are presented. Finally, a section regarding existing approaches is identified and presented.

3.1 Production

The production industry has played a critical role in the development of human civilization (Mokyr 1992). From the earliest forms of tool-making to the advanced manufacturing technologies of today, this industry has been responsible for shaping the world we live in (Basalla 1988). Production is the process of converting raw materials into finished goods, and it involves a wide range of activities, from design and engineering to fabrication and assembly (Groover 2020). The industry encompasses a diverse range of sectors, from automotive and aerospace to food, beverages and pharmaceuticals (Pil and Holweg 2004), and it is a significant contributor to economic growth and development (Szirmai 2012).

According to research, the production industry is undergoing a significant transformation driven by digital technologies, such as the Internet of Things (IoT), robotics, artificial intelligence (AI), and cloud computing (Lu 2017). A paper by Groover (2020) stated that this shift towards smart manufacturing is leading to new levels of efficiency, quality, and customization. The adoption of these technologies is enabling manufacturers to optimize their processes, reduce costs, and improve product quality (Groover 2020).

However, the production industry is also facing several challenges, including the increasing complexity of products, rising customer expectations, and the need to reduce environmental impact (Kiel et al. 2017). A study by Bjørnbet et al. (2021) highlighted the importance of sustainability in the production industry and suggested that companies need to adopt circular economy principles to minimize waste and maximize resource efficiency. Furthermore, the paper suggested that digital technologies can be used to enable closed-loop production systems, which can reduce the environmental impact of production processes (Bjørnbet et al. 2021).

3.1.1 Production characteristics

The characteristics of a production environment will be strongly linked to the characteristics of the supply chain in which the production environment is a part of (Romsdal 2014). The supply chain characteristics include product, market and production system characteristics (Romsdal 2014). These characteristics describe factors such as the product's perishability, complexity and variability, but also lead time expectations, demand uncertainty and inventory management (Romsdal 2014). The following section highlights the characteristics of the production environment.

Customer order decoupling point (CODP)

The customer order decoupling point is a point in the production process where the production is no longer driven by customer demand, but rather by a forecast or a planned production schedule (J. Strandhagen 2021). It is the point where the production process is decoupled from the variability in customer demand, allowing for more efficient production planning and control (J. Strandhagen 2021). This point can vary depending on the industry, product characteristics, and supply chain configuration (J. Strandhagen 2021).

Different production management strategies can be described by the customer order decoupling point (J. Strandhagen 2021). The level of customization and customer influence on the production can be established by the placement of the customer order decoupling point, as seen in figure 3.1 (J. Strandhagen 2021). The alternatives for production management strategies, or production environmental characteristics, are separated into four main strategies (J. Strandhagen 2021).

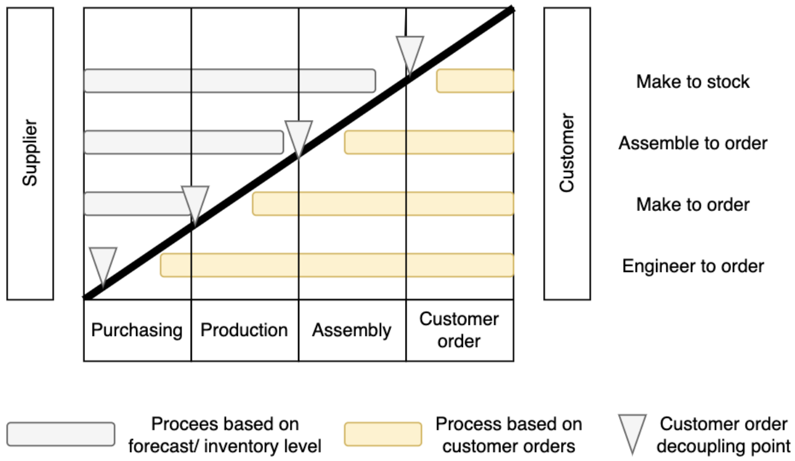


Figure 3.1: Customer Order Decoupling Points (J. Strandhagen 2021)

Make to stock

Make-to-stock (MTS) involves forecasting product demand to determine production and products from past production cycles are used to meet future orders (Chapman 2006; Soman et al. 2004). MTS, a push-type operation, begins with anticipated demand and pushes raw materials through the production process (Karrer 2012). It is common in areas like food production and retail where products are fully finished and stocked before customer orders are received (Chapman 2006).

Assemble to order

Assemble-to-order (ATO) production involves quick and partially customizable manufacturing in response to customer orders (Chapman 2006). Basic parts are pre-made and assembled only when needed (Song and Zipkin 2003). This approach allows customers to choose from various options for final assembly (Chapman 2006). The computer industry is a common example (Chapman 2006).

Make to order

Make-to-order (MTO) production maintains no finished goods inventory due to varied customer demands and production processes initializes only after specific orders are received (Chapman 2006). This approach allows customization of the final product, provided standard materials are used (Chapman 2006).

Engineer to order

Engineer-to-order (ETO) refers to products engineered and manufactured after the order is received (William J Stevenson et al. 2014). It allows for, and supports, customer specifications, regardless ff standard material is used or not (Chapman 2006). ATO typically involves unique, complex machinery or large-scale projects and ETO often includes designing new products with the customer (Rahim and Baksh 2003).

Process Categories

Process selection impacts facilities arrangement, equipment choices, capacity planning and design of work system (William J Stevenson et al. 2014). Each process type has distinct capacity ranges, making process changes challenging, costly, and time-consuming (William J Stevenson et al. 2014). There are five fundamental process categories, although practical applications often involve hybrid types (William J Stevenson et al. 2014).

Project A project process is employed for tasks that are not routine, possessing a distinct set of goals to be achieved within a specified time period (William J Stevenson et al. 2014). The projects are typically large in scope and will often be managed by teams of individuals with a very particular set of skills, brought together to work on a specific one-time activity (Chapman 2006). Large infrastructure projects like building a bridge is an example of a project process (William J Stevenson et al. 2014).

Job process Job processes operate on a smaller scale and are ideal when demand is low, but with a wide range of variety (William J Stevenson et al. 2014). The workflow is discontinuous and comprises numerous unique tasks, requiring flexible equipment and skilled workers (William J Stevenson et al. 2014). Job processes are often used in Engineer-to-Order (ETO) or Make-to-Order (MTO) environments (Chapman 2006).

Batch production Batch production suits medium volume production and it can handle a moderate diversity in product or service variety (William J Stevenson et al. 2014). Specialized equipment can perform a significant part of the production and there is less need for skilled a workforce compared to job shops (Chapman 2006). In batch production, identical products are produced simultaneously in customized batches which is common in the food and beverage industry (William J Stevenson et al. 2014, Chapman 2006).

Repetitive manufacturing Repetitive processes is suitable when there is a need for a high-volume, limited-variety demand and costly, specialized equipment (William J Stevenson et al. 2014). Such process reduces labor needs, spreading the equipment investment over a large production volume, resulting in a lower cost per item (Chapman 2006). Repetitive processes are often used for Make-to-Stock (MTS) designs and examples include production and assembly lines for items like pencils, computers, and televisions (William J Stevenson et al. 2014, Chapman 2006).

Continuous production Continuous production typically serves very high-volumes and highly uniform, non-discrete output demands (William J Stevenson et al. 2014). Given the uniform output, equipment flexibility is minimal, and labor skills become less significant (William J Stevenson et al. 2014). This is common in high-volume chemical processes and petroleum refining, where the process maintains constant and uninterrupted (William J Stevenson et al. 2014).

3.2 Production Planning

3.2.1 Introduction

Production planning involves determining the most efficient and effective way to utilize an organization's resources to meet customer demand through production (Chapman 2006). It encompasses various activities, such as forecasting demand, coordinating production schedules, managing inventory levels, and optimizing resource and capacity utilization to ensure that products or services are produced in a timely and cost-effective manner (Kiran 2019). Production planning plays a vital role in balancing supply and demand, minimizing inventory holding costs, optimizing production processes, and achieving customer satisfaction (Kiran 2019).

An effective production planning process requires close coordination among different departments within an organization, such as sales, marketing, operations, and supply chain management (Chopra and Meindl 2001). It involves strategic decision-making, utilizing tools and techniques such as sales and operations planning (S&OP), master scheduling, material requirements planning (MRP), and various planning methods (Chopra and Meindl 2001). Production planning is critical for organizations to meet customer demand efficiently, minimize costs, optimize resource utilization, and maintain a competitive edge in today's dynamic business environment (Kiran 2019).

3.2.2 Sales and Operations Planning

Sales and Operations Planning (S&OP) is a comprehensive approach that merges all business plans, including those for sales, marketing, development, manufacturing, sourcing, and finance into a unified and cohesive set of plans (Thomé et al. 2012). It is conducted over the long term and routinely undergoes aggregate level reviews by management (Thomé et al. 2012). By ensuring the reconciliation of supply, demand, and new product plans at both detailed and aggregate levels, S&OP aligns with the overall business plan, serving as the definitive statement of a company's near to intermediate-term operational strategy (Thomé et al. 2012).

Commonly referred to as aggregate planning, S&OP shifts its focus towards product groups rather than individual products (Cox and Blackstone 2002; Thomé et al. 2012). This process provides a horizon that facilitates resource planning and supports the annual business planning process (Thomé et al. 2012). When executed effectively, S&OP establishes a direct link between the strategic plans of the business and their implementation, allowing for a review of performance measures and promoting continuous improvement (Thomé et al. 2012). Consequently, S&OP emerges as a crucial component in the efficient operation and success of any business, enabling a well-coordinated and strategic approach to operational planning (Thomé et al. 2012).

3.2.3 Master Production Schedule

The master schedule turns the S&OP into specific final products. It outlines the amounts and timings for each product, rather than just giving a general view of product groups (Jacobs et al. 2011). This is why the S&OP is a key input for the master schedule (Chapman 2006). It helps in adding more details to the production processes (Chapman 2006).

In each step of creating the master schedule, an updated forecast is needed (Chapman 2006). This helps in making a definite plan for the final product (Chapman 2006). A crucial part of this plan is the bill of materials (BOM) (Chapman 2006). The BOM lists all the parts needed to make the product, including their amounts and how they connect to each other (Chapman 2006). To make sure the whole production plan is covered, the schedule's timeframe should be as long as or longer than the total time it takes to make the product (Chapman 2006).

3.2.4 Material Requirements Planning and Enterprise Resource Planning

Material Requirements Planning (MRP) is a strategic tool used in the assembly process of production, providing critical guidance throughout the manufacturing cycle (William J Stevenson et al. 2014). Leveraging data from the master schedule, including the quantity and planned completion date of final products, MRP formulates a detailed production plan (William J Stevenson et al. 2014). This plan outlines the necessary quantities and timings for subassemblies, components, and raw materials, ensuring a seamless production process (William J Stevenson et al. 2014). By integrating lead times, MRP establishes the ideal timeline for procurement, production, and assembly (William J Stevenson et al. 2014). Further, it provides an intricate insight into the product's components, right down to the smallest level of assembly (William J Stevenson et al. 2014). Therefore, MRP becomes an invaluable tool for strategizing various facets of a company's operations, such as marketing, finance, and engineering (William J Stevenson et al. 2014).

Enterprise Resource Planning (ERP), on the other hand, is a system designed to enhance coordination within a company (William J Stevenson et al. 2014). Comprised of various business applications or modules, ERP serves to interconnect different units within an organization, promoting efficient and logical operations, as well as facilitating a common platform for information sharing (Beheshti 2006). ERP systems can be tailored and employed in a multitude of ways, such as for production, ordering, procurement, distribution, design, finance, and human resources, based on the specific needs and requirements of the organization (William J Stevenson et al. 2014).

3.2.5 Capacity Planning

Capacity planning is a process aimed at equating a system's present capacity with the necessary capacity to effectively manage customer orders (Chapman 2006). This process requires a planner to quantify both the demand in terms of customer orders and the available capacity (Chapman 2006). Upon determining the gap, adjustments are required either by modifying the available capacity or altering the demand (Chapman 2006). However, businesses often favor adjusting the available capacity over modifying demand to maintain service levels and accommodate customer requirements (Slack, Brandon-Jones, and Johnston 2013, Chapman 2006).

Operating below maximum capacity is not unusual for many companies, and this can be attributed to factors such as fluctuating demand, insufficient demand, or a strategic decision to preserve flexibility for accommodating new orders (Slack, Brandon-Jones, and Johnston 2013). It is quite common to have parts of the production process operating at full capacity while others remain underutilized, thereby establishing a capacity constraint for the entire production system (Slack, Brandon-Jones, and Johnston 2013). In such circumstances, the concept of *spare capacity* can be adopted, which essentially designates a volume of capacity beyond the projected demand (Slack, Brandon-Jones, and Johnston 2013). This is particularly utilized in scenarios marked by uncertainties concerning demand forecasts (William J Stevenson et al. 2014).

A holistic overview of the entire production process is crucial when strategizing capacity planning (William J Stevenson et al. 2014). Otherwise, it may result in an imbalanced system, potentially causing bottlenecks, which are situations where a single process or segment of production possesses have lower capacity than the rest (William J Stevenson et al. 2014). This imbalance could negatively impact the entire operation and throughput, leading to extended lead times, transportation issues, processing delays, and queuing (William J Stevenson et al. 2014, L. Li et al. 2007). The root causes of bottlenecks can vary significantly, including factors such as insufficient demand, limited financial resources, unreliable supply chains, and regulatory constraints (L. Li et al. 2007).

3.2.6 Production Activity Control

Production Activity Control (PAC) functions as an execution arm of the master production schedule and the material requirements plan intro operational actions (Arnold 2020). It manages the detailed planning of order flows throughout the manufacturing process, the execution of these plans, and the monitoring of work as it progresses toward completion (Arnold 2020). Furthermore, PAC coordinates the utilization of labor and machinery, the management of work-in-process inventory, and the maintenance of customer service (Arnold 2020).

The activities within PAC can be divided into three primary functions: planning, implementation, and control (Arnold 2020). The planning function focuses on arranging the flow of work to meet predetermined delivery timelines, ensuring that necessary materials, personnel, and information are accessible when required (Arnold 2020). Following the planning stage, PAC transitions into the implementation phase, predominantly through issuing work orders, a procedure known as dispatching (Arnold 2020). The control function then oversees the performance of these work orders, comparing actual results to planned schedules, and begins necessary measures to correct if discrepancies are identified (Arnold 2020).

3.2.7 Inventory Management

Inventory Management involves the planning and control of inventory from raw materials to finished goods (Arnold 2020). It includes various planning levels: production planning, which pertains to the overall inventory; master scheduling, which focuses on end items; and material requirements planning, which deals with component parts and raw materials (Arnold 2020). The inventory is categorized based on its stage in the production environment, namely raw materials, work-in-process (WIP), finished goods, distribution inventories, and maintenance, repair, and operational supplies (MROs) (Arnold 2020).

In assessing the efficiency of inventory management, the inventory turnover rate is often calculated and considered critical in the manufacturing industry (Kwak 2019). This metric is a ratio indicating the number of times an organization has sold and replaced its inventory within a specified period (Kwak 2019). The inventory turnover rate serves as an indicator of operational efficiency, liquidity of inventory items, and the accuracy of inventory levels in relation to sales volumes (Kwak 2019). A high turnover rate may suggest strong sales, effective inventory management practices, and minimal obsolescence risk (Kwak 2019). On the other side, a low turnover rate could imply overstocking, underperforming sales activity, or high holding costs (Kwak 2019). What the actual ratio should be in order to be considered good can vary between industries, but an inventory turnover ratio between 5-10 is considered good for most industries (Jenkins 2022, Kwak 2019). However, a higher inventory turnover ratio is usually desired when working with perishable goods, such as food and beverages are, to avoid spoilage and inventory losses (Jenkins 2022, Kwak 2019).

Inventory management systems can be categorized broadly into two types: manual and digital (Atieh et al. 2016). Manual systems often involve the use of spreadsheets or paper-based tracking, and while they have been traditionally employed, they are prone to human errors and lack the capability to provide real-time visibility into inventory levels (Wu et al. 2020). These challenges could potentially lead to inaccuracies in the inventory records, which might, in turn, influence production planning and customer service levels (Chapman 2006). On the other hand, digital inventory management systems utilize information and communication technology (ICT) to automate and streamline the inventory management process (Mashayekhy et al. 2022). These digital systems offer improved accuracy, real-time tracking, and robust data analysis capabilities, which are beneficial in mitigating the limitations associated with manual inventory management systems (Mashayekhy et al. 2022).

3.3 Production Control

Production control is an essential aspect of managing company operations, encompassing both planning and control activities (Slack and Brandon-Jones 2018). While the precise distinction between these terms may not always be clear, planning involves formalizing intentions for the future, while control involves responding to unexpected changes and ensuring the operation achieves its objectives, even when plans don't unfold as expected (Slack and Brandon-Jones 2018). Furthermore, production control will aim to ensure optimal use of resources, quality management and cost efficiency (Slack and Brandon-Jones 2018). Within these main goals, activities such as the regulation of inventory management and organization of production schedules are aspects of production control (Kağnıcıoğlu et al. 2019).

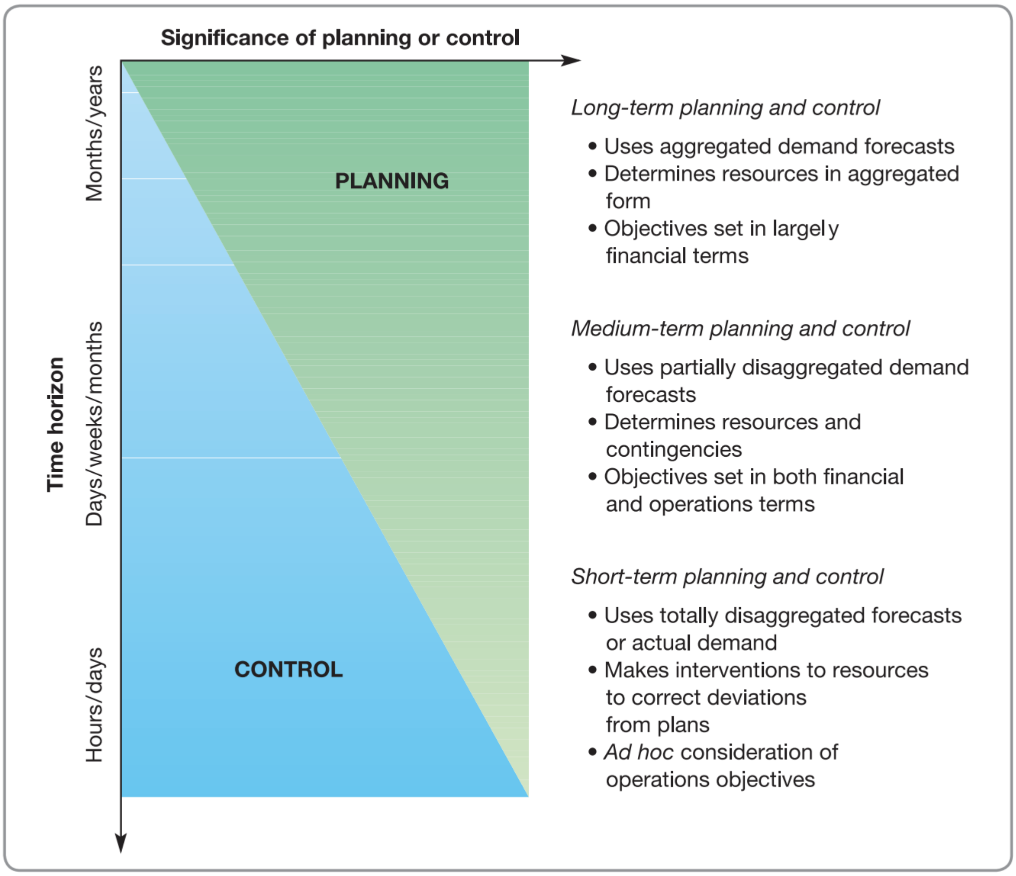


Figure 3.2: Relationship between planning and control (Slack and Brandon-Jones 2018).

In the long term time frame, the primary focus of operations managers is on strategic planning, goal setting, resource identification, and action plans, as product control is not a relevant factor during this stage (Slack and Brandon-Jones 2018). The medium-term perspective makes a balance between planning and control necessary, both of which are integral components of PPC (W. Stevenson 2018). In this phase, operations managers take a forward-looking approach to evaluate total demand and plan accordingly, while also paying attention to control activities (William J Stevenson et al. 2014). However, in the short term, resources have already been allocated, thereby complicating the process of introducing changes (Romsdal 2014). Consequently, PPC tasks during the short term mainly focus on operational-level control (Romsdal 2014).

The aim of the planning and control function is to ensure efficient and effective operations by aligning supply and demand in relation to time, quantity, and quality (Heizer et al. 2016). The PPC system is composed of multiple functions such as material requirements planning (MRP), demand management, capacity planning, and job scheduling and sequencing (Chapman 2006). These functions embody both planning and control principles, given their interconnected nature and the fact that most PPC applications are relevant to both aspects (Slack and Brandon-Jones 2018).

3.4 Supply Chain Management

Supply chain management (SCM) is recognized as a central element in the functioning of modern operations management and in the planning of strategic business development (Christopher 2016). Its importance has been increasingly recognized in both academia and industry as it plays a critical role in enhancing competitiveness and improving organizational performance (Chopra and Meindl 2001; Simchi-Levi et al. 2000).

The concept of supply chain management extends beyond the traditional view of managing operations within a business or organization, also called intra-organizational management (Mentzer et al. 2001). It encapsulates a broader perspective, involving the coordination of all activities related to the flow and transformation of goods from the raw material stage to the end customer (Mentzer et al. 2001). As stated by Mentzer et al. (2001), *Supply chain management is the systemic, strategic coordination of the traditional business functions and the tactics across these business functions within a particular company and across businesses within the supply chain, for the purposes of improving the long-term performance of the individual companies and the supply chain as a whole.*

Supply chain management includes a range of activities including procurement, production, inventory management, logistics, and customer service, among others (Chopra and Meindl 2001; Christopher 2016). The aim is to create seamless and efficient processes across all these areas, which leads to cost reductions, improved service levels, and overall value creation for the organization and its stakeholders (Christopher 2016).

Furthermore, supply chain management can be influenced by various external factors such as market conditions, customer demands, technological innovations, and regulatory requirements (Simchi-Levi et al. 2000). Understanding these factors and how they work with the internal operations of a company is crucial for managing the supply chain effectively (Simchi-Levi et al. 2000). Hence, this section will further explore aspects of Supply Chain Management and its effects on operational performance.

3.4.1 Supply Chain Planning

Supply chain planning (SCP) is an important process within supply chain management that involves the systematic and strategic coordination of resources and operations to meet demand (Chopra and Meindl 2001). SCP includes a range of different activities, such as demand forecasting, inventory management, production planning and aggregate planning (Chopra and Meindl 2001).

Supply Chain Inventory Management

Supply chain inventory management (SCIM) is the systematic approach of ordering, storing, and using a company's inventory at different stages of the supply chain, spanning from raw materials, work-in-process inventory, to finished goods (Chopra and Meindl 2001). SCIM is an important part of supply chain planning, with the main goal of maintaining an optimal balance of inventory (Chopra and Meindl 2001). Doing so can help with avoiding overstocking costs and potential product expiry, while also preventing stock outs that could disrupt the supply chain and lead to potential losses in sales and decreased customer service levels (Chopra and Meindl 2001).

One of the key concepts in SCIM is multi-echelon inventory management (Chopra and Meindl 2001). This approach accounts for the interconnected and hierarchical nature of modern supply chains that consist of multiple levels, or "echelons" (Chopra and Meindl 2001). These echelons range from suppliers, manufacturers, and distributors, to retailers (Chopra and Meindl 2001). Multi-echelon inventory theory focuses on determining the optimal inventory levels at each stage of the supply chain, with the objective of minimizing overall costs while maintaining sufficient service levels (Chopra and Meindl 2001). This theory helps organizations and companies make informed decisions about how much stock should be held at different points in the supply chain to ensure cost-effectiveness and efficiency (Clark and Scarf 1960).

Another important aspect of SCIM is the inventory location strategy (Chopra and Meindl 2001), which focuses on the physical placement of inventory within the supply chain (Chopra and Meindl 2001). The location of inventory can significantly affect several elements, including transportation costs, handling costs, and the risk of stockouts (Chopra and Meindl 2001). Decisions regarding inventory location should consider factors such as lead times, transportation costs, demand variability, and service level requirements (Melkote and Daskin 2001). Selecting the right inventory location strategy could mean the difference between a lean, cost-effective supply chain and one that experiences delays, excess costs, and lower service levels (Chopra and Meindl 2001; Daskin et al. 2002).

Aggregate Production Planning (APP)

Aggregate production planning (APP) is a critical component of SCP that involves creating a production schedule over an intermediate time horizon, typically between 3 to 18 months (Chopra and Meindl 2001). This strategy aims to balance demand and supply at an aggregate level, thus ensuring that resources are utilized efficiently while meeting customer demand (Chopra and Meindl 2001). The process involves determining the ideal levels of capacity, production, subcontracting, inventory, stockouts and pricing (Chopra and Meindl 2001). APP can thus be considered important when deciding on topics such as making new and large investments (Chopra and Meindl 2001).

Several strategies can be employed in APP, each with its own set of advantages and potential drawbacks (Chopra and Meindl 2001). These strategies include:

1. **Chase Strategy** The chase strategy aims to match demand precisely by adjusting production rates (Simchi-Levi et al. 2000). This could mean hiring or laying off workers, varying production hours or incorporating overtime to meet fluctuations in demand (Chopra and Meindl 2001). While this strategy can prevent overproduction and excess inventory, it may lead to increased costs associated with hiring, training, and laying off workers and can cause disruption in the workforce (Chopra and Meindl 2001).
2. **Level Strategy:** The level strategy aims to maintain a steady rate of output while using inventories or backlog to absorb fluctuations in demand (Simchi-Levi et al. 2000). This approach can lead to more stable production rates and workforce levels, but it also risks excessive inventory during periods of low demand or stockouts during periods of high demand (Chopra and Meindl 2001).
3. **Flexible Strategy:** The flexible strategy uses a combination of worker levels, work hours, and inventory levels to respond to demand changes (Simchi-Levi et al. 2000). This strategy provides a balance between the chase and level strategies, allowing the company to adjust to demand fluctuations while maintaining some level of stability in production and workforce levels (Chopra and Meindl 2001).
4. **Mixed Strategy:** The mixed strategy employs a combination of the above strategies to balance costs and service levels (William J Stevenson et al. 2014). For example, a company might use a level strategy during periods of stable demand and switch to a chase or flexible strategy during periods of high demand variability (Chopra and Meindl 2001). This approach requires careful planning and management but can provide a balance between cost control, customer service levels, and workforce stability (Chopra and Meindl 2001).

These strategies illustrate the complexities involved in aggregate production planning (Chopra and Meindl 2001). The choice of strategy depends on a company's specific context, including its production capabilities, workforce flexibility, customer demand patterns, and strategic objectives (Chopra and Meindl 2001).

3.4.2 Coordination and Information Sharing

Information sharing and information technology

In the context of supply chain management, coordination refers to the alignment and synchronization of activities, policies, and decisions within an organization and across different actors in the supply chain (Simatupang and Sridharan 2002). Proper coordination is essential for enhancing supply chain performance, and one of the critical enablers of effective coordination is information sharing (Chopra and Meindl 2001).

Information sharing in SCM involves the exchange of data and insights related to the demand, production, inventory, and other crucial aspects across the supply chain (Chopra and Meindl 2001). This information transparency leads to a more comprehensive understanding of the supply chain system as a whole, enhancing decision-making and operational efficiency (S. Li and B. Lin 2006).

Information technology (IT) has emerged as an indispensable tool in facilitating information sharing and coordination in supply chains (Chopra and Meindl 2001). IT systems and applications enable the capture, storage, processing, and dissemination of information in a timely and cost-efficient manner (Chopra and Meindl 2001). They improve the visibility of supply chain activities, thereby enabling all stakeholders to monitor operations and make informed decisions (Chopra and Meindl 2001).

Several IT applications such as enterprise resource planning (ERP) systems, cloud computing, and software as a service (SaaS) have revolutionized supply chain processes (Bharadwaj et al. 2013). ERP systems provide an integrated platform that enhances internal coordination and information sharing within an organization (Chopra and Meindl 2001). On the other hand, cloud computing and SaaS solutions enable inter-organizational coordination and information sharing by providing a shared digital space accessible by all supply chain participants (Chopra and Meindl 2001).

Advancements in mobile technology and social media have also contributed significantly to improving supply chain coordination (Chae 2015). Mobile technologies provide real-time access to critical supply chain information and foster quick decision-making, while social media platforms facilitate communication and collaboration among supply chain participants (Chopra and Meindl 2001).

However, the utilization of IT in SCM is not without its risks, such as technical failures and security issues (Chopra and Meindl 2001). These risks necessitate prudent strategies for IT implementation and operation, such as incremental implementation, parallel running of old and new systems, and stringent data security measures (Chopra and Meindl 2001).

To summarize, coordination, information sharing, and IT are closely interlinked elements that contribute to the effectiveness and efficiency of supply chain operations (Chopra and Meindl 2001). As the business environment continues to evolve, organizations need to continuously adapt their coordination mechanisms (Seifert 2003), information sharing strategies (Raghunathan 2001), and IT systems to maintain a competitive edge in their supply chains (Manyika et al. 2015).

The Role of Industry 4.0 in Supply Chain Management

Industry 4.0 is an emerging paradigm that introduces disruptive technologies such as the Internet of Things (IoT), machine learning, and big data analytics, all of which have significant potential for transforming SCM (Schwab 2017).

A significant role of Industry 4.0 in SCM is the improvement of manufacturing logistics (J. Strandhagen 2021). This revolves around the application of real-time monitoring and control systems, which are more compatible with repetitive production environments characterized by higher automation and standardized processes (J. W. Strandhagen et al. 2017). Auto-ID technologies, for instance, are key enablers for real-time monitoring and control, although their implementation might be challenging in a high product variety environment (J. W. Strandhagen et al. 2017).

Consequently, the complexity of material flows makes it beneficial with tracking and tracing of products, making Industry 4.0 a valuable asset in such scenarios (J. W. Strandhagen et al. 2017).

However, the application of Industry 4.0 technologies is not a "one-size-fits-all" solution (J. W. Strandhagen et al. 2017). The specific characteristics of the production environment play a significant role in determining the applicability of these technologies (J. W. Strandhagen et al. 2017). Companies should conduct an in-depth analysis of their production environment before applying industry 4.0 technologies to their supply chain management practices and activities (J. W. Strandhagen et al. 2017).

Industry 4.0 also has the potential to optimize inventory and demand management (Kiel et al. 2017). Real-time data collection and analysis can enable companies to better forecast demand and adjust their production schedules accordingly, thus reducing inventory costs and improving customer service (Kiel et al. 2017).

Furthermore, the traceability and transparency of the supply chain can be greatly enhanced by technologies associated with Industry 4.0 (Shiyong Wang et al. 2016). IoT devices and sensors enable real-time tracking and monitoring of components in the supply chain, providing valuable information for decision-making processes and risk management (Shiyong Wang et al. 2016).

Overall, Industry 4.0 technologies have the potential to significantly improve the efficiency, responsiveness, and adaptability of supply chain management processes. However, their application requires careful consideration of the unique characteristics of each company's production environment (J. W. Strandhagen et al. 2017).

3.4.3 Distribution and Sourcing

Distribution

In supply chain management, distribution involves the processes used to move and store a product from the supplier stage to the customer stage (Chopra and Meindl 2001). This occurs between each pair of stages in the supply chain, influencing both supply chain costs and customer experience (Chopra and Meindl 2001).

The distribution network design is a crucial aspect of supply chain distribution (Simchi-Levi et al. 2000). It involves determining the most efficient and cost-effective way to get products from the point of origin to the point of consumption (Chopra and Meindl 2001). This includes decisions about the number and location of distribution centers, the choice of transportation modes, and the routes to be used (Simchi-Levi et al. 2000).

Transportation is another key element of supply chain distribution. It involves the movement of goods from one location to another, typically from a manufacturing or storage facility to a distribution center, and then to retail outlets or directly to customers (Chopra and Meindl 2001). The choice of transportation mode (e.g., truck, rail, air, sea) depends on various factors, including the nature of the goods, the distance to be covered, cost considerations, and service level requirements (Chopra and Meindl 2001).

The importance of effective distribution can be seen in companies like Wal-Mart and Seven-Eleven Japan, both of which have built their business models around superior distribution design and operations (Chopra and Meindl 2001). Wal-Mart uses distribution to offer high availability of common products at low costs, while Seven-Eleven Japan leverages effective distribution to achieve high customer responsiveness at a reasonable cost (Chopra and Meindl 2001).

The design of the distribution network can significantly impact customer service levels and costs (Chopra and Meindl 2001). Inadequate networks can harm customer service and inflate costs, leading to a negative impact on firm profitability (Chopra and Meindl 2001). Therefore, choosing an appropriate distribution network is key to meeting customer needs at the lowest possible cost (Chopra and Meindl 2001).

Sourcing

Sourcing is the term that describes the entire set of processes that a business goes through when acquiring goods, such as raw materials, or services (Chopra and Meindl 2001). It involves the process of identifying, evaluating, and contracting with suppliers of goods and services (Chopra and Meindl 2001). The goal of supply chain sourcing is to ensure that the organization has the necessary resources to meet its operational requirements and strategic objectives (Chopra and Meindl 2001).

Supplier selection is a critical component of supply chain sourcing (Chopra and Meindl 2001). It involves assessing potential suppliers based on various criteria, such as price, quality, delivery performance, and capacity (Chopra and Meindl 2001). The use of decision-making tools, such as data envelopment analysis (DEA), can help in this process by providing a systematic approach for evaluating the relative efficiencies of potential suppliers (Vörösmarty and Dobos 2020).

Strategic sourcing goes beyond transactional procurement to include a more comprehensive approach that considers the total cost of ownership, long-term supplier relationships, and strategic alignment with the organization's goals (Chopra and Meindl 2001). It involves a continuous process of analyzing the organization's spend and supply market, developing a sourcing strategy, executing the sourcing strategy through a structured supplier selection process, and managing the supplier relationships to achieve the desired outcomes (Chopra and Meindl 2001).

3.5 Digitalization

Digitalization has become a buzzword in the business world, referring to the transformation of traditional processes and systems from manual into digital ones (Aggarwal and Aggarwal 2017). With the arrival of new technologies, such as the Tnternet of Things (IoT), big data analytics, and artificial intelligence (AI), digitalization is transforming the way businesses operate, including those in the manufacturing and production industry (Aggarwal and Aggarwal 2017; Albach 2015; R. Y. Zhong et al. 2017).

The manufacturing and production industry has traditionally been reliant on manual labor and physical processes (Brynjolfsson and Hitt 2000). However, the rise of digital technologies has enabled manufacturers to automate and streamline their operations, resulting in increased efficiency, reduced costs, and improved quality (Aggarwal and Aggarwal 2017; J. Lee, Kao, et al. 2014; R. Y. Zhong et al. 2017. For example, sensors and IoT devices can monitor machinery and equipment in real-time, allowing for predictive maintenance and reduced downtime (Aggarwal and Aggarwal 2017. Additionally, data analytics can help manufacturers identify inefficiencies in their processes and optimize them for better performance (Aggarwal and Aggarwal 2017).

Moreover, digitalization can also enable manufacturers to enhance their product offerings, such as by adding new features or customizing products to meet specific customer demands (Haleem and Javaid 2019). Digitalization can aid this in several ways, for example with big data analytics by gaining insight into customer preferences and behavior which can guide the development of new products and features (Thompson et al. 2005). Another example is increased customization by the use of manufacturing technologies such as 3D printing, which makes it possible to tailor products to customer needs and preferences at a higher rate (Dabbagh et al. 2021; Lyu et al. 2021). This can help manufacturers differentiate themselves from competitors and increase customer satisfaction (Haleem and Javaid 2019).

Despite the potential benefits of digitalization, implementing digital technologies in the manufacturing and production industry can be challenging (Horváth and Szabó 2019). This is because many manufacturers have outdated legacy systems and may lack the necessary expertise to implement new digital technologies (Horváth and Szabó 2019). Additionally, there may be concerns around data security and privacy (Albach 2015; Horváth and Szabó 2019).

Overall, digitalization has the potential to revolutionize the manufacturing and production industry by improving efficiency, reducing costs, and enhancing product offerings (Aggarwal and Aggarwal

2017; R. Y. Zhong et al. 2017). However, it is important for manufacturers to carefully consider the challenges and risks associated with digitalization before initiating these transformations to their business or organization (Horváth and Szabó 2019).

3.5.1 Industry 4.0

Industry 4.0, also known as the fourth industrial revolution, is a term used to describe the advanced digitalization and automation of manufacturing enterprises through the use of digital technologies (Qin et al. 2016; Thoben et al. 2017). The term originated in 2011 from a German government program aimed at promoting competitiveness in the nation's manufacturing companies (J. Strandhagen 2021). Driven by the demand for shorter delivery times, more efficient and automated processes, higher quality and customized products, Industry 4.0 envisions a future of modular and efficient manufacturing systems, where products control their own manufacturing process (Qin et al. 2016).

Unlike Industry 3.0, which introduced computers into the manufacturing process, Industry 4.0 focuses on connecting those computers with each other and making the processes of manufacturing "smart", where machines can communicate and decisions can be made automatically (Thoben et al. 2017). The need for Industry 4.0 has been driven by the challenges of quickly changing customer preferences, demand uncertainty and disruptions, national security interests, trade barriers, and logistics disruptions that are pushing businesses to find alternatives to globalized supply chains (Qin et al. 2016).

Today, Industry 4.0 has evolved to rely on a combination of Cyber-Physical-Systems (CPS), Internet of Things (IoT), and a spectrum of technologies such as big data, cloud technology, artificial intelligence, and others (Zheng et al. 2021). These technologies facilitate the real-time exchange of data, merging the physical and virtual worlds, and enabling the mass production of highly customized products (Leitão et al. 2016; Stojmenovic and Wen 2014; Zheng et al. 2021; R. Y. Zhong et al. 2017).

3.5.2 Internet of Things

The Internet of Things (IoT) plays a crucial role in Industry 4.0, enabling interconnected devices to exchange data over the Internet or private networks, transforming factories into "smart" environments where machines communicate and make decisions autonomously (B. Chen et al. 2017; Khan and Javaid 2022; J. Lin et al. 2017). In particular, manufacturers have been able to significantly improve production processes through the real-time collection and analysis of data from IoT sensors (Rüßmann et al. 2015; H. Xu et al. 2018).

IoT technologies offer multiple benefits, such as effective monitoring of production processes and resources, autonomous control of machinery, and enhanced workflow adjustment in response to changes or breakdowns (Cohen et al. 2019; Qi and Tao 2018; Tao and M. Zhang 2017; H. Xu et al. 2018; R. Y. Zhong et al. 2017). Furthermore, IoT can improve product quality control and facilitate more effective production scheduling and resource allocation when coupled with machine learning algorithms (Cohen et al. 2019; Gonzalez Rodriguez et al. 2020; Y. Wang et al. 2017). Also, real-time product progression information can optimize production planning and capacity utilization (Gonzalez-Neira et al. 2017; Ran 2021; Seitz and Nyhuis 2015; William J Stevenson et al. 2014).

However, the implementation of IoT does not come without challenges (Mahmoud et al. 2015). As interconnected devices increase, so does their vulnerability to cyber attacks, necessitating robust security measures to protect sensitive data and maintain system integrity (Atzori et al. 2010; Mahmoud et al. 2015; Shah and Sengupta 2020). Furthermore, the lack of interoperability standards between IoT devices can hinder integration, limiting the system's efficiency and effectiveness (Albouq et al. 2022).

System scalability also becomes a critical challenge as the number of devices increases due to the growing data flow and processing demands (J. Lee, Bagheri, et al. 2015). Managing large volumes of data, which must be collected, analyzed, and effectively managed to extract valuable information, poses another challenge (Gubbi et al. 2013).

3.5.3 Big Data Analytics and Artificial Intelligence

Big data analytics (BDA) and artificial intelligence (AI), particularly machine learning and deep learning, are transforming the manufacturing industry (Dubey et al. 2020). Both technologies are integrated into these sectors for process optimization, cost reduction, quality improvement, and managing the vast amount of data generated (Dubey et al. 2020; Jan et al. 2022).

BDA, powered by AI and IoT, allows real-time monitoring and predictive maintenance, reducing downtime and enhancing productivity (Moyne and Iskandar 2017). Similarly, AI-driven predictive maintenance uses machine learning algorithms to predict required maintenance based on data patterns, leading to increased efficiency (Çınar et al. 2020).

In supply chain management, BDA improves operations by analyzing data from various stakeholders, reducing costs, and improving delivery times (I. Lee and Mangalaraj 2022). AI's role in autonomous equipment operation, such as robots and production lines, can increase precision and reduce labor costs, although it presents challenges such as data privacy, the need for skilled professionals, and ethical considerations (Azambuja et al. 2023; Peres et al. 2020; Soori et al. 2023a; W. Wang and Siau 2019).

Both AI and BDA enhance process optimization and energy efficiency. They identify production inefficiencies, facilitate informed decision-making, and lead to cost savings (Soori et al. 2023a; Sheng Wang et al. 2018). In customer interaction, understanding preferences through BDA allows for product personalization, improving customer satisfaction and competitive advantage (Anshari et al. 2019; Perera et al. 2018). However, human interaction with autonomous systems presents unique challenges such as loss of situational awareness or overconfidence in automated systems (Endsley 2017).

3.5.4 Smart PPC

Smart PPC is a digital manufacturing technology that uses data analytics and artificial intelligence to optimize the production process in real-time (Bueno et al. 2020). It aims to improve efficiency, reduce costs, and enhance overall productivity by creating a seamless flow of information throughout the manufacturing chain (Bueno et al. 2020).

Industry 4.0, on the other hand, is a broader concept that encompasses the integration of the physical and digital worlds in manufacturing to create smart factories that are connected, flexible, and efficient (K. Zhou et al. 2015). While Smart PPC is a specific application of Industry 4.0, it plays a crucial role in achieving the goals of Industry 4.0 (Bueno et al. 2020). By providing real-time data analytics and predictive information, such as predictive capacity planning, maintenance and forecasting, Smart PPC enables manufacturers to make informed decisions and respond quickly to changing market demands (Bueno et al. 2020). Furthermore, with technologies such as Big Data Analytics and Artificial Intelligence, more flexibility in for example capacity management can be introduced through the implementation of production resource monitoring and material-flow optimization (Bueno et al. 2020).

One study conducted by researchers in Brazil found that Smart PPC can lead to significant improvements in production efficiency and flexibility in the manufacturing process (Bueno et al. 2020). The researchers found that Smart PPC can result in increased productivity in manufacturing by approximately 30 % (Bueno et al. 2020). Furthermore, the same study found that the integration of Smart PPC and Industry 4.0 technologies can help manufacturers achieve greater sustainability in their production processes (Bueno et al. 2020). The study found that the real-time monitoring and optimization of production processes can lead to a reduction in energy consumption and waste generation, ultimately resulting in a more sustainable manufacturing system (Bueno et al. 2020).

Overall, Smart PPC is a critical application of Industry 4.0 that enables manufacturers to optimize their production processes in real-time (Bueno et al. 2020). By providing real-time data analytics and predictive information, Smart PPC can help manufacturers make informed decisions and respond quickly to changing market demands, leading to improvements in efficiency, sustainability, and productivity (Bueno et al. 2020).

3.5.5 Benefits of Digitalization

The emergence of digitalization has created several benefits for the manufacturing industry (Björkdahl 2020), transforming traditional manufacturing processes and systems into more efficient, productive, and sustainable operations (Akundi et al. 2022). This section discusses the key benefits of digitalization in manufacturing.

Increased Productivity and Efficiency

Digitalization has been important in enhancing productivity and efficiency in manufacturing (Buer, Semini, et al. 2021). By optimizing processes, reducing waste, and improving throughput, manufacturers can significantly increase their productivity (Buer, Semini, et al. 2021). Predictive analytics, for instance, can be used to identify and address production bottlenecks before they occur, thereby preventing delays and inefficiencies in the production process (G. Wang et al. 2016). Moreover, the implementation of real-time monitoring and control systems can optimize machine utilization and reduce downtime, further enhancing productivity and efficiency (Soori et al. 2023b).

Improved Quality Control

Quality control is another area that has greatly benefited from digitalization (Parviainen et al. 2017). The implementation of real-time monitoring and control systems has enabled manufacturers to detect defects and deviations in the production process in a timely manner (C. Wang et al. 2015). This allows for immediate interventions, preventing defective products from reaching the customer and ensuring the delivery of high-quality products (C. Wang et al. 2015).

Enhanced Supply Chain Visibility

Digitalization has also improved supply chain visibility, allowing manufacturers to track and analyze the movement of goods and materials throughout the supply chain (Ahmed et al. 2021). This increased transparency and traceability can help identify inefficiencies, reduce costs, and improve the accuracy of demand forecasting (Ahmed et al. 2021). Furthermore, it enables manufacturers to respond more effectively to changes in demand, ensuring that supply aligns with customer needs (Ahmed et al. 2021).

Increased Flexibility and Agility

The flexibility and agility of manufacturing operations have been significantly enhanced through digitalization (Isaksson et al. 2018). Manufacturers can now quickly adapt to changing market conditions and customer demands, thanks to the use of modular production systems that enable rapid reconfiguration of the production line (Bortolini et al. 2018). This allows manufacturers to react to changes in product design or volume, ensuring that they can meet customer needs in a timely and efficient manner (Butollo 2021).

Improved Safety and Sustainability

Digitalization has also contributed to safer and more sustainable production processes (Ghobakhloo 2020). Machine learning algorithms, for instance, can be used to predict and prevent accidents, thereby improving safety in the workplace (B. Zhong et al. 2020). Additionally, the implementation of energy management systems can help reduce energy consumption and carbon emissions, contributing to more sustainable manufacturing practices (Kanchev et al. 2011).

In conclusion, digitalization offers numerous benefits to the manufacturing industry. By leveraging digital technologies, manufacturers can enhance their productivity and efficiency, improve quality control, increase supply chain visibility, and improve safety and sustainability. However, it is important to note that the successful implementation of digitalization requires a strategic approach that takes into account the specific needs and capabilities of the organization.

3.5.6 Summary

- PPC is crucial for coordinating all the activities related to the production process, and managing it effectively leads to enhanced productivity, lower costs, and better customer satisfaction.
- PPC has evolved over time from manual, paper-based systems to automated and digital ones, allowing more efficient planning, scheduling, and real-time adjustments.
- The challenges in implementing effective PPC include dealing with uncertainty in demand and supply, managing complex and global supply chains, and adapting to rapidly changing market conditions.
- Digital transformation in PPC has its own set of challenges, like the need for significant capital investment, the requirement of skilled labor, and cybersecurity concerns.
- The digital transformation of PPC has been driven by advances in technology and growing competition, and it enables greater responsiveness to market changes and overall improved efficiency.
- Industry 4.0 is transforming manufacturing, leading to streamlined processes, increased productivity, and enhanced competitiveness.
- Industry 4.0, underpinned by technologies such as IoT, Big Data Analytics, AI, and Smart PPC, is revolutionizing the manufacturing landscape, driving efficiency, and enabling customization at scale.
- While digitalization brings immense benefits like increased productivity, improved quality control, and better supply chain visibility, it requires strategic considerations to mitigate challenges related to cybersecurity, data management, and the development of appropriate skills.

3.6 Machine Learning

3.6.1 Introduction

This section covers an overview of available theory on machine learning, its suitable areas of implementation and prominent predictive algorithms used for regression. The information presented is mainly based on academic books and papers. However, machine learning is a broad topic in continuous development (K. P. Murphy 2022), and we have therefore chosen to focus on algorithms and models relevant to our available data and problem statement. Machine learning algorithms are basically rooted in mathematical principles (Deisenroth et al. 2020). In this part, we'll give a quick overview of the key formulas and equations used in the algorithms, but we will not go into a detailed examination of the mathematics involved.

In the context of machine learning, it is important to distinguish between the terms method, algorithm, system and model as they each refer to distinct concepts (Hastie et al. 2009). A method refers to a systematic procedure used to address a given problem (Hastie et al. 2009). An algorithm represents a precise set of instructions or steps designed to solve a specific problem (Z.-H. Zhou 2021). A system includes not just the algorithm, but also the implementation details and computational resources utilized (Badillo et al. 2020). Finally, a model can be described as a trained and instantiated representation of a specific algorithm and undergoes optimization and fine-tuning (Badillo et al. 2020).

In the current age of the fourth industrial revolution, the digital world has a wealth of data, such as Internet of Things data, cybersecurity data, mobile data, business data, social media data, and health data (Sarker 2021). To intelligently analyze that data and develop smart solutions the knowledge of artificial intelligence, particularly, machine learning is the key (Sarker 2021).

3.6.2 Data types

In the field of machine learning, we work with data and datasets (Hastie et al. 2009). A dataset consists of numerous data points, each representing an entity or observation that we intend to analyze (Hastie et al. 2009). These data points can represent various entities, such as a patient or a reading from a sensor (Hastie et al. 2009). To construct a dataset, measurements and data collection efforts are employed to gather a set of features (Hastie et al. 2009). These features represent the data's characteristics and can encompass categorical, ordinal, or numerical data types (Badillo et al. 2020). An example of a dataset can be seen in table 3.1.

ID	Name	Age	Gender	Height	Weight
1	John	25	Male	180.5	75.2
2	Emily	30	Female	165.2	63.8
3	Michael	35	Male	175.7	80.4
4	Sarah	28	Female	162.9	55.1
5	David	32	Male	190.1	90.6

Table 3.1: An example dataset, made by author

The quality and availability of data are considered the key constructs of a machine learning model (Sarker 2021). A machine learning model is only as good as the data you provide it with (El Naqa and M. J. Murphy 2015). Data appear in different forms and structures. McCallum (2005) briefly explains the different types of data:

Type of Data	Description
Structured	Structured data follows a specific format, aligns with an established data model, and is systematically arranged for easy accessibility and utilization by individuals or software applications. Such data are usually arranged in a table-like format in systems like relational databases. Examples of structured data include details such as names, dates, addresses, and credit card numbers.
Unstructured	In contrast, unstructured data lacks a pre-determined arrangement or format, which makes its capture, processing, and analysis more challenging. It predominantly consists of text and multimedia content. Instances of unstructured data may include sensor readings, emails, blog posts, wikis, or web pages..
Semi-structured	Semi-structured data, while not maintained in a relational database as structured data, do possess certain organizing characteristics that simplify its analysis. Instances of semi-structured data can include documents in HTML, XML, or JSON formats, as well as NoSQL databases
Metadata	The difference between data and metadata lies in their roles. Data represents raw information that is capable of quantifying, describing, or recording certain aspects related to an organization's data properties. Metadata, in contrast, provides additional information about the data itself, thereby enriching its contextual relevance for the data users. For instance, in the case of a document, metadata could include elements such as the author's name, the file size, or the creation date of the document.

Table 3.2: Different types of data, according to McCallum (2005)

3.6.3 Machine learning types

Machine learning can be divided into different types, depending on the data and what information is available (Bishop and Nasrabadi 2006). There are four kinds of machine learning techniques: supervised, unsupervised, semi-supervised and reinforcement learning (Smola and Vishwanathan 2008).

Supervised learning

Supervised learning is a type of machine learning where an algorithm learns from labeled training data, and makes predictions based on that data (Bishop and Nasrabadi 2006). The "supervisor" refers to the provided labels that guide the learning process (Hastie et al. 2009). The goal is to train a model that maps inputs to outputs, so one can predict the output for new unseen data, where we have observed input values but not their associated output (Badillo et al. 2020). The output can be numerical or categorical. Supervised methods look to learn the function f which maps the input x to the output y, based on a labeled set of input-output pairs D (Russell 2010):

$$f(x|D = x_i, y_i) = y \tag{3.1}$$

Unsupervised learning

In unsupervised learning, the algorithm attempts to identify natural relationships and groupings within the data without reference to any outcome or the correct answer (Duda, Hart, et al. 2006). Unlike in supervised learning, the data is not labeled, and the model tries to detect patterns in the input data known as knowledge discovery (Biamonte et al. 2017). The most common unsupervised problems are feature engineering, dimensionality reduction, clustering, density estimation, and anomaly detection (Sarker 2021).

Semi-supervised learning

Semi-supervised learning models are trained using both labeled and unlabeled datasets (Sarker 2021). The process of data labeling, particularly for large datasets, can be both cost and labor-intensive (Sarker 2021). To augment the efficiency of model training, semi-supervised learning leverages a significant volume of unlabeled data in conjunction with a smaller set of labeled data, with the ultimate objective of enhancing the model's performance (Zhu and Goldberg 2009). Some application areas where semi-supervised learning is used include machine translation, fraud detection, labeling data and text classification (Sarker 2021).

Reinforcement learning

In reinforcement learning, systems learn a task over time through trial and error (Richard S Sutton and Barto 2018). Reinforcement learning techniques take an iterative approach to learn by obtaining positive or negative feedback based on the performance of a given task on some data and then self-adapting and attempting the task again on new data (Richard S Sutton and Barto 2018). Reinforcement learning is neither supervised nor unsupervised as it does not require labeled data or a training set (R. Sutton 1998). It relies on the ability to monitor the response to the actions of the learning agent (Dojo 2023). Reinforcement learning has shown great success with games, such as chess, but is also used in robotics (Silver et al. 2018).

Case-Based reasoning

Case-based reasoning means using old experiences to understand and solve new problems (Kolodner 1992). In case-based reasoning, a reasoner remembers a previous situation similar to the current one and uses that to solve the new problem (Kolodner 1992). A case is a collection of previous cases and based means that the reasoning is based on earlier cases (Richter and Weber 2016). Reasoning means that the approach is intended to draw conclusions using cases, given a problem to be solved (Richter and Weber 2016). Case-based reasoners solve new problems by retrieving stored 'cases' describing similar prior problem-solving episodes and adapting their solutions to fit new needs (De Mantaras and Plaza 1997).

3.6.4 Machine learning problems

Machine learning algorithms can also be categorized based on the type of problem they solve (Edgar and Manz 2017). The different problems are classification, clustering, regression and anomaly detection (Edgar and Manz 2017).

Regression

Regression is a type of supervised learning problem that involves predicting a continuous number from one or more input features (Han et al. 2022). The goal is to construct a mathematical model that can best describe the relationship between the input features and the outcome (Han et al. 2022). This model is then used to predict the outcome for unseen data (Russell 2010). Examples of applications include predicting house prices based on features like the number of bedrooms or predicting stock prices based on past trends and other economic indicators (Adetunji et al. 2022).

Classification

Classification is another type of supervised learning problem where the aim is to predict the category or class of an instance based on its input features (Hastie et al. 2009). Unlike regression, the outcome variable in a classification problem is discrete rather than continuous (Han et al. 2022). The algorithm is trained on a dataset where the class labels are known, and the goal is to generalize from this training set to unseen data (Hastie et al. 2009). Common applications include email spam detection, image recognition, and medical diagnosis (Z.-H. Zhou 2021, Sarker 2021, De Villiers et al. 2004).

Clustering

Clustering is an unsupervised learning task that involves grouping similar instances together (Badillo et al. 2020). The objective is to partition the dataset into clusters such that instances within the same cluster are more similar to each other than they are to instances in other clusters (Badillo et al. 2020). Unlike classification, clustering does not require labeled data and the number of clusters may not be known in advance (Sarker 2021). This technique is commonly used in market segmentation and social network analysis (Madhulatha 2012).

Anomaly detection

Anomaly detection, sometimes referred to as outlier detection, is an unsupervised learning problem that involves identifying unusual patterns or outliers in the data that deviate significantly from the expected behavior (Aggarwal and Aggarwal 2017). The goal is to detect instances that are significantly different from the majority of the other instances (Chandola et al. 2009). These could be due to errors, noise, or genuinely anomalous behavior (Aggarwal and Aggarwal 2017). Anomaly detection has important applications in fraud detection, network security, and fault detection (Chandola et al. 2009), and is often used in time series in order to detect noise in the data (Blázquez-Garcıa et al. 2021).

3.6.5 Time series

Time series are sequences of data observed over a period of time (Brockwell and Davis 2002). When multiple variables are used as features, the time series is referred to as multivariate, with one it is called univariate (Cryer and Kellet 1991). When data is measured at a defined sampling rate, the time series is discrete, which is different from continuous measuring (Brown 2004). A time series is deemed stationary when the average value and variability stay roughly the same over the course of time. (Cryer and Kellet 1991). This is the case for time series without trends or seasonal components, which is normally not the case in practice (Brockwell and Davis 2002). Table 3.3 shows an example of how time-series data can be structured. Figure 3.3 illustrates an example of a time series with trend and seasonality.

Date	**Sales** (NOK)	**Customers**
04.11.22	1900	15
05.11.22	3500	23
06.11.22	1250	8
07.11.22	760	4

Table 3.3: Example of time series data, produced by author.

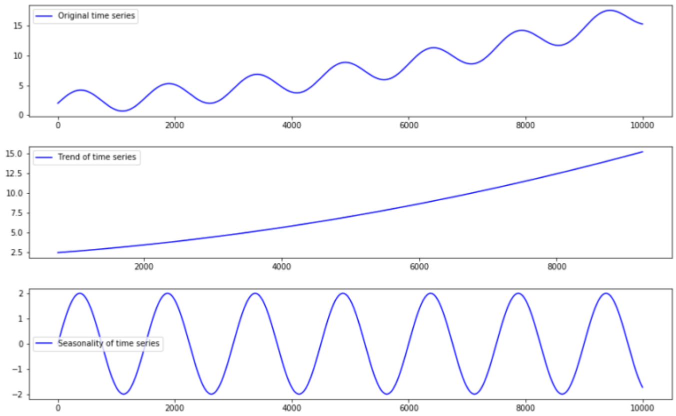


Figure 3.3: Visualization of time series data, and its corresponding trend and seasonality (Haber 2023)

One of the primary objectives of building a model for a time series is to be able to forecast the values for that series at future times (Cryer and Kellet 1991). Of equal importance is the assessment of the precision of those forecasts (Cryer and Kellet 1991). Time-series forecasting is an important process in many fields, including economics, finance, and climate science, where understanding and predicting temporal patterns yields significant insights (R. J. Hyndman and Athanasopoulos 2018). Traditional statistical methods for time-series forecasting, such as Autoregressive Integrated Moving Average (ARIMA) models and state-space models, remain widely used due to their robustness and interpretability (R. J. Hyndman and Athanasopoulos 2018). These methods capture linear dependencies and deal with stationary data, but they are insufficient when encountering complex non-linear patterns or when the underlying data-generating process changes over time (Cryer and Kellet 1991).

In recent years, machine learning approaches have been increasingly applied to time-series forecasting (Masini et al. 2023). These methods are capable of modeling non-linear relationships and can be particularly effective when large volumes of data are available (Masini et al. 2023). However, they often require careful tuning and may lack the interpretability of traditional models (Goodfellow et al. 2016).

Structuring a time-series forecasting problem is very important, and various approaches can be employed depending on the nature of the data and the specific task at hand (Goodfellow et al. 2016). There have recently been developed sophisticated techniques such as encoder-decoder architectures, which can be utilized for more complex time-series data (Platen 2023). These architectures, often based on Recurrent Neural Networks (RNNs) or Long Short-Term Memory (LSTM) networks, encode past observations (input sequence) into a fixed-size vector and then decode this vector to forecast future values (output sequence) (Goodfellow et al. 2016). This approach allows for the modeling of complex temporal dependencies (Goodfellow et al. 2016).

Another common method involves transforming the time-series forecasting problem into a supervised learning problem by creating lagged features or variables (R. J. Hyndman and Athanasopoulos 2018). Lag functions are used to save information in a data point about previously time-stamped data point's information (R. J. Hyndman and Athanasopoulos 2018). This way we can use advanced machine learning algorithms designed for supervised learning on time series data (Brown 2004). An example of this approach can be seen in table 3.4. In the table, we can see the target value *Electricity Price*. Each row, which represents a date, holds features from previous timestamps. In this example, each row can be fed to a machine learning algorithm, with the target feature, and the time-series forecasting problem can be solved using traditional supervised machine learning algorithms, such as regression, decision trees or artificial neural networks (Masini et al. 2023).

Time	Electricity Price (NOK)	Price t-1	Price t-2	...	Price t-n
04.11.22	82	-	-	...	X
05.11.22	80	82	-	...	X
06.11.22	97	80	82	...	X
07.11.22	85	97	80	...	X

Table 3.4: An example table, made by author, of how time series data can be structured, in order to save information from previous timestamps.

The traditional approach to time series in machine learning is to divide the time series into a training and a test set (R. J. Hyndman and Athanasopoulos 2018). A model tries to learn the patterns, mainly the trend and seasonality of training data in order to predict the upcoming data (the test data) (R. J. Hyndman and Athanasopoulos 2018). An example taken from Brockwell and Davis (2002), seen in Figure 3.4, shows the number of accidental deaths in a certain region. The black part of the graph shows the data used for training/learning and the gray dotted line is the data they want to predict. While this is the traditional approach (Brockwell and Davis 2002), it is also possible to train a model on multiple time series (Montero-Manso and R. Hyndman 2021).

According to Herzen (2023), the most popular time series forecasting techniques were focusing on isolated time series; that is, predicting the future of one time series considering the history of this series alone. She further states that the evolving technologies in deep learning are bringing many interesting new applications to time series forecasting. For instance, it is now possible to create models that capture patterns and work on multi-dimensional and trained on multiple related time series (Herzen 2023). A multi-dimensional time series, also called multivariate, is a sequence of observations where each observation consists of multiple related variables recorded over a period of time (Brockwell and Davis 2002). Multiple time series refers to separate time series that are somehow related to each other (Montero-Manso and R. Hyndman 2021). An example of multiple time series is the demand for multiple products offered by a retailer (Montero-Manso and R. Hyndman (2021)). They further state that forecasting of groups of time series can be approached locally, by considering each time series as a separate regression task and fitting a function to each, or globally, by fitting a single function to all time series in the set. Their study concludes that global methods can outperform local for groups composed of similar time series.

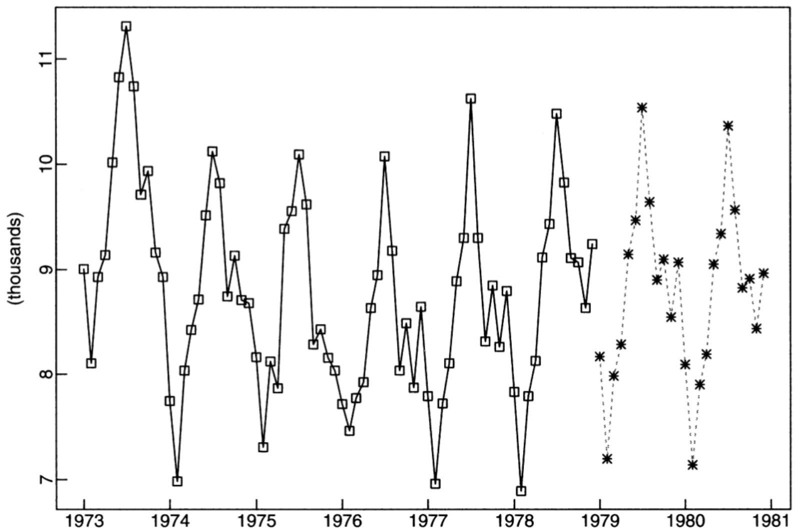


Figure 3.4: Visualization of how training and test data can be divided in a time series (Brockwell and Davis 2002).

3.6.6 Machine learning models

As mentioned earlier in this section and as we can see in section 3.7.3, time series forecasting can be approached in several different ways. In our thesis, we have decided to try solving the challenge of time series forecasting with machine learning. Therefore this section mainly focuses on machine learning algorithms that are suitable for supervised learning with a structured dataset. This section goes through the ideas and application of several algorithms, without analyzing them in depth from a mathematical standpoint.

Model formulation involves judgment, experience, trial, and error (Davison 2003). It is crucial for models to accurately mirror the system they are studying, and have the ability to be applied to similar datasets (Freedman 2009). A principle commonly used regarding machine learning models is Occam's Razor, which says that the problem-solving principle recommends searching for explanations constructed with the smallest possible set of elements (Rasmussen and Ghahramani 2000). In machine learning, this means that we favor simple models over complex ones that fit our data equally well (Rasmussen and Ghahramani 2000). Every machine learning problems usually have differences in available data, input, noise, volume, information and the desired goal (Wongsuphasawat et al. 2019). The problems may also have different requirements in terms of runtime and accuracy (Wongsuphasawat et al. 2019). A big part of the development of models is therefore to test and compare models, but also to explore the available data (Wongsuphasawat et al. 2019).

Regression

Regression can be used to predict the outcome for unseen objects, given a set of input and output measurements (Han et al. 2022). In the field of machine learning, regression analysis is a powerful tool for predicting continuous outcomes and is commonly used for prediction and forecasting (Draper and Smith 1998). The simplest form of regression, linear regression, attempts to draw a line that comes closest to the data by finding the slope and intercept that minimize the sum of the squares of the errors made in the results of every point in the line (Hastie et al. 2009). If only one independent variable is the input, the line can be formulated as in equation 3.2 (Draper and Smith 1998).

$$Y = \beta_0 + \beta_1 X + \epsilon \tag{3.2}$$

Where Y is the predicted target value, β_0 is the intercept constant of the line, β_1 is a constant that represents increase per time unit, and ϵ is a random variable with zero as mean and variance representing the scatter about the trend (Hastie et al. 2009).

Linear regression may also be used when there are multiple independent variables (Freedman 2009). In that case, the formula looks like this:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \epsilon \tag{3.3}$$

Where each X is a vector that represents a feature in the input space (Hastie et al. 2009). The constants are typically found through the minimization of some errors function, also referred to as a loss function (Hastie et al. 2009). In regression problems, the loss function often uses the difference between the desired and the predicted output (Hastie et al. 2009). Error functions containing the square of this difference are most commonly used in practice, such as in the residual sum of squares (RSS) (Hastie et al. 2009):

$$\text{RSS} = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 \tag{3.4}$$

The RSSE is often divided by the number of samples to obtain the mean squared error, which is more comparable between different datasets of different sizes (Chollet 2021):

$$\text{MSE} = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 \tag{3.5}$$

To ensure that the error has the same unit as the original output feature, root mean squared error is also a common loss function (Bishop and Nasrabadi 2006):

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \tag{3.6}$$

When the residual error for regression has a Gaussian distribution with zero mean and constant variance, minimizing the sum of squared residuals will find an optimal solution (Goodfellow et al. 2016). However, datasets may contain outliers that are hard to detect and remove, which means the assumptions made for the Gaussian distribution do not hold (Goodfellow et al. 2016). This can result in a poor fit when using RMSE as a loss function (Hastie et al. 2009). To make the model more robust to outliers, one can transform the data to make the distribution become more similar to Gaussian, or one can change the error metric to for instance the mean absolute error (Bishop and Nasrabadi 2006):

$$\text{MAE} = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i| \tag{3.7}$$

Another metric used in regression is the R-squared score (Hastie et al. 2009). It is a measure of how good our model fit the training data, indicating the proportion of the variance in the target variable that can be predicted from the training features (Hastie et al. 2009). The values for R^2 range between 0 and 1, where a higher value indicates that a larger proportion of the variance in the outcome can be explained by the features used in training (Russell 2010). According to Tayo (2022) R^2 score is one of the most important metrics for evaluating a continuous target regression model, but should be used in combination with other metrics such as RMSE and MSE. The equation for R^2 can be seen below:

$$\begin{aligned}
SS_{\text{res}} &= \sum_i (y_i - \hat{y}_i)^2 \\
SS_{\text{tot}} &= \sum_i (y_i - \bar{y})^2 \\
R^2 &= 1 - \frac{SS_{\text{res}}}{SS_{\text{tot}}}
\end{aligned}$$

Several other machine learning algorithms use regression principles to predict future outcomes (Hastie et al. 2009). For instance, ridge regression, and lasso regression are variants of regression that use regularization to handle overfitting (Hoerl and Kennard 1970, Tibshirani 1996). Generalization refers to the ability of a model to make accurate predictions on new, unseen data (Emmert-Streib and Dehmer 2019). Training and test data do not follow the same distribution (Bishop and Nasrabadi 2006). Ridge and lasso regression introduce a penalty term for large weights, that prevents the model from overfitting and improves the model's generalization for new data (Bishop and Nasrabadi 2006).

Decision trees

Decision trees are an essential building block for many of the state-of-the-art machine learning algorithms (Dobra 2009). It is a supervised learning approach used in machine learning, in order to learn from observed data and predict on unseen data (Kingsford and Salzberg 2008). A decision tree is called a classification tree if the target value takes a discrete set of values (Kingsford and Salzberg 2008). If the target value is continuous, the decision tree can also perform regression tasks (Rokach and Maimon 2005).

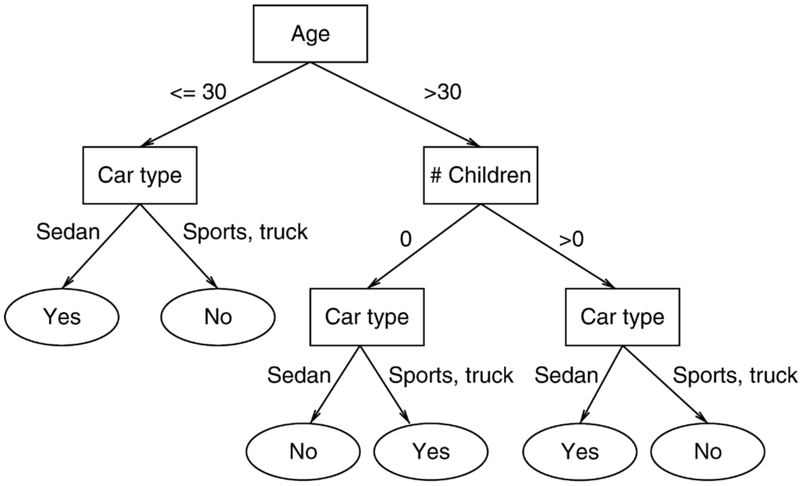


Figure 3.5: A visualization of a decision tree (Dobra 2009)

A decision tree consists of attribute nodes linked to two or more subtrees and decision nodes, called leaves, with a class or number representing the decision (Podgorelec et al. 2002). The test node computes an outcome based on the attribute values of an instance, where each possible outcome leads to a leaf-node or a new subtree (Podgorelec et al. 2002). The target value for an instance is computed by starting at the tree's root node, and continuing downwards until a leaf-node is reached (Rokach and Maimon 2005). The leaf-nodes value or class is assigned to the current instance (Rokach and Maimon 2005).

The construction of a decision tree is mathematically calculated based on the entropy of the dataset (Podgorelec et al. 2002). Entropy is a measure of disorder in the dataset (Wehrl 1978). For instance, if a dataset contains equally many instances of two classes, it has high disorder (Wehrl 1978). Conversely, if a majority of instances belong to a single class, the dataset has a lower level of disorder (Wehrl 1978). The formula for calculating entropy:

$$\mathrm{E(S)} = \sum_{i=1}^{c} -p_i \log_2 p_i \tag{3.8}$$

In the formula, p_i represents the probability of a class in a dataset (Wehrl 1978). When deciding which split to choose, we want to use the split that gives us the most information gain (Russell 2010):

$$\text{Information Gain} = Entropy_{parent} - Entropy_{children} \tag{3.9}$$

When the splits in a decision tree are calculated, the potential different datasets after a split are compared, and we choose the split that gives the minimum disorder (G. Zhang and Gionis 2023). The objective is to select the split that minimizes the overall entropy, thereby maximizing the purity of the resulting subsets (Wehrl 1978). The process is repeated recursively for each subset, creating branches in the decision tree, until a stopping criterion is met (Podgorelec et al. 2002). The stopping criterion could be a predefined tree depth, a minimum number of instances required to split a node, or a threshold on the decrease of entropy (Podgorelec et al. 2002).

Decision tree regression is a variant of a decision tree classifier that can be used to approximate continuous values (M. Xu et al. 2005). The construction of a regression tree is also based on binary recursive partitioning, which splits the data into partitions (M. Xu et al. 2005). It differs from a classification task since its final output is a real value and it, therefore, computes the average in the leaf-node, instead of the mode number as it does in classification (Sayad 2023).

Decision trees have an advantage over "black-box" models, such as artificial neural networks, in terms of comprehensibility (Kotsiantis 2013). The logical rules followed by a decision tree are much easier to interpret than the numeric weights of the connections between the nodes in a neural network (Kotsiantis 2013). Decision makers in companies tend to feel more comfortable using models that they can illustrate and understand (Kotsiantis 2013).

Artificial neural networks

Artificial neural networks (ANN), or just neural networks, have become a popular and helpful model for classification, clustering, pattern recognition and prediction in many disciplines (Dave and Dutta 2014). An ANN is a data processing model based on the way biological nervous systems, such as the brain, process data (Dastres and Soori 2021). Many artificial intelligence experts believe that artificial neural networks are currently the best method for designing intelligent models that can interpret and learn patterns (Dastres and Soori 2021). Deep learning is a very promising sub-science of machine learning, and it refers to the training of deep neural networks (LeCun et al. 2015). The term *deep* signifies the utilization of multiple layers in a neural network (LeCun et al. 2015), illustrated in figure 3.6.

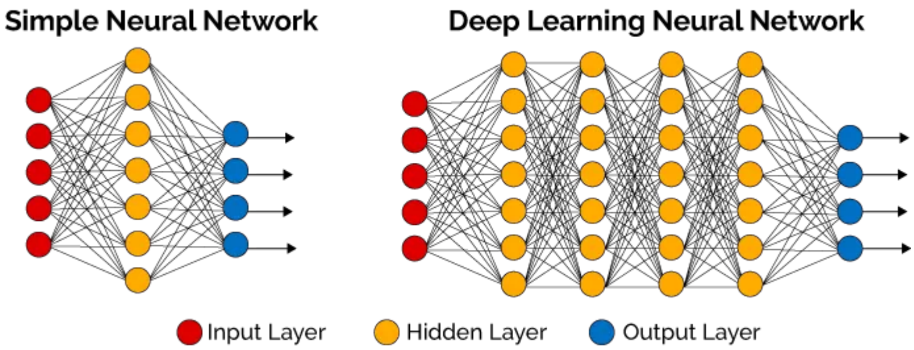


Figure 3.6: A visualization of a single and a deep ANN (Santos et al. 2021)

.

As the name implies, a neural network is a network with several connected neurons (Haykin 2009). A simple neuron takes inputs, does mathematical calculations on them and produces one output (Dastres and Soori 2021). In the mathematical calculation, the inputs are multiplied with a weight and a bias term is added (Dastres and Soori 2021), shown in equation 3.10. The bias parameter is added so that the output is biased towards being equal to the bias in the absence of any input, which is essential to ensure successful learning (Haykin 2009).

$$\sum_{i=1}^{N} w_i x_i + b_i \tag{3.10}$$

Each neuron has a structure similar to the illustration in figure 3.7 (Haykin 2009). As we can see in figure 3.6, the network consists of multiple layers. Units in the input layer are the plain features, while units in the other layers are neurons with the described structure (Santos et al. 2021). A dense network is a network where each unit is connected to every unit in the following layer (Goodfellow et al. 2016). All the layers between the input and output layer are referred to as hidden layers (Santos et al. 2021). The number of layers can be modified to suit the problem to be solved (Goodfellow et al. 2016). The term *network depth* refers to the total number of layers, and *layer width* denotes the number of neurons within each layer (Dastres and Soori 2021). The layer width can differ across layers (Dastres and Soori 2021).

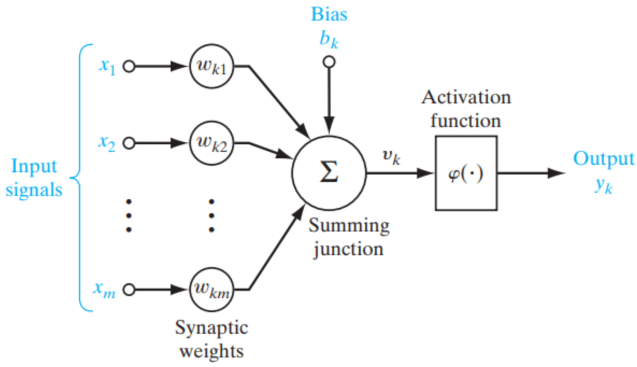


Figure 3.7: Composition of the neuron used in ANN (Haykin 2009)

The activation function is responsible for transforming the weighted sum of the input in every neuron into a format that can be used by the next layer (Santos et al. 2021). It introduces non-linearity into the output of a neuron, enabling the network to learn from complex patterns and make accurate predictions (Goodfellow et al. 2016). Nonlinear activation functions are required in the hidden neuron networks in order to learn nonlinear transformations (Haykin 2009). The most used activation functions are the following (Santos et al. 2021):

$$\sigma(x) = \frac{1}{1 + e^{-x}} \tag{3.11}$$

$$\tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}} \tag{3.12}$$

$$\text{ReLU}(x) = \max(0, x) \tag{3.13}$$

$$\text{LeakyReLU}(x) = \begin{cases} x & \text{if } x > 0 \\ 0.01x & \text{otherwise} \end{cases} \tag{3.14}$$

In order to train the model we start by initializing all the weights in the network to small random numbers (Krogh 2008s). For each training example we send the data through the network and to begin with, we receive a random output (Haykin 2009). We measure the squared difference between this output and the desired output - the correct class or value (Krogh 2008). The sum of all these numbers over all training examples is called the total error of the network (Krogh 2008). The smaller the total error, the better the network predicts the target value (Krogh 2008). By choosing the weights that minimize the total error, one obtains an optimal neural network, for the problem at hand (Haykin 2009).

In order to minimize the given error function, neural networks use an algorithm called gradient descent (Haykin 2009). In an ANN the values for weights and biases are randomly initialized, hence the output value is likely to be far off to begin with (Goodfellow et al. 2016). To optimize the weights and biases in the neural network, the derivatives of these parameters in relation to the error function are calculated for each layer (Jain et al. 1996). This enables the modification of the weights and biases by moving in the direction of their respective gradients (Lek and Park 2008). The lengths of this movement are regulated by the learning rate (Lek and Park 2008). This process is called backwards propagation as it propagates the error backward in the neural network (Lek and Park 2008). Repeated iterations of forward and backward propagation result in weights and biases that give an output that is ideally, or nearly ideally, aligned with the cost function (Goodfellow et al. 2016).

The weights are adjusted as we train the normal network (Haykin 2009). In order to improve the algorithm even further, there are important hyperparameters to tune in an ANN (Feurer and

Hutter 2019). The learning rate determines how much the weight and biases are adjusted at each iteration (Haykin 2009). The number of layers and the number of neurons in each layer are also considered to be hyperparameters that need to be tuned carefully (Feurer and Hutter 2019).

According to (Haykin 2009) there are several critical issues and problems using neural networks. First off, he states that neural networks are hard to interpret. There is no explicit explanation given for the decision or prediction the neural network makes (Haykin 2009). Second, he states that they are prone to overfitting, due to the very large number of parameters of the neural network. In fact, if we allowed the neural network to be of sufficient size, it could fit any arbitrary dataset (Goodfellow et al. 2016). Finally, (Haykin 2009) states that neural networks work best when the dataset is sufficiently large and the data are reasonably highly structured.

3.6.7 Ensemble learning

Regularly, as the complexity of machine learning models increases, there will be a reduction in error due to lower bias in the model (Feurer and Hutter 2019). However, after a particular point, the model will start suffering from high variance, so-called overfitting (Dong et al. 2020). The goal of ensemble methods is to overcome this problem by combining the predictions of several estimators to improve generalizability and robustness over a single estimator (G. Zhang and Gionis 2023). The primary concern is which base learners to combine and how to combine them (Sagi and Rokach 2018).

Because the error surface of neural network models often contains many local minima, models trained on the same data may benefit from the use of ensemble models (Sagi and Rokach 2018). Most models used for state-of-the-art machine learning models in competitions, use very large ensemble models (Chollet 2021).

Bootstrap Aggregating

construct multiple instances of a model on randomly selected subsets of the original training set, and aggregate their predictions to create a final prediction (Sagi and Rokach 2018). The training of the models takes place in parallel before the results of all the learners are joined by taking the average (Dong et al. 2020). By introducing randomization of the data each model trains on and by averaging this approach causes a reduction in variance. (Dong et al. 2020)

Random Forest

Random Forest is a regression tree technique that uses bootstrap aggregation and randomization of predictors to achieve a high degree of predictive accuracy (Rigatti 2017). This algorithm constructs a variety of decision trees, each based on random subsets of the dataset's features and subsamples (Rigatti 2017). The collective output of these trees is then averaged, a process that improves performance and mitigates the risk of overfitting (Liaw, Wiener, et al. 2002). This technique reduces high variance at the cost of a slightly increased bias (Rigatti 2017). Random Forest is a popular option due to its short training duration, avoiding the need for normalization of the input data, and few hyperparameters to tune (Thorn 2023).

Boosting

Boosting, similar to bagging, uses a collection of base learners to generate the final prediction (Russell 2010). However, the distinguishing characteristic of boosting methods is their sequential structure (Russell 2010). The idea is to fit models iteratively, such that the training of a model at a given step depends on the previously fitted models (Rocca 2023). After each iteration, the next model will focus on improving where the previous model made mistakes (Rocca 2023). This is done by emphasizing the weights for misclassified data (Maclin and Opitz 1997). The concept is that the models are becoming more significant at classifying these cases so that the ensemble is ultimately predicting more accurately (Drucker 1997). Each model's prediction is weighted averaged based on their predictive power, in order to make a final prediction (Maclin and Opitz 1997). A visualization of the difference can be seen in Figure 3.8.

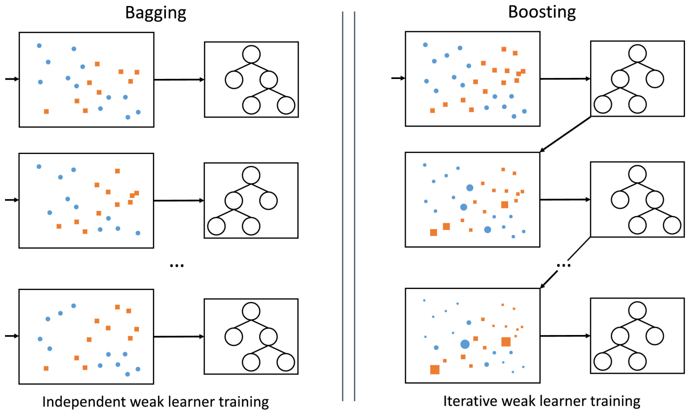


Figure 3.8: Representation of the workflows of Bagging and Boosting strategies (Gonzalez Rodriguez et al. 2020).

Gradient Boosting

There are several alternatives for boosting, which variate how the weights are determined during the training phase of the models (Natekin and Knoll 2013). Gradient boosting machines (GBM) is one of those offsprings (Natekin and Knoll 2013).

In GBM the boosting process is guided by the gradient of the loss function, hence the term *gradient boosting* (Natekin and Knoll 2013). This means that each new model in the sequence is built to reduce the residual errors of the previous model, as defined by the gradient of the loss function (Natekin and Knoll 2013). This approach allows GBMs to optimize arbitrary differentiable loss functions, making them more flexible and adaptable to a variety of problems compared to traditional boosting methods (Natekin and Knoll 2013).

Moreover, GBMs incorporate regularization techniques such as shrinkage and subsampling to prevent overfitting, which is a common problem in traditional boosting (Singh 2023). Therefore, while both GBMs and traditional boosting methods aim to improve the model's performance by combining multiple weak learners, GBMs offer a more flexible and robust approach to model building and optimization (Singh 2023).

Several popular machine learning models have been developed based on the principles of GBM (Singh 2023). XGBoost, short for extreme gradient boosting, is a machine learning algorithm designed to be highly efficient and flexible (T. Chen et al. 2015). It is implemented with a library and offers several advantages such as regularized boosting to prevent overfitting, parallel processing for speed and handling of missing values (T. Chen et al. 2015). CatBoost, is quite similar to XGBoost, but also manages to automatically handle categorical features (Hancock and Khoshgoftaar 2020).

3.6.8 Hyperparameter tuning

In addition to the parameters that are tuned in a model during the training process, numerous machine learning models also require tuning of their hyperparameters (Paleyes et al. 2022). These hyperparameters could for instance include the depth of a decision tree, the quantity of hidden layers in a neural network, or regularization parameters (Paleyes et al. 2022). The process of determining the optimal configuration of these hyperparameters is known as hyperparameter optimization (HPO) or hyperparameter tuning (Paleyes et al. 2022).

Many hyperparameter tuning approaches require the user to define a complete search space of the parameters to optimize, as in theory, these can be unlimited (Paleyes et al. 2022). However, in practical scenarios, this is often unfeasible due to a lack of domain knowledge about the problem (Shahriari et al. 2016). Consequently, setting the bounds for hyperparameter optimization remains an obstacle for making state-of-the-art models (Shahriari et al. 2016). The default hyperparameter settings in a model may not guarantee optimal performance of machine-learning techniques and additional attention should be directed to this critical step (Schratz et al. 2019). There are several HPO techniques and a selection of the usual techniques are described briefly:

Grid search

Grid search is a traditional method used for HPO (Bergstra and Bengio 2012). It operates by searching through a predefined set of hyperparameters and evaluating the model performance for each combination (Bergstra and Bengio 2012). The model's performance is usually measured by using cross-validation on the training set, each fold with different hyperparameters (Feurer and Hutter 2019). While grid search is simple and easy to implement, it can be computationally expensive and time-consuming, especially when dealing with a large number of hyperparameters or when the model takes a long time to train (Baruch 2023).

Random search

Random search is another method for HPO that addresses some of the limitations of grid search (Feurer and Hutter 2019). Instead of exploring all possible combinations of parameters, it selects a subset of hyperparameters to evaluate until a certain budget for the search is exhausted (Feurer and Hutter 2019). This approach can be much more efficient, and works better than grid search when some hyperparameters are much more important than others (Bergstra and Bengio 2012, Hutter et al. 2012).

Bayesian optimization

Bayesian optimization, as described by (Feurer and Hutter 2019), is a step-by-step process with two main parts: a model that estimates probabilities and a function that chooses the next point to check. In each step, the model is updated based on all previous observations (Feurer and Hutter 2019). The function then uses this model to decide the usefulness of different potential points, balancing between trying new areas and focusing on known good ones (Snoek et al. 2012).

3.6.9 Supporting libraries

This section presents a selection of modern technologies that offer solutions intended to simplify aspects of implementing and optimizing machine learning models. In recent years, a variety of software libraries have been released, which significantly ease and accelerate the research and application of machine learning models (Pang et al. 2020). We present some of the most commonly used ones, focusing on the ones implemented and used in the model development section. The theory in this subsection is mostly based on documentation from their website in addition to machine learning blogs and discussions.

Scikit-learn

Scikit-learn, often referred to as *Sklearn*, is a popular machine learning library for the Python programming language, that provides a comprehensive range of tools and algorithms to facilitate the implementation of machine learning models (Scikit-Learn 2023). It offers a diverse set of functionalities, including data preprocessing, feature selection, model evaluation, and model selection (Scikit-Learn 2023). Sklearn makes it efficient to implement a variety of machine learning models, without the need of developing them from scratch (Scikit-Learn 2023). For instance, several supervised regressor models such as decision trees and regression models can easily be implemented (Brownlee 2023a).

Keras and TensorFlow

TensorFlow is an open source deep learning software library for defining, training and deploying machine learning models (Goldsborough 2016). Keras is a high-level neural network library that runs on top of TensorFlow (Brownlee 2023c). Keras has been designed with user-friendliness and supports almost all types of neural architectures and provides tools for building a network with the architecture desired for the actual problem (Keras 2023). It is intuitive syntax and flexibility to quickly design and test models have made it a favored tool among researchers and developers in the machine learning community (Brownlee 2023c).

H2o

The H2O ML library is a powerful machine learning platform that aims to simplify the development and deployment of advanced models (Pandey 2019). H2O offers a user-friendly interface and a wide range of algorithms, allowing users to leverage state-of-the-art techniques without requiring in-depth knowledge of the underlying mathematics (Pandey 2019). The library supports both supervised and unsupervised learning tasks, including classification, regression, clustering, and anomaly detection (Pandey 2019). H2O ML differs from the other libraries with the opportunity to do automatic machine learning (Pandey 2019). They offer an easy-to-use interface that automates the process of training a large selection of candidate models on a given dataset (H2O auto 2022). This way it is possible to automatically tune and a handful of state-of-the-art machine learning models, without using time developing them (H2O auto 2022). The H2o Auto ML library also offers ensemble learning methods (H2O ensemble 2022).

3.6.10 Summary

- Supervised learning occurs when the target variable is known beforehand, and makes it possible to train a model on labeled data.
- Time series data can be reshaped into a supervised learning problem by using prior observations as features together with the target variable.
- With the recent development of new technology, training and prediction on multiple time series is an interesting and promising field.
- For regression problems, decision trees, random forests, and artificial neural networks have shown great performance, due to their flexibility and ability to model complex relationships.
- Ensemble learning, which aggregates the output of several models has shown great performance by leveraging the strengths and compensating for the weaknesses of each model.

3.7 Previous Research on PPC and Machine Learning in the food industry

3.7.1 Introduction

This section presents a review of academic papers discussing factors that can improve PPC and methods used for forecasting in fermentation processes. The first part presents findings from studies that focus on the improvement of PPC with traditional methods in breweries and other food-related industries. The second part presents studies that focus on forecasting fermentation processes. This process happens in several food industries, and studies regarding other industries using fermentation, other than breweries, are therefore also presented. The third part discusses research papers that use digitalization in order to improve PPC, and lastly studies exploring predictions with multiple time series are presented.

3.7.2 PPC improvement in breweries

Monroy and Vallejo (2012) implemented a discrete event simulation in a brewery in order to analyze the effects on production planning and resource scheduling. The study states the disadvantages of resource scheduling done manually in MS Excel and presents an improvement by implementing a simulation for production planning and scheduling. The planning tool automatically generates the production schedule on the basis of current stocks, master production schedule and minimum lot sizes. The study shows how the user can configure plant parameters, stock levels and weekly production orders and generate a production schedule in a really short time which makes good decisions in order to optimize plant utilization, stock levels and throughput times.

Knight (1991) developed a Holt-Winters algorithm to forecast future sales data of breweries in South Africa. The model leverages exponential smoothing in order to identify and forecast seasonal patterns and trends. Furthermore, the study integrated sales forecasting together with production planning. Their system ensured the optimal balance between labor utilization levels, stockholdings, production, warehouse and distribution capacities through analysis of trade-offs between production, marketing and operations variables. This analysis, together with a reduced level of buffer inventory because of the accurate sales forecast, led to cost minimization. The study also showed the ability to make production more efficient as they were able to extend production runs of a pack type, reducing the downtime for pack changeovers.

Georgiadis, Elekidis, et al. (2021) made a mixed-integer linear programming (MILP) model in order to minimize the total production cost in a brewery. The study focused on optimizing production planning and scheduling and minimization of cost while reducing energy needs and waste. The model was successfully applied to a Greek brewery and resulted in near-optimal production schedules. Georgiadis, Elekidis, et al. (2021) states that a major challenge in this study and in food production in general, is how to handle the uncertainty in fermentation processes. According to him, the fermentation process is the bottleneck in brewing and was solved in this problem by introducing a constraint that limits the amount of batches getting fermented at the same time. In Monardes et al. (2017) study, a MILP model is also developed in order to assist the winemaker in making an optimal production schedule. A challenge identified in this study, which also resulted in weaknesses in the model, is the uncertainty of the fermentation process in wine. There are several other studies using optimization approaches in order to improve food processes, such as Weston Jr et al. (1999), Georgiadis, Pampın, et al. (2020) and Sáenz-Alanıs et al. (2016). These studies mainly focus on determining the right time to start and stop processes or batches, and they do not pay as much attention to how these decisions impact PPC.

3.7.3 Fermentation forecasting with advanced methods

Defernez et al. (2007) presented three different approaches of characterizing beer fermentation, with the aim of predicting the likelihood of the target variable *present gravity* (PG) being reached within a given time window. Fermentation data were obtained from three different beer types from UK breweries and consisted of PG measurements taken every 12 hours. The first approach uses a smooth analytic function to model the observed PG as a function of time. The second method uses a version of the algorithm *nearest neighbors* (NN). The argument is that if a batch is similar to historical batches at the start of the fermentation, it is likely to stay similar and reach the target PG at the same time. The third approach develops reference centile charts that depict the normal behavior of PG during the course of fermentation. The study concludes that no fully-automated system exists to reliably predict the future course of a fermentation process. Their models show the best performance when used together, but the models are struggling to predict correctly when a batch behaves abnormally. Further, it concludes that in order to improve the models' performance, more data and a higher sample rate are needed.

Thibault et al. (1990) attempted to predict biomass and substrate concentrations in a fermentation process by developing a neural network and comparing its performance with a more traditional approach - Kalman Filters. The study showed how neural networks were able to learn very complex relationships without requiring the knowledge of the model structure. The input data to the neural network were all previous measurements from a certain timestamp. The goal of the model was to predict the key fermentation variables in the next time stamp ($\tau + 1$). The data used in the study was created by Thibault et al. (1990).

Montague et al. (2008) presented a case study demonstrating how forecasting algorithms can be used to assess future bioprocess conditions, in order to support decision making in production. They used a case based reasoning (CBR) procedure to make forecasts. The CBR model finds the closest batch based on historical batches and a distance metric and calculates the upcoming batch graph based on the values of the most similar batch. Data were retrieved from 100 beer batches where PG and temperature were manually logged with a sampling rate varying between 8 and 24 hours. Due to outliers and temperature control issues 40% of the batches were removed (Montague et al. 2008). Due to the large impact batch outliers had on the learning phase and problems with manual temperature measurements, the model was unable to consistently predict accurately when batches were finished (Montague et al. 2008). When it comes to PPC the study states that early prediction of the fermentation endpoint is valuable information allowing tighter scheduling of downstream operations such as bottling and canning, reduced turnaround time for fermentation vessels and ultimately, an increase in annual plant throughput. However, it is not discussed how to utilize this potential in PPC.

Syu et al. (1994) made an ANN to predict the effects of varying input conditions on hypothetical fermentations. The neural network was constructed by inputting initial free amino nitrogen concentration, initial oxygen concentration and initial viable cell count in order to produce an ethanol profile of a seven days long fermentation. The study demonstrated how neural networks are able to recognize and learn from patterns in data. However, it did not account for more data becoming available as fermentation progresses. Ekpenyong et al. (2021) developed an ANN to model and optimize the fermentation conditions for the production of a cell. The input variables of the model were temperature, pH, agitation speed and fermentation time from 21 different samples and the target value was biomass and glycolipopeptide concentration. The neural network consisted of 9 hidden layers.

Speers et al. (2003) developed a non-linear regression model of beer fermentations. The data in the study consisted of initial Plato values, the time and volume of wort, the number of *pitchings* of each yeast crop as well as temperatures and Plato values manually recorded every day at 3:00 and 15:00. They created a four parameter logistic model to fit the data. Speers et al. (2003) also identified that the number of times a yeast was repitched had no effect on the fermentation and that the starting temperature of the batch increased the fermentation rate while decreasing the time to the midpoint of the fermentation. In addition, the initial temperature influenced the time final gravity (FG) occurred positively. Their model does not predict the end time of a batch while it ferments, but learns from previously finished batches and predicts how the Plato curve will look like based on its initial values.

Reid et al. (2021) showed great results by developing sigmoidal models to predict the rate of density decline. The models were compared by fitting them to real industrial data from lagers and whisky fermentations. Especially the 5-parameter logistic equation is interesting, as it allows the s-curved shape to be mirrored. According to Reid et al. (2021) almost every instance of density decline in fermentations follows a sigmoidal curve. They identified that the 5-parameter model is dependent on a high number of observations, in order to reduce the root squared error when an extra parameter is added. By creating prediction bands the study revealed where 99% of the data were expected to lie.

Trelea et al. (2001) proposed three dynamic models for the beer fermentation process. The models take into account fermentation temperature, the maximum pressure that occurred, and initial yeast concentration with the aim to predict the current wort density, residual sugar concentration, the released CO_2 and the ethanol concentration. Trelea et al. (2001) installed sensors in the fermentation tanks that could measure temperature with an accuracy of ± 0.07°C and top pressure with ± 10 mbar accuracy. This, in addition to wort density, residual sugar concentration and ethanol concentration was given from samples taken every 12 hours. The first knowledge-based model made mathematical equations based on theoretical knowledge regarding fermentation. The second was an empirical model and the third was a neural network. The predictions gave a lower error of 5% on CO_2, 10% on wort density and 10% on ethanol concentration. Trelea et al. (2001) states that the most used descriptor of the fermentation process is the density, but they identify CO_2 as a better parameter for online measurement.

Bowler et al. (2021) investigated how an ultrasonic sensor combined with machine learning is able to predict the alcohol concentration during beer fermentation. The sensors, together with mathematical calculations, are able to create the following features: energy, the peak-to-peak amplitude, the standard deviation for the energy in the brew, *time of flight* (which is a measure of the speed of sound in the wort that is dependent on its density and compressibility (Henning and Rautenberg 2006)) and the current temperature. This data was deployed into a Long short-term network to predict the corresponding alcohol level at the current time. The network was hyperparameter tuned by a 5-fold cross-validation procedure. The model showed great accuracy and potential to accurately predict the endpoint of fermentation using ultrasonic sensors. The authors further state that the largest barrier to implementing sensors and ML in the industry is to obtain labeled data. The study concludes that IoT, cloud computing and ML combined can cause huge beneficial effects on automatic information gathering and support in decision making in operations.

There are several studies that have a more chemical approach to the problem of predicting fermentation processes. For instance, B. Li et al. (2021) made a Long-Short Term Memory Network (LSTM) in combination with a mechanistic model to predict when cream cheese is finished fermenting. The features inputted to this model consisted of time series, mainly changes in biomass, lactose level, and lactic acid over time. The model was able to correctly predict the finishing time to within 3 minutes for 6 of the 7 batches. Garcıa et al. (1994) predicts the diacetyl (a liquid that occurs naturally in beer (Shibamoto 2014)) production throughout the industrial beer fermentation process from a kinetic equation. The equation uses parameters such as fermentable sugar, amino acid valine, diacetyl level, yeast concentration, temperature and pH level in order to learn from observed data.

3.7.4 Smart PPC

Usuga Cadavid et al. (2020) presented how state-of-the-art of machine learning can aid PPC by systematically analyzing 93 research papers focusing on ML in PPC. In addition to showing which ML techniques was the currently most used, the study also addresses how these techniques improve PPC. The most addressed use cases were smart planning, smart scheduling and time estimation. Further, inventory control, distribution control and smart design of products and processes were rarely considered. It is necessary to enable data availability, continuity and sharing over the design, logistics and production departments in order to improve the integration between PPC, logistics and design (Usuga Cadavid et al. 2020). To utilize ML in PPC intra-organizational systems such as PLM, ERP and MES are needed (Usuga Cadavid et al. 2020). The results suggest that 75% of the possible research domains in machine learning and PPC combined remain untouched. Usuga Cadavid et al. (2020) identifies the complexity of using Internet of Things technologies to collect data and the challenges of updating an ML model to adapt it to manufacturing system changes.

Bueno et al. (2020) conducted a systematic literature review to develop an analytic framework that explains how PPC in the context of Industry 4.0 is influenced by smart capabilities. The main focus of the study is the technologies IoT, cyber-physical systems, big data, artificial intelligence and additive manufacturing and how they influence manufacturing system performance indicators and environmental factors. In particular, the study goes through how Industry 4.0 can contribute to demand forecasting, capacity planning and control and inventory planning and control. It also analyzes its impact on sales and operations planning, smart production scheduling and material requirements planning. A key takeaway from all these categories is how smart manufacturing positively influences PPC by providing smart data collection, real-time monitoring, forecasting, optimized scheduling and automation.

Numerous other research studies investigated the link between industry 4.0 and PPC. For instance, Buer, J. O. Strandhagen, et al. (2018) presented, through a systematic literature review, how Industry 4.0 affects lean manufacturing. Ryback et al. (2019) presents how forecasting from machine learning models can improve production planning. They state that the production plan often gets modified subsequent execution due to bad data quality, inappropriate planning and control systems, or unforeseeable events. Ryback et al. (2019) developed forecasting techniques that increased the quality of the production plan which can lead to less need of production buffers. Oluyisola et al. (2022) presents how smart PPC can be used in a sweet and snacks manufacturing company in order to harness insights from the data allowing dynamic and near real-time action to the production system.

3.7.5 Multiple time series

Various studies demonstrate how machine learning models can train and learn from multiple time series. The idea is to train one model one multiple similar time series instead of fitting a model to each time series (Hewamalage et al. 2022). The principle is called global forecasting models (GFM) (Hewamalage et al. 2022). Hewamalage et al. (2022) gives a detailed survey of how advanced algorithms performs on different multiple time series. Modern approaches, such as recurrent neural networks, feed-forward neural networks, pooled regression models and boosting models are developed as GFM and are compared to how traditional models trained on a single time series perform. The study demonstrates how GFM outperforms traditional models and are suitable for short series, heterogeneous time series, and when minimal domain knowledge exists. However, GFM are also very sensitive to how different the multiple time series are. Cerqueira (2023) states the importance that the multiple time series are somehow related to each other, and present a boosting algorithm to learn and forecast on multiple electricity demand time series. Montero-Manso and R. Hyndman (2021) investigation revealed how global linear models provide competitive accuracy with far few parameters than the local method. Even though these studies share the idea that global forecasting is promising, none of them undertake training on complete time series, nor attempt to forecast the remainder of a series within a test set.

3.7.6 Summary

- Several breweries in different areas of the world are eager to optimize their PPC. The section shows several research papers using traditional approaches such as event simulation, MILP and Holt-Winters algorithm in order to forecast sales and the fermentation process in breweries. These studies acknowledge the impact this could have on PPC, without discussing it further.
- Previous researchers have used features such as CO_2, amino concentration, oxygen concentration and initial cell count in order to retrieve insights regarding the fermentation process.
- There have been used several modern approaches in order to forecast the fermentation process in the wine and brewing industry such as artificial neural networks, case based reasoning and nearest neighbors algorithm. These approaches, however, focus on predicting an output in the current timestamp, and not for the rest of the fermentation process.
- In order to improve the developed models, the researchers state that more data with a higher sample rate is needed.
- There exist several approaches to creating models that learn from multiple time series. These approaches, however, do not train on the whole time series. Studies regarding training and predicting on multiple multivariate time series are therefore scarce.

Chapter 4

Empirical background

This chapter presents the empirical background of the thesis. Initially, the beer production industry is introduced. The section concerning beer production will first present an overview of the processes of beer production. This is followed by a description of the craft beer supply chain and production planning and control in beer production. Furthermore, the chapter presents a multiple case study conducted with five craft breweries in the fall of 2022.

4.0.1 The brewing industry

The beer industry has witnessed remarkable growth over recent decades, particularly in the United States, with the number of breweries undergoing a substantial increase over the past decades, as seen in figure 4.1. This growth is largely fueled by an increase in customer demand for varied beer offerings, which has in turn stimulated the rise of the craft brewing sector (Alfaro 2022). From a mere 37 craft breweries in 1987, the number has now surpassed 9000 (Bart Watson et al. 2021). The industry is divided into craft breweries and macrobreweries, where the former are independent businesses that often embrace traditional brewing processes (Bart Watson et al. 2021), while the latter are enormous multinational corporations (Alfaro 2022).

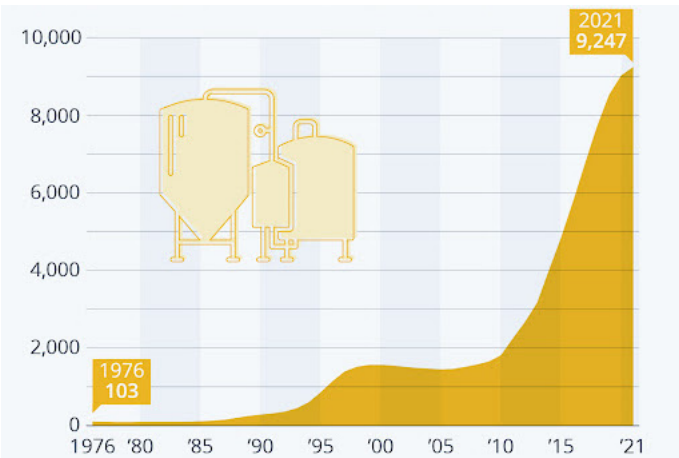


Figure 4.1: Number of breweries in the US from 1976 to 2021
(Armstrong 2022)

This thesis concentrates on the craft beer segment within the broader beer industry. The American Brewers Association (2015) defines a craft brewery as: *A small and independent American brewer producing 6 million barrels of beer (702,000,000 liters) or less annually (approximately 3 percent of U.S. annual sales).*

The American Brewers Association (2019) further categorizes craft breweries into four subgroups, outlined as follows:

Microbrewery: *A brewery producing under 15,000 barrels of beer (1,755,000 liters) per year, with 75 percent or more of its beer sold off-site. Microbreweries distribute their products through the traditional three-tier system (brewer to wholesaler to retailer to consumer), the two-tier system (brewer acting as wholesaler to retailer to consumer), and directly to consumers via carry-out and/or on-site taproom or restaurant sales.*

Brewpub: *A restaurant-brewery that sells at least 25 percent of its beer on-site, offering substantial food services. The beer is primarily brewed for sale within the restaurant and bar and is often served directly from the brewery's storage tanks. Brewpubs may also sell beer to-go or distribute to off-site accounts, where legally permitted.*

Taproom brewery: *A professional brewery selling at least 25 percent of its beer on-site without providing significant food services. The beer is primarily brewed for sale in the taproom and is often dispensed directly from the brewery's storage tanks. Taproom breweries may also sell beer to-go or distribute to off-site accounts, where legally permitted.*

Regional brewery: *A brewery with an annual beer production ranging from 15,000 barrels (1,755,000 liters) to 6,000,000 barrels (702,000,000 liters).*

4.0.2 The beer production process

Beer production encompasses several processes, which can be broadly divided into three primary and distinct production stages: wort preparation, fermentation and packaging (Stewart et al. 2016). The processes of wort preparation and fermentation can also be referred to as liquid preparation (Georgiadis, Elekidis, et al. 2021).

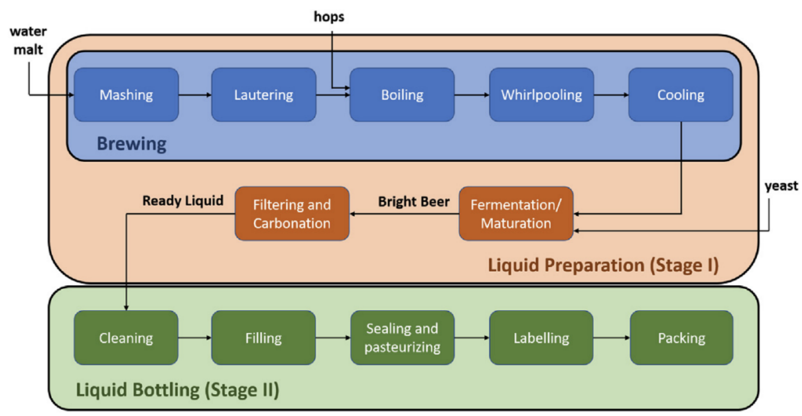


Figure 4.2: Beer production process map (Georgiadis, Elekidis, et al. 2021)

Wort preperation

Wort preparation, or brewing, involves the processes of mashing, lautering, boiling, clarification, and cooling (Georgiadis, Elekidis, et al. 2021). Mashing converts starch into fermentable sugar, while lautering filters solids from the mixture (Georgiadis, Elekidis, et al. 2021). Hops are added during boiling, and the resulting mixture, or wort, is subsequently clarified through whirlpooling and then cooled (Georgiadis, Elekidis, et al. 2021). This concludes the brewing process (Georgiadis, Elekidis, et al. 2021).

Fermentation

Fermentation, in which yeast is added to the wort, is a biological process that stimulates the production of alcohol and carbon dioxide, giving beer its characteristic taste and aroma (Alba-Lois and Segal-Kischinevzky 2010). This stage is often considered a significant bottleneck in beer production due to its time-consuming nature (Supekar et al. 2019) and variable duration, which can fluctuate by several days (Koulouris et al. 2021). Lastly, all equipment used during brewing and fermentation goes through detailed and advanced cleaning and sanitation routines (Loeffler 2006). This is considered to be of significant importance to ensure that the beer produced is of high quality (Loeffler 2006).

Packaging

In this thesis, the term *packaging* encompasses the processes of bottling, canning, or kegging, depending on the type of container used for the beer. The use of packaging allows for a streamlined reference to all three methods. As the final stage in beer production (Dıaz et al. 2022), packaging involves filling kegs, cans, or bottles with the liquid, followed by placement in finished goods inventory (Brewery 2019). The associated sub processes include cleaning, filling, sealing, pasteurizing, and ultimately labeling the containers (Georgiadis, Elekidis, et al. 2021).

4.0.3 Characteristics of the supply chain

In the United States, the beer supply chain typically follows a three-tier system that consists of the producer, in this case, the craft brewer, the distributor, and the retailer (Grassel et al. 2021, Sorini 2017). This structure is a standard model across the industry, affecting both craft breweries and macrobreweries (Grassel et al. 2021). However, these regulations can vary state by state, resulting in some flexibility in the application of this system (Durkin 2006). For instance, in Ohio, craft breweries are not obliged to strictly follow the three-tier system, providing them with an alternative operating framework within the supply chain (Durkin 2006).

Macrobreweries generally own distribution centers, which gives them a certain degree of control over their supply chain (Grassel et al. 2021). In contrast, craft breweries largely depend on third-party distributors (Grassel et al. 2021). This dependence places craft breweries at a disadvantage when it comes to obtaining direct customer sales data from retailers (Grassel et al. 2021). Furthermore, craft beer has an average shelf life of 120 days, compared to the 180 days of macrobrewery beer (Grassel et al. 2021). This difference is partly due to the limited access to information that craft breweries experience (Grassel et al. 2021). The lack of supply chain visibility impacts their ability to manage inventory effectively, leading to potential inefficiencies and additional challenges in maintaining optimal supply chain operations (Grassel et al. 2021).

Figure 4.3: Beer production process map (Sorini 2017)

Key actors in the craft brewery supply chain

There are several key actors in the craft beer supply chain. These include suppliers of raw materials, the breweries themselves, and customers (Grassel et al. 2021). Customers can range from distributors and retailers to hotels, restaurants, cafes (HORECA), and even the end consumers themselves (Grassel et al. 2021). Furthermore, the Three Tier System, a common supply chain model, doesn't account for brewpubs and taproom breweries (Grassel et al. 2021). These types of breweries often sell directly to customers, as described in section 4.0.1, bypassing traditional distribution channels, at least with that particular sales channel. Despite the diversity in beer types, breweries around the world tend to rely on similar types of suppliers for their ingredients (Salanță et al. 2020). These include suppliers of grains, hops, and yeast, which are essential for wort preparation and fermentation (Salanță et al. 2020), and suppliers of packaging materials like metal cans, glass bottles, and kegs (Martin et al. 2022). The final piece of the supply chain is the distribution and sale of beer, which can vary widely (Brand et al. 2007). It depends a lot on where you are in the world, as each country has its own set of rules and regulations around alcohol production, distribution, and sales (Brand et al. 2007).

Product attributes

Beer, being a food and beverage product, has a limited shelf life (Salanță et al. 2020). Generally, craft beer can be kept for about 120 days (Grassel et al. 2021). Furthermore, beer is a highly diverse product with a virtually limited number of product variants (Clemons et al. 2006). However, most derive from the same basic ingredients: water, grains, yeast, and hops (Clemons et al. 2006). This, along with the fact that the same beer can be packed in various containers, makes beer a product that has many versions but one base, leading to many end products, also called many-to-one-to-many.

Regardless of beer type, key stages such as wort preparation, fermentation, and packaging are universally applicable (Aroh 2019). However, the specific subprocesses may differ across different types of beer (Aroh 2019). The transition from batch production to continuous production during the move from wort preparation and fermentation to packaging contributes additional complexity (Axtman 2006).

Craft breweries drive innovation in the beer industry (Alfaro 2022). They offer a wide variety of products, maintaining successful ones while discontinuing less popular options (Dodd et al. 2021). However, they also produce long-standing, consistent styles like lager and ale (Rodhouse and Carbonero 2019).

Craft breweries distinguish themselves from macrobreweries through continuous innovation (Alfaro 2022), maintaining an expansive and dynamic product portfolio (Dodd et al. 2021). This not only includes a spectrum of innovative and unique products, but also features consistent, traditional styles such as lagers and ales (Rodhouse and Carbonero 2019).

Market attributes

Frequent orders from suppliers and deliveries to customers is necessary because beer is a perishable product (Donselaar et al. 2006). Craft breweries face stiff competition from macrobreweries and other craft breweries (Craig 2021). Responsiveness to customer demand and maintaining high service levels are thus critical for a craft brewery's success (Slocum et al. 2018).

Craft breweries typically have core products with stable demand and seasonal products with fluctuating demand (Bert Watson 2014). However, many craft breweries also produce new items or those based on current trends, which have higher demand uncertainty due to the lack of historical sales data (Alfaro 2022). Furthermore, as seen in figure 4.4, seasonality is common within the beer industry, with peak during the summer and Christmas time.

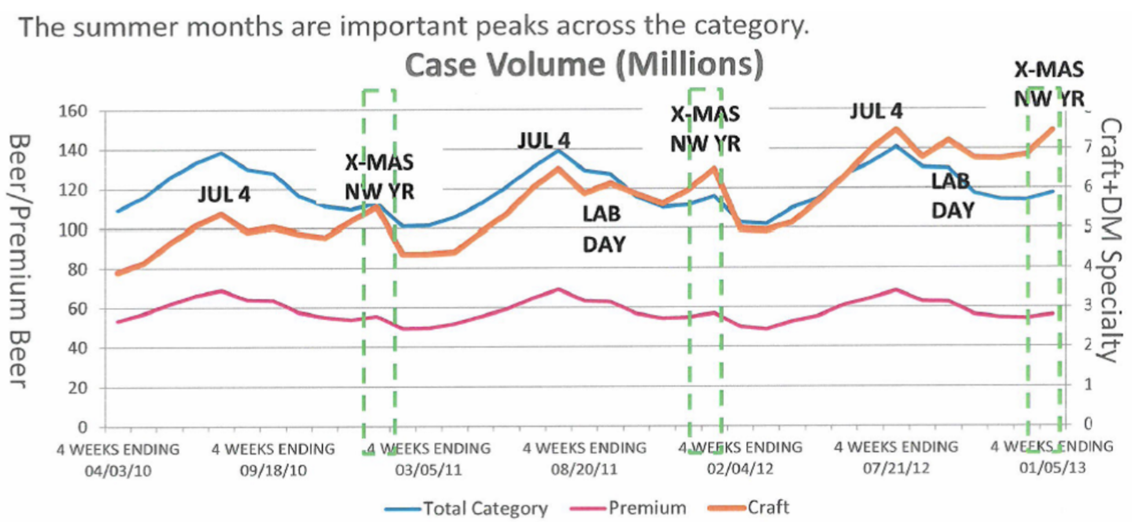


Figure 4.4: Seasonality in demand for beer (Watson 2014)

Product system characteristics

A significant factor in the beer production process is the equipment involved (Thesseling et al. 2019). The machinery required for beer production is not only specialized and expensive but also demands substantial space (Thesseling et al. 2019). To maintain the quality of the beer, it's crucial to keep this equipment clean, which often results in long setup times (Atwell et al. 2017). Furthermore, batch sizes in beer production are typically large, and the rate of production is usually determined by the capacity of the brewing facilities, similar to other food production industries (Donk 2000).

The length of the beer production process and the resulting lead times differ considerably between beer styles (Hartmeier and Reiss 2011). One key factor determining these variations is the fermentation process (Rodman and Gerogiorgis 2016). Specifically, ales typically ferment for about a week or less , while lagers may require fermentation from one week up to as long as 12 weeks (Rodman and Gerogiorgis 2016, Supplies 2019). Other stages in the beer production cycle, not including fermentation, commonly range from 3 to 7 days. This results in total lead times that can span from as short as 10 days to as long as 13 weeks (Silva Nunes et al. 2016).

4.0.4 Production planning and control in the brewing sector

In the context of beer production, with wort preparation, fermentation and packaging are the primary processes, breweries represent complex, multi-product, and multi-stage production facilities, employing a mix of batch and continuous processes to produce a variety of beer products. (Georgiadis, Elekidis, et al. 2021). Due to the nature of the food and beverage sector, breweries typically follow a make-to-stock strategy to fulfill to the market's demand for short lead times. (Gupta and Starr 2014. Ross Ackerman 2020).

Breweries often rely on manual execution of production planning and control tasks by key decision-makers, though brewery management software systems, a form of ERP systems, are increasingly being adopted. (Georgiadis, Elekidis, et al. 2021, Baker 2018). Master production schedules (MPS) and material resource plans (MRP) are generated either within the ERP system or through manual spreadsheet input (Georgiadis, Elekidis, et al. 2021). Comprehensive planning is essential due to the long lead times, and coordinating the various subprocesses is difficult because of the significant variations in their durations and the unpredictable nature of beer fermentations (Georgiadis, Elekidis, et al. 2021).

To control production, breweries follow a standardized plan (Brewing Chemists 2019). Precise measurements during wort preparation and fermentation are essential to ensure product quality, with temperature and density being the main variables to monitor. (Bokulich and Bamforth 2013, Brewing Chemists 2019, Thesseling et al. 2019). The control frequency and parameters can differ across the processes, but density remains a crucial factor during wort boiling and fermentation, influencing both the quality and completion of the process. (2019, Thesseling et al. 2019). Given its significance, fermentation necessitates a more stringent control system with daily monitoring (2019). Various factors can be regulated, including specific gravity (density), pH, ethanol, and CO2 levels (Bokulich and Bamforth 2013). Insufficient temperature management can significantly impact production, making temperature control essential (Bhonsale et al. 2021).

4.0.5 Summary

- The U.S. beer industry, divided into craft breweries and macrobreweries, has seen a massive growth with more than 9000 craft breweries emerging since 1987 due to increased customer demand for varied beer offerings.
- Beer production follows three main stages: wort preparation, fermentation, and packaging. Variations in subprocesses and transition from batch to continuous production introduce complexity.
- The beer supply chain follows a three-tier system involving the craft brewer, the distributor, and the retailer, with variations allowed in some states. Craft breweries often face challenges due to limited supply chain visibility and shorter product shelf life.
- Beer is a highly diverse product with a wide range of variants, driven by the innovative efforts of craft breweries. Beer production requires specialized, expensive equipment, and the length of the production process can vary greatly based on beer styles.
- Production planning and control in breweries are often manually executed. The long lead times and unpredictability of fermentation processes necessitate comprehensive planning and control measures.

4.1 Multiple Case Study

A case study on production planning and control at five different craft breweries was conducted during the fall of 2022 as part of the project thesis (Heum and Hjort 2022). The interview guide used in interviews with the five craft breweries can be seen in appendix A. The following paragraphs will summarize that case study and its findings.

Three American craft breweries, whereas one of them is Lock 27 Brewing Company, and two Norwegian craft breweries participated in the case study as case companies. Relevant stakeholders at each craft brewery were interviewed. An overview of key information about each brewery can be seen in table 4.1.

Brewery	**Norwegian Brewer 1**	**Norwegian Brewery 2**	**Lock 27 Brewing Company**	**US Brewery 2**	**US Brewery 3**
Brewery type	Micro	Micro	Micro	Micro	Regional
Annual production volume	60.000 L (500 BBL)	100.000 L (850 BBL)	140.000 L (1200 BBL)	240.000 L (2000 BBL)	1.400.000 liters (12.000 BBL)
Batches produced annually	65	40	40	170	250
Annual production capacity	200.000 L (1700 BBL)	300.000 L (2560 BBL)	234.000 L (2000 BBL)	351.000 L (3000 BBL)	2.000.000 liters (17.000 BBL)
Production employees	1	2	4	7	17
Fermenter vessels	13	19	4	18	20
Sales channels	Vinmonopolet, bars, restaurants and one festival	Retail stores, Vinmonopolet, bars and restaurants	Retail stores, Bars, Restaurants	Wholesale, direct to end-customer, distributor, bars and restaurants	Distributors
Tap room	No	No	Yes (2)	Yes (1)	No

Table 4.1: Overview of production in the breweries

4.1.1 Current situation at five craft breweries

At the core of the operations of the five case breweries, Lock 27 Brewing Company, Norwegian brewery 1 and 2, and US brewery 2 and 3, is a make-to-stock strategy. This approach relies on their ability to forecast demand effectively, typically using historical sales data. Intrinsic forecasting methods, such as a 2-3 month moving average, are commonly employed. However, Norwegian breweries 1 and 2 utilize reorder points, initiating new production cycles when inventory levels hit a predetermined threshold.

Seasonality and market trends also significantly influence their operations. All breweries regularly introduce new products to keep up with market trends and customer preferences. The demand for these innovative products is forecasted based on qualitative assessments and the popularity of similar products in the market. Seasonal variations are also an important factor, with higher sales recorded during the summer months, Christmas and geographically influenced demand changes, such as the start of new college semesters and the active baseball season. On a few occasions, Norwegian brewery 1 and US brewery 2 adapt a make-to-order strategy for large sales orders with acceptable lead times.

Production control and inventory management is essential aspects of their operational strategy. Production control involves monitoring the quality and material flow using manual samples and Plaato's real-time fermentation monitoring system. Inventory management targets the optimization of finished goods levels due to high carrying costs and the products limited shelf life. To meet these objectives, raw materials are ordered based on production plans and current inventory levels. Different tools, such as Brewery Management Software Systems and spreadsheets, are also utilized across the breweries to support these planning processes.

Capacity management is a crucial element of their operations. Notably, all breweries typically operate below their maximum capacity to provide flexibility and resilience against the unpredictability of the fermentation process. The transition from wort preparation and fermentation to packaging has been identified as a common bottleneck. To mitigate these challenges, the breweries incorporate spare capacity into their planning, effectively building buffer time to accommodate fermentation process variability. The duration of these buffers varies among the breweries as follows:

- **Norwegian Brewery 1:** 7 days
- **Norwegian Brewery 2:** 14 days
- **Lock 27 Brewing Company:** 7 days
- **US Brewery 2:** Dependent on each recipe. On average 4 days.
- **US Brewery 3:** 9 days for ales and 7 days for lagers

All the breweries reported to fully understand the severe implications of a batch fermenting longer than planned and hence tend to prioritize ensuring enough time rather than optimizing for efficiency.

4.1.2 Summary of PPC findings in the craft brewing industry

The aspect of capacity utilization was identified as a critical concern in the craft brewing industry during the course of all our interviews. The wish to optimize the use of costly production equipment to secure profitable returns on investment was also highlighted, but affected by the fermentation uncertainty, as mentioned in 4.1. The case breweries reported a relatively low capacity utilization; US brewery 2 showed the highest utilization rate at 80% during peak demand periods, whereas Norwegian breweries 1 and 2 reported less than 50% utilization.

The seasonal fluctuations in demand presented another layer of complexity to capacity utilization. All breweries, reflecting the broader craft brewing industry, witnessed increased demand during the summer months and Christmas, and variations influenced by localized factors. Adapting production capacity to meet demand was considered impractical, and implementing a leveling strategy was unfeasible due to the perishable nature of beer. Hence, maintaining a high capacity utilization during off-peak seasons remains a challenge.

Further, our investigation revealed that the actual capacity utilization at each brewery was lower than initially reported, given the unpredictability in fermentation duration. For instance, Lock 27 Brewing Company, which produces approximately 140,000 liters annually with a reported capacity of 234,000 liters, incorporated a seven-day spare capacity. Reducing the spare capacity to two days would increase their capacity to as much as 350,000 liters, according to their operations manager, suggesting their true capacity utilization is closer to 40%, as opposed to the reported 60%.

The planned annual spare capacity days for each case company, based on the interview data, are as follows:

- **Norwegian Brewery 1:** 910 days
- **Norwegian Brewery 2:** 280 days
- **Lock 27 Brewing Company:** 280 days
- **US Brewery 2:** 680 days
- **US Brewery 3:** 1875 days

The figures represent an approximation of the annual spare capacity days each brewery had planned, calculated by multiplying the number of spare capacity days per batch by the total batches produced annually.

A notable bottleneck occurs during the transition from batch to continuous production. The wort preparation and fermentation phases of beer production involved creating batches, ranging from 1,000 to 50,000 liters. However, the packaging process required the batch-produced liquid to be transferred into smaller containers, from 0.33 to 25 liters in capacity, which created bottlenecks in the production, thus increasing the production lead time or throughput time.

Another challenge highlighted by the case breweries was demand forecasting. Core products, being a permanent part of the product portfolio, showed stable demand and a recurring seasonal pattern. In contrast, newly launched products had shorter product lifecycle (PLC) due to vast demand fluctuations. Accurate demand forecasting for new products was problematic due to limited access to historical data and data from other supply chain actors.

The complexities of inventory management were also apparent among the case breweries. They held inventory at several stages of production. Raw materials like barley, yeast, hops, and aromas were stored before initiating the wort preparation process. Ideally, the next storage phase would involve moving the product to the finished goods storage area. However, bottlenecks in the production process sometimes necessitated an intermediate work-in-process storage. Given beer's finite shelf life, this could potentially escalate spoilage costs, thereby adding an additional layer of challenge to inventory management.

4.1.3 Summary

- The five case breweries, including Lock 27 Brewing Company, operate primarily on a make-to-stock strategy, using historical sales data and seasonality to forecast demand.
- Breweries operate below their max capacity, buffering for the unpredictable fermentation process with buffers lasting 4 to 14 days. They prioritize ensuring sufficient time for fermentation over optimizing for efficiency.
- Capacity utilization is a critical concern, with reported utilization rates less than 80% even during peak periods.
- The transition from batch to continuous production during packaging creates a common bottleneck, increasing production time.
- Breweries face challenges in demand forecasting and inventory management, especially for new products and due to beer's finite shelf life.

Chapter 5

Model development

This chapter presents the approach and results in developing machine learning models. First, the problem formulation is presented. Followed by a description of the sensor data and how it was prepared in order to fit a machine learning model. Then we present different machine learning models in addition to how they performed on the prediction task. Finally, the results and performance of the model are discussed.

5.1 Problem Formulation

Collaboration with Plaato gives us access to a large database with information about batches from several different breweries around the world. Plaato installs IoT sensors into fermentation tanks that measure the current frequency, temperature and density in the batch every 30 minutes. The values are recorded by an inline density meter that is in direct contact with the liquid through a Tri Clamp port. Advanced tuning-fork technology is used to measure the density of the liquid, by energizing the stainless steel probe to its own resonating frequency which is directly proportional to the density of the liquid (Plaato 2022). Information regarding the start and end of the fermentation, as well as the yeast used in the batch are also recorded. Data from every brewery is automatically uploaded through their local WiFi to a database that we have access to. Hence, we are able to analyze and retrieve insights about batches behavior throughout the fermentation process. An illustration of how the data is perceived by the customers is shown in figure 5.1

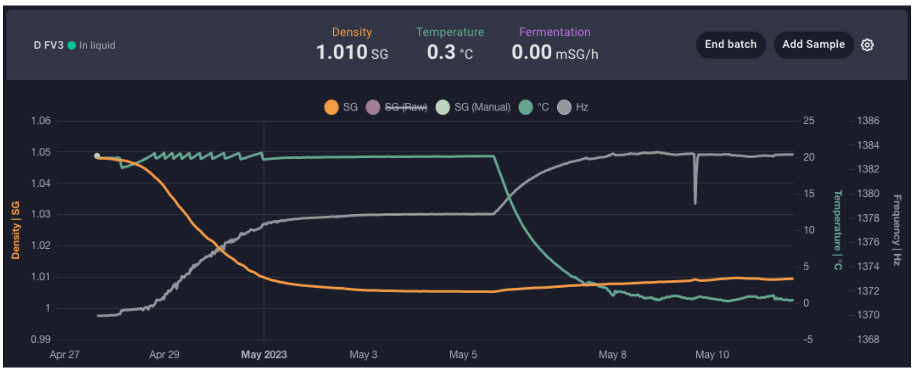


Figure 5.1: Graphs of density, temperature and frequency in batch (Plaato 2022).

As described in chapter 4 the fermentation process can vary from batch to batch which introduces a level of uncertainty for production planners. Since the time it takes to finish a batch is subject to variability, production planners keep spare capacity (Supplies 2019, Supekar et al. 2019). Hence, the brewery is not utilizing its production potential. Considering the available fermentation data we have access to combined with the existing challenges associated with fermentation we want to utilize the data to increase the predictability of the fermentation. We want to provide insights, during fermentation, about the finish time of the batch. In particular, we want to predict the finish time of a batch hours before it has finished.

In the process of beer production, density is used as a way to measure the amount of fermentable sugars in the wort, and subsequently, to determine when fermentation has finished (Thesseling et al. 2019). The density measurement, often referred to as specific gravity (SG) is typically taken before, after and throughout fermentation (Thesseling et al. 2019). The final gravity (FG) is lower than the original gravity (OG) because sugars have been consumed by yeast and turned into alcohol and CO_2, which have a lower density than sugar (Thesseling et al. 2019). If the gravity stops decreasing and remains steady, it indicates that the yeast has consumed all available sugars and that the fermentation is complete (Thesseling et al. 2019). This phenomenon can be seen as the orange graph in figure 5.1. Once the density measured over time starts flattening out, the batch is considered completed. Factors such as tank temperature, yeast type and vitality, and sugar quantity influence the trajectory of the density curve (Şener et al. 2007).

As mentioned in section 3.6, machine learning models are able to capture and learn patterns in data unrecognizable for the human eye. We therefore want to utilize the large amounts of data we have available combined with machine learning in order to reduce the uncertainty in fermentation. In particular, we want to feed machine learning models with data recorded during fermentation and predict the remaining time until the batch is fully fermented. Machine learning models may perform differently based on the data provided and the objective of the task. Hence, we will develop several of the current state-of-the-art machine learning models on a supervised learning problem and evaluate their performance. This development consists of cleaning the recorded data and making new features serve as input to the models.

The concept of turning the problem into a traditional supervised problem is mainly because of the amount of good supervised machine learning algorithms easily available. Many of the models also have identical data format requirements, which facilitates the testing and comparison of multiple models without incurring significant overhead. Within this supervised learning framework, there is a single target variable: the completion time of the fermentation. This approach greatly simplifies the process of comparing and evaluating the performance of different models. The data and machine learning pipeline can be seen in figure 5.2.

In this section, we frequently use the term *training stop*. It is the hour into the fermentation process that we stopped the training. If the *training stop* is 40, it means we used all the information recording before hour 40 in the fermentation to train the machine learning model. The terms "model" and "algorithm" are frequently used and is explaining the same concept, namely the machine learning model that is developed in order to predict the time when a beer fermentation is finished. Additionally, the terms *recordings* and *measurements* are often used, referring to the information gathered by the sensors.

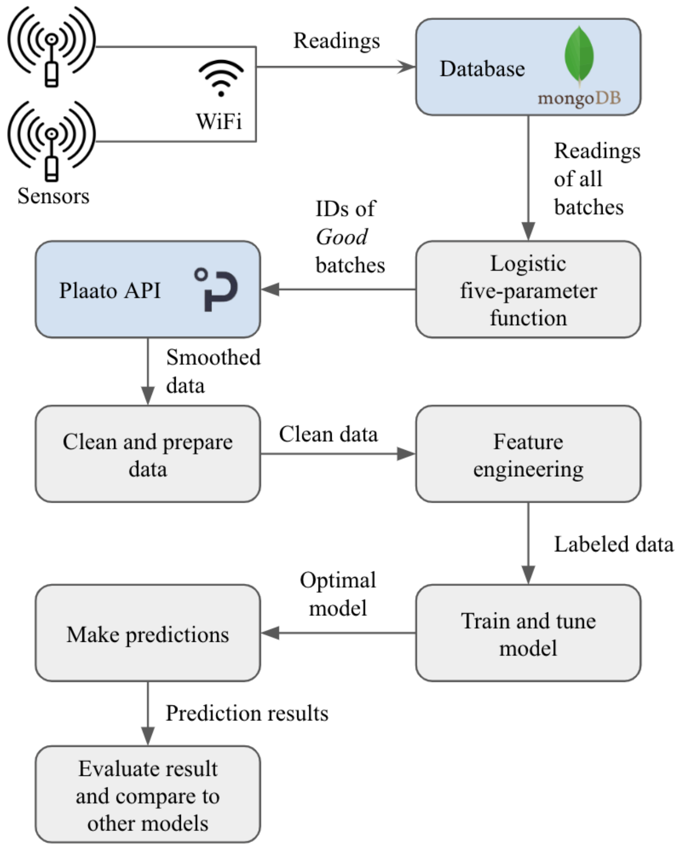


Figure 5.2: An overview of the data ingestion pipeline.

5.2 Data Description

Raw sensor data from numerous breweries are available in a MongoDB database. The database at our disposal comprises three distinct types of data which, in combination, offer comprehensive insights about a specific batch. Firstly, there's the *reading* data type, containing a reading ID, batch ID and measure of frequency, degree and density in the batch taken at the timestamp. *Readings* get taken in 30-minute intervals throughout the fermentation process. Given the presence of a timestamp, this data type qualifies as time series data described in section 3.6.5. Each reading is linked to a unique batch through the batch ID.

Reading_id	Timestamp	Frequency	Degree	Density	Batch_id
Randomly made unique identifier	Time of when datapoint was measured	Measured frequency of fork vibrations	Temperature measured	Density calculated based on frequency and temperature	Id of the batch

Table 5.1: Reading

The batch data type consists of a unique identifier, original gravity (OG), final gravity (FG), alcohol by volume (ABV), attenuation, start, end and a recipe id. OG is the measure of density in the batch before the fermentation begins. FG is the density level in the brew when the fermentation activity has declined and stabilized. ABV is a measure of the amount of ethanol contained in the beer. It is calculated by using the measured OG and FG. Attenuation is a measure of how much of the sugar in the wort is converted to alcohol and carbon dioxide during fermentation (Trelea et al. 2001). Every individual reading with the same batch identifier will be connected to the same batch. The recipe identifier is a link to the last datatype, the recipe. The recipe's most important feature is the ingredients, which give information about the yeast type used in the batch. However, it is optional for breweries to add ingredients and this information is therefore not always available.

Batch_id	FG	OG	ABV	Attenuation	Start	End	Recipe_id
Randomly made unique identifier	Estimated final gravity of the batch	Estimated density at start of fermentation	Alcohol by volume	The amount of sugar converted to alcohol and CO_2	Start time of the batch	End time of the batch	Id of the used recipe

Table 5.2: Batch

Recipe_id	Name	OG	FG	Ingredients
Randomly made unique identifier	Name of the recipe	Preferred original gravity	Preferred final gravity	All ingredients used in this batch

Table 5.3: Recipe

Each reading uploaded to the database has a timestamp and a batch id. This information makes it possible to arrange data according to specific batches and when the fermentation has occurred. The data structure in the database enables the execution of aggregation queries. For instance, one can extract data that contains all the measurements for batches that use a particular type of yeast. This facilitates comparative analyses of the fermentation process across different batches from different breweries.

We decided to retrieve all batches recorded from all breweries, regardless of yeast type or geographical location, from 2022 and onward. The physical location of a brewery will not significantly affect the fermentation process and if we want to analyze batches with different yeast types, this can be sorted afterward. In 2022 Plaato developed new hardware for their fork sensors, which makes the measurements more reliable and accurate. Even though the database consists of readings from 2019 to this date, we decided to exclude the older batches with bad accuracy. Older and newer readings could be analyzed separately, but for simplicity and to easier evaluate performance we decided to focus on newer batches with more accurate readings. We retrieved a total of 1406 batches. A visualization of the temperature and density is given in figure 5.3. In the density curve, one can see the noise in the graph due to inaccurate readings taken by the sensors. Nevertheless, the trajectory of the curve is still visible.

Batch filtering with five-parameter logistic function

A machine learning model learns patterns from the inputted data. In our approach, we are going to develop models that predict upcoming events. In order for the predictions to be accurate we therefore have to ensure that the data is reliable and without too much variation. The quality of the data delivered by the sensors can consist of much noise and have batches with abnormal deviation. Factors such as changing environmental conditions, interference from nearby electronic devices and physical disruptions can produce unreliable data (C. Wang et al. 2015). The sensors are dependent on consistent WiFi connection through the fermentation process, in order to upload the recorded data to the database. Disruptions in the WiFi connection result in batches with gaps and errors. Many of the batches recorded are taken from breweries that are not yet familiar with the usage of sensor technologies, which in some cases causes user errors. The recording for certain batches is not started at the same time as the fermentation actually starts. The same problem occurs when the batch ends, leaving for instance the sensors turned on in liquid while the

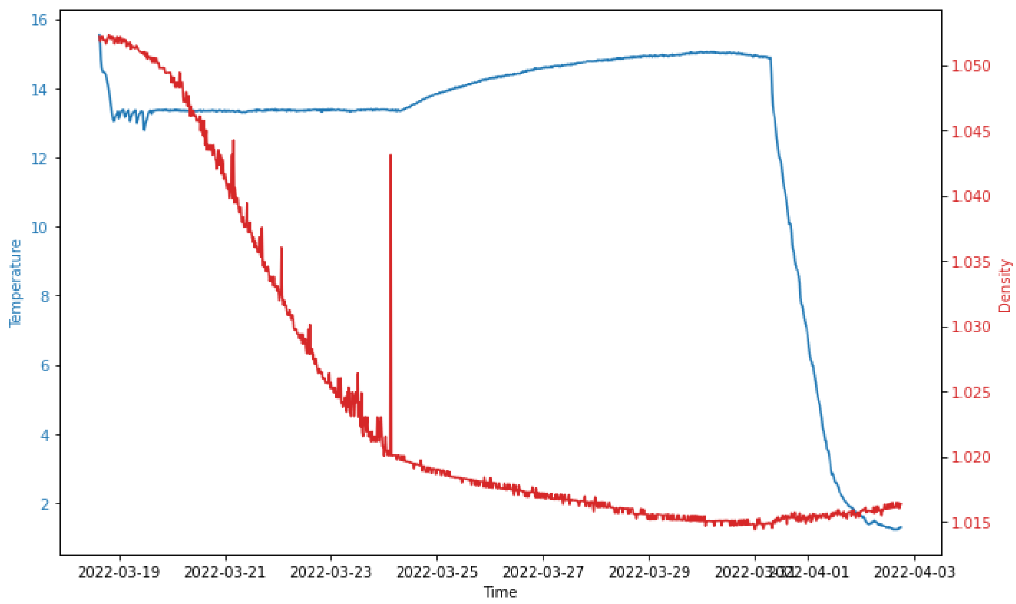


Figure 5.3: A visualization of the development of temperature and density in a batch.

batch actually is finished. This causes several of the retrieved batches to be unfavorable to use for machine learning model training.

In order to compensate for deviations made by the sensor technology and user error we have implemented a five-parameter logistic function to assure that all batches used for model development are somewhat similar and of good quality. We have taken inspiration from Speers et al. (2003) and Reid et al. (2021) which developed logistic functions in order to forecast fermentation activity. Instead of checking the quality of all batches manually, we fit a five-parameter logistic function to each batch:

$$f(t_i) = p_e + \frac{p_i - p_e}{\left(1 + s \cdot e^{-B \cdot (t_i - M)}\right)^{\frac{1}{s}}} \tag{5.1}$$

The function outputs the density level at time t_i. In the equation, M is the point where the rate of extract decline is maximal. P_i is the upper asymptote and P_e is the lower asymptote, where time is approaching ∞ (Reid et al. 2021). B is a representation of the line which is the slope at M. The S in this model describes the asymmetric behavior and is an adjustable parameter that permits fitting asymmetric attenuation curves such as those demonstrating a substantial lag or slow finishing attenuation (Reid et al. 2021). A visualization of how the parameters are retrieved is shown in figure 5.4.

In order to find the optimal parameters for each batch we use a Python function called *curve fit* (*Scipy - Curve Fit* 2023). The code can be seen in appendix C.1. We start with random initial guesses for the parameters s, B and M. p_i is calculated by taking the median of the first four readings of the batch. p_e is the median of the four smallest measured densities in the batch. It then calculates the residual sum of squares (RSS), which is the difference between the actual data and the model's predictions. The algorithm then uses gradient descent to find a new set of parameters that will reduce the RSS. This process is repeated until the parameters stop changing significantly for each turn. After the function is optimized for a batch, we evaluate the estimated error of the parameters. Following the criteria in equation 5.2, we determine the quality of the batch. The criteria are retrieved by experts in Plaato that have empirical evidence regarding the optimal distribution of good, bad and mediocre batches. As seen in table 5.4, the algorithm categorizes 492 batches as good. In this study we will only continue with batches characterized as good.

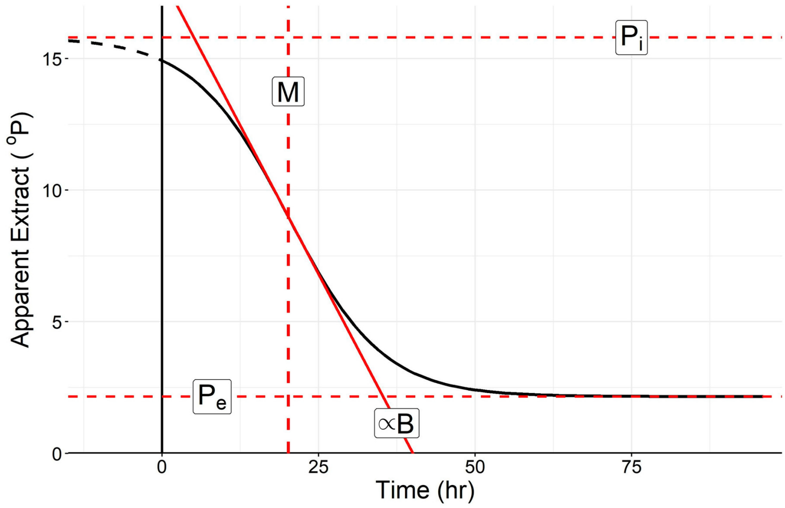


Figure 5.4: A visualization of the parameters in the logistic function (Reid et al. 2021). The *Apparent Extract* is the density in the brew.

$$\text{Batch} = \begin{cases} \text{Bad}, & \text{if } B_{\text{err}} > 0.000025 \lor M_{\text{err}} > 5000 \lor s_{\text{err}} > 1.5, \\ \text{Mediocre}, & \text{if } 0.000025 > B_{\text{err}} > 0.00001 \lor 5000 > M_{\text{err}} > 3000 \lor 1.5 > s_{\text{err}} > 0.75, \\ \text{Good}, & \text{otherwise.} \end{cases} \tag{5.2}$$

Good	Mediocre	Bad	Could not retrieve	Total
492	180	549	185	1406

Table 5.4: Overview of the amount of batches in each categorization.

5.3 Data preparation

Even though we only have retrieved the batches considered as *good* there still exists noise and gaps in the data. The ID of good batches are sent to Plaato's API which returns processed information about the batch. The processing consists of two main steps. First, the API removes spikes in the data by removing outliers that deviate greatly from surrounding values. This is typically caused by particles and bubbles passing through and sticking to the sensor. Additionally, the API refines the data by considering both preceding and subsequent information. The value at each point is adjusted according to the trend. This procedure smooths out differences caused by smaller particles and bubbles in the liquid, as well as the natural variations in density. The smoothed data is more suitable for machine learning as it reduces the noise, removes the outliers and makes the batches more similar.

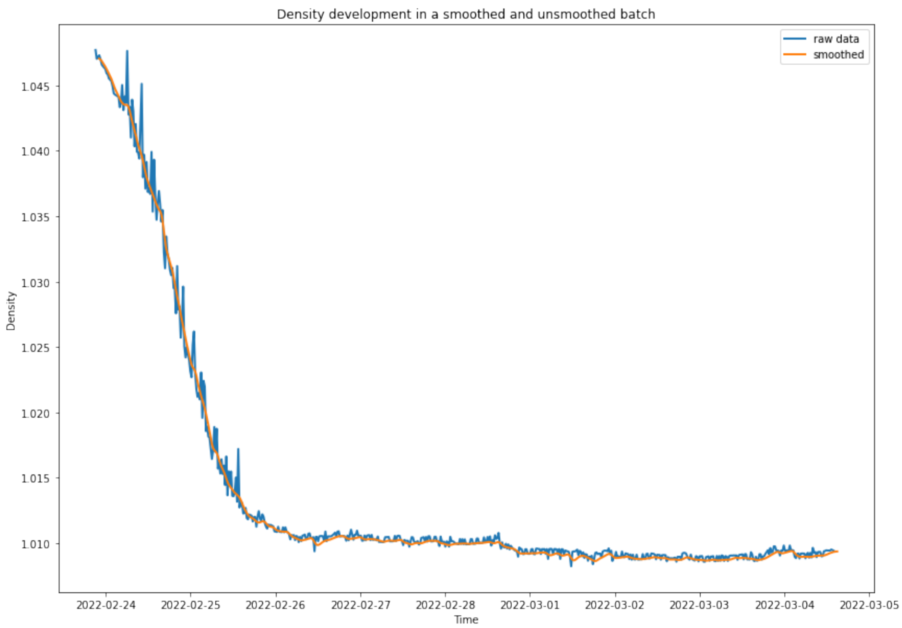
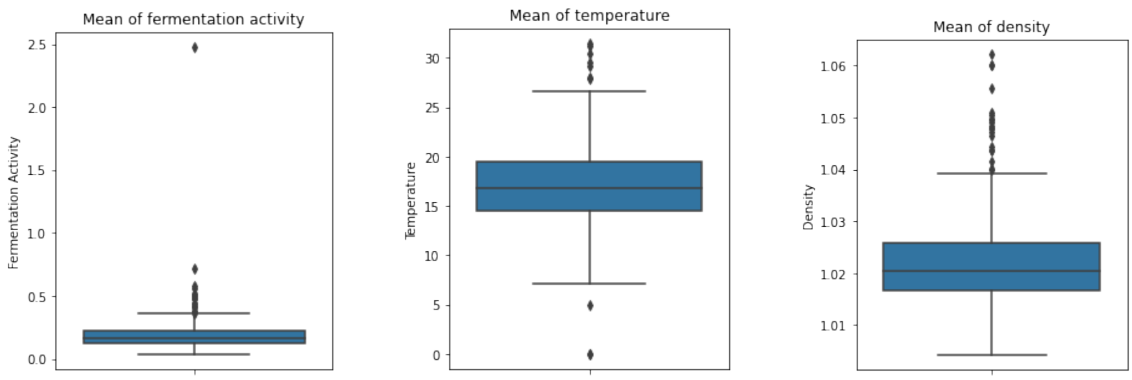


Figure 5.5: Visualization of how the data get smoothed.

Furthermore, we transformed the timestamp associated with each reading to denote the duration of fermentation, thus replacing the specific timestamp with the elapsed time into the batch, expressed in hours. This was done in order to compare batches depending on how long they have fermented. In addition, we added a new feature called *fermentation activity* which is the derivative of the density at each point. The derivative indicates the rate at which the density is changing which provides valuable information regarding the end time of the fermentation process.

5.4 Data Cleaning

It occurs that the sensors record readings with inconsistent or missing information. Some readings, therefore, have *None* values. The missing values are replaced by using linear interpolation, Which assumes a constant rate of change between two points, allowing it to predict absent values. When interpolating from both sides, we use the adjacent known values on both sides of the missing value, ensuring a more informed estimation that takes into account trends from both directions. We remove batches entirely if they have more than 40 readings without information.



(a) Fermentation activity (b) Temperature (c) Density

Figure 5.6: Boxplots of the mean values at each batch.

To clean the data further we identified several outliers through the boxplot visualizations, seen in figure 5.6. We removed batches that had mean values that deviated significantly from the interquartile range. Outliers can influence the models' learning process and may cause misleading predictions and interpretations. By removing the outliers we are later able to create a more robust model that learns from data that represents the majority.

Because of user errors, certain batches experience a delayed initiation. This primarily occurs when the brewers prematurely activate the sensors, substantially prior to the onset of the fermentation process This implies that the fermentation activity remains static for several hours after the fermentation process has started. In order to compensate for this, we identified all the batches with a slow start and removed the beginning of their fermentation process. In particular, we removed all the readings that occurred before the fermentation activity reached a certain threshold. Our goal is to understand the fermentation curve. If each batch begins at varying times, it would pose significant challenges for the model in learning the underlying patterns.

In our study, we chose to remove the latter part of the fermentation curve. We did this because, after the density curve levels off, there is not much useful information being provided. This last part, where the density barely changes, could add unnecessary noise to our model without offering helpful insights into the fermentation process. It also requires additional data storage to save and apply calculations to this data. Furthermore, it is a common practice among brewers to initiate a *cold crash* at the culmination of the fermentation process, resulting in alterations to the fermentation activity (Lordan et al. 2019). It is important to exclude this forced change from our models, as it is not a universal phenomenon across all batches but rather a brewer-specific intervention (Lordan et al. 2019). Specifically, the readings at the end of the fermentation that has activity below a certain threshold are removed. The code is available in the appendix. It is very important to note that the fermentation process is not finished, just because we removed the tail - the brew is still in the tank for a couple of days, but in terms of machine learning this part is irrelevant.

After the process of cleaning the batches, we are left with 406 good batches that can be further used for machine learning. An illustration of how the fermentation activity looks in some arbitrary batches after the cleaning process has been applied can be seen in figure 5.7. The threshold is used to identify when the batch is finished. As indicated, once the batch's metrics fall to the threshold and persist below this level for an extended duration, we can assert that the batch has reached its endpoint.

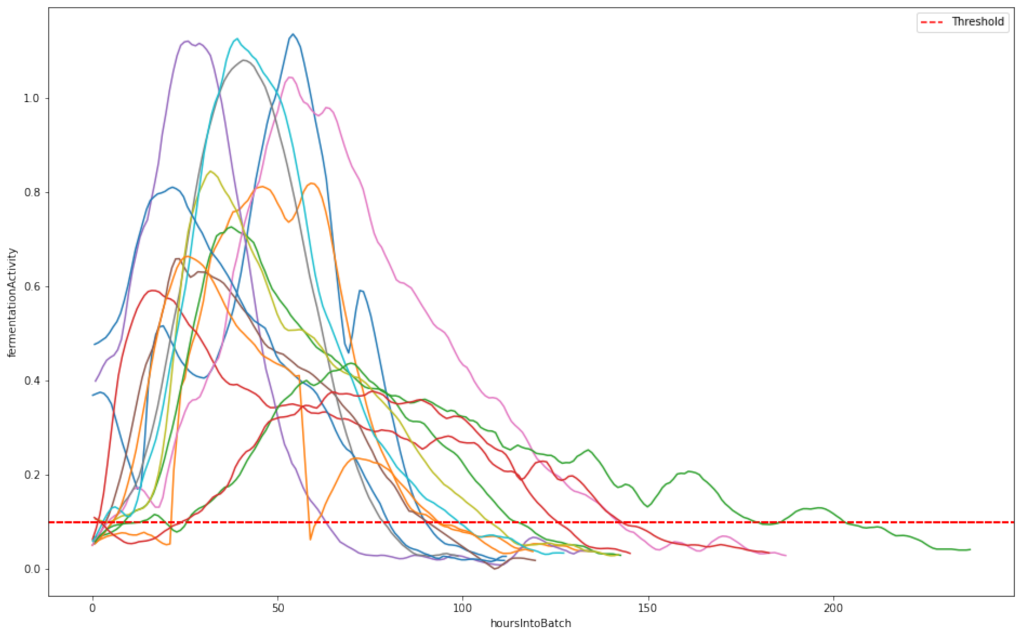


Figure 5.7: Fermentation activity in random batches after cleaning has been applied.

5.5 Feature engineering

The sensors record readings every 30 minutes over the duration of several days. We are not interested in every measurement as the density in the brew does that change that frequently. Therefore, we aggregated the readings into hourly averages. This reduces the dimensionality of the data. Even though a certain degree of information is lost by aggregating, our primary objective remains to determine when a batch is complete. From a production planning and control perspective, an hourly estimation of batch completion is sufficient.

The target variable, that we are going to predict on not finished batches, is defined as the exact hour the fermentation process ends. The definition of when a batch ends may vary from brewer to brewer and the information is not available in the retrieved dataset. The hour of completion is determined as when the fermentation activity is below 0.1 for 12 consecutive readings. The fermentation activity must have reached 0.3 once in order to have a completion time. A histogram of the target variable can be seen in figure 5.8. The lowest possible finish time in the dataset of batches is after 50 hours and the highest is after 241 hours.

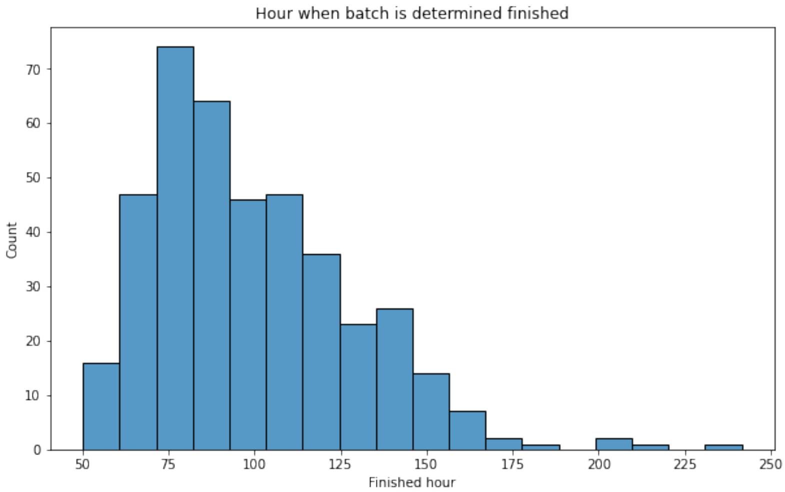


Figure 5.8: Histogram of the target variable, the finish time.

For each batch, we computed an additional feature which is the average temperature at the beginning of the fermentation process. This measure allows us to categorize the batch into either lager or ale. This classification is important because lagers and ales are brewed at different temperatures, which causes the density curve to vary between the types. With this feature, an ML model can determine what type of batch is being brewed. In particular, we categorized the brew as an ale if the average of the first 20 readings was above 15. Otherwise, we classified the brew as a lager.

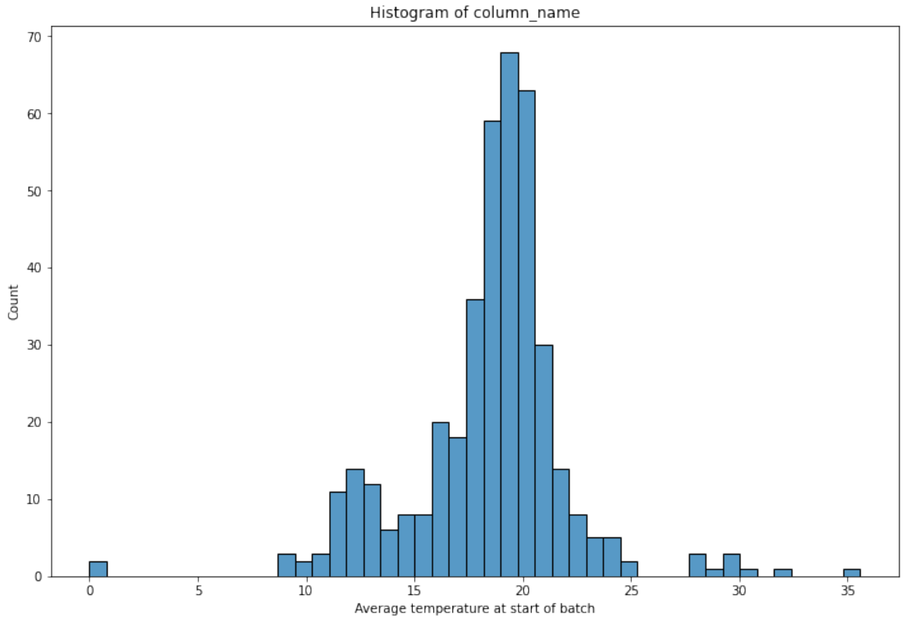


Figure 5.9: Average temperature at the start of each batch.

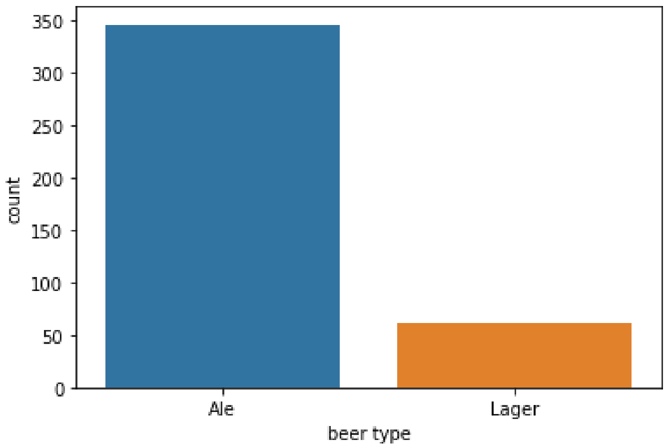


Figure 5.10: Amount of different beer types.

Another developed feature was the hour into the fermentation process where each batch had the highest fermentation activity. This feature, marking the point where the density changes the quickest provides important information about the future trajectory of the fermentation process in the subsequent hours. In our machine learning model, we also included prior density, fermentation activity and temperature from earlier readings. An illustration can be seen in figure 5.11. The blue filled area represents the amount of data used to train the model and the red cross represents the target variable, which is when the batch is finished.

All the features, except the target variable, were scaled by using Scikit learn's StandardScaler library (Pedregosa et al. 2011). It modifies all the values so that they have a mean of zero and a standard deviation of one. Scaling is important in machine learning because algorithms perform better when the numerical input variables are scaled to a standard range (Pedregosa et al. 2011). This primarily helps to prevent variables with larger scales from dominating over those with smaller scales in the computational algorithm, resulting in a fair contribution from all features which can improve the model's accuracy (Goodfellow et al. 2016).

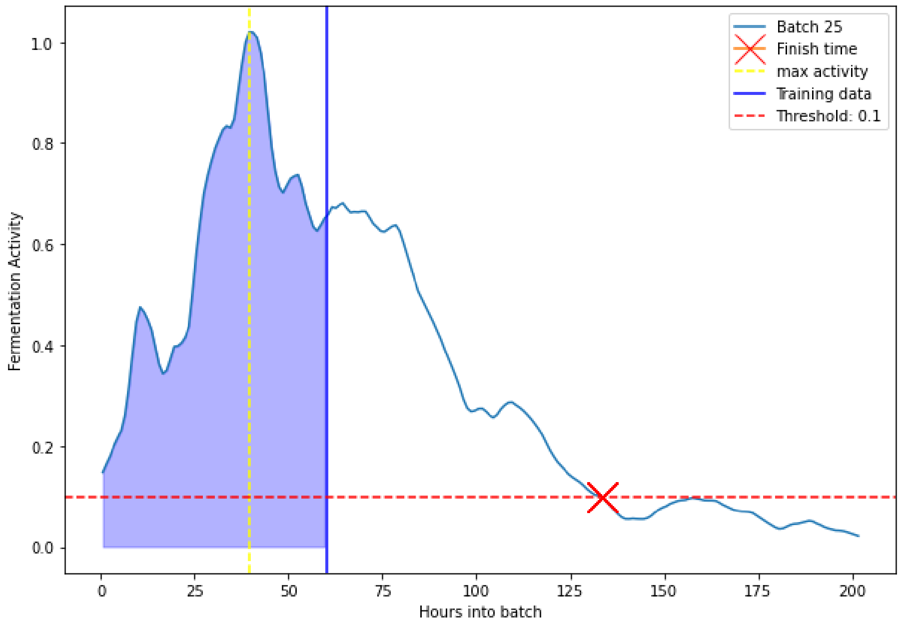


Figure 5.11: Training data with a marked target value.

5.6 Model Descriptions

This section presents a brief description of each of the models applied. Each model's functionality and benefits are described. However, we do not go into detail in terms of the mathematical operations and in depth of their most suitable areas. The section also presents how we found the optimal hyperparameters and their value that gave the best results. For the models to be comparable, we trained them on the exact same data, by using the same seed for random operations. This way, the same splits are made across the algorithms. A validation set, consisting of 10% of the training data is utilized to analyse how the model fits unseen data in the training process. 20% of the original batches are saved to test the accuracy of the models on unseen data.

5.6.1 XGBoost

XGBoost is short for extreme gradient boosting and is a powerful machine learning algorithm based on the gradient boosting framework (T. Chen et al. 2015). It is widely used by data scientists to achieve state-of-the-art results on regression and classification tasks (T. Chen et al. 2015). XGBoost creates a sequence of multiple decision trees that learn and correct errors made by its predecessor (T. Chen et al. 2015). As mentioned in section 3.6.7 it is a boosting model that combines the predictive power of multiple weaker models, in order to create one powerful predictor.

The key parameters in an XGBoost model are, according to (T. Chen et al. 2015):

n_estimators: The number of gradient boosted decision trees to be constructed. Too many trees result in overfitting, while too few result in underfitting. The default value is 100.

max_depth: This parameter controls the maximum depth of each tree. It can be tuned to control the complexity of the model.

learning_rate: This parameter is the scaling factor applied to the correction made by each tree. It controls the step size in the gradient procedure.

There are other hyperparameters in an XGBoost model, such as the maximum number of leaves in each tree, minimum loss reduction required to make a further partition on a leaf node of the tree and gamma and how many parallel threads to be used. However, to reduce the complexity and solution space of the problem we restricted the hyperparameter tuning to n_estimators, max_depth and the learning_rate.

To find the optimal hyperparameters we used RandomizedSearchCV from the scikit-learn library. That uses cross-validation with random combinations of the parameters. To compare the results of different combinations of hyperparameters, RMSE was used. The optimal parameters can be seen in table 5.5. In figure 5.12 the RMSE for each epoch during the training of the model can be seen. The error rates in both the validation and the training sets show similar initial decreases. However, the decline for the validation error stops at approximately 75 epochs.

Training Stop	n_estimators	max_depth	learning_rate
40	200	6	0.05
60	100	3	0.05
80	200	4	0.05

Table 5.5: Overview of the optimal hyperparameters for each training set for XGBoost.

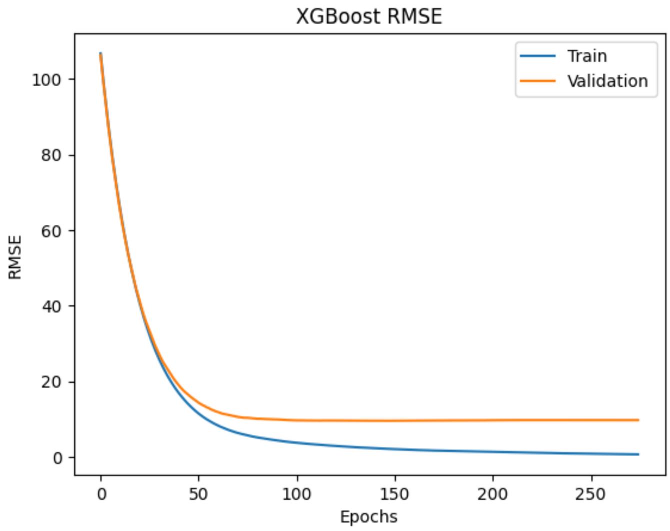


Figure 5.12: Training and validation loss with XGBoost. Trained on the first 60 hours.

5.6.2 CATboost

CatBoost, similar to XGBoost, employs gradient boosting on decision trees. In CatBoost the trees become balanced or symmetric, which means the splitting condition is consistent across all nodes at the same depth of the tree (Wong 2020). XGBoost however, makes asymmetric trees, which means that the splitting conditions differ for each node across the same depth (Wong 2020). CatBoost also employs ordered boosting, which focuses on reducing bias and variance, leading to less overfitting and better accuracy.

The most important features of CatBoost are the following (*CatBoost* 2023):

depth: Depth of each tree.

l2_leaf_reg: Ciefficient at the L2 regularization term of the cost function.

iterations: The maximum number of trees that can be built. Similar to n_estimators.

learning_rate: The learning rate of the gradient descent optimizer. Same as in XGBoost.

CatBoost has several other hyperparameters such as which loss function to be used, bootstrap type, bagging temperature and sampling frequency (*CatBoost* 2023). However, for simplicity, we excluded them in the tuning and focused on the most dominant ones. We applied Randomized-SearchCV from the scikit-learn library, which resulted in the optimal parameters seen in table 5.6. As can be seen in figure 5.13 the models training loss is declining, but it is struggling to reduce the validation error after 30 iterations. This means the model is overfitting the training data.

Training Stop	depth	l2_leaf_reg	iterations	learning_rate
40	5	3	250	0.1
60	4	3	300	0.1
80	4	2	300	0.1

Table 5.6: Overview of the optimal hyperparameters for each training set for CatBoost.

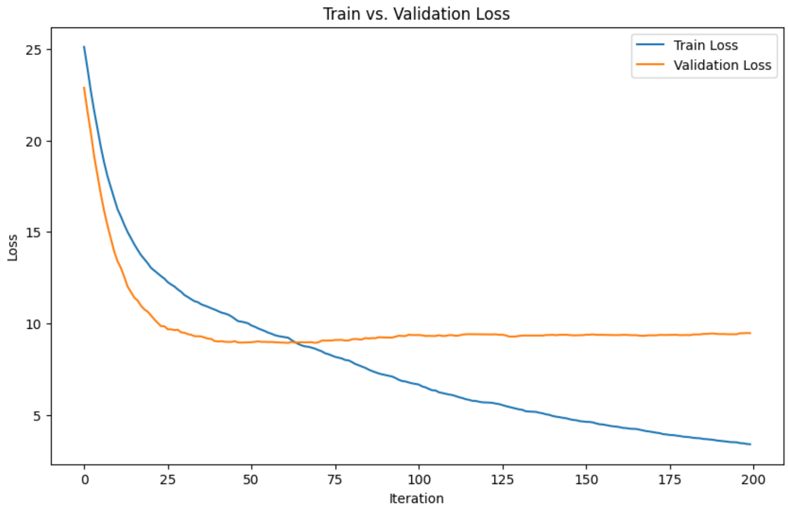


Figure 5.13: Training and validation loss with CatBoost. Trained on first 60 hours.

5.6.3 Artificial neural network

Artificial neural networks (ANN) form a major class of models widely used the recent years in machine learning. As mentioned in section 3.6.6, they are inspired by the human brain by mimicking the way the biological neurons communicate. The networks have revolutionized the field of artificial intelligence (Sharma 2023). They have shown remarkable success in image recognition, self-driving cars, stock market predictions and medical diagnosis (Sharma 2023). ANNs are typically used in supervised learning (Dave and Dutta 2014). Their flexible, adaptive nature allows them to learn and model non-linear and complex relationships, which is a major advantage over traditional linear models (Haykin 2009). The main disadvantage of ANNs is that it is challenging to interpret how the model has achieved a certain output (Haykin 2009). They are also very prone to overfitting on the training data if the hyperparameters are not tuned correctly (Haykin 2009).

Their most important features are, according to (Hovden 2019):

Number of hidden layers: Number of layers between the input and output layer. Too few may lead to underfitting, while too many leads to overfitting.

Number of neurons in each layer: The number of neurons in each hidden layer. Similar to the number of layers, too many neurons cause overfitting and too few cause underfitting. Overfitting occurs when the neural network has so much information processing capacity that the limited amount of information contained in the training set is not enough to train all of the neurons in the hidden layers (Heaton 2017). Many neurons also increase the time it takes to train the network.

According to (Heaton 2017), the number of neurons in the hidden layers should be between the size of the input layer and the size of the output layer. Additionally, the number of hidden neurons should be 2/3 the size of the input layer, plus the size of the output layer (Heaton 2017).

Learning rate: This is a crucial parameter for the optimization algorithm that determines how much the weights are updated during training. A smaller learning rate might require more training epochs, whereas a larger learning rate might lead to unstable results. (Brownlee 2023b)

Activation function: Decides whether a neuron in the network should be activated or not. It will decide whether the neuron's input to the network is important or not in the process of prediction using simpler mathematical operations (Sharma 2023).

Epochs: An epoch is a single pass through the network with the entire training set. The number of epochs determines how many times the learning algorithm will work through the entire training dataset (Brownlee 2023b). A large network with many hidden layers and neurons requires more epochs in order to fit the data.

To implement a neural network we utilize Keras, a powerful open-source neural network library written in Python. Keras provides a convenient and flexible toolset for building and designing different deep learning models (Keras 2023). We use Kera's sequential model which allows us to create models layer-by-layer, but with the limitation that it is not possible to share layers or have multiple inputs or outputs. The loss function used to optimize the weights in the neural network is mean squared error. As an optimizer, which calculates how the weights should be changed, we applied Adam deep learning optimizer. It updates the learning rate for each network weight individually and uses the second moment of the gradients as well when adapting the learning rates.

In order to hyperparameter tune our neural network we applied a random search tuner. The objective of this search is to minimize the validation loss, which is the loss the model has on a validation set of 10 % of the training data. The search is set to run for a maximum of 50 trials with 5 executions per trial. This random search strategy tries out a fixed number of hyperparameter settings sampled from defined probability distributions. After several attempts of adding additional layers to our network, we concluded that it is more efficient to only tune the number of neurons in each layer and the learning rate. The accuracy of our model did not improve with additional hidden layers, and we, therefore, excluded this in the hyperparameter tuning. This resulted in limitation of search space, which made the hyperparameter search more efficient. The code can be seen in appendix C.4. The optimal hyperparameters for each set of training data can be seen in table 5.7.

Training Stop	**neurons 1st layer**	**neurons 2nd layer**	**learning_rate**
40	300	250	0.01
60	120	64	0.01
80	200	150	0.01

Table 5.7: Overview of the optimal hyperparameters for each training set for a neural network.

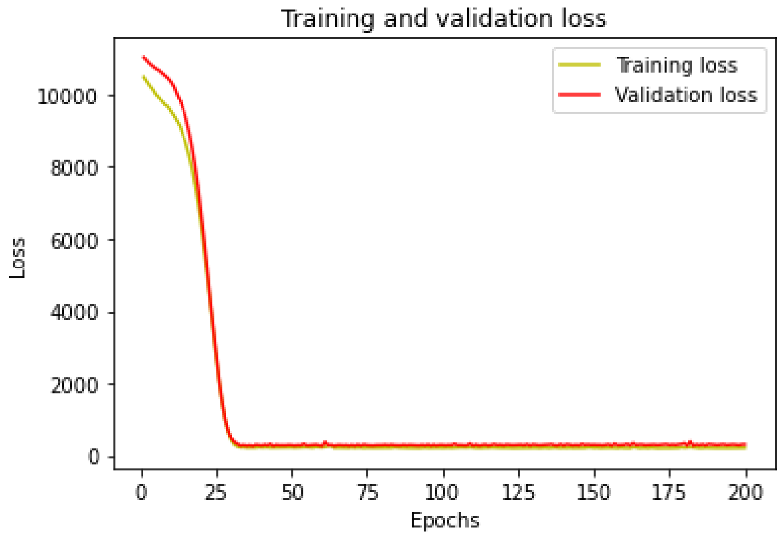


Figure 5.14: Training and validation loss with a neural network. Trained on first 60 hours.

5.6.4 h2o automatic model

As described in section 3.6.9, H2O AutoML is a highly efficient automated machine learning tool developed by H2O.ai. It's designed to abstract the complexity of machine learning pipelines, allowing users with varying levels of expertise to develop and deploy highly accurate predictive models (H2O auto 2022). H2O AutoML uses a series of machine learning algorithms to train a variety of models.

In our application of the h2o autoML library we only need to provide a dataset and a target variable and the tool automatically handles tasks such as data preprocessing, feature engineering, model selection and hyperparameter tuning (H2O auto 2022). After several models have been trained, tuned and ensembled, the library returns a leaderboard of the models that achieved the best result according to a set metric. We used root mean squared error as a metric, which resulted in the optimal models in table 5.8. The names of each model make it possible to retrieve and apply only that model.

Training Stop	Optimal h2o model
40	StackedEnsemble_BestOfFamily_1_AutoML4_20230604_210808
60	GBM_3_AutoML_1_20230604_205814
80	StackedEnsemble_BestOfFamily_1_AutoML_1_20230604_211405

Table 5.8: Overview of the optimal hyperparameters for each training set for a neural network.

5.7 Results

After the models had been trained and the hyperparameters tuned on the training data, we made predictions on a separate test set. We trained the models on three different datasets consisting of

40, 60 and 80 hours of information about fermentation. For each dataset we calculated the R^2, RMSE and the average amount of hours the predictions are deviating from the target variable. The R^2 score is a measure used to determine the proportion of variance in a dependent variable that can be predicted by an independent variable (Davison 2003). R^2 values range from 0 to 1, with a value of 1 indicating that the model perfectly predicts the outcome variable, and a value of 0 indicating that the model does not explain any of the variability of the outcome data around its mean (Davison 2003). The R^2 score does not penalize model complexity.

As an additional metric to evaluate and compare the models' performance, we also used the root mean squared error (RMSE). It is the squared root of the average of the squared differences between the predictions and the actual values. RMSE penalizes large errors more heavily than small ones, as a result of the squaring of the differences (Russell 2010). RMSE is therefore a useful metric when particularly large errors are undesirable. RMSE also has the same units as the variable we are predicting (Davison 2003). In other words, the RMSE represents the average amount of hours we are deviating from the target variable. An RMSE of 20 tells us that the model's predictions are deviating from the target variable, which is the hour of the finish time of a batch, with an average of 20 hours. In terms of PPC, the most valuable insight from the fermentation process is information regarding when a batch is finished and how accurate this prediction is. Therefore, we use RMSE as the main metric in identifying the model with better performance.

To get a better understanding of how much additional insight our model performs we create a baseline prediction to compare our predictions with. The baseline model predicts the average of the target variable on all data in the training set on the test data. This is a traditional approach, that our model should aim to outperform.

Dataset	**Model**	R^2	**RMSE**	**Baseline hours off**
40 hours	XGBoost	0.706	14.5	21.67
	CatBoost	0.737	13.87	
	ANN	0.6623	15.659	
	H2o	0.605	17.08	
60 hours	XGBoost	0.665	13.3	20.73
	CatBoost	0.719	13.388	
	ANN	0.654	14.8	
	H2o	0.71	14.635	
80 hours	XGBoost	0.7499	11.642	20.8
	CatBoost	0.831	9.9	
	ANN	0.705	12.78	
	H2o	0.818	9.533	

Table 5.9: Summary of model performance.

In table 5.9 an overview of the model's performance is given. The best models for each dataset are highlighted. For the models that were trained on the first 40 hours of information, the models' prediction varies from 13.87 to 17.08 RMSE. The best model for this dataset is the CatBoost with 13.87, beating the baseline model with almost 8 hours. For the dataset consisting of information from the 60 first hours of fermentation, the XGBoost had the best performance in terms of RMSE, but the CatBoost model had a better R^2 score. Therefore the performance of the XGBoost and the CatBoost model are evaluated as very similar. With 20 hours of additional information, the model only predicted 0.5 hours better compared to the first dataset. For the dataset with 80 hours of information, H2o had the best model with an RMSE of 9.533. This model had over twice as good predictions as the average baseline model. The CatBoost model only performed 0.4 hours worse than the H2o model. Overall the CatBoost model shows the best ability to adapt to different datasets which resulted in the best performance across all datasets.

In figure 5.15 we can see a scatterplot of predictions and the actual value of a CatBoost prediction with 60 hours of information. The X-axis is the actual values and Y-axis is the predicted value.

The red diagonal line represents the optimal pattern the dots should lie on if the predictions were perfect. From the figure, we can tell that the predictions become less accurate the longer time the beer is fermenting. In particular, if the batch is finished fermenting after 100 hours, our model is struggling more to recognize the pattern and make accurate predictions. The scatter plot also identifies some predictions that are very far away from the real value. These predictions could be analyzed further to identify why the model is struggling to learn from them. One reason is that the training data from these predictions deviate significantly from other batches, which means they could be marked as outliers and removed from the training data. The plot shows that model is able to predict accurately if the batch is finished within 60 to 100 hours.

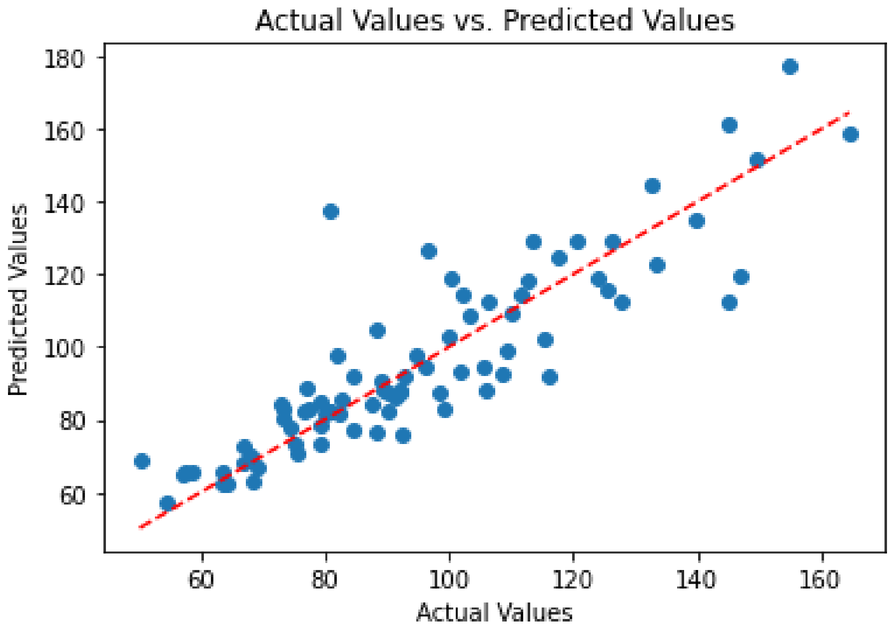
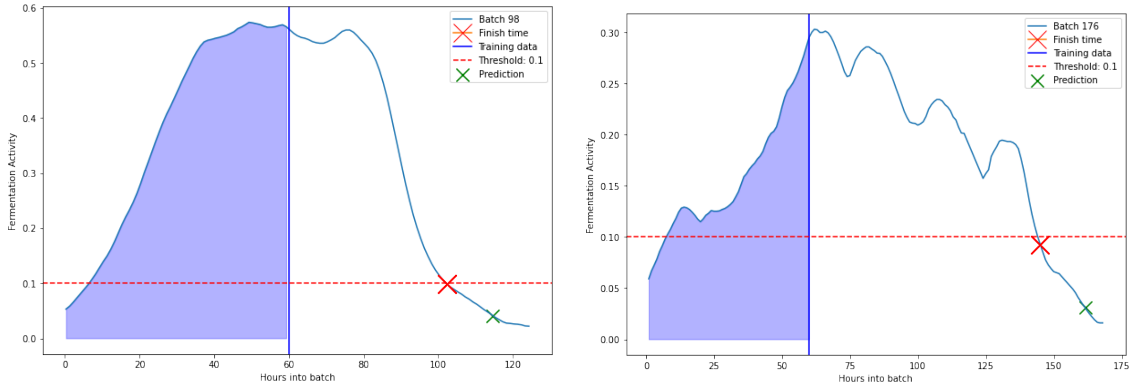


Figure 5.15: Scatterplot of actual vs predicted value from a CatBoost model.

In figure 5.16 we have illustrated how the training data and the prediction looks like from a CatBoost model. The red cross represents the actual time the batch is finished and the green cross represents the prediction made by the model. If we look at the blue part of each graph, which is the information used for training the machine learning model, both graphs act quite similarly, with slightly different fermentation activity values on the Y-axis. However, after 60 hours of fermentation, the curve of the fermentation activity deviates from each other. In the left graph, it quickly declines and the fermentation stops at around 100 hours. On the right, the graph descends much slower and finishes around 150 hours. This illustrates the challenges of learning from prior fermentation data to predict the coming hours. It is challenging for the model to identify that the batch on the right will take almost 50 hours longer to finish than the left one when the training data is so similar.



(a) Prediction of batch 25 (b) Prediction of batch 176.

Figure 5.16: Deviations in how the training data look like.

Figure 5.17 displays the performance of a model when the number of batches available changes. The y-axis represents the RMSE. The X-axis represents the number of batches used to train the model. A CatBoost model was used for this experiment. We can see that error decreases significantly from training on 10 to around 100 batches. When more than 100 batches are used for training the error starts jumping between 17.5 and 12.5 hours. While the error does not decrease significantly when more batches are being used for training, the model becomes much more robust and it will create more accurate predictions across varying datasets. Furthermore, when more batches are used for training, there is also more data to learn from. As more batches are added, the model must adapt to new training data which does not necessarily look like data used for training earlier.

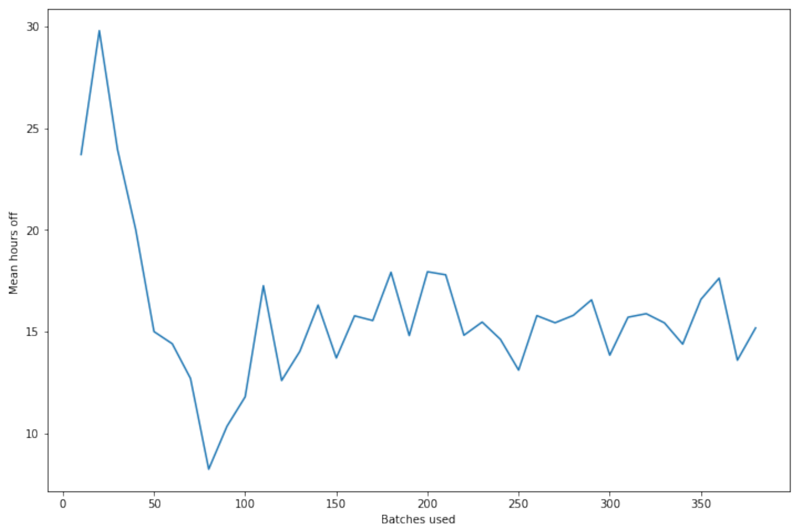


Figure 5.17: The mean hours off when training on different amount of batches.

5.8 Evaluation of Model Development

In the following section, we discuss and compare the results obtained from the implementation of our machine learning models. We explore the reliability and value gain of the predictions and their practical utility. This leads to a discussion regarding the quality of the data and the challenges with data gathering.

5.8.1 Model performance

Overall the performance of the models was quite similar. All models were able to perform better than the baseline prediction, with CatBoost having slightly better overall accuracy than the other ones. XGBoost and CatBoost had similar RMSE with a deviation of just a couple of hours. Both models have an equivalent approach to machine learning with the usage of boosted gradient descent with several decision trees. However, CatBoost is more modern and has more optimal default hyperparameters and a built-in approach to avoid overfitting. It also handles categorical features, which XGBoost does not. This explains CatBoost's slightly better performance.

Despite the RMSE being only slightly worse for the ANN compared to the gradient boosted models, the performance of the ANN is disappointing. The main advantage of ANN's is to capture complex patterns unnoticeable by the human eye. Our model was not able to capture the pattern of the fermentation process satisfactorily. ANN's performance is heavily influenced by how its parameters are tuned (Feurer and Hutter 2019), which may explain the bad performance. Despite implementing a random search optimizer to find the optimal number of neurons and the optimal learning rate, we could have explored the structure of the ANN more. We implemented a simple one-hidden layer sequential neural network. There are several alternative neural networks such as recurrent neural network (RNN), long short-term memory (LSTM) and autoencoders that have shown great results in learning and prediction on complex time series (Manaswi and Manaswi 2018). We did develop a transformer-based encoder-decoder model inspired by Hugging Face's implementation (Platen 2023). However, the model was not able to learn from multivariate, multiple time series in order to predict one outcome.

The H2o automatic machine learning framework had mediocre performance when training on 40 and 60 hours, but showed slightly better results than the CatBoost model when training on 80 hours of fermentation. The implementation of the automatic machine learning framework was to be able to detect whether a model could give good predictions, without hyperparameter tuning errors or wrongly designed models. The results show that the automatic framework does not perform significantly better than our implemented models. However, it gives insights regarding what type of machine learning approach gave the best predictions, such as a stacked ensemble of trees and gradient boosting machine models. Even though the framework automatically hyperparameter tunes the models, it could be beneficial to explore their suggested models even further and hyperparameter tune them further.

5.8.2 Prediction accuracy

After 60 hours of batch fermentation, our best model is able to predict when the batch is finished with an average error of 13.3 hours. While this may look like a large error, it is important to note that when we evaluate the batch as finished it is not regarded as finished in the brewery. Usually, the batch is in the tank for 4-5 additional days. As mentioned in section 6.3.3, this is because production plans and schedules are set in advance and is a static future oriented process. We predict and evaluate when the fermentation activity has stopped, meaning that the fermentation process has ended. Therefore, even with some deviation error, gaining insights regarding when a batch has stopped fermenting several days before the next process steps is regarded as valuable information.

Table 5.9 illustrates that the best RMSE after 40 hours is 13.87 and the best RMSE after 60 hours is 13.3. Surprisingly, the models do not improve significantly with additional 20 hours of information from the fermentation process. This illustrates that the 20 additional hours do not capture any additional useful information. If the process variables during these hours do not show a clear trend or pattern that can be linked to the end of fermentation, feeding this data to the model might introduce more noise than signal. The lack of improvement with additional information might suggest that the model has reached its capacity to learn from the data. In other words, the current model might not be complex enough to capture additional information provided after 40 hours. This could lead to an investigation into trying different machine learning algorithms, tuning hyperparameters, or exploring more complex models.

As observed in the scatter plot in figure 5.15 it is clear that the short-term predictions are more precise than the ones made for the longer horizon. This is reasonable as it is easier to predict when a batch is finished fermenting if that point is nearby in time. Figure 5.15 also illustrates a few outliers in the prediction. Due to the use of RMSE as a metric these outliers punishes the score of our model significantly. In general, our models are struggling to predict batches where the fermentation activity deviates significantly from other batches. While removing batches with abnormal behavior could enhance performance, it would weaken the goal of the machine learning model which is to capture the uncertainty in the fermentation process. If all batches behaved similarly, it would be a better approach to develop sigmoidal or dynamic models, as mentioned in section 3.7.3. A weakness of our model is, therefore, its inability to detect patterns, learn, and predict batches with significantly abnormal behavior.

5.8.3 Available data

We received a total of 1406 batches from Plaato. 492 batches were identified as *good* through the implementation of the five-parameter logistic function. While it is good for machine learning performance to remove outliers and have a dataset with similarities, the removal of 65% of the data is a huge loss. Furthermore, we only continue the model development with the *good* batches. For better performance, the *mediocre* could be included as well. A disadvantage of our approach is that the first filter was through the five-parameter logistic function. If we had cleaned and smoothed the data properly prior to fitting the logistic function, we would potentially have more batches available to train on, as more batches would be categorized as *good*.

While figure 5.17 does not illustrate model improvement with more batches available, we still evaluate the lack of data as a main reason for not achieving better results. By adding more batches, the model gets exposed to fermentation processes with varying curves. By adding even more batches we believe the models could learn to identify similar ones and improve the predictions. For instance, batches are brewed with different types of yeast which results in different types of fermentation curves. At first, we tried to train the models only on batches with the same type of yeast. However, this resulted in just 40-60 *good* batches of each yeast type, which is insufficient for machine learning. Currently, Plaato is gaining access to approximately 50 new batches every week, with varying yeast types. Training the models on thousands of batches with the same yeast type would increase their accuracy significantly.

Noise and deviations in the data may also hinder optimal model performance. Due to particles and bubbles getting stuck on the sensor, it could provide false and noisy readings. This noise, if not properly managed or filtered, can complicate the interpretation of the data, making it more challenging to discern genuine trends or events from these measurements. Another major concern with the available data is the errors caused by users. They are responsible for starting and ending the fermentation, and deciding when the sensors should record. This could cause batches to have wrong timestamps or start at the wrong time, which has a negative influence on machine learning models.

As mentioned in section 3.7.3, many researchers apply additional data such as CO_2, amino concentration, oxygen concentration and initial cell count from the brew in their machine learning models. Adding this additional data to our models could potentially boost accuracy even further. However, the sensors used to retrieve the data we used do not measure these features.

5.9 Summary

- Retrieved smoothed fermentation data from Plaato's API.
- After cleaning the data through a sigmoidal function we continued with 492 of 1406 batches.
- As features in the ML models we used density, temperature, fermentation activity, max fermentation activity and average start temperature from 40, 60 and 80 hours of fermentation data.

- The target variable is when the fermentation is finished, which is when the fermentation activity has been below 0.1 for 12 consecutive readings.
- The accuracy of the models were measured on a separate test set (20% of all batches).
- The models deviated on average by around 13 hours when trained on 40 and 60 hours of data.
- The best model trained on 80 hours of data deviated with an average of 9.5 hours.

Chapter 6

Case study

This case study focuses on production planning and control at one American craft brewery and how prediction data with information about the completion of beer fermentations can be utilized in production planning and control at the craft brewery in question. The chapter introduces the craft brewery, also referred to as the case company, and a description of the current situation of their production planning and control activities. Furthermore, an analysis of their production planning and control activities are conducted before suggestions for improvement through the use of prediction data are outlined and discussed. The information was gathered through three in-depth interviews with the operational manager at Lock 27 Brewing Company

6.1 Introduction to the Case Company

Lock 27 Brewing Company is an American craft brewery from Miamisburg, Ohio. They were founded in 2012 with one location and have since expanded by adding another location in Dayton, Ohio. Lock 27 Brewing Company started selling their beer through their own bar, also called tap room, in 2012, but began wholesale of kegs and cans to local bars, restaurants and retailers in Ohio in 2018.

Lock 27 Brewing Company now operates two tap rooms, both with significant food serving and is thus classified as a Brewpub, as per section 4.0.1. They produce approximately 140.000 liters (1200 BBL) yearly, consisting of 40 batches, with 4 fermenter vessels and an annual production capacity of 234.000 liters (2000 BBL). Lock 27 Brewing Company has a production team of four employees. This case study focuses on Lock 27 Brewing Company's production facility, and not their test facility in Dayton, Ohio. The test facility works as its own entity that focuses on small batch production of a large variety of test recipes. That said, the taproom in Dayton, Ohio is included as its own entity, and not a part of the brewing location there.

The production facility at Lock 27 Brewing Company consists of a sales department, production management department, brewing department and logistics department. The sales department is responsible for order handling and forecasting. The production management department is responsible for purchasing and production planning. The brewing department is responsible for carrying out the production plan. Lastly, the logistics department is responsible for delivering orders to customers.

6.2 Current Situation

The description of the current situation at Lock 27 Brewing Company provides a perspective of production planning and control in a craft brewery setting. The section delves into the company's customer base and market in which it operates, product range, variants and portfolio, production planning and control, supply chain, material flow, information flow and inventory management.

6.2.1 Actors in the supply chain

Lock 27 Brewing Company collaborates with two suppliers, one for raw materials, such as yeast and grains, and one for cans. Their cans suppliers also supply the service of canning beer to Lock 27 Brewing Company because they do not possess the necessary equipment in house, called a canning line. Lock 27 Brewing Company's direct customers are their own taprooms, grocery stores, restaurants and bars, all located in Ohio. The state of Ohio does not require craft breweries to follow the three tier system and allows breweries to self-distribute. As a result Lock 27 Brewing Company does not sell their products through any distributor. The number of end customers remains uncertain because Lock 27 Brewing Company sells their products to grocery stores, bars and restaurants, and does not monitor those numbers.

6.2.2 Production planning and control

The brewery itself serves as the single control area within Lock 27 Brewing Company's production facility, as seen in figure 6.1 and they follow a make-to-stock (MTS) approach, with the customer order decoupling point (CODP) situated within the finished goods inventory. This approach means that their production planning and control activities are based on demand forecasts and current inventory levels, characterizing a material resource planning (MRP) strategy.

Long term planning at Lock 27 Brewing starts with the sales department creating quarterly forecasts based on long term sales data from the same period in previous years and sales data from the preceding quarter. Quarterly forecasts are then communicated to the production manager in a quarterly sales and operations meeting. The company's production plan is created on a quarterly basis as well and are based on the quarterly demand forecasts provided by the sales department. The responsibility of creating the production plan lies solely with the production management department, led by the production manager. The responsibility for material requirements planning also lies with the production management department, which also generates this document quarterly. In addition, the production management department is tasked with the creation of quarterly production schedules.

In the short term, the sales manager and the production manager hold weekly production activity control meetings, as elaborated in section 3.2.5, where they monitor the production schedule for the coming week. Inventory levels, both raw materials, work in process inventory and finished goods inventory are investigated in this meeting and compared to sales orders and the quarterly demand forecast to determine if changes to the weekly schedule are needed to fulfill customer demand.

The production manager updates the material resource plan on a weekly basis, based on the output of the production activity control meeting. The production manager then orders the needed raw materials from their supplier on a weekly basis. This can be done through the supplier's website, but Lock 27 Brewing Company usually calls in the order by phone to get real-time information about inventory and delivery times. The company maintains a high frequency of procurement for all supplies, with the exception of hops, which are purchased in bulk to minimize order costs. Other raw materials are kept at low inventory levels due to the limited storage capacity of the raw material storage area, making the high frequency of purchasing necessary. The order quantity is adjusted between each order, based on the production schedule and current inventory level and the delivery lead time is generally between one and three days.

Lock 27 Brewing Company is using spreadsheets for production planning documentation and a specialized software for tracking kegs. They have yet to implement any additional Information and Communication Technology (ICT) systems, such as an Enterprise Resource Planning (ERP) system, to support these activities. Consequently, the quarterly production plan is sent out to the brewing department over email, and potential adjustments from the weekly production activity control meeting on Mondays are communicated manually.

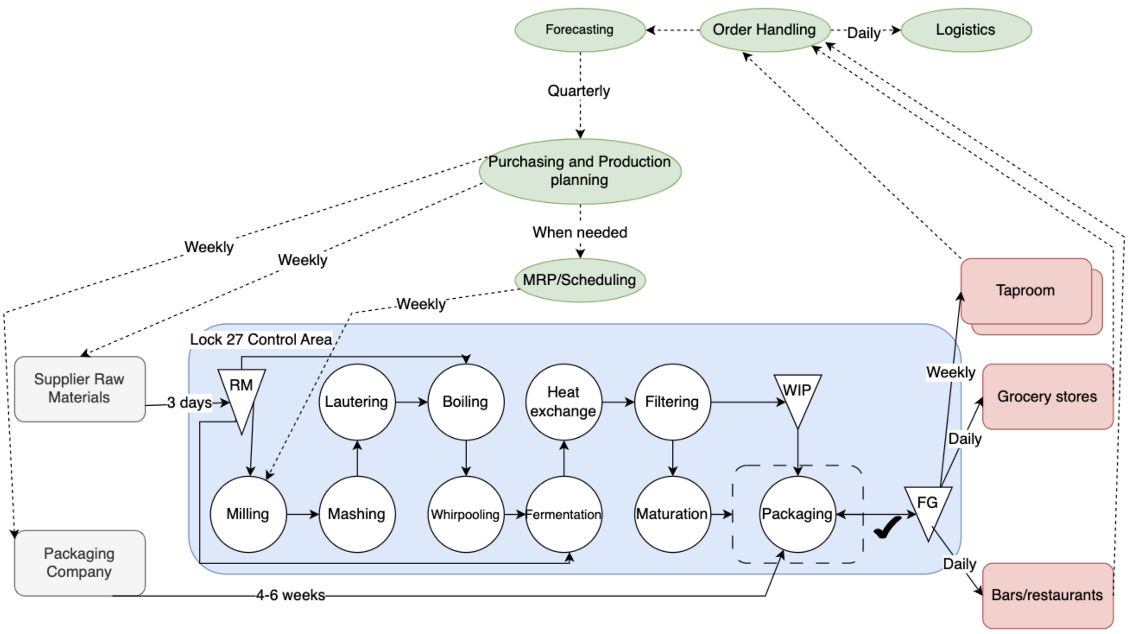


Figure 6.1: Control Model: Lock 27 Brewing Company

6.2.3 Products

Lock 27 Brewing Company produces beer of different styles. Their products are, as seen in 6.2, standard, but with different variants. They produce some core and flagship products, some seasonal products and some innovative variants and limited edition products. However, the production brewery focuses on core products, flagships products and seasonal products. Their products are many-to-one-to-many and the lead time in production is long.

Figure 6.2: Product variants at Lock 27 Brewing Company

6.2.4 Product variants

Lock 27 Brewing Company's product portfolio is divided into three categories: flagship, core, and seasonal products, each with its own set of product variants. The flagship category consists of a Citra Pale Ale, a Belgian Wit, and an India Pale Ale. The core category includes a Golden Ale, a Brown Ale, and an American Beer. The seasonal category includes a German Lager, a Double IPA, a Coffee Stout, and an Irish Red Ale.

The flagship and core products are always available and produced throughout the year. The portfolio of flagship and core products are only updated with products that have been introduced to the market as limited editions and proved to have demand over longer time periods. The seasonal products are available only during specific periods of the year, as seen in table 6.1 below.

Season	Months	Product Name	Product Type
Summer	July - September	Loktoberfest	German Lager
Fall	August - November	Stumbling Nag	Double Inida Pale Ale
Christmas	Oktober - January	Dirty Bean Stout	Coffee Stout
Winter	December - February	Kiefaber Street Red	Irish Red Ale

Table 6.1: Seasonal Products Lock 27 Brewing Company

Each type of beer is sold in 35 cl (12 oz) cans and kegs. Kegs are sold to restaurants, bars and their own tap rooms. Cans are sold to the aforementioned as well as grocery stores and other retailers. Lock 27 Brewing normally operates with one size of cans, benign the 35 cl ones, but will occasionally produce seasonal beers in different sized cans. However, that is not accounted for in this thesis due to the inconsistency in when its produced.

Considering the information above, the total number of product variants produced at Lock 27 Brewing's production facility annually is 10 types of beer in two types of containers, giving a total number of 20 product variants.

6.2.5 Material flow and processes

As seen in section 6.3 the beer production process at Lock 27 Brewing Company starts with the transfer of barley and grains to milling, which takes about 30-40 minutes. This is done with a counterbalanced lift truck. Malt milling, which involves crushing the malted grains to produce grist, takes another 30-40 minutes. The grist is then transferred to the mash tun for mashing through pipes. This takes approximately an hour. During mashing, the starches in the grist are converted into fermentable sugars. Lautering and sparging, which take approximately two hours, involve separating the liquid wort from the spent grain and rinsing it to extract all the sugars.

After lautering and sparging, the wort is transferred to the kettle, which takes approximately two hours. The wort is brought up to a boil for 10-15 minutes before boiling for 90 minutes. After its finished boiling, the wort is whirlpooled for 30 minutes to remove any solids or trub. The wort is then transferred to the fermenter vessel, which takes about 40-45 minutes to do. At the end of the day, the brewing equipment is cleaned, which takes approximately one hour. The process of fermentation takes between one and two weeks, depending on the style of beer, the recipe and desired flavor profile.

Fermentation accounts for 70 % to 90 % of the total process time at Lock 27 Brewing Company. The duration varies with different beer's recipes, but it can also vary between batches of the same recipe, as per section 4.0.2. To account for this uncertainty, Lock 27 Brewing Company adds one additional week to the planned process lead time of fermentation in their production plan and schedule compared to the average duration of the process. Once fermentation is complete, the beer is transferred to a bright tank for conditioning. The bright tanks also works as storage tanks where beer are stored until packaging starts. The cleaning cycle for the fermentation equipment involves three steps. The first step is a caustic cycle that takes about 30-40 minutes to clean the equipment of any residue. The second step is a rinse that takes 15 minutes, followed by a sandy cycle that takes 15 minutes to ensure that all the equipment is clean.

The duration the beer is situated in the bright tank depends on the packaging method. If the beer is kegged, it takes about two days because Lock 27 Brewing Company can proceed with kegging themselves right after conditioning and do not need to rely on other actors for this process. If the beer is planned to be canned, it can take up to three weeks waiting for the canning supplier.

The canning process at Lock 27 Brewing Company is performed by a company that brings the necessary equipment to Lock 27 Brewing Company's production facility, connects it to their equipment and proceeds with canning. This must be booked by Lock 27 Brewing Company 4 weeks in advance and is the responsibility of the production management department.

Process	**Duration**
Malt transfer from RM inventory to milling station	30 minutes
Malt milling	30 minutes
Transfer from mill to mash tun	60 minutes
Mashing	60 minutes
Transfer from mash tun to lauter tun	30 minutes
Lautering	120 minutes
Transfer from lauter tun to kettle	120 minutes
Boiling	100 minutes
Whirlpooling	20 minutes
Transfer from kettle to fermenter vessel	45 minutes
Fermentation	Ales: 7 days Lagers: 14 days
Post fermentation WIP	7 days
Transfer from fermenter vessel to bright tanks	45 minutes
Conditioning	2 days
Pre packaing storage	Kegged beers: 0 days. Canned beers: 0-21 days
Packaging	1 days

Table 6.2: Production Process lead times

In conclusion, the beer production process at Lock 27 Brewing Company involves several steps that require specific time durations for each process, as seen in table 6.2 and figure 6.3. Additionally, proper cleaning of the equipment is necessary to ensure the quality and consistency of the beer. Understanding the time required for each step and ensuring that the equipment is properly cleaned is critical to the brewery's success in producing high-quality beer, as desrbided in section 4.0.2.

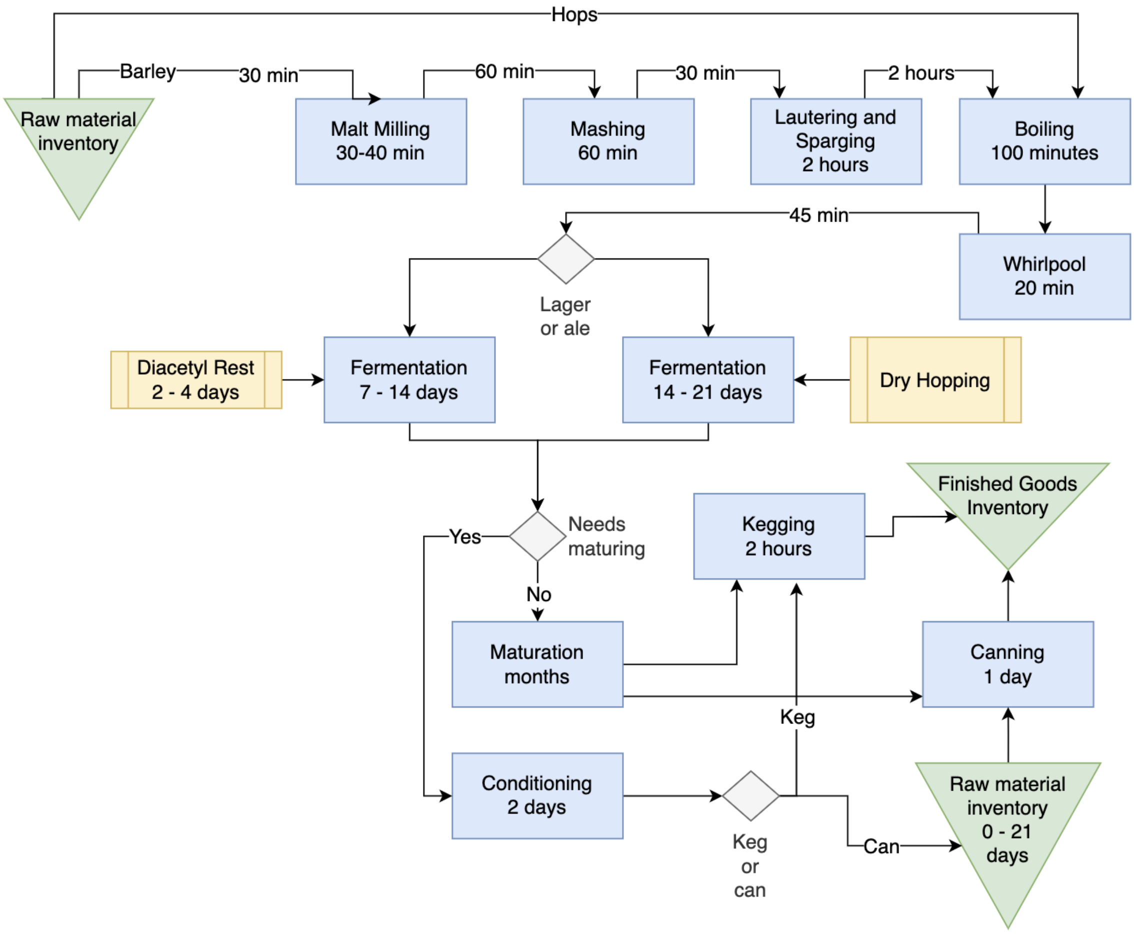


Figure 6.3: Process Flow Map Lock 27 Brewing Company

6.2.6 Information flow

External customers, such as grocery stores, bars and restaurants, place sales orders directly with the sales representative from Lock 27 Brewing Company responsible for their customer account. This is done either via phone or email. Lock 27 Brewing Company's tap rooms also place sales order through phone or email, but they do not have a designated sales representative and will thus place order directly with anyone available at the sales department.

6.2.7 Inventory management

Lock 27 Brewing maintains a raw material inventory, work in process (WIP) inventory between fermentation and packaging and a finished goods inventory. The WIP inventory consists of 3 tanks with a capacity equal to 75 % of the total fermentation capacity, considering that each can hold the volume as each of the fermenter vessels. Consequently, as seen in figure 6.4, Lock 27 Brewing Company can have a WIP inventory almost as large as their current, total production volume. The duration of which their product can be held in WIP inventory varies between 2 days and up to 3 weeks. The duration depends on when fermentation, which is the prior process, is complete, what type of packaging container they will use and when they have booked the next round of canning.

Lock 27 Brewing Company maintains a substantial finished goods inventory, as seen in figure 6.4, to ensure they can meet customer demand between canning rounds. The necessity for this amount of inventory arises primarily due to the reliance on an external company for canning, which requires a four-week advance booking. This arrangement results in less control and limited flexibility in the canning process, making it essential for Lock 27 Brewing to have enough finished goods on hand to cover customer demand until the next packaging round.

	Fermentation Work In Progress	Canning Queue	Finished Goods Inventory
Average Days in Inventory	7 days	14 days	7 days
Average volume in inventory	6825 L	5850 L	23984 L

Figure 6.4: Inventory overview at Lock 27 Brewing Company

Lock 27 Brewing Company employs a combination of digital and manual inventory management systems to track and manage their inventory across different stages of the beer production process. For their keg inventory, they utilize a specialized application called KegID. This app uses barcodes to monitor all their kegs, providing detailed information about the contents of each keg, the quantity of kegs available, and their specific locations.

The brewery maintains its raw material inventory and finished goods inventory, specifically cans, using spreadsheets. The spreadsheets are updated regularly by the production manager to reflect changes in inventory levels due to production activities and sales activities. Additionally, Lock 27 Brewing Company holds work-in-process inventory in fermentation vessels and bright tanks. Information about these intermediate stages of the beer production process is also maintained in spreadsheets.

6.3 Case Study Analysis of Production Planning and Control

The following section presents an analysis of production planning and control at Lock 27 Brewing Company. The analysis is conducted to better create an understanding of relevant challenges at the case company, investigate potential opportunities for improvement and justify the application of information from a predictive machine learning model as a tool to reduce the challenges associated with fermentation lead time uncertainty. The whole control area at Lock 27 Brewing Company is subject for the analysis.

6.3.1 Actors in the supply chain

The supply chain that Lock 27 Brewing Company is a part of consists of few actors upstream, with only two suppliers; one for all ingredients and the other for cans. Downstream, however, the amount of customers can be extensive and variable, considering that Lock 27 Brewing sells directly to retailers, bars and restaurants, as opposed to selling through a distributor.

Despite the simplicity of the upstream supply chain, Lock 27 Brewing Company's reliance on a single supplier for all ingredients can lead to significant challenges. Issues can occur that cause large disruptions in their ability to continue production. For instance, a week-long delay in maintenance of their supplier's ordering system resulted in an equally long production delay at Lock 27 Brewing Company. Their dependence on one supplier and their limited raw material storage capacity, makes it challenging for Lock 27 Brewing Company to reduce the impacts of such delays. Potential opportunities to mitigate that risk would for example be to build relationships with alternative suppliers or investigate the opportunity of expanding their storage areas, thus being able to build a larger raw material inventory.

Craft Breweries in Ohio, US are exempted from the three tier system, which is described in section 4.0.2, meaning that Lock 27 Brewing self distribute and work with a larger number of customers than breweries in other states that only work with distributors. A large number of direct customers makes Lock 27 Brewing less vulnerable because they are not solely dependent on single customers or a small number of customers. That said, a larger customer base can increase the complexity with regards to demand forecasting, order handling and transportation.

6.3.2 Seasonality

Demand fluctuations due to seasonality is common at Lock 27 Brewing Company, as was discovered in section 4.0.3 to be a general phenomenon in the American beer industry. Demand for beer is at its peak during the summer months, but there is also a spike during the Christmas period.

The seasonality in demand for beer means that it is necessary for Lock 27 Brewing Company to build up inventory outside of peak season, making the current Make to Stock strategy a natural choice, and alternatives such as a Make to Order strategy less ideal. Beer is, however, a perishable product that can not be produced too far ahead of consumption. As a result Lock 27 Brewing Company must plan to produce more in and around the seasons of peak demand, necessitating the need for some spare capacity the rest of the year.

6.3.3 Material Flow and Processes

The throughput time at Lock 27 Brewing varies between 11 days and 37 days, as seen in table 6.9, and depends on the recipe, fermentation lead time, on the packaging material and on when the next round of canning is planned. Customers expected delivery lead time is between 1 and 2 days and that particular relation is another reason for following a Make to Stock strategy with the CODP in the finished goods storage.

Lock 27 Brewing Company's reliance on an external canning company introduces complexity to their operations. The requirement to book canning services four weeks in advance limits the brewery's control over this stage of the production process. The lack of control reduces the company's flexibility to respond to changes in demand or production schedules, leading to a need for a significant WIP inventory between fermentation and canning, as seen in figure 6.4. Furthermore, the arrangement with an external provider of canning makes it necessary to maintain a substantial finished goods inventory to ensure that customer demand can be met between rounds of canning. This ties up capital and increases the risk of product spoilage and waste, due to beers perishability.

However, even if Lock 27 Brewing Company had an in-house canning line, maintaining sizable levels of inventory would still be necessary. This is because the variable fermentation lead times present a significant challenge in the production process at Lock 27 Brewing Company. As fermentation accounts for 70 % to 90 % of the total process time, variability in this stage significantly impacts the overall production timeline. As described, the uncertainty makes the addition of an extra week to the fermentation process necessary in the production plan, which potentially can cause inefficiencies by introducing idle time. This variability can complicate production planning and also make it difficult to synchronize it with subsequent process stages, such as packaging. As a result, finished beer can be stored in fermentation vessels for several days before available space has opened up in the bright tanks, practically creating another point of WIP inventory.

Moreover, the shift from batch production in brewing and fermentation to continuous production in packaging introduces additional complexities and potential bottlenecks, as described in section 4.1. This transition requires time and coordination, further increasing the need for a buffer of finished goods to ensure a steady supply for customers. By maintaining a sufficient finished goods inventory, Lock 27 Brewing Company can manage the uncertainties and complexities associated with the fermentation process and the transition from batch to continuous production, ultimately ensuring a reliable supply of products for their customers.

6.3.4 Inventory management

Lock 27 Brewing's use of a specialized digital system for inventory management of kegs allows for real-time updates and efficient management of their keg inventory. This allows them to ensure improved accuracy in inventory records, which can reduce the risk of stock outs or overstocks. It also enhances decision-making capabilities at Lock 27 Brewing Company, as the company has immediate access to accurate and up-to-date information about their keg inventory and location of kegs. This could potentially lead to cost savings, as the company can make more informed decisions about when to produce more beer or when to slow down production to avoid excess inventory, as well as time production with the availability of kegs.

The use of spreadsheets for managing raw material and finished goods inventory, as well as work-in-process inventory, might have both positive and negative effects. On the positive side, spreadsheets are a simple and straightforward tool for recording and tracking inventory data. They are easy to use and can be customized to meet the specific needs of the company. However, on the downside, manual systems like spreadsheets can be prone to errors and may not provide the same level of real-time visibility as digital systems, as elaborated in section 3.2.6. This could potentially lead to inaccuracies in inventory records, which could in turn affect production planning and customer service levels.

Moreover, the use of separate systems for different types of inventory might lead to a lack of integration and coordination in inventory management. This could potentially result in inefficiencies and missed opportunities for optimization. For example, if the company could integrate their raw material, work-in-process, and finished goods inventory systems, they might be able to better coordinate their production and distribution activities to minimize costs and maximize customer service levels.

In conclusion, while the current inventory management policies of Lock 27 Brewing Company may serve their needs to a certain extent, there could be potential for improvement and optimization. This could involve exploring more sophisticated and integrated inventory management systems, as well as continually reviewing and updating their inventory policies to align with best practices and the evolving needs of their business.

6.3.5 Operational performance

In order to further analyze production planning and control at Lock 27 Brewing Company, an analysis of operational performance was conducted. Operational performance was for this particular case grouped based on available data and information from the case company and is as following:

1. Capacity Utilization
2. Inventory Turnover

Lock 27 Brewing Company does not monitor metrics for the aforementioned 2 categories for operational performance. However, they do have relevant data and information available to perform the analysis based on estimates. Consequently, information could not be extracted directly from any sort of ICT system, but through interviews and production plans, the relevant calculations could be performed and analyzed.

Capacity utilization

Lock 27 Brewing Company estimates to have an annual production capacity of 234.000 L (2000 BBL) if uncertainty regarding fermentation process time is accounted for, as seen in table 4.1. However, if they did not have to account for the uncertainty of fermentation lead time, they estimated that they would have a production capacity of 351.000 L (3000 BBL) with their current production set up, as per section 4.0.2.

To better be able to thoroughly analyze and understand the capacity utilization at Lock 27 Brewing Company, we investigated the relation between actual process time, as seen in table 6.2, and production throughput time. This analysis investigates time based efficiency performance and exposes the duration product sits in WIP inventory.

If we consider the average fermentation duration stated by Lock 27 Brewing Company as the actual fermentation process time, it results in 7 days for ales and 14 days for lagers. However, the production plan will state 14 days for ales and 21 days for lagers due to the uncertainty of the fermentation process time.

The production processes and its process times is simplified in figures 6.5, 6.6, 6.7, 6.8 below. For the purpose of visualization, the time beer sits in WIP inventory in bright tanks are set to 7 days. As a result, the conditioning process in figures 6.5, 6.6, 6.7, 6.8 are set to either 2 or 9 days. The reason for this simplification is that processes are spread between working days and working hours. Lock 27 Brewing Company will do all processes which constitutes brewing on 1 day. Fermentation will start the day after and run until finished. After 2 or 3 weeks, depending on the style of beer, it will be moved to bright tanks for conditioning. The calculations seen in section 6.3.5 describes the production throughput time of both ales and lagers, packed in both kegs and cans. These are based on figures 6.5, 6.6, 6.7, 6.8 to make them more realistic, compared to excluding the working days and working hours as relevant factors.

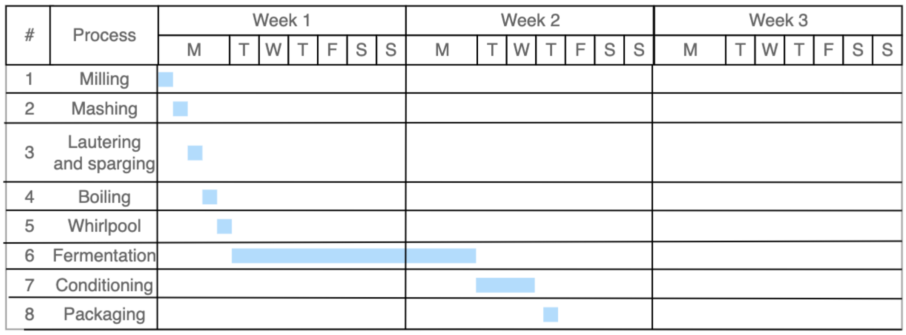


Figure 6.5: Process Gantt Chart Kegged Ales

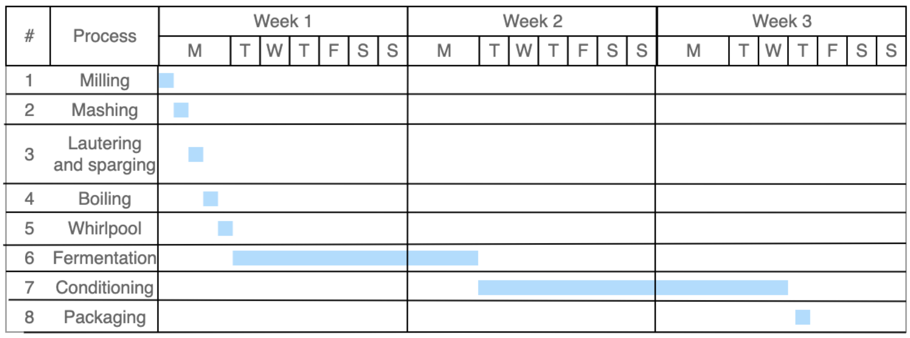


Figure 6.6: Process Gantt Chart Canned Ales

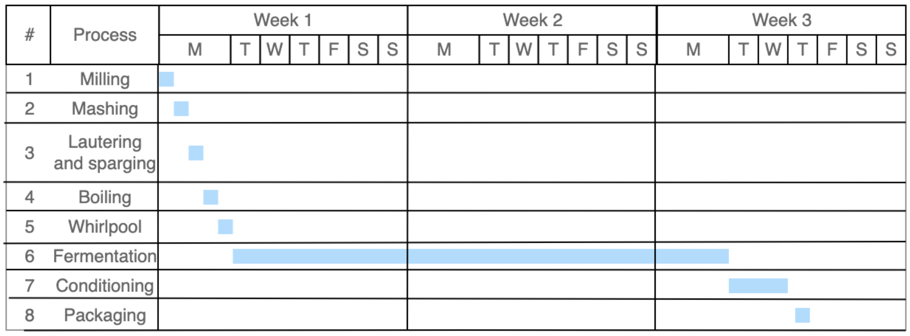


Figure 6.7: Process Gantt Chart Kegged Lagers

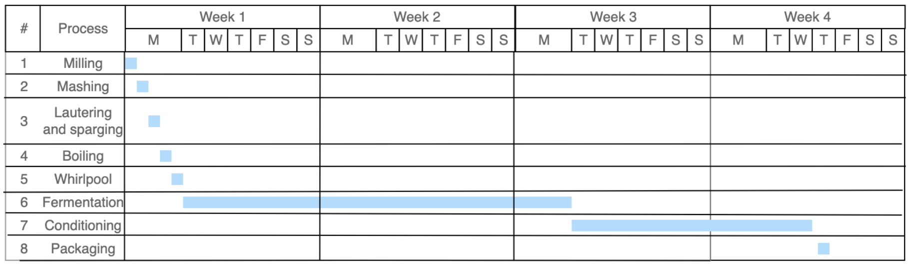


Figure 6.8: Process Gantt Chart Canned Lagers

Ale packed in kegs:
1 day (Brewing) + 7 days (fermentation) + 7 days (WIP) + 2 days (conditioning) 1 day (packaging) = 18 days

Ale packed in cans:
1 day (Brewing) + 7 days (fermentation) + 7 days (WIP) + 2 days (conditioning) + 21 days (WIP) + 1 day (packaging) = 39 days

Lager packed in kegs:
1 day (Brewing) + 14 days (fermentation) + 7 days (WIP) + 2 days (conditioning) 1 day (packaging) = 25 days

Lager packed in cans:
1 day (Brewing) + 14 days (fermentation) + 7 days (WIP) + 2 days (conditioning) 1 day + 21 days (WIP) + (packaging) = 46 days

The total process time excludes WIP from the calculations, and is as following:

Ale packed in kegs:
1 day (Brewing) + 7 days (fermentation) + 2 days (conditioning) 1 day (packaging) = 11 days

Ale packed in cans:
1 day (Brewing) + 7 days (fermentation) + 2 days (conditioning) + 1 day (packaging) = 11 days

Lager packed in kegs:
1 day (Brewing) + 14 days (fermentation) + 2 days (conditioning) 1 day (packaging) = 18 days

Lager packed in cans:
1 day (Brewing) + 14 days (fermentation)+ 2 days (conditioning) 1 day + (packaging) = 18 days

The relation between production throughput time and actual process time, seen in 6.2 is thus as following:

	Througput time	Process time	Ratio
Ale packed in kegs	18 days	11 days	1.63
Ale packed in cans	39 days	11 days	3.55
Lager packed in kegs	25 days	18 days	1.32
Lager packed in cans	46 days	18 days	2.56

Figure 6.9: Throughput Time and Process Time

The data on throughput time and process time reveals significant differences in the ratio between the two metrics for both ales and lagers packaged in kegs and cans. The most significant are both types of beers packed in cans, which is naturally caused by the WIP inventory situated before canning. However, these findings indicate that there is a considerable amount of idle time within the production process, particularly in the canning process for both ales and lagers. However, the ratio between throughput time and process time for kegged beers indicates that idle time is significant throughout the production at Lock 27 Brewing Company, also when disregarding the WIP inventory caused by the lack of control of the canning process. These findings further emphasize that the packaging arrangement and the variable fermentation lead times are significant challenges for production planning and control at Lock 27 Brewing Company.

Inventory turnover

Lock 27 Brewing's PPC approach is MTS and orders are thus fulfilled from the finished goods inventory. As mentioned in section 6.2.6, inventory is used to buffer against uncertainties in production, such as the process time of fermentation, and cycles between each round of canning. In addition to those factors, seasonality in demand variations is common throughout the craft brewing industry. Safety stock is thus important to ensure that Lock 27 Brewing can meet customer demand throughout the year.

Inventory turnover (IT) is a commonly used measure to assess a company's performance, and inventory turnover ratios are critical in the manufacturing industry, as described in section 3.2.6.

$$Inventory\ Turnover = \frac{Annual\ cost\ of\ goods\ sold}{Value\ of\ average\ inventory} \tag{6.1}$$

We do not have access to financial data from the case company, such as value of inventory, annual revenue or cost of goods sold. Consequently we have decided to adjust the inventory turnover formula and base it on volume instead of monetary value.

$$Inventory\ Turnover\ in\ volume = \frac{Annual\ volume\ of\ goods\ sold}{Average\ inventory\ volume} \tag{6.2}$$

As seen in figure 6.4 Lock 27 Brewing Company keeps on average a total volume of 6,825 L in the fermenter vessels as work in process inventory, 5,850 L in bright tanks as work in process inventory and 23,984 L in the finished goods inventory, bringing the total average inventory up to 36,659 L. The annual volume of goods sold is 140,400 L, resulting in an inventory turnover of 3.83. This means that Lock 27 Brewing Company on average will sell out their inventory approximately every 3 months.

As described in section 3.2.6 an inventory turnover ratio between 5-10 is considered good for most industries. However, a higher inventory turnover ratio is usually desired when working with perishable goods, such as beer is, to avoid spoilage and inventory losses.

The inventory turnover analysis further supports the significant challenges that the variable fermentation lead times and reliance on an external canning provider is for production planning and control at Lock 27 Brewing Company. The company's Make to Stock approach, coupled with the uncertainties in the fermentation process and the inflexibility of the canning schedule results in a need for maintaining high levels of inventory. This is particularly concerning for a company dealing with perishable goods like beer, where a higher inventory turnover ratio is preferred to avoid spoilage and inventory losses. Therefore these high inventory levels tie up capital and increase the risk of product waste.

6.4 Suggestions for improvement

Previously in this thesis, the development of a predictive model that predicts beer fermentations time of completion was described. In light of our findings, we propose to enhance Lock 27's production planning and control through the implementation of a predictive model such as this for the beer fermentation process. Our predictive model has been developed and trained to predict the day of completion of a beer fermentation process 40 hours, 60 hours and 80 hours after the batch has started. As described in 5.6.4, the level of accuracy to the predictions does not significantly increase when predicting longer into the fermentation. As a result, we have decided to base our suggestions on a theoretical implementation of predictive information made available 40 hours into the batch. This prediction capability provides a significantly improved level of control and visibility over the fermentation process compared to the current situation, potentially allowing for better planning and utilization of resources.

Fermentation is a critical step in the beer production process, yet it is one that has traditionally been hard to predict with precision due to its biological nature. This uncertainty can result in suboptimal capacity utilization, longer throughput times, and thus increased work-in-progress (WIP) inventory, among other issues, as identified in section 6.3.5. By addressing this key area of uncertainty, we can potentially work towards a more efficient and effective production at Lock 27 Brewing Company.

Our analysis of Lock 27 Brewing Company's operations highlighted several other challenges, with the reliance on an external packaging service potentially being a key issue overall. This arrangement, which requires booking services four weeks in advance, significantly limits the company's flexibility and control over a critical stage of the production process. It also makes it necessary to maintain high levels of finished goods inventory, tying up capital and increasing the risk of product spoilage.

Furthermore, other challenges such as variable demand due to seasonality, managing a diverse customer base, and dependency on a single supplier for ingredients were also identified to impact production planning and control. However the packaging issue stands out due to its direct impact on production planning, inventory management, and overall operational efficiency. Although the proposed predictive information of the fermentation process can potentially enhance production planning and control, addressing these challenges may require different strategies. Given the scope of this thesis, we will not provide suggestions with the aim of mitigating these particular challenges, considering that they are not directly caused by fermentation uncertainty.

6.4.1 Using predictive information in production planning and control

In the following section, we will explore in detail how predictive information from the predictive model, developed in chapter 4.1.3, can be utilized within Lock 27 Brewing Company's existing operations through concrete suggestions on potential improvements.

Dynamic scheduling system

Based on our analysis of material flow and processes at Lock 27 Brewing Company, seen in section 6.3.3, we suggest the development and implementation of a dynamic scheduling system. This system would leverage the predictive information about the fermentation process' duration to create a more flexible and responsive production schedule, consequently reducing idle time and total production throughput time.

The dynamic scheduling system could be implemented in several ways. One approach could be to introduce an additional planning meeting once the predictive information about the fermentation process becomes available. This would allow the production and sales management departments to adjust the production plan and schedule based on the predicted fermentation completion times. Alternatively, the predictive information could be integrated into a scheduling software system. This system could be programmed to automatically update the schedules for post-fermentation activities, such as conditioning, packaging and maturation based on the predicted fermentation completion times. This would not only reduce the manual effort required to update the schedules but also ensure that the schedules are updated as soon as the predictive information becomes available.

The goal of implementing a dynamic scheduling system is to create a more agile and responsive production process that can better adapt to changes and uncertainties. The agility and responsiveness of the dynamic scheduling system would contribute to a decrease in the total production throughput time by allowing Lock 27 Brewing Company to move beer from the fermenter vessel to subsequent processes faster, thus reducing idle time. This is particularly important for a perishable product like beer, as it reduces the risk of spoilage and waste. Moreover, it can also allow Lock 27 Brewing Company to reduce the amount of beer situated in work-in-process storage in the fermenter vessels. As a result, they can potentially increase their production volumes because the duration of each production cycle can be reduced. Furthermore, it could allow them to save on inventory carrying costs, such as energy usage due to cooling stored beer.

With the current brewery setup, Lock 27 Brewing Company adds an extra seven days to each batch during their planning phase to account for the uncertainty in the fermentation process. However, with the predictive information, they would know the when each batch will be finished 40 hours into fermentation, or two and a half days into fermentation, with an estimated deviation of 13 hours. This would allow them to plan for the next week in advance, rather than just reacting to which batches are finished and which tanks are ready at their Monday production activity control meetings. This means that they potentially can be able to eliminate the work in process (WIP) inventory that is situated in the fermenter vessels, which on average sits there for 7 days. This this results in a potential to reduce their total production lead time from 18 to 11 days for ales packed in kegs, from 39 days to 32 days for ales packed in cans, from 25 to 18 days for lagers packed in kegs, and from 46 to 39 days for lagers packed in cans, by following the simplified schedule seen in figures 6.6, 6.5, 6.8 and 6.7. This also means that they will be able to cool beers in the fermenter vessels for 7 days less.

With the implementation of a dynamic scheduling system at Lock 27 Brewing Company having the potential of reducing throughput time with up to 7 days, there exists a potential to significantly enhance their capacity utilization. Currently, Lock 27 produces 40 batches annually, each of which includes 7 days of idle time due to storage in the fermenter vessels. This idle time amounts to 280 days annually, which represents a substantial underutilization of their production capacity. This would free up capacity in their production schedule, allowing them to produce additional batches. The 280 days of idle time represents the total production time of 25 batches of kegged ales, if they were to utilize the predictive information about the fermentation process. As a result they could potentially, with the same amount of days they currently use, produce an additional 25 batches of kegged ales annually. This represents a significant increase in their production capacity and could potentially lead to increased revenues and profitability.

It's important to note that this improvement in capacity utilization is achievable without addressing other challenges identified in our analysis, such as the reliance on an external canning provider. By focusing on optimizing their fermentation process through the use of predictive information, Lock

27 Brewing Company can make significant strides toward improving their production efficiency and capacity utilization.

Inventory management system

In addition to the dynamic scheduling system, we also suggest the implementation of a more sophisticated inventory management system at Lock 27 Brewing Company. This system would incorporate predictive information about the fermentation process' duration, providing more accurate and real-time inventory data.

A key feature of this system would be its ability to track the predicted day of completion for all beers currently in the fermentation process. This predictive information would enable real-time adjustments to inventory levels, reflecting not only the current stock of finished goods but also the products that will soon be ready for packaging and order fulfillment.

This system could significantly reduce work-in-process (WIP) inventory, particularly in the fermenters post-fermentation and bright tanks post-conditioning. By providing accurate estimates of when the beer in the fermenter vessels will be ready to move to the next stage, the time that beer spends as WIP inventory could be minimized. This would free up capacity in these critical resources, allowing for more efficient use of the fermenter vessels and bright tanks.

With the predictive information provided by the proposed system, Lock 27 Brewing Company would have a more accurate estimate of how much beer will be ready for packaging at any given time. This would enable them to more precisely match their WIP inventory levels to their packaging needs, reducing the need to maintain excess inventory. Not only could this lead to cost savings through reduced inventory carrying cost, but it could also reduce the risk of product spoilage or waste.

Importantly such a system can also be beneficial towards reducing the total production throughput time and its ratio to the actual processing time. As with dynamic scheduling, reducing the total production throughput time, the proposed system could also help to reduce the risk of spoilage and waste. Considering that beer is a perishable product, any reduction in the time it spends in production could potentially extend its shelf life and improve its quality.

6.4.2 Summary

- Lock 27 Brewing Company is an American craft brewery based in Ohio, established in 2012. It operates two tap rooms and uses a make-to-stock principle for production, with demand forecasts driving activities.
- The company collaborates with two suppliers: one for raw materials and another for cans and canning services. They directly supply their products to their taprooms, grocery stores, restaurants, and bars within Ohio.
- Lock 27's product portfolio includes flagship, core, and seasonal beers, totaling 10 types in two kinds of containers, creating 20 product variants. Their production and inventory management relies on spreadsheets and a keg tracking software.
- Their production process is intricate, involving several steps each with specific time durations. The duration of the beer in fermentation and bright tanks can range from days to weeks, dependent on the beer type and packaging method.
- The company maintains inventory at different production stages. The necessity for substantial finished goods inventory arises due to reliance on an external canning company requiring advanced booking, making large finished goods inventory crucial for meeting customer demand.
- The supply chain of Lock 27 Brewing Company, while simple upstream, is complex downstream. Reliance on a single ingredient supplier and limited storage capacity can pose risks. Mitigation could include alternative suppliers or expanded raw material storage.

- The variability in fermentation times and external canning services scheduling introduces complexities in production and leads to substantial inventory buildup at Lock 27 Brewing Company.
- While there's no formal monitoring of performance metrics for PPC, high idle times in production and high inventory levels suggest operational inefficiencies and risks tied to beer spoilage.
- A dynamic scheduling system could utilize predictive data to decrease idle time and increase capacity, potentially reducing production lead time by up to 7 days and enabling additional annual production.
- An improved inventory management system, using predictive fermentation data, could lower work-in-process inventory and costs, reduce spoilage risk, and enhance beer quality by shortening total production time.

Chapter 7

Discussion

The purpose of this chapter is to connect and discuss the findings from the theoretical background, model development and the case study, in order to answer the research questions. The chapter is divided according to the research questions.

7.1 How ML can be used in order to retrieve insights about the fermentation process in breweries (RQ1)

Our implemented models are able to learn patterns from the fermentation data and predict when a fermentation process will be finished to a certain degree. When trained on the first 80 hours of data, the model gives valuable predictions. When training on 40 and 60 hours the predictions do not give that valuable information. However, some of the predictions from these datasets are very accurate and they deviate from the true value by merely a few hours. The model is able to give very accurate predictions on fermentation processes that follow normal behavior. However, it is struggling to predict accurately on abnormal batches, which are the batches causing uncertainty that we want to be able to detect. By including a metric regarding how certain the model is about its predictions would make it possible to detect batches that deviate from abnormal behavior and handle them accordingly. This would cause a substantial improvement in the value of our research. The study also shows that the most important factor in order to retrieve insights about the fermentation process is to have a large amount of data with high quality. By developing the model to also include certainty combined with additional data this study shows that it is possible to gain insights regarding the fermentation process.

7.1.1 Relation to previous research

Determining the completion of fermentation processes is a challenging task due to several factors (Syu et al. 1994). Firstly, fermentation is a biological process governed by numerous variables including temperature, pH, CO_2 level, amount of amino concentration and oxygen concentration, which all can affect the duration and efficiency of the process (Syu et al. 1994). Small variations in these parameters can lead to significant differences in fermentation outcomes. Secondly, the behavior of fermenting organisms, such as yeast in beer brewing, can be inconsistent and somewhat unpredictable, adding another layer of complexity (Defernez et al. 2007). Lastly, the methods for measuring important parameters in a brew, such as the density or the temperature, can be imprecise.

Several previous researchers investigate how the fermentation process can become less uncertain. For instance, Thibault et al. (1990) predicted biomass and substrate concentrations in fermentation by developing a neural network. Montague et al. (2008) presents a case study demonstrating how forecasting algorithms can be used to assess future bioprocess conditions. Bowler et al. (2021)

developed a machine learning model that uses data from ultrasonic sensors in order to predict the alcohol concentration during beer fermentation. While these studies showed great results in their respective research scope, their goals differ from ours. The related work does not retrieve insights regarding when a batch is finished, multiple hours ahead. Rather, the focus appears to be on forecasting outcomes for the current or the immediately following time-step.

Previous research presents approaches where traditional methods such as sigmoidal models, non-linear regression models or case based reasoning is used to model the fermentation process in a brewery (Reid et al. 2021, Speers et al. 2003, Montague et al. 2008). Other researchers use more modern approaches, such as neural networks and kNN (Trelea et al. 2001, Thibault et al. 1990, Defernez et al. 2007). In general, the studies reveal how machine learning is able to learn complex patterns, and give valuable insight about fermentation. Bowler et al. (2021) and Defernez et al. (2007) reveal lack of substantial data with high quality as a major challenge. Thibault et al. (1990) attempted to predict biomass and substrate concentrations in fermentations with a neural network, but the data used in the development was constructed by the author.

The previous research showed great success by using fermentation data combined with machine learning. This, combined with the recent development of advanced machine learning models, revealed in section 3.6.6, inspired us to try neural networks and gradient boosted methods. By using data with a higher volume and of better quality compared to previous research, we aim to get improved predictions. Our research also focuses on predicting the end time of when a fermentation process is finished, rather than just predicting the next measurement. Our study deviates from previous research because of the data we have available and the target value for predictions. Additionally, we examine the potential to forecast fermentation utilizing solely features like density, temperature, and fermentation activity, thereby avoiding reliance on chemical properties such as oxygen and CO_2 levels.

7.1.2 Prediction ability

By utilizing fermentation data recorded every 30 minutes from several different breweries, we developed several machine learning models in order to predict when the fermentation activity ended. The models were trained on data from the first 40, 60 and 80 hours of prediction. On the first 40 hours the best model was able to predict when the batch was finished with an average error of 13.87 hours. On the first 60 hours the best model predicted with an average error of 13.3 hours and on the first 80 hours of data the best model predicted with an average error of 9.533 hours.

The forecasting error obtained by the machine learning models was higher than what is common in short-term time-series forecasting, bringing into question whether the applied models are the most suitable for the task, and to what degree they are able to predict when a fermentation is finished. Compared to Defernez et al. (2007) which predicted a similar process of when a fermentation process obtains a certain density, our model had a better prediction average error of 12.6 hours. While that research used a completely different dataset, the comparison still shows that our model was able to predict a fermentation event in the future better than previous research.

As noted in section 5.8.2, the beer stays in the tank for several days after the fermentation activity has stopped. This is because brewers decide to keep it there considering maturation, conditioning, carbonation and quality assurance and testing. However, the density level and therefore also the fermentation activity remain stable during this process. This means that a prediction of when a batch is finished fermenting actually gives insights into the whole fermentation process which usually lasts between 5-14 days. Even though the average error of prediction seems high, we consider 13 hours of deviation of when the fermentation activity stops as reasonable since the process actually lasts for several days.

7.1.3 Model applicability and improvements

It is important to note that the deviation of 13.3 hours is the average error of all the predictions. From figure 5.15, which displays a scatterplot of the predictions versus the actual values, we can see that the prediction error increases with time. In particular, batches where the fermentation activity is finished approximately between 60-110 hours, have a much better average error than 13.3. On the other side, batches that ferment for more than 120 hours have a much larger error. This reasoning can be perceived as obvious, since batches that are finished closer to the point where the model's data is from, would be easier to predict. While this is true, the majority of the batches are finished in this time gap, and other batches can therefore be classified as outliers. The fermentation process is hard to evaluate, as it is full of uncertainties. The aim of our research is to capture and identify this uncertainty. In our approach, we only predicted when the batch is finished fermenting. There exists a variety of machine learning methodologies that not only predict a target outcome but also provide a measure of confidence or uncertainty associated with the predictions. Incorporating this supplemental information can provide an enriched understanding of the fermentation process. Early in the fermentation process, it would be possible to recognize outliers, by looking at the confidence of the prediction, and further take necessary actions for the brewers.

The accuracy of the predictions improved as more data from the fermentation process was utilized. While we did not train the models beyond 80 hours, our approach can be slightly modified to predict at any point during the fermentation process. This adjustment opens new opportunities for gaining insights into fermentation. First of all, one can continuously make predictions from the start of the fermentation. This will give insights into how accurate the model is from the beginning and at what timestamps the predictions are the most reliable. Secondly, our model showed that predictions became substantially better after 80 hours of training. By predicting continuously and on more data than 80 hours, it will give more accurate predictions about normal batches but also on the batches that act abnormal.

We implemented a baseline prediction that is the average finish time of all the batches in the training data set. If we compare our predicted results with the baseline prediction, our model shows great results when trained on 80 hours of fermentation. In fact, it averagely predicts more than twice as good as the baseline prediction. For models trained on 40 and 60 hours of data, the predictions are only 7-8 hours better compared to the baseline. The baseline model was rapidly constructed, primarily to serve as a comparative reference. If we had spent more time developing a baseline model, by for instance implementing one of the approaches mentioned in the related work section, we would get a better overview of the performance of the developed machine learning models. The fact that the developed machine learning models did not perform significantly better than the baseline model implies that the available data in this study is the most important factor. As can be seen in section 5, a significant part of this model development is to retrieve, clean, and prepare the data.

We had the advantage of substantial data at our disposal, collected via a network of IoT sensors. However, we were unable to use a large part of the data because of its quality. Primarily due to inaccuracies and inconsistencies stemming from the nature of sensor technology, combined with user error caused data with bad quality. We had to disregard 63% of the available batches. Despite these challenges, the potential for more data with sufficient quality is substantial. New batches are being recorded on a weekly basis, continuously expanding the volume of data available for analysis. This continuous data generation offers an exciting opportunity for future improvements in the predictive accuracy of our models. Even though figure 5.17 does not show a significant improvement by applying more batches, we believe adding more batches would improve the accuracy. The figure only shows the RMSE from 0 to 350 batches. By using thousands of batches, one would be able to classify and train the batches according to the yeast type used. Batches with similar yeast types have a much more similar density curve, which could therefore potentially increase accuracy.

7.1.4 Generalizability

The particular data used in this study was exclusively provided by our collaborating company. The data consists of information about density and temperature from batches produced by all their customers. This data is evaluated as of better quality compared to data used in previous research. The reason for this is the large number of batches and the high-frequency rate of measuring.

Despite the exclusivity of the data, the methodology we used in developing our machine learning models and the insights gained from the process can be generalized and applied to other breweries. The pipeline we established including data cleaning, feature extraction, model selection, hyperparameter tuning, and performance evaluation is a process that can be utilized on another set of fermentation data.

The type of data we used could also be gathered from different breweries or even from other industries that use sensors to track fermentation processes. While our data is specific to our case study, our approach is quite flexible and can be used in many situations. The models we've developed could potentially be used or adjusted to fit a variety of settings.

7.1.5 Weaknesses

We did not discuss how the models could be practically implemented and used by a brewery. The application of such models in a real-world scenario involves several logistical and technical considerations. For instance how to integrate the models into breweries' existing IT systems and how to ensure the stability and robustness of models over time. In our study, we were dependent on data from several different breweries. A single brewery may not have access to the same volume of data. Additionally, our study did not consider the costs associated with implementation. Implementing such models could involve substantial investments, including the cost of integrating the models into the existing systems, the potential need for additional hardware or software, ongoing maintenance costs, and possibly training costs for brewery staff. The challenge of manually labeling batches is another expensive operation. Without a clear picture of these costs, it is difficult to fully assess the benefits of implementing the models from an economic standpoint. The cost-effectiveness of the models is a crucial aspect that would need to be evaluated in a more comprehensive study.

An additional limitation of our study concerns the features used in our model development. Research in this field has demonstrated noticable success when incorporating features such as CO_2 levels, amino concentration, and oxygen levels. However, the sensors deployed in our study did not measure these particular parameters. With the inclusion of such features alongside our current dataset, the predictive performance of our models could potentially be enhanced. This suggests an avenue for future work, exploring the impact of a richer set of features on the predictive accuracy of the models.

Furthermore, the rapid evolution of machine learning technology presents yet another limitation in our study. The field is consistently offering new and more advanced machine learning models that could potentially better address the challenges we faced. For instance, Recurrent Neural Networks (RNNs) and Long Short-Term Memory (LSTM) networks are specifically designed to handle sequence prediction problems and have shown promise in similar scenarios. Fresh research also has begun to explore how to use LSTMs and RNNs in order to learn from multiple, multivariate time-series. These types of networks could potentially capture temporal dependencies in the fermentation data better than the models we employed.

7.1.6 Summary

- Our model gives relevant insights regarding when a batch is finished to a certain degree.
- The model is inaccurate on batches with abnormal activity.
- The model gives great predictions on batches that are finished from 60-110 hours, but is struggling on batches that are finished beyond 120 hours.
- The most valuable resource in gaining insights about fermentation is the data, with a focus on volume and quality.
- By including a probability of the accuracy on the predictions, the insights would become much more valuable.

7.2 How increased fermentation predictability can mitigate PPC challenges in breweries (RQ2)

To investigate how increased predictability regarding the duration of beer fermentations can mitigate production planning and control challenges in breweries, findings from the analysis of operational challenges and performance from the case study are discussed. Findings from the empirical background and the potential impact of the suggestions provided in the case study are also discussed. Initially, the PPC challenges that the findings suggest fermentation uncertainty leads to, and whether there might be other contributing factors to these challenges, are discussed. This discussion is crucial as understanding the root causes of PPC challenges in breweries is a prerequisite to evaluating how increased predictability could potentially mitigate these challenges. Following this, the suggestions from the case study are discussed, focusing on the extent to which these suggestions could mitigate the identified challenges. Furthermore, the applicability of these suggestions is discussed. Collectively, this approach allows for a systematic and structured response to the research question. In conclusion, the generalizability of the findings and potential weaknesses are discussed.

Increased predictability regarding beer fermentation durations will allow craft breweries to be more dynamic in their planning and scheduling activities and more responsive to changes, because they will know the lead time of their main production process, and thus the whole production cycle, several days earlier than they currently do. They will not only know the current state of fermentation, but also the expected time of completion. This can enhance their ability to manage the flow of material accurately, and thus increase capacity utilization. As a result of more efficient flow of material, idle time, and thus the WIP inventory situated throughout production, can be reduced. Furthermore, breweries gain the ability to integrate expected completion times into their inventory management system, enabling them to replenish inventory based on a production throughput time that has less idle time. Ultimately, this can reduce the time period they plan for current inventory levels to fulfill demand.

7.2.1 Production Planning and Control Challenges in Craft Breweries

As described in section 4 the craft brewing industry consists of thousands of independent breweries, where many are considered small or medium businesses. Production planning and control will vary between them, and thus the challenges they experience. Several challenges were identified, both in the empirical background, which included a multiple case study, and in the single case study. The different challenges, and their connection to the fermentation process, specifically the uncertainty connected to the fermentation process will be discussed.

Managing the flow of material

The uncertainty of the fermentation process time presents a significant challenge in managing the flow of material in a craft brewery like Lock 27 Brewing Company. This uncertainty makes the addition of a buffer period to the production plan and schedule necessary, which in turn leads to two key issues: an accumulation of work-in-progress (WIP) inventory within the fermenter vessels and a decrease in capacity utilization.

The need for a buffer period, or days of spare capacity, in the case of Lock 27 Brewing Company, 7 days, results in an extended production throughput time. This leads to an excess of WIP inventory, as the beer remains in the fermenter vessels for longer than necessary. This excess inventory represents an inefficient use of resources and can lead to increased carrying costs and potential spoilage and product waste.

All five craft breweries that participated in the multiple case study produce significantly less beer annually compared to the amount their system is capable of, as seen in section 4.1. Section 6.3.5 delves further into capacity utilization and uncovered findings, specifically regarding Lock 27 Brewing Company, that suggest that the uncertainty of the fermentation process time is a significant contributor to the low capacity utilization. As section 6.4.1 describes, the uncertainty of the fermentation process time contributes to 280 days of idle time which limits Lock 27 Brewing Company's ability to increase their production output, thus reducing their capacity utilization. However, section 4.0.3 and section 6.3.1 uncover that seasonality is also an important factor contributing to the low capacity utilization seen across all case companies, in both section 4.0.5 and chapter 6. The seasonality factor in the brewing industry indicates that breweries should strive to build inventory during seasons of lower demand, but the perishability of beer limits the opportunity to fully do so. As a result, striving toward the complete removal of spare capacity is likely unattainable.

Inventory Management

The case study revealed that Lock 27 Brewing Company has inventory situated in several locations; a raw material inventory, work-in-process inventory in the fermenter vessels, work-in-process inventory in bright tanks and a finished goods inventory. As discovered in section 6.3.5 the different inventory positions accumulate to a substantial amount of inventory, resulting in a less than desired inventory turnover rate.

The work-in-process inventory located in the fermenter vessels constitutes more than 18 % of their total inventory, which is caused by the 7 day buffer added to account for the uncertainty of the fermentation process time. Furthermore, such buffers were also added in the other breweries participating in the multiple case study, as described in section 4.0.5. Consequently, these findings suggest that the uncertainty of the fermentation process time is a contributing factor to the challenge of managing inventory effectively.

Lock 27 Brewing Company holds the majority of its inventory in the finished goods inventory and in bright tanks due to the inflexibility caused by their reliance on an external canning provider and their decision to follow an MTS approach. Consequently, this indicates that the uncertainty of the fermentation process time is not the sole factor contributing to a low inventory turnover rate. However, an MTS approach was identified as the only viable approach for craft breweries in section 4.0.3 and 4.0.3 and the reliance on an external canning supplier can be solved by investing in a canning line. Furthermore, the bottleneck that arises because of the transition from batch to continuous production, as described in section 4.1, introduces an additional layer of complexity.

The challenge of inventory management in craft breweries is a multifaceted issue with potential implications for operational efficiency and cost management, as seen in section 6.3.5. Maintaining high levels of inventory, particularly work-in-process inventory, can tie up significant capital, increase carrying costs, and lead to product waste due to the perishable nature of beer, as seen in section 3.2.6.

The uncertainty of the fermentation process time is a contributing factor to the challenge of inventory management in craft breweries. The need to maintain buffers to account for this uncertainty can lead to work-in-process inventory situated in the fermenter vessels because the planned duration exceeds the actual duration. This will increase the amount of work-in-process inventory overall and prolong production throughput times, as seen in section 6.3.5. However, while addressing the uncertainty of the fermentation process time could potentially lead to improvements in inventory management, it is also important for craft breweries to consider other aspects of their operations and supply chain management.

The challenges caused by fermentation uncertainty seem to have similar impact in increasing the amount of work-in-process inventory, idle time and production throughput time. It is therefore likely that potential solutions also will result in similar contributions. With a dynamic scheduling system and an inventory management system, the potential benefits is in this case to reduce work-in-process inventory, idle time and production throughput time. The proposed solutions aims to mitigate different challenges, but the benefits of doing so will largely have similar impact, thus indicating that the proposed solutions can amplify each other

Production Planning and scheduling

It is very hard to accurately create production schedules when lead time is uncertain. Decisions regarding when to plan for downstream activities such as packaging, when to order new supplies and thus be ready to start the next round of production and how to effectively manage the raw material inventory, and procurement of suppliers, relies on accurate lead time estimations.

Production planning involves determining the most efficient and effective way to utilize resources to meet customer demand through production. Further downstream, in the planning process, master production scheduling focus on the quantity and timing of specific end items in production. Consequently, the duration of production processes must be a crucial input when planning and more specifically, scheduling.

The uncertainty of the fermentation process time introduces a significant challenge to this planning process. To account for this uncertainty, craft breweries often resort to adding capacity buffers, or days of spare capacity, as described in section 4.1.1, which consequently increase the production throughput time, while the actual process time remains unchanged. This results in the introduction of idle time, a period of non-productivity that hinders the efficiency and effectiveness of the production process. This idle time stands as a hinder to the goals of production planning, which strives to achieve efficiency and effectiveness. Therefore, addressing the uncertainty of the fermentation process time could potentially lead to improvements in production planning and scheduling in craft breweries.

7.2.2 How challenges caused by the uncertainty of the fermentation process time can be mitigated

As outlined in section 6.4, predictive information that can provide more certainty about the fermentation process duration has the potential to be integrated into the production planning and control systems of craft breweries, with potentially positive effects. This could mitigate the challenges caused by the uncertainty of the fermentation process time, leading to improvements in several areas. For instance, it could enhance the accuracy of production schedules, reduce the need for capacity buffers, and improve inventory management. Furthermore, it could potentially increase capacity utilization and reduce production throughput time. The following suggestions was proposed in section 6.4:

Dynamic Scheduling System

Predictive information from our model can be incorporated into scheduling activities in breweries, creating a dynamic scheduling system, as described in section 6.4.1. This can enhance craft breweries ability to accurately manage the flow of material by allowing them to plan for the actual completion of fermentaions instead of reacting to it after it occurs.

However, predictive information from our model can only be made available to the production planner with acceptable accuracy 40 hours into the batch. As a result, the level of predictability will not be increased during planning phases and when the production starts, because predictive information is not available at that time. Considering the importance of planning with a sufficient time frame described in section 4.1 where fermentations never exceeds the time allocated, spare capacity would still be needed. Furthermore, our model has an estimated average error of 13 hours. Schedules updated by the dynamic scheduling system must therefore account for an estimated error of 13 hours. Consequently, a buffer, or spare capacity, must also be incorporated in schedules updated 40 hours into fermentation.

The original buffer, or days of spare capacity, is significantly longer than the models estimate of error, as seen in section 5.6.4. As a result they can reduce the buffer significantly during fermentation, as opposed to have it in the schedule throughout fermentation. Furthermore, in order to move beer out of the fermenter vessel, downstream processes, and the necessary equipment, must be prepared and have the necessary capacity. This solution will enable a proactive approach, by allowing breweries to initiate the required actions in order to ensure that the necessary capacity is available by the time fermentation is complete, as supposed to doing it when they have recognized that fermentation is complete. Consequently, they will be able to move from the traditional reactive approach and plan for downstream activities days in advance and with a significantly higher degree of accuracy, potentially making the flow of material much more efficient.

Inventory Management System

In terms of inventory management, predictive information can help breweries to maintain more accurate and efficient inventory levels by implementing a inventory management system as described in section 6.4.1. By knowing when a batch of beer will be ready for packaging and distribution, breweries can avoid overstocking products. This can reduce the risk of product spoilage, reduce storage costs, and ensure that breweries are better able to meet customer demand without too excessive inventory levels.

An inventory management system with predictive information incorporated will provide breweries with enhanced insights regarding when different inventory stages will be replenished. If such a system automatically updates every time prediction information from the fermentation prices becomes available, it will allow the system to track and monitor future inventory levels with a timeline. With the current fermentation uncertainty, they must account for the buffer, or days of spare capacity, when future inventory is estimated. The proposed solution will allow this information to be updated with an average error estimate of 13 hours, instead of the current buffer. Thus can the production schedule be updated so it is more in tact with the forecast for customer demand. As a result, it will allow breweries to reduce their WIP inventory levels throughout production as well as reduce their finished goods inventory.

As with a dynamic scheduling system, the inventory management system will be limited by the fact that prediction information is not reliable until 40 hours into fermentation. Having access to prediction information before the production cycle is initiated would give breweries significantly more flexibility and allow them to accelerate production in times of high demand or delay in times of lower demand to avoid excess inventory.

Applicability of the proposed solutions

The reliability of predictive information is crucial for its effective integration into the production planning and control systems of breweries. As outlined in section 6.3.5, craft breweries must schedule their operations with a significant lead time to ensure smooth production flow and the ability to meet customer demand. If the predictive information about the fermentation process is not consistently accurate, it could lead to serious disruptions in the production process. For instance, if a batch of beer requires more fermentation time than predicted, it could potentially halt the entire production line. This could lead to increased idle time, inefficient use of resources, and even missed deadlines for product delivery. Such disruptions could enhance the challenges breweries face due to fermentation uncertainty, rather than mitigating them. Breweries need to have confidence in the accuracy of the predictive information, knowing that it will provide a true reflection of the fermentation process time, batch after batch. This trust in the predictive information is crucial for it to be effectively utilized in managing the flow of materials, inventory management, and production planning and scheduling.

While predictive information about the fermentation process time can potentially mitigate some of the challenges faced by craft breweries, it's important to recognize that uncertainty about the fermentation process time is not the sole reason for their challenges. The challenges of managing the flow of materials, inventory management, and production planning and scheduling are multifaceted, and are influenced by a range of factors beyond the uncertainty of the fermentation process time. For instance, bottlenecks in production, particularly during the transition from batch to continuous production, can significantly impact the efficiency of operations. Even with more accurate predictions about the fermentation process time, these bottlenecks will exist. Consequently, while predictive information about the fermentation process time can be a valuable tool for mitigating some of the challenges faced by craft breweries, it is not a standalone solution. Breweries must consider a holistic approach that addresses the various factors contributing to these challenges in order to fully optimize their production planning and control.

7.2.3 Generalizability

The generalizability of this study's findings is influenced by the characteristics of the participating breweries in the multiple case study and the subsequent in-depth single case study. The multiple case study involved five craft breweries, which providing a broader perspective than a single case study, but were relatively similar in size. As seen in 4.1 four of these breweries fall into the category of smaller breweries with regards to production volume, with only one classified as a regional brewery. This similarity in size could potentially limit the applicability of the findings to larger or smaller breweries.

The influence of local laws and regulations on breweries' operations also suggests that the geographical diversity of the sample could impact the generalizability of the findings. Including breweries from a wider range of locations could have provided a more comprehensive understanding of how different regulatory environments affect production planning and control. Moreover, conducting in-depth case studies with more than one brewery from the multiple case study could have further enhanced the generalizability of the study by providing a more diverse range of perspectives and experiences. Despite these limitations, it's important to note that all the case companies reported similar challenges, many of which were linked, at least in part, to the uncertainty of the fermentation process. This suggests that the findings of this study could be relevant to a broad range of craft breweries facing similar challenges.

However, the specific strategies for addressing these challenges may vary depending on the unique circumstances of each brewery. For instance, Lock 27 Brewing Company's reliance on an external canning provider presents a unique challenge that may not be applicable to other breweries. Nevertheless, the potential benefits of implementing predictive information in production planning and control, as demonstrated in the case of Lock 27 Brewing Company, suggest that this approach could be beneficial in mitigating challenges and improving operational efficiency in a variety of craft brewery contexts.

7.2.4 Weaknesses

While this study provides valuable insights into the challenges faced by craft breweries due to the uncertainty of the fermentation process time, it is important to acknowledge its limitations. The sample size and diversity, particularly in the single case study, limit the generalizability of the findings. The study primarily focuses on a single brewery, Lock 27 Brewing Company, which may not be representative of all craft breweries. The multiple case study includes a broader range of breweries, but they are still relatively similar in size and operation.

Geographical limitations also may affect the generalizability of the results. All the breweries in the study are located in the United States or Norway, and the practices and challenges they face may be influenced by local factors such as regulations, market conditions, and cultural preferences. This geographical focus, combined with the limited sample size and diversity, means that the findings may not apply to breweries in other regions or to breweries of different sizes or operational styles.

The study relies mostly on the interpretation of qualitative data, which can be subjective. While efforts was made to ensure an objective and rigorous analysis, different researchers might interpret the data in different ways. Additionally, the study is limited by its temporal scope. It provides a snapshot of the craft breweries at a specific point in time, and the practices and challenges they face may evolve over time.

The reliance on self-reported data from the breweries is another limitation. While this data provides valuable insights, it may be subject to biases and inaccuracies. For instance, the breweries might overestimate or underestimate certain figures, or they might unintentionally leave out important information. Most of our case companies rely on a single person for production management, and they use different methods for documentation, which can affect the quality of the data we received. This means that the information in our analysis, and particularly the numbers presented, might not be an absolute reflection of the actual truth.

The study lacks experimental control, observing the breweries as they are, without manipulating any variables. This means that it can identify correlations, but it cannot establish definitive causal relationships. Furthermore, the study does not actually implement predictive information into production planning and control at any brewery. The potential benefits of such implementation are discussed, but they are theoretical and would need practical implementation to be tested and validated.

Finally, it's important to consider the feasibility of implementing predictive information systems. Many craft breweries are small businesses, and investing heavily in such systems might not be feasible or cost-effective for them. Despite these limitations, this study provides valuable insights into the challenges faced by craft breweries due to the uncertainty of the fermentation process time, and it suggests potential strategies for mitigating these challenges.

7.2.5 Summary

- Craft breweries face challenges in managing material flow, inventory, and production planning due to fermentation process time uncertainty.
- Predictive information about fermentation duration can potentially mitigate these challenges, enhancing operational efficiency.
- Mitigating these challenges can increase capacity utilization, reduce throughput time, improve product shelf life, and save energy.
- The study's findings may be applicable to a broad range of craft breweries, but strategies may vary based on each brewery's unique circumstances.
- Limitations of the study include limited sample size and diversity, geographical limitations, reliance on self-reported data, and lack of experimental control.

Chapter 8

Conclusion

This chapter summarises the key findings of the thesis, evaluates the fulfillment of the research objectives, and presents concluding reflections. A description of how the thesis contributed to knowledge is presented. Further, limitations of the research process are reflected upon, followed by suggestions for further work.

8.0.1 Summary of findings

This research has explored how machine learning can be used to predict the fermentation process, and how this prediction affects PPC activities in breweries. The purpose of this study was to resolve the existing research gap and answer the research objectives.

The first research objective concerned finding the most suitable forecasting approaches in order to forecast fermentation processes. This research revealed that previous studies had shown great success in implementing machine learning models. In addition, the literature study revealed that recently developed ML models are suitable for learning complex patterns and give accurate predictions, if they have access to sufficient data. We identified that supervised machine learning on time series was the most suitable approach. However, using only the finish time of the fermentation as the target variable should be reevaluated, as brewers might benefit from additional relevant information.

The second research objective was to develop and evaluate machine learning models. With inspiration from previous work we developed a neural network, two gradient boosted models and an automatic machine learning model. The models were trained on 40, 60 and 80 hours of batch information. The best models were able to predict the end of fermentation to a certain degree. Notably, the predictions are highly accurate for batches that complete fermentation within 60-110 hours, but less accurate for batches that extend beyond 120 hours or display abnormal activity. The models trained on 60 hours did not show any particular improvement compared to those trained on 40 hours. However, training on 80 hours of information showed great ability to predict the end fermentation time. By training models on more data and adding a metric presenting the certainty of the model's predictions, sufficient value would be added to the research.

The third research objective focus on identifying challenges caused by uncertainty in the fermentation process. Through literature and in-depth analysis of interviews and information, key challenges was identified. One of the challenges found was managing the flow of material. Uncertainty in fermentation causes breweries to add a buffer period to the production plan. This leads to an accumulation of work-in-progress inventory within the fermenter vessels and a decrease in capacity utilization. Secondly, the case study revealed that inventory mangement challenges further amplifies the effects of increased work-in-process inventory as well as idle time. Lastly, it is challenging to create production schedules when lead time is uncertain. To compensate for this, breweries add capacity buffers, which increase the production throughput time while the actual process time remains the same.

The fourth research objective concerned how the identified challenges caused by the uncertainty can be mitigated. Our study suggests two changes in breweries' PPC activities that can reduce the effect fermentation uncertainty causes. Firstly, predictive information from the ML models can be incorporated to create a dynamic scheduling system. This allows breweries to implement a proactive approach, as opposed to the current reactive approach. As a result, plans can be updated days in advance of completion, which can allow breweries to more efficiently proceed with downstream processes. Additionally, predictive information can be used to maintain a more accurate and efficient inventory level. By integrating predictive information in inventory management systems. This will allow breweries to estimate future inventory levels with more accuracy, which in turn reduces the need for work-in-process inventory along with production. It can also allow breweries to reduce the amount of finished goods in inventory. The reason is that they will know the data inventory will be replenished, instead of basing it on worst-case scenarios.

The last research objective was to investigate which benefits the mitigation of challenges caused by fermentation could have for craft breweries. Through the literature and case study we identify that increased accuracy in fermentation predictions can cause additional benefits. The reduction in idle time and WIP inventory can lead to an increase in potential capacity utilization. Predictive results can also be utilized to monitor the quality of the beer, which can cause a reduction in spoilage and product waste. To conclude, our study theoretically identifies values that increased fermentation accuracy possibly can generate by increasing production capacity, less spoilage, and energy reduction.

8.0.2 Knowledge contribution

Previous literature has either focused on developing models with the aim of forecasting fermentation or investigated how PPC activities can be improved in breweries. This research, however, combines these topics by using real fermentation data to develop forecasting models and investigating how this information can be applied in a real brewery. Compared to previous research that also has developed machine learning models on real fermentation data, this study stands out by using more data with better quality and with a higher frequency of measuring. Hence, this research provides valuable information indicating how to increase fermentation predictability and the potential effects it has on PPC activities with the usage of a large dataset.

8.0.3 Limitations

The aim of the study was to investigate how fermentation activity can be predicted. Several state-of-the-art models were applied. However, as mentioned in section 3.6.6 there exist and it is currently being developed more modern approaches that could be suitable for the problem. Even though the models implemented have been hyperparameter tuned, tuning of parameters is a topic of its own and could be expanded. Moreover, although the models developed were useful, they were relatively simple and have not yet been implemented in a real brewery setting, meaning the benefits identified are, at this stage, mainly theoretical.

The study is largely conceptual in nature, and the proposed strategies of incorporating prediction information in production planning and control activities have not been practically implemented or tested within a craft brewery setting. Therefore, we lack concrete evidence that these measures would be effective or beneficial when applied in practice. Secondly, our findings are mostly based on qualitative data, and these have not been validated with quantitative data. This lack of cross-verification leaves room for potential inaccuracies in the data used, which could potentially influence the validity of our conclusions.

8.0.4 Future work

Several directions for further research have been identified. Firstly, the exploration of other machine learning techniques, such as RNN, LSTM, and encoder-decoder sequential models, may lead to improvements in the predictive models.

Further, hyperparameter tuning of the neural network and collection of additional high-quality data could be beneficial. Most importantly, implementing the developed models in a real brewery and monitoring their effects on PPC activities is recommended. This would provide a practical understanding of the benefits of the approach, creating a solid foundation for the improvement and expansion of the models.

Future research could build on this study by addressing its limitations, for instance by including a larger and more diverse sample of breweries, by conducting longitudinal research to observe changes over time, or by implementing and evaluating predictive information systems in practice. In particular, the practical implementation of predictive information into production planning and control activities within a craft brewery would be an promising research direction. This approach would allow for the gathering of evidence about the potential benefits of such systems, offering a better insight into their real-world applicability and effectiveness. Furthermore, a practical application would not only test the concepts and models developed in this study but also pave the way for refining them based on direct industry experience.

8.0.5 Recommendations for Plaato

This study revealed the challenges of getting accurate predictions based on the available data. Compared to other studies, the data used in this research was of higher volume and with a higher rate of sampling. However, a factor that restricted optimal model performance was the data quality. We lost 64% of our data because of errors. The study revealed that data preparation, cleaning and access to large amounts of data is more important than the development of machine learning models. We, therefore, advise exploring how the quality of the data can be improved. For instance, a particular challenge was how the fermentations started at different times. Improvement of data quality combined with the large amount of new data Plaato receives weekly, the forecasts could become more accurate.

Our study revealed that previous researchers had great success in fermentation forecasting by using features such as amino concentration, CO_2 and oxygen level. In this study, we only had access to density and temperature measured in a batch. We therefore suggest exploring the possibilities of measuring additional features.

Through interviews and analysis of Plaato's customers, we identified a loyal customer base with a genuine interest in how machine learning can be used to improve activities in their brewery. This interest can be utilized to retrieve valuable insights about how machine learning can be used in order to gain the most value for the breweries. For instance, we developed models that predicted the end of fermentation. By including breweries in discussions one could identify what should be predicted in order to gain value. The best way to investigate how accurate the developed models are, and how they mitigate PPC challenges is to actually implement the models in some breweries.

References

Adetunji, Abigail Bola et al. (2022). "House price prediction using random forest machine learning technique". In: *Procedia Computer Science* 199, pp. 806–813.

Aggarwal, Charu C and Charu C Aggarwal (2017). *An introduction to outlier analysis.* Springer.

Ahmed, Shehzad et al. (2021). "Towards supply chain visibility using internet of things: A dyadic analysis review". In: *Sensors* 21.12, p. 4158.

Akundi, Aditya et al. (2022). "State of Industry 5.0—Analysis and identification of current research trends". In: *Applied System Innovation* 5.1, p. 27.

Alba-Lois, L and C Segal-Kischinevzky (2010). "Beer & wine makers". In: *Nature.*

Albach, Horst (2015). *Management of permanent change.* Springer.

Albouq, Sami S et al. (2022). "A survey of interoperability challenges and solutions for dealing with them in IoT environment". In: *IEEE Access* 10, pp. 36416–36428.

Alfaro, Martin (2022). "Beer leaders against craft brewers: will Goliath end up defeating David in the US?" In: *Applied Economics* 54.5, pp. 554–567.

Anshari, Muhammad et al. (2019). "Customer relationship management and big data enabled: Personalization & customization of services". In: *Applied Computing and Informatics* 15.2, pp. 94–101.

Armstrong, M (2022). *America's Brewery Boom.*

Arnold, JR Tony (2020). "Introduction to materials management". In.

Aroh, Kenechukwu (2019). "Beer production". In: *Available at SSRN 3458983.*

Atieh, Anas M et al. (2016). "Performance improvement of inventory management system processes by an automated warehouse management system". In: *Procedia Cirp* 41, pp. 568–572.

Atwell, Charlotte et al. (2017). "Optimization of cleaning detergent use in brewery fermenter cleaning". In: *Journal of the Institute of Brewing* 123.1, pp. 70–76.

Atzori, Luigi, Antonio Iera, and Giacomo Morabito (2010). "The internet of things: A survey". In: *Computer networks* 54.15, pp. 2787–2805.

Axtman, B (2006). "Ripe opportunities". In: *Progressive Grocer* 85.4, pp. 76–80.

Azambuja, Antonio João Gonçalves de et al. (2023). "Artificial Intelligence-Based Cyber Security in the Context of Industry 4.0—A Survey". In: *Electronics* 12.8, p. 1920.

Badillo, Solveig et al. (2020). "An introduction to machine learning". In: *Clinical pharmacology & therapeutics* 107.4, pp. 871–885.

Baker, Pam (2018). *Why craft beer makers are tapping into ERP for craft brewery software.* URL: https://www.techtarget.com/searcherp/feature/Why-craft-beer-makers-are-tapping-into-ERP-for-brewery-software (visited on Nov. 12, 2022).

Baruch, Jack Raifer (2023). *Day 19 of DataScience28: Hyperparameter Optimization.* Accessed: 2023-05-25. URL: https://jackraiferbaruch.medium.com/day-19-of-datascience28-hyperparameter-optimization-c4ac2f9fd622.

Basalla, George (1988). *The evolution of technology.* Cambridge University Press.

Beheshti, Hooshang M (2006). "What managers should know about ERP/ERP II". In: *Management Research News* 29.4, pp. 184–193.

Bergstra, James and Yoshua Bengio (2012). "Random search for hyper-parameter optimization." In: *Journal of machine learning research* 13.2.

Bharadwaj, Anandhi et al. (2013). "Digital business strategy: toward a next generation of insights". In: *MIS quarterly*, pp. 471–482.

Bhonsale, Satyajeet, Wannes Mores, and Jan Van Impe (2021). "Dynamic Optimisation of Beer Fermentation under Parametric Uncertainty". In: *Fermentation* 7.4, p. 285.

Biamonte, Jacob et al. (2017). "Quantum machine learning". In: *Nature* 549.7671, pp. 195–202.

Bishop, Christopher M and Nasser M Nasrabadi (2006). *Pattern recognition and machine learning*. Vol. 4. 4. Springer.

Björkdahl, Joakim (2020). "Strategies for digitalization in manufacturing firms". In: *California Management Review* 62.4, pp. 17–36.

Bjørnbet, Marit Moe et al. (2021). "Circular economy in manufacturing companies: A review of case study literature". In: *Journal of Cleaner Production* 294, p. 126268.

Blázquez-Garcıa, Ane et al. (2021). "A review on outlier/anomaly detection in time series data". In: *ACM Computing Surveys (CSUR)* 54.3, pp. 1–33.

Bokulich, Nicholas A and Charles W Bamforth (2013). "The microbiology of malting and brewing". In: *Microbiology and Molecular Biology Reviews* 77.2, pp. 157–172.

Bortolini, Marco, Francesco Gabriele Galizia, and Cristina Mora (2018). "Reconfigurable manufacturing systems: Literature review and research trend". In: *Journal of manufacturing systems* 49, pp. 93–106.

Bowler, Alexander et al. (2021). "Predicting alcohol concentration during beer fermentation using ultrasonic measurements and machine learning". In: *Fermentation* 7.1, p. 34.

Brand, Donald A et al. (2007). "Comparative analysis of alcohol control policies in 30 countries". In: *PLoS medicine* 4.4, e151.

Brewery, Fuller's (2019). *How To Make Beer, According To A Brewer*. URL: https://www.fullers.co.uk/blog/beer-articles/how-to-make-beer (visited on Oct. 20, 2022).

Brewing Chemists, American Society of (2019). *ASBC Sampling Plan*. URL: https://www.asbcnet.org/lab/samplingplan/Pages/default.aspx#Malt (visited on Dec. 12, 2022).

Brockwell, Peter J and Richard A Davis (2002). *Introduction to time series and forecasting*. Springer.

Brown, Robert Goodell (2004). *Smoothing, forecasting and prediction of discrete time series*. Courier Corporation.

Brownlee, Jason (2016). "Time series forecasting as supervised learning". In: *Machine Learning Mastery. Recuperado de de: https://machinelearningmastery. com/time-series-forecasting-supervised-learning*.

— (2023a). *A Gentle Introduction to Scikit-Learn: A Python Machine Learning Library*. Accessed: 2023-05-25. URL: https://machinelearningmastery.com/a-gentle-introduction-to-scikit-learn-a-python-machine-learning-library/.

— (2023b). *Understand the Dynamics of Learning Rate on Deep Learning Neural Networks*. Accessed: 2023-06-05. URL: https://machinelearningmastery.com/understand-the-dynamics-of-learning-rate-on-deep-learning-neural-networks/ (visited on June 5, 2023).

— (2023c). *Your First Deep Learning Project in Python with Keras Step-By-Step*. Accessed: 2023-05-25. URL: https://machinelearningmastery.com/tutorial-first-neural-network-python-keras/.

Brynjolfsson, Erik and Lorin M Hitt (2000). "Beyond computation: Information technology, organizational transformation and business performance". In: *Journal of Economic perspectives* 14.4, pp. 23–48.

Bueno, Adauto, Moacir Godinho Filho, and Alejandro G Frank (2020). "Smart production planning and control in the Industry 4.0 context: A systematic literature review". In: *Computers & Industrial Engineering* 149, p. 106774.

Buer, Sven-Vegard, Marco Semini, et al. (2021). "The complementary effect of lean manufacturing and digitalisation on operational performance". In: *International Journal of Production Research* 59.7, pp. 1976–1992.

Buer, Sven-Vegard, Jan Ola Strandhagen, and Felix TS Chan (2018). "The link between Industry 4.0 and lean manufacturing: mapping current research and establishing a research agenda". In: *International journal of production research* 56.8, pp. 2924–2940.

Butollo, Florian (2021). "Digitalization and the geographies of production: Towards reshoring or global fragmentation?" In: *Competition & Change* 25.2, pp. 259–278.

CatBoost (2023). URL: https://catboost.ai/en/docs/ (visited on Jan. 5, 2023).

Cerqueira, Victor (2023). *Introduction to Global Forecasting Models*. https://towardsdatascience.com/introduction-to-global-forecasting-models-3ca8e69a6524. Accessed: yyyy-mm-dd.

Chae, Bongsug Kevin (2015). "Insights from hashtag# supplychain and Twitter Analytics: Considering Twitter and Twitter data for supply chain practice and research". In: *International Journal of Production Economics* 165, pp. 247–259.

Chandola, Varun, Arindam Banerjee, and Vipin Kumar (2009). "Anomaly detection: A survey". In: *ACM computing surveys (CSUR)* 41.3, pp. 1–58.
Chapman, Stephen N (2006). *The fundamentals of production planning and control.* Pearson/Prentice Hall Upper Saddle River, NJ, USA.
Chen, Baotong et al. (2017). "Smart factory of industry 4.0: Key technologies, application case, and challenges". In: *Ieee Access* 6, pp. 6505–6519.
Chen, Tianqi et al. (2015). "Xgboost: extreme gradient boosting". In: *R package version 0.4-2* 1.4, pp. 1–4.
Chollet, Francois (2021). *Deep learning with Python.* Simon and Schuster.
Chopra, Sunil and Peter Meindl (2001). "Strategy, planning, and operation". In: *Supply Chain Management*, pp. 13–17.
Christopher, Martin (2016). *Logistics & supply chain management.* Pearson Uk.
Çınar, Zeki Murat et al. (2020). "Machine learning in predictive maintenance towards sustainable smart manufacturing in industry 4.0". In: *Sustainability* 12.19, p. 8211.
Clark, Andrew J and Herbert Scarf (1960). "Optimal policies for a multi-echelon inventory problem". In: *Management science* 6.4, pp. 475–490.
Clemons, Eric K, Guodong Gordon Gao, and Lorin M Hitt (2006). "When online reviews meet hyperdifferentiation: A study of the craft beer industry". In: *Journal of management information systems* 23.2, pp. 149–171.
Cohen, Yuval et al. (2019). "Assembly systems in Industry 4.0 era: a road map to understand Assembly 4.0". In: *The International Journal of Advanced Manufacturing Technology* 105.9, pp. 4037–4054.
Conway, J (2021). *Beer production worldwide from 1998 to 2020.*
Cox, James F and John H Blackstone (2002). *APICS dictionary.* APICS Educational Society for Resource Management.
Craig, Christopher A (2021). "Managing a Craft Brewery: Competitive, Legal, and Environmental Forces". In: *Entrepreneurship Education and Pedagogy* 4.3, pp. 492–499.
Cryer, Jonathan D and Natalie Kellet (1991). *Time series analysis.* Springer.
Dabbagh, Sajjad Rahmani et al. (2021). "3D-printed microneedles in biomedical applications". In: *Iscience* 24.1, p. 102012.
Daskin, Mark S, Collette R Coullard, and Zuo-Jun Max Shen (2002). "An inventory-location model: Formulation, solution algorithm and computational results". In: *Annals of operations research* 110.1-4, pp. 83–106.
Dastres, Roza and Mohsen Soori (2021). "Artificial neural network systems". In: *International Journal of Imaging and Robotics (IJIR)* 21.2, pp. 13–25.
Dave, Vachik S and Kamlesh Dutta (2014). "Neural network based models for software effort estimation: a review". In: *Artificial Intelligence Review* 42.2, pp. 295–307.
Davison, Anthony Christopher (2003). *Statistical models.* Vol. 11. Cambridge university press.
De Mantaras, Ramon Lopez and Enric Plaza (1997). "Case-based reasoning: an overview". In: *AI communications* 10.1, pp. 21–29.
De Villiers, Ethel-Michele et al. (2004). "Classification of papillomaviruses". In: *Virology* 324.1, pp. 17–27.
Defernez, M et al. (2007). "Modelling beer fermentation variability". In: *Journal of food engineering* 83.2, pp. 167–172.
Deisenroth, Marc Peter, A Aldo Faisal, and Cheng Soon Ong (2020). *Mathematics for machine learning.* Cambridge University Press.
Dequin, Sylvie (2001). "The potential of genetic engineering for improving brewing, wine-making and baking yeasts". In: *Applied microbiology and biotechnology* 56, pp. 577–588.
Dıaz, Ana Belén et al. (2022). "From the Raw Materials to the Bottled Product: Influence of the Entire Production Process on the Organoleptic Profile of Industrial Beers". In: *Foods* 11.20, p. 3215.
Dobra, Alin (2009). *Decision Tree Classification.*
Dodd, Sarah et al. (2021). "Crafting growth together". In: *Entrepreneurship & Regional Development* 33.9-10, pp. 789–815.
Dojo, Data Science (2023). *Machine Learning 101.* Accessed: 2023-05-25. URL: https://datasciencedojo.com/blog/machine-learning-101/.
Dong, Xibin et al. (2020). "A survey on ensemble learning". In: *Frontiers of Computer Science* 14, pp. 241–258.

Donk, Dirk Pieter van (2000). "Customer-driven manufacturing in the food processing industry". In: *British food journal*.
Donselaar, Karel van et al. (2006). "Inventory control of perishables in supermarkets". In: *International journal of production economics* 104.2, pp. 462–472.
Draper, Norman R and Harry Smith (1998). *Applied regression analysis*. Vol. 326. John Wiley & Sons.
Drucker, Harris (1997). "Improving regressors using boosting techniques". In: *Icml*. Vol. 97. Citeseer, pp. 107–115.
Dubey, Rameshwar et al. (2020). "Big data analytics and artificial intelligence pathway to operational performance under the effects of entrepreneurial orientation and environmental dynamism: A study of manufacturing organisations". In: *International journal of production economics* 226, p. 107599.
Duda, Richard O, Peter E Hart, et al. (2006). *Pattern classification*. John Wiley & Sons.
Durkin, Gregory E (2006). "What Does Granholm v. Heald Mean for the Future of the Twenty-First Amendment, the Three-Tier System, and Efficient Alcohol Distribution". In: *Wash. & Lee L. Rev.* 63, p. 1095.
Edgar, Thomas W and David O Manz (2017). "Chapter 6-machine learning". In: *Research methods for cyber security*, pp. 153–173.
Ekpenyong, Maurice et al. (2021). "Statistical and artificial neural network approaches to modeling and optimization of fermentation conditions for production of a surface/bioactive glyco-lipo-peptide". In: *International journal of peptide research and therapeutics* 27, pp. 475–495.
El Naqa, Issam and Martin J Murphy (2015). *What is machine learning?* Springer.
Emmert-Streib, Frank and Matthias Dehmer (2019). "Evaluation of regression models: Model assessment, model selection and generalization error". In: *Machine learning and knowledge extraction* 1.1, pp. 521–551.
Endsley, Mica R (2017). "From here to autonomy: lessons learned from human–automation research". In: *Human factors* 59.1, pp. 5–27.
Feurer, Matthias and Frank Hutter (2019). "Hyperparameter optimization". In: *Automated machine learning: Methods, systems, challenges*, pp. 3–33.
Freedman, David A (2009). *Statistical models: theory and practice*. cambridge university press.
Garcıa, Ana I, Luis A Garcıa, and Mario Dıaz (1994). "Modelling of diacetyl production during beer fermentation". In: *Journal of the Institute of Brewing* 100.3, pp. 179–183.
Georgiadis, Georgios P, Apostolos P Elekidis, and Michael C Georgiadis (2021). "Optimal production planning and scheduling in breweries". In: *Food and Bioproducts Processing* 125, pp. 204–221.
Georgiadis, Georgios P, Borja Mariño Pampın, et al. (2020). "Optimal production scheduling of food process industries". In: *Computers & Chemical Engineering* 134, p. 106682.
Ghobakhloo, Morteza (2020). "Industry 4.0, digitization, and opportunities for sustainability". In: *Journal of cleaner production* 252, p. 119869.
Goldsborough, Peter (2016). "A tour of tensorflow". In: *arXiv preprint arXiv:1610.01178*.
Gonzalez Rodriguez, German, Jose M Gonzalez-Cava, and Juan Albino Méndez Pérez (2020). "An intelligent decision support system for production planning based on machine learning". In: *Journal of Intelligent Manufacturing* 31.5, pp. 1257–1273.
Gonzalez-Neira, Eliana Maria et al. (2017). "A biased-randomized simheuristic for the distributed assembly permutation flowshop problem with stochastic processing times". In: *Simulation Modelling Practice and Theory* 79, pp. 23–36. ISSN: 1569-190X. DOI: https://doi.org/10.1016/j.simpat.2017.09.001. URL: https://www.sciencedirect.com/science/article/pii/S1569190X17301338.
Goodfellow, Ian, Yoshua Bengio, and Aaron Courville (2016). *Deep Learning*. http://www.deeplearningbook.org. MIT Press.
Goodman, Leo A (1961). "Snowball sampling". In: *The annals of mathematical statistics*, pp. 148–170.
Gopal, CV, JRM Hammond, and JM Pye (1993). "Interactive control of beer fermentation: Demonstration at the laboratory scale". In: *Technical quarterly (USA)*.
Grassel, Joshua et al. (2021). "The Craft Beer Game and the Value of Information Sharing". In: *International Conference on Computational Logistics*. Springer, pp. 222–236.
Groover, Mikell P (2020). *Fundamentals of modern manufacturing: materials, processes, and systems*. John Wiley & Sons.

Gubbi, Jayavardhana et al. (2013). "Internet of Things (IoT): A vision, architectural elements, and future directions". In: *Future generation computer systems* 29.7, pp. 1645–1660.
Gupta, Sushil and Martin Starr (2014). *Production and operations management systems*. CRC Press.
H2O auto (2022). *H2O Documentation: Automatic Machine Learning (AutoML)*. Accessed: 2023-05-25. URL: https://docs.h2o.ai/h2o/latest-stable/h2o-docs/automl.html.
H2O ensemble (2022). *H2O Documentation: Stacked Ensembles*. Accessed: 2023-05-25. URL: https://docs.h2o.ai/h2o/latest-stable/h2o-docs/data-science/stacked-ensembles.html.
Haber, Aleksandar (2023). *Decomposition of time series in Python - trend and seasonal periodic components*. Accessed: 2023-05-25. URL: https://aleksandarhaber.com/decomposition-of-time-series-in-python-trend-and-seasonal-periodic-components/.
Haleem, Abid and Mohd Javaid (2019). "Industry 5.0 and its applications in orthopaedics". In: *Journal of Clinical Orthopaedics & Trauma* 10.4, pp. 807–808.
Han, Jiawei, Jian Pei, and Hanghang Tong (2022). *Data mining: concepts and techniques*. Morgan kaufmann.
Hancock, John T and Taghi M Khoshgoftaar (2020). "CatBoost for big data: an interdisciplinary review". In: *Journal of big data* 7.1, pp. 1–45.
Hartmeier, Winfried and Monika Reiss (2011). "Production of beer and wine". In: *Industrial Applications*. Springer, pp. 59–77.
Hastie, Trevor et al. (2009). *The elements of statistical learning: data mining, inference, and prediction*. Vol. 2. Springer.
Haykin, Simon (2009). *Neural networks and learning machines, 3/E*. Pearson Education India.
Heaton, Jeff (2017). "Heaton research: The number of hidden layers". In: *Availablefrom: https://www. heatonresearch. com/2017/06/01/hidden-layers. html [13/3/2019]*.
Heizer, J, B Reder, and C Muson (2016). *Operations Management: Sustainability and Supply Chain Management JAY*.
Henning, Bernd and Jens Rautenberg (2006). "Process monitoring using ultrasonic sensor systems". In: *Ultrasonics* 44, e1395–e1399.
Herzen, Julien (2023). *Training Forecasting Models*. Accessed: 2023-05-25. URL: https://unit8.com/resources/training-forecasting-models/.
Heum and Hjort (2022). "Improvement of Production Planning and Control in Craft Breweries With Real-Time Fermentation Data". In.
Hewamalage, Hansika, Christoph Bergmeir, and Kasun Bandara (2022). "Global models for time series forecasting: A simulation study". In: *Pattern Recognition* 124, p. 108441.
Hoerl, Arthur E and Robert W Kennard (1970). "Ridge regression: Biased estimation for nonorthogonal problems". In: *Technometrics* 12.1, pp. 55–67.
Horváth, Dóra and Roland Zs Szabó (2019). "Driving forces and barriers of Industry 4.0: Do multinational and small and medium-sized companies have equal opportunities?" In: *Technological forecasting and social change* 146, pp. 119–132.
Hovden, I (2019). *Optimizing Artificial Neural Network Hyperparameters and Architecture*.
Hutter, Frank, Holger H Hoos, and Kevin Leyton-Brown (2012). "Parallel algorithm configuration". In: *Learning and Intelligent Optimization: 6th International Conference, LION 6, Paris, France, January 16-20, 2012, Revised Selected Papers*. Springer, pp. 55–70.
Hyndman, Rob J and George Athanasopoulos (2018). *Forecasting: principles and practice*. OTexts.
Isaksson, Alf J, Iiro Harjunkoski, and Guido Sand (2018). "The impact of digitalization on the future of control and operations". In: *Computers & Chemical Engineering* 114, pp. 122–129.
Jacobs, F Robert et al. (2011). *Manufacturing planning and control for supply chain management: APICS/CPIM Certification Edition*. McGraw-Hill Education.
Jain, Anil K, Jianchang Mao, and K Moidin Mohiuddin (1996). "Artificial neural networks: A tutorial". In: *Computer* 29.3, pp. 31–44.
Jan, Zohaib et al. (2022). "Artificial Intelligence for Industry 4.0: Systematic Review of Applications, Challenges, and Opportunities". In: *Expert Systems with Applications*, p. 119456.
Jenkins, A (2022). *Inventory Turnover Ratio Defined: Formula, Tips, Examples*.
Johnson, CR, KJ Burnham, and DJG James (1998). "Extended simulation model for a brewery fermentation process". In: *International Conference on Simulation'98.(Conf. Publ. No. 457)*. IET, pp. 259–262.
Kaggle: Your Machine Learning and Data Science Community (2023). Accessed: 2023-04-07. URL: https://www.kaggle.com/.

Kağnıcıoğlu, C.H et al. (2019). *Operations Management*. T.C. Anadolu University Publication.
Kanchev, Hristiyan et al. (2011). “Energy management and operational planning of a microgrid with a PV-based active generator for smart grid applications”. In: *IEEE transactions on industrial electronics* 58.10, pp. 4583–4592.
Karlsson, Christer (2010). *Researching operations management*. Routledge.
Karrer, Christoph (2012). *Engineering production control strategies: a guide to tailor strategies that unite the merits of push and pull*. Springer Science & Business Media.
Keras (2023). *Keras: The Python Deep Learning API*. Accessed: 2023-05-25. URL: https://keras.io/.
Khan, Ibrahim Haleem and Mohd Javaid (2022). “Role of Internet of Things (IoT) in adoption of Industry 4.0”. In: *Journal of Industrial Integration and Management* 7.04, pp. 515–533.
Kiel, Daniel et al. (2017). “Sustainable industrial value creation: Benefits and challenges of industry 4.0”. In: *International journal of innovation management* 21.08, p. 1740015.
Kingsford, Carl and Steven L Salzberg (2008). “What are decision trees?” In: *Nature biotechnology* 26.9, pp. 1011–1013.
Kiran, DR (2019). *Production planning and control: A comprehensive approach*. Butterworth-heinemann.
Knight, Trevor David (1991). “The development of an integrated sales forecasting and production planning system for the brewing industry”. MA thesis. University of Cape Town.
Kolodner, Janet L (1992). “An introduction to case-based reasoning”. In: *Artificial intelligence review* 6.1, pp. 3–34.
Kothari, Chakravanti Rajagopalachari (2004). *Research methodology: Methods and techniques*. New Age International.
Kotsiantis, Sotiris B (2013). “Decision trees: a recent overview”. In: *Artificial Intelligence Review* 39, pp. 261–283.
Koulouris, Alexandros, Nikiforos Misailidis, and Demetri Petrides (2021). “Applications of process and digital twin models for production simulation and scheduling in the manufacturing of food ingredients and products”. In: *Food and Bioproducts Processing* 126, pp. 317–333.
Krogh, Anders (2008). “What are artificial neural networks?” In: *Nature biotechnology* 26.2, pp. 195–197.
Kwak, Jin Kyung (2019). “Analysis of inventory turnover as a performance measure in manufacturing industry”. In: *Processes* 7.10, p. 760.
LeCun, Yann, Yoshua Bengio, and Geoffrey Hinton (2015). “Deep learning”. In: *nature* 521.7553, pp. 436–444.
Lee, In and George Mangalaraj (2022). “Big data analytics in supply chain management: a systematic literature review and research directions”. In: *Big Data and Cognitive Computing* 6.1, p. 17.
Lee, Jay, Behrad Bagheri, and Hung-An Kao (2015). “A cyber-physical systems architecture for industry 4.0-based manufacturing systems”. In: *Manufacturing letters* 3, pp. 18–23.
Lee, Jay, Hung-An Kao, and Shanhu Yang (2014). “Service innovation and smart analytics for industry 4.0 and big data environment”. In: *Procedia cirp* 16, pp. 3–8.
Leitão, Paulo, Armando Walter Colombo, and Stamatis Karnouskos (2016). “Industrial automation based on cyber-physical systems technologies: Prototype implementations and challenges”. In: *Computers in industry* 81, pp. 11–25.
Lek, S and YS Park (2008). “Artificial neural networks”. In: *Encyclopedia of Ecology, Five-Volume Set*. Elsevier Inc., pp. 237–245.
Li, Bing et al. (2021). “Application of mechanistic modelling and machine learning for cream cheese fermentation pH prediction”. In: *Journal of Chemical Technology & Biotechnology* 96.1, pp. 125–133.
Li, Lin et al. (2007). “Bottleneck detection of manufacturing systems using data driven method”. In: *2007 IEEE international symposium on assembly and manufacturing*. IEEE, pp. 76–81.
Li, Suhong and Binshan Lin (2006). “Accessing information sharing and information quality in supply chain management”. In: *Decision support systems* 42.3, pp. 1641–1656.
Liaw, Andy, Matthew Wiener, et al. (2002). “Classification and regression by randomForest”. In: *R news* 2.3, pp. 18–22.
Lin, Jie et al. (2017). “A survey on internet of things: Architecture, enabling technologies, security and privacy, and applications”. In: *IEEE internet of things journal* 4.5, pp. 1125–1142.
LLC, MultiMedia (2015). *Craft Brewer Definition*. URL: https://www.brewersassociation.org/statistics-and-data/craft-brewer-definition (visited on Nov. 26, 2022).

LLC, MultiMedia (2019). *National Beer Sales Production Data.* URL: https://www.brewersassociation.org/statistics-and-data/national-beer-stats/ (visited on Nov. 26, 2022).

Loeffler, D (2006). "Modern brewery sanitation". In: *Brewing.* Elsevier, pp. 308–334.

Lordan, Ronan et al. (2019). "Total, neutral, and polar lipids of brewing ingredients, by-products and beer: Evaluation of antithrombotic activities". In: *Foods* 8.5, p. 171.

Lu, Yang (2017). "Industry 4.0: A survey on technologies, applications and open research issues". In: *Journal of industrial information integration* 6, pp. 1–10.

Lyu, Zhiyang et al. (2021). "Design and manufacture of 3D-printed batteries". In: *Joule* 5.1, pp. 89–114.

Maclin, Richard and David Opitz (1997). "An empirical evaluation of bagging and boosting". In: *AAAI/IAAI* 1997, pp. 546–551.

Madhulatha, T Soni (2012). "An overview on clustering methods". In: *arXiv preprint arXiv:1205.1117.*

Mahmoud, Rwan et al. (2015). "Internet of things (IoT) security: Current status, challenges and prospective measures". In: *2015 10th international conference for internet technology and secured transactions (ICITST).* IEEE, pp. 336–341.

Manaswi, Navin Kumar and Navin Kumar Manaswi (2018). "Rnn and lstm". In: *Deep Learning with Applications Using Python: Chatbots and Face, Object, and Speech Recognition With TensorFlow and Keras*, pp. 115–126.

Manyika, James et al. (2015). "Unlocking the Potential of the Internet of Things". In: *McKinsey Global Institute* 1.

Martin, Michael et al. (2022). "From Circular to Linear? Assessing the Life Cycle Environmental and Economic Sustainability of Steel and Plastic Beer Kegs". In: *Circular Economy and Sustainability*, pp. 1–24.

Mashayekhy, Yasaman et al. (2022). "Impact of Internet of Things (IoT) on inventory management: A literature survey". In: *Logistics* 6.2, p. 33.

Masini, Ricardo P, Marcelo C Medeiros, and Eduardo F Mendes (2023). "Machine learning advances for time series forecasting". In: *Journal of economic surveys* 37.1, pp. 76–111.

Master Brewers Association of the Americas: Publications (2023). Accessed: 2023-06-07. URL: https://www.mbaa.com/publications/Pages/default.aspx.

McCallum, Andrew (2005). "Information extraction: Distilling structured data from unstructured text". In: *Queue* 3.9, pp. 48–57.

Medium (2023). *Medium - Machine learning.* Accessed: 2023-03-07. URL: https://medium.com/tag/machine-learning.

Melkote, Sanjay and Mark S Daskin (2001). "Capacitated facility location/network design problems". In: *European journal of operational research* 129.3, pp. 481–495.

Mentzer, John T et al. (2001). "Defining supply chain management". In: *Journal of Business logistics* 22.2, pp. 1–25.

Mokyr, Joel (1992). *The lever of riches: Technological creativity and economic progress.* Oxford University Press.

Monardes, Carlos et al. (2017). "Production Planning Model for the assignment of Fermentation Tanks at Wineries". In: *Proceedings of the 1st International Conference on Agro Big Data and Decision Support Systems in Agriculture*, pp. 191–194.

Monroy, Diego Fernando Zuluaga and Cristhian Camilo Ruiz Vallejo (2012). "Production planning and resource scheduling of a brewery with plant simulation". In: *Use Cases of Discrete Event Simulation: Appliance and Research*, pp. 321–330.

Montague, Gary A, Elaine B Martin, and Christopher J O'Malley (2008). "Forecasting for fermentation operational decision making". In: *Biotechnology progress* 24.5, pp. 1033–1041.

Montero-Manso, P. and R.J. Hyndman (2021). "Principles and algorithms for forecasting groups of time series: Locality and globality". In: *International Journal of Forecasting* 37.4, pp. 1632–1653.

Moser, Albine and Irene Korstjens (2018). "Series: Practical guidance to qualitative research. Part 3: Sampling, data collection and analysis". In: *European journal of general practice* 24.1, pp. 9–18.

Moyne, James and Jimmy Iskandar (2017). "Big data analytics for smart manufacturing: Case studies in semiconductor manufacturing". In: *Processes* 5.3, p. 39.

Murphy, Kevin P (2022). *Probabilistic machine learning: an introduction.* MIT press.

Natekin, Alexey and Alois Knoll (2013). "Gradient boosting machines, a tutorial". In: *Frontiers in neurorobotics* 7, p. 21.

Oluyisola, Olumide Emmanuel et al. (2022). “Designing and developing smart production planning and control systems in the industry 4.0 era: a methodology and case study”. In: *Journal of Intelligent Manufacturing* 33.1, pp. 311–332.

Paleyes, Andrei, Raoul-Gabriel Urma, and Neil D Lawrence (2022). “Challenges in deploying machine learning: a survey of case studies”. In: *ACM Computing Surveys* 55.6, pp. 1–29.

Pandey, Parul (2019). *Democratising Machine Learning with H2O.* Accessed: 2023-05-25. URL: https://towardsdatascience.com/democratising-machine-learning-with-h2o-7f2f79e10e3f.

Pang, Bo, Erik Nijkamp, and Ying Nian Wu (2020). “Deep learning with tensorflow: A review”. In: *Journal of Educational and Behavioral Statistics* 45.2, pp. 227–248.

Parviainen, Päivi et al. (2017). “Tackling the digitalization challenge: how to benefit from digitalization in practice”. In: *International journal of information systems and project management* 5.1, pp. 63–77.

Pedregosa, F. et al. (2011). “Scikit-learn: Machine Learning in Python”. In: *Journal of Machine Learning Research* 12, pp. 2825–2830.

Perera, WKR, KA Dilini, and T Kulawansa (2018). “A review of big data analytics for customer relationship management”. In: *2018 3rd International Conference on Information Technology Research (ICITR).* IEEE, pp. 1–6.

Peres, Ricardo Silva et al. (2020). “Industrial artificial intelligence in industry 4.0-systematic review, challenges and outlook”. In: *IEEE Access* 8, pp. 220121–220139.

Pil, Frits K and Matthias Holweg (2004). “Linking product variety to order-fulfillment strategies”. In: *Interfaces* 34.5, pp. 394–403.

Plaato (2022). URL: https://plaato.io/pages/plaato-pro (visited on Nov. 13, 2022).

Platen, Patrict von (2023). *Introducing the Encoder-Decoder Framework.* Accessed: 2023-05-25. URL: https://huggingface.co/blog/encoder-decoder.

Podgorelec, Vili et al. (2002). “Decision trees: an overview and their use in medicine”. In: *Journal of medical systems* 26, pp. 445–463.

Pöllänen, Jermu, Juho Rousu, and Jukka Kronlöf (2001). “A Neural Network Tool for Brewery Fermentations”. In.

Qi, Qinglin and Fei Tao (2018). “Digital twin and big data towards smart manufacturing and industry 4.0: 360 degree comparison”. In: *Ieee Access* 6, pp. 3585–3593.

Qin, Jian, Ying Liu, and Roger Grosvenor (2016). “A categorical framework of manufacturing for industry 4.0 and beyond”. In: *Procedia cirp* 52, pp. 173–178.

Raghunathan, Srinivasan (2001). “Information sharing in a supply chain: A note on its value when demand is nonstationary”. In: *Management science* 47.4, pp. 605–610.

Rahim, Abd Rahman Abdul and Mohd Shariff Nabi Baksh (2003). “The need for a new product development framework for engineer-to-order products”. In: *European Journal of Innovation Management* 6.3, pp. 182–196.

Ran, Hailan (2021). “Construction and optimization of inventory management system via cloud-edge collaborative computing in supply chain environment in the Internet of Things era”. In: *Plos one* 16.11, e0259284.

Rasmussen, Carl and Zoubin Ghahramani (2000). “Occam’s razor”. In: *Advances in neural information processing systems* 13.

Reid, Struan J et al. (2021). “Predicting Fermentation Rates in Ale, Lager and Whisky”. In: *Fermentation* 7.1, p. 13.

Richter, Michael M and Rosina O Weber (2016). *Case-based reasoning.* Springer.

Rigatti, Steven J (2017). “Random forest”. In: *Journal of Insurance Medicine* 47.1, pp. 31–39.

Rocca, Joseph (2023). *Ensemble Methods: Bagging, Boosting, and Stacking.* Accessed: 2023-05-25. URL: https://towardsdatascience.com/ensemble-methods-bagging-boosting-and-stacking-c9214a10a205.

Rodhouse, Lindsey and Franck Carbonero (2019). “Overview of craft brewing specificities and potentially associated microbiota”. In: *Critical reviews in food science and nutrition* 59.3, pp. 462–473.

Rodman, Alistair D and Dimitrios I Gerogiorgis (2016). “Dynamic simulation and visualisation of fermentation: effect of process conditions on beer quality”. In: *IFAC-PapersOnLine* 49.7, pp. 615–620.

Rokach, Lior and Oded Maimon (2005). “Decision trees”. In: *Data mining and knowledge discovery handbook*, pp. 165–192.

Romsdal, Anita (2014). "Differentiated production planning and control in food supply chains". In.

Ross Ackerman, Bud Dunn (2020). *Selling Beer: Forecasting Through Uncertainty.* URL: https://brewingindustryguide.com/selling-beer-forecasting-through-uncertainty/ (visited on Nov. 29, 2022).

Russell, Stuart J (2010). *Artificial intelligence a modern approach.* Pearson Education, Inc.

Rüßmann, Michael et al. (2015). "Industry 4.0: The future of productivity and growth in manufacturing industries". In: *Boston consulting group* 9.1, pp. 54–89.

Ryback, Thomas et al. (2019). "Improving the planning quality in production planning and control with machine learning". In.

Sáenz-Alanıs, César A, M Angélica Salazar-Aguilar, and Vincent Boyer (2016). "A parallel machine batch scheduling problem in a brewing company". In: *The International Journal of Advanced Manufacturing Technology* 87, pp. 65–75.

Sagi, Omer and Lior Rokach (2018). "Ensemble learning: A survey". In: *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery* 8.4, e1249.

Salanță, Liana Claudia et al. (2020). "Non-alcoholic and craft beer production and challenges". In: *Processes* 8.11, p. 1382.

Santos, Iria et al. (2021). "Artificial neural networks and deep learning in the visual arts: A review". In: *Neural Computing and Applications* 33, pp. 121–157.

Sarker, Iqbal H (2021). "Machine learning: Algorithms, real-world applications and research directions". In: *SN Computer Science* 2.3, pp. 1–21.

Sayad, Saed (2023). *Decision Tree - Regression.* Accessed: 2023-05-25. URL: https://www.saedsayad.com/decision_tree_reg.htm.

Schratz, Patrick et al. (2019). "Hyperparameter tuning and performance assessment of statistical and machine-learning algorithms using spatial data". In: *Ecological Modelling* 406, pp. 109–120.

Schwab, Klaus (2017). "The Fourth Industrial Revolution, Crown Business". In: *New York* 192.

Scikit-Learn (2023). *Scikit-learn: Machine Learning in Python.* Accessed: 2023-05-25. URL: https://scikit-learn.org/stable/.

Scipy - Curve Fit (2023). https://docs.scipy.org/doc/scipy/reference/generated/scipy.optimize.curve_fit.html. Accessed: 2023-06-09.

Seifert, Dirk (2003). *Collaborative planning, forecasting, and replenishment: How to create a supply chain advantage.* AMACOM Div American Mgmt Assn.

Seitz, Kai-Frederic and Peter Nyhuis (2015). "Cyber-physical production systems combined with logistic models–a learning factory concept for an improved production planning and control". In: *Procedia Cirp* 32, pp. 92–97.

Şener, Aysun, Ahmet Canbaş, and M ÜNAL (2007). "The effect of fermentation temperature on the growth kinetics of wine yeast species". In: *Turkish Journal of Agriculture and Forestry* 31.5, pp. 349–354.

Shah, Yash and Shamik Sengupta (2020). "A survey on Classification of Cyber-attacks on IoT and IIoT devices". In: *2020 11th IEEE Annual Ubiquitous Computing, Electronics & Mobile Communication Conference (UEMCON).* IEEE, pp. 0406–0413.

Shahriari, Bobak, Alexandre Bouchard-Côté, and Nando Freitas (2016). "Unbounded Bayesian optimization via regularization". In: *Artificial intelligence and statistics.* PMLR, pp. 1168–1176.

Sharma, Rounak (2023). *AI/ML Neural Networks.* Accessed: 2023-04-05. URL: https://emeritus.org/in/learn/ai-ml-neural-networks/ (visited on June 5, 2023).

Shibamoto, Takayuki (2014). "Diacetyl: occurrence, analysis, and toxicity". In: *Journal of agricultural and food chemistry* 62.18, pp. 4048–4053.

Silva Nunes, Caique Emanuel da et al. (2016). "Demand Forecasting and Production Scheduling for a Craft Brewery Company: A Case Study". In: *2016 IEMS OFFICERS*, p. 106.

Silver, David et al. (2018). "A general reinforcement learning algorithm that masters chess, shogi, and Go through self-play". In: *Science* 362.6419, pp. 1140–1144.

Simatupang, T.M and R Sridharan (2002). "The collaborative supply chain". In: *The international journal of logistics management* 13.1, pp. 15–30.

Simchi-Levi, David et al. (2000). "Designing and managing the supply chain". In.

Singh, Harshdeep (2023). *Understanding Gradient Boosting Machines.* Accessed: 2023-05-25. URL: https://towardsdatascience.com/understanding-gradient-boosting-machines-9be756fe76ab.

Slack, Nigel and Alistair Brandon-Jones (2018). *Essentials of operations management.* Pearson UK.

Slack, Nigel, Alistair Brandon-Jones, and Robert Johnston (2013). *Operations Management (7th edition)*. English. Pearson.

Slocum, Susan L, Carol Kline, and Christina T Cavaliere (2018). *Craft Beverages and Tourism: Volume 2-Environmental, Societal, and Marketing Implications*. Springer.

Smola, Alex and SVN Vishwanathan (2008). “Introduction to machine learning”. In: *Cambridge University, UK* 32.34, p. 2008.

Snoek, Jasper, Hugo Larochelle, and Ryan P Adams (2012). “Practical bayesian optimization of machine learning algorithms”. In: *Advances in neural information processing systems* 25.

Snyder, Hannah (2019). “Literature review as a research methodology: An overview and guidelines”. In: *Journal of business research* 104, pp. 333–339.

Soman, Chetan Anil, Dirk Pieter Van Donk, and Gerard Gaalman (2004). “Combined make-to-order and make-to-stock in a food production system”. In: *International journal of production economics* 90.2, pp. 223–235.

Song, Jing-Sheng and Paul Zipkin (2003). “Supply Chain Operations: Assemble-to-Order Systems”. In: *Supply Chain Management: Design, Coordination and Operation*. Vol. 11. Handbooks in Operations Research and Management Science. Elsevier, pp. 561–596. DOI: https://doi.org/10.1016/S0927-0507(03)11011-0. URL: https://www.sciencedirect.com/science/article/pii/S0927050703110110.

Soori, Mohsen, Behrooz Arezoo, and Roza Dastres (2023a). “Artificial intelligence, machine learning and deep learning in advanced robotics, A review”. In: *Cognitive Robotics*.

— (2023b). “Internet of things for smart factories in industry 4.0, a review”. In: *Internet of Things and Cyber-Physical Systems*.

Sorini, Marc (2017). *Understanding the Three-Tier System: Its impacts on U.S. Craft Beer and You*. URL: https://www.craftbeer.com/craft-beer-muses/three-tier-system-impacts-craft-beer.

Speers, R Alex, Peter Rogers, and Bruce Smith (2003). “Non-linear modelling of industrial brewing fermentations”. In: *Journal of the Institute of Brewing* 109.3, pp. 229–235.

Stevenson, William J, Mehran Hojati, and James Cao (2014). *Operations management*. McGraw-Hill Education New York.

Stevenson, WJ (2018). *Operations management 13rd Edition*.

Stewart, G.G, D.L Maskell, and A Speers (2016). “Brewing Fundamentals – Fermentation”. In: *Technical quarterly (USA)* 53.1, p. 760.

Stojmenovic, Ivan and Sheng Wen (2014). “The fog computing paradigm: Scenarios and security issues”. In: *2014 federated conference on computer science and information systems*. IEEE, pp. 1–8.

Strandhagen, J (2021). *Produksjonslogistikk 4.0*. Fagbokforlaget.

Strandhagen, Jo Wessel et al. (2017). “The fit of Industry 4.0 applications in manufacturing logistics: a multiple case study”. In: *Advances in Manufacturing* 5, pp. 344–358.

Supekar, Sarang D et al. (2019). “A framework for quantifying energy and productivity benefits of smart manufacturing technologies”. In: *Procedia CIRP* 80, pp. 699–704.

Supplies, Midwest (2019). *How Long is the Fermentation Process?* URL: https://www.midwestsupplies.com/blogs/bottled-knowledge/how-long-is-the-fermentation-process (visited on Dec. 2, 2022).

Sutton, Richard S and Andrew G Barto (2018). *Reinforcement learning: An introduction*. MIT press.

Sutton, RS (1998). *Reinforcement learning: An introduction/Richard S. Sutton and Andrew G. Barto. Adaptive computation and machine learning*.

Syu, Mei-Jywan et al. (1994). “Neural network modeling for predicting brewing fermentations”. In: *Journal of the American Society of Brewing Chemists* 52.1, pp. 15–19.

Szirmai, Adam (2012). “Industrialisation as an engine of growth in developing countries, 1950–2005”. In: *Structural change and economic dynamics* 23.4, pp. 406–420.

Tao, Fei and Meng Zhang (2017). “Digital twin shop-floor: a new shop-floor paradigm towards smart manufacturing”. In: *Ieee Access* 5, pp. 20418–20427.

Tayo, Benjamin Obi (2022). *What Really is R2 Score in Linear Regression?* Accessed: 2023-04-03. URL: https://benjaminobi.medium.com/what-really-is-r2-score-in-linear-regression-20cafdf5b87c.

Thesseling, Florian A et al. (2019). “A hands-on guide to brewing and analyzing beer in the laboratory”. In: *Current Protocols in Microbiology* 54.1, e91.

Thibault, Jules, Vincent Van Breusegem, and Arlette Chéruy (1990). “On-line prediction of fermentation variables using neural networks”. In: *Biotechnology and Bioengineering* 36.10, pp. 1041–1048.

Thoben, Klaus-Dieter, Stefan Wiesner, and Thorsten Wuest (2017). ""Industrie 4.0" and smart manufacturing-a review of research issues and application examples". In: *International journal of automation technology* 11.1, pp. 4–16.

Thomé, Antônio Márcio Tavares et al. (2012). "Sales and operations planning: A research synthesis". In: *International journal of production economics* 138.1, pp. 1–13.

Thompson, Debora Viana, Rebecca W Hamilton, and Roland T Rust (2005). "Feature fatigue: When product capabilities become too much of a good thing". In: *Journal of marketing research* 42.4, pp. 431–442.

Thorn, James (2023). *Random Forest Hyperparameters and How to Fine-tune Them.* Accessed: 2023-05-25. URL: https://towardsdatascience.com/random-forest-hyperparameters-and-how-to-fine-tune-them-17aee785ee0d.

Tibshirani, Robert (1996). "Regression shrinkage and selection via the lasso". In: *Journal of the Royal Statistical Society: Series B (Methodological)* 58.1, pp. 267–288.

Towards Data Science: Your home for data science. A Medium publication sharing concepts, ideas, and codes (2023). Accessed: 2023-02-04. URL: https://towardsdatascience.com/.

Trelea, Ioan Cristian et al. (2001). "Predictive modelling of brewing fermentation: from knowledge-based to black-box models". In: *Mathematics and Computers in Simulation* 56.4-5, pp. 405–424.

Usuga Cadavid, Juan Pablo et al. (2020). "Machine learning applied in production planning and control: a state-of-the-art in the era of industry 4.0". In: *Journal of Intelligent Manufacturing* 31, pp. 1531–1558.

Vörösmarty, Gyöngyi and Imre Dobos (2020). "A literature review of sustainable supplier evaluation with Data Envelopment Analysis". In: *Journal of Cleaner Production* 264, p. 121672.

Voss, Chris, Nikos Tsikriktsis, and Mark Frohlich (2002). "Case research in operations management". In: *International journal of operations & production management* 22.2, pp. 195–219.

Wang, Chen, Hoang Tam Vo, and Peng Ni (2015). "An IoT application for fault diagnosis and prediction". In: *2015 IEEE International Conference on Data Science and Data Intensive Systems.* IEEE, pp. 726–731.

Wang, Gang et al. (2016). "Big data analytics in logistics and supply chain management: Certain investigations for research and applications". In: *International journal of production economics* 176, pp. 98–110.

Wang, Sheng et al. (2018). "Big Data enabled Intelligent Immune System for energy efficient manufacturing management". In: *Journal of cleaner production* 195, pp. 507–520.

Wang, Shiyong et al. (2016). "Implementing smart factory of industrie 4.0: an outlook". In: *International journal of distributed sensor networks* 12.1, p. 3159805.

Wang, Weiyu and Keng Siau (2019). "Industry 4.0: ethical and moral predicaments". In: *Cutter business technology journal* 32.6, pp. 36–45.

Wang, Yi et al. (2017). "Industry 4.0: a way from mass customization to mass personalization production". In: *Advances in manufacturing* 5.4, pp. 311–320.

Watson, B (2014). *The Timing of Craft Sales.*

Watson, Bart, Paul Gatza, and Ann Obenchain (2021). *National Beer Sales Production Data.* Brewers Association.

Watson, Bert (2014). *The Timing of Craft Sales.* URL: https://www.brewersassociation.org/insights/timing-of-craft-sales/ (visited on Dec. 10, 2022).

Wehrl, Alfred (1978). "General properties of entropy". In: *Reviews of Modern Physics* 50.2, p. 221.

Weston Jr, FC, Fred Herrmann, and Peter H Davidoff (1999). "Capacity planning and process analysis a simulation study of a microbrewery". In: *Production and Inventory Management Journal* 40.2, p. 48.

Wong, Kay Jan (2020). *CatBoost vs LightGBM vs XGBoost.* URL: https://towardsdatascience.com/catboost-vs-lightgbm-vs-xgboost-c80f40662924 (visited on May 6, 2022).

Wongsuphasawat, Kanit, Yang Liu, and Jeffrey Heer (2019). "Goals, process, and challenges of exploratory data analysis: An interview study". In: *arXiv preprint arXiv:1911.00568.*

Wu, Wei et al. (2020). "An IoT-enabled real-time logistics system for a third party company: a case study". In: *Procedia Manufacturing* 49, pp. 16–23.

Xu, Hansong et al. (2018). "A survey on industrial Internet of Things: A cyber-physical systems perspective". In: *Ieee access* 6, pp. 78238–78259.

Xu, Min et al. (2005). "Decision tree regression for soft classification of remote sensing data". In: *Remote Sensing of Environment* 97.3, pp. 322–336.

Zhang, Guangyi and Aristides Gionis (2023). “Regularized impurity reduction: accurate decision trees with complexity guarantees”. In: *Data mining and knowledge discovery* 37.1, pp. 434–475.

Zheng, Ting et al. (2021). “The applications of Industry 4.0 technologies in manufacturing context: a systematic literature review”. In: *International Journal of Production Research* 59.6, pp. 1922–1954.

Zhong, Botao et al. (2020). “Hazard analysis: A deep learning and text mining framework for accident prevention”. In: *Advanced Engineering Informatics* 46, p. 101152.

Zhong, Ray Y et al. (2017). “Intelligent manufacturing in the context of industry 4.0: a review”. In: *Engineering* 3.5, pp. 616–630.

Zhou, Keliang, Taigang Liu, and Lifeng Zhou (2015). “Industry 4.0: Towards future industrial opportunities and challenges”. In: *2015 12th International conference on fuzzy systems and knowledge discovery (FSKD)*. IEEE, pp. 2147–2152.

Zhou, Zhi-Hua (2021). *Machine learning*. Springer Nature.

Zhu, Xiaojin and Andrew B Goldberg (2009). “Introduction to semi-supervised learning”. In: *Synthesis lectures on artificial intelligence and machine learning* 3.1, pp. 1–130.

Appendix A

Interview Guide Multiple Case Study

Interviews with representatives were held in the following manner and at the following dates:

- **Beer Flag Brewery:** Matt Thomas (Head Brewer). Physical interview conducted at 11.12.2022
- **Sagene Brewery:** Kristian Johnsen (Head Brewer). Physical interview conducted at 15.11.2022
- **Lock 27 Brewing Company:** Spencer Moore (Operations Manager). Digital interview conducted at 21.11.2022
- **Ponysaurus Brewing Company:** Peter Hergruth (Packaging Manager and Cellar Specialist). Digital Interview conducted at 30.11.2022
- **Yee Haw Brewing Company:** Chase Shippy (Operations Manager): Digital interview conducted at 06.12.2022

A.1 Production

- How large is the brewhouse?
- How many fermenter vessels do you have?
- How much do you produce annually?
- How many batches do you produce annually?
- What is your production capacity?
- Do you use any ERP systems?
- How many employees are working with production?
- What is your capacity utilization?

A.2 Products and demand forecasting

- What types of beer do you produce?
- Which sales channels do you sell through?

- Where (geographically) are your products sold?
- How do you forecast demand?
- Do you experience variations in demand between products?
- Do you experience seasonal demand?
- Are you affected by trends?
- Do you experience any challenges with demand forecasting?

A.3 Production planning and control

- How do you plan and schedule your production?
- Are you following a make-to-stock strategy?
- Are you experiencing any challenges with planning and scheduling?
- How do you perform production control?
- Do you experience any challenges with production control?
- How do you plan the lead time of fermentation?
- Do you find fermentation to be challenging?

A.4 Inventory

- Where do you keep inventory and how much do you keep?
- How often do you replenish?
- Packaging
- Do you own your own packaging line?
- How often do you package?
- For how long do you package?
- How much do you package each time?
- Are you experiencing any challenges with packaging?

A.5 Other questions

- What are the main challenges you experience with regards to production planning and control?

Appendix B

Interview Guide Single Case Study

Interviews with the operations manager at Lock 27 Brewing Company were held in the following manner and at the following dates:

- **Interview 1:** Spencer Moore (Operations Manager). Digital interview conducted at 26.04.2023
- **Interview 2:** Spencer Moore (Operations Manager). Phone interview conducted at 15.05.2023
- **Interview 3:** Spencer Moore (Operations Manager). Phone interview conducted at 06.06.2023

B.1 Procurement

- What is the order frequency?
- What is the delivery lead time from your suppliers?
- How many suppliers do you have?
- Do you use an automated or manual ordering system?
- Do you order lot-for-lot?
- Do you use a fixed order quantity?
- Do you use an economic order quantity?
- Do you use a periodic order quantity?
- Do you use a order point system

B.2 Inventory Management

- Do you use aggregate or item inventory managment?
- Do you have a fixed safety stock?
- What is the item cost?
- What is your carrying cost?
- What is your stock out costs?
- What is your capacity-associated costs?

- What is your inventory turnover rate?
- What is your days of supply?
- Do you have an inventory control strategy or use any inventory control tools?

B.3 Forecasting and Demand Management

- Do you experience any demand variations caused by trends, seasonality, random variations or cycles?
- What is your error estimates in your demand forecasts?
- Do you forecast for single items or product families?
- What is your forecasting span?
- What type of forecasting techniques do you use? For example qualitative or quantitative suchs as historical data
- How do you track the forecasts?

B.4 Production Planning

- Do you create a strategic business plan?
- How do you create the production plan?
- Do you use any strategies such as chase strategy, production levelling, subcontracting or a hybrid strategy?
- Do you follow a MTS or MTO approach?
- Do you create a master production schedule? If so, how often and what is the input
- Do you use MRP?
- Do you do production activity control?
- Do you do sales and operations planning?
- Do you use any type of ERP system?
- Which production processes does your production consist of?
- What type of equipment is involved?¨
- How does material flow through production
- WHere is inventory located?

B.5 Supply Chain Management

- What are your sales channels
- How do you distribute and transport goods to customers
- Who are your end customers?

Appendix C

Code

C.1 Curve fit

```
good_batches = []
mediocre_batches = []
bad_batches = []

for id_of_batch in all_batches:
    #batch = batchId
    #x = timestamps
    #y = densities
    #temp = temperature
    batch, x, y ,temp= get_batch_and_readings_by_batch_id(id_of_batch)

    p_i = median(y[:4])
    p_e = median(nsmallest(4, y))

    def model(t, B, M, s):
        return [
            (p_e + ((p_i - p_e) / math.pow(1 + s * math.pow(math.e,
            (-1 * B * (t_i - M))), (1 / s)))) for t_i in t]

    t_0 = x[0]
    #p0 = the initial parameters
    popt, pcov = curve_fit(model, x, y, p0=[-0.000154, t_0, 6],
    bounds=((-0.001, 0, 1),(0, np.inf, 200)))
    #a,b,c = the optimal parameters that minimizes RSS
    a, b, c = popt
    #perr = standard error of optimal parameters
    perr = np.sqrt(np.diag(pcov)).tolist()
    B_err = perr[0]
    M_err = perr[1]
    s_err = perr[2]

    if float(B_err) > 0.000025 or float(bMerr) > 5000 or float(S_err) > 1.5:
        bad_batches.append(batch)
        continue

    if 0.000025 > float(B_err) > 0.00001 or 5000 > float(M_err) > 3000 or 1.5 >
    float(S_err) > 0.75:
        mediocre_batches.append(batch)
```

```
            continue
        else:
            good_batches.append(batch)
```

C.2 Remove slow start

```
def remove_slow_start(group):
    #slow_start_idx = the index of when the batch's fermentation
    activity is first above 0.05
    slow_start_idx = group.loc[group["fermentationActivity"] >= 0.05].index.min()

    #Removes all the readings occured before the slow start index
    if slow_start_idx is not None:
        num_removed_rows = len(group) - len(group.loc[slow_start_idx:])
        group = group.loc[slow_start_idx:]

        group["hoursIntoBatch"] = group["hoursIntoBatch"] - num_removed_rows
        group["hours_when_FG_happened"] = group["hours_when_FG_happened"]
        num_removed_rows
    return group

df = df.groupby("batchId").apply(remove_slow_start).reset_index(drop=True)
```

C.3 Remove tail

```
for group_name, group_df in grouped:
    # Reset the consecutive hours counter
    consecutive_hours = 0
    # Get the original indices of the rows in the batch
    original_indices = group_df.index

    # Check if the batch has more than 100 readings
    if len(group_df) > 100:
        # Iterate over each row in the group
        for index, row in group_df.iterrows():
            # Skip the first 100 rows
            if index in original_indices[:100]:
                filtered_df= pd.concat([filtered_df,
                pd.DataFrame([row])], ignore_index=False)
                continue

            if row["fermentationActivity"] < 0.05:
                # Increment the consecutive hours counter
                consecutive_hours += 1
            else:
                # Reset the consecutive hours counter if
                "fermentationActivity" is above 0.1
                consecutive_hours = 0
            filtered_df = pd.concat([filtered_df,
            pd.DataFrame([row])], ignore_index=False)

            if consecutive_hours >= 24:
                # Remove the last 12 rows from the filtered dataframe
                filtered_df = filtered_df[:-12]
```

```
            # Exit the loop to skip remaining rows in the group
            break
    else:
        #If it has less than 100 readings, we do not remove the tail
        filtered_df = filtered_df.append(group_df)
```

C.4 Hyperparameter tune ANN

```
def build_model(hp):
    model = Sequential()
    model.add(Dense(hp.Int('units1', min_value=32, max_value=500, step=32),
    input_dim=59, activation='relu'))
    model.add(Dense(hp.Int('units2', min_value=32, max_value=500, step=32),
    activation='relu'))
    model.add(Dense(1, activation='linear'))

    model.compile(optimizer=keras.optimizers.Adam(hp.Choice('learning_rate',
    values=[0.1, 0.05, 0.01,0.005])),loss='mean_squared_error',
    metrics=['mse'])

    return model

tuner = RandomSearch(
    build_model,
    objective='val_loss',
    max_trials=50,
    executions_per_trial=5)

tuner.search_space_summary()

tuner.search(X_train_scaled, y_train, epochs=5, validation_split=0.1)

tuner.results_summary()

# Retrieve the best model.
best_model = tuner.get_best_models(num_models=1)[0]
model = best_model
# Train the best model.
history = best_model.fit(X_train_scaled, y_train, epochs=200, validation_split=0.1)



optimal_hyperparameters = best_model.get_config()
```

C.5 Program code

The whole pipeline of our data including retrieval, cleaning, feature engineering and hyperparameter tuning can be seen in the .zip archive attached to this study. Exactly how the data is smoothed in the API is not included due to competitive circumstances. The attached code will not run properly, as the access keys to the databases are removed.

NTNU

Norwegian University of
Science and Technology