Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 239 (2024) 1778–1783

Procedia Computer Science

www.elsevier.com/locate/procedia

CENTERIS – International Conference on ENTERprise Information Systems / ProjMAN – International Conference on Project MANagement / HCist – International Conference on Health and Social Care Information Systems and Technologies 2023

How Complex Systems Sometimes Follow Murphy's Law: Train Delay Prediction at a Station Using Delays at Previous Stops as the Features

Pranjal Mandhaniya[a]*, Nils O. E. Olsson[a], Anders S. Larsen[a], Caroline Skjøren[a]

[a]*Norwegian University of Science and Technology, Department of Mechanical and Industrial Engineering, Trondheim, Norway*

Abstract

Prediction or estimation of the delay of a train is essential to analyze customer behavior. In case of longer delays, there is a high probability that the customer will abort their travel or change the mode/time of travel. The delay propagation from previous stations to the next station creates a chain of these types of customer reactions. Thus, using delays at the previous station to predict the next station's delay is a good approach to align with customer behavior. The present research attempts to predict the delays at a station using the delays at previous stations. The previous delays at stations are generated by creating lags in the original delays and creating them as one of the prediction features. The present research uses the delay data acquired from Bane NOR from 1 January 2021 to 28 February 2023. This data contains the scheduled and actual departure and arrival times of different trains between Oslo and Trondheim (up and down the line) in the specified period. The machine learning models based on neural networks were used on the data in the present research. Different prediction algorithms, i.e., recurrent neural network (RNN), gated recurrent unit (GRU), and long short-term memory (LSTM), were used. The prediction results are compared to look at the insights of the train delays in the given period. In conclusion, this study highlights how extreme feature engineering can negatively affect the output of a model.

* Corresponding author. Tel.: +4747629680;
E-mail address: pranjal.mandhaniya@ntnu.no

1877-0509 © 2024 The Authors. Published by Elsevier B.V.

10.1016/j.procs.2024.06.357

Keywords: Delay prediction; railway management; long short-term memory networks; predictive modeling; gated recurrent unit; recurrent neural network.

1. Introduction

It is well understood from the literature that the data quality is an upper constraint on a machine learning model's performance [1]. Poor data quality can significantly decrease the effectiveness of the related data applications [2]. Ensuring good data quality is, therefore, essential. Data can be characterized as physical entities that can be stored, retrieved, elaborated upon, and sent through a network [3]. Goverde and Hansen [4] declared that it is possible to analyze delay propagation and conflicts using detailed information on event times associated with train services from data records of the Dutch train describer TNV-system (Telecommunication Network Voltage). Conte and Schobel [5] demonstrated how data-based approaches might be used to conduct an organized analysis of dependencies between delays. The study aimed to learn about the relationships between delays to identify their root and describe how they spread throughout the system. The research by Olsson and Haugland [6] addresses and discusses the primary causes of railway delays and the resulting comparison findings. They highlight crucial factors to consider in advance of improving railway punctuality based on empirical findings from Norwegian research.

According to the literature, managing boarding and alighting passengers is a critical success element for punctuality on local and regional trains in congested locations. However, as stated by Sorensen et al. [7], getting actual data on train ridership is challenging. Their study investigated how mobile phone data can be used as a different source of information to analyze the number of travelers on trains and their travel patterns. Data-driven methods for train delay prediction are single-step forecasts that utilize simulation or historical system data to assess the robustness of a complex system, often producing deterministic predictions [8]. Adjetey-Bahun et al. [9] suggests a simulation-based model that measures passenger delay and passenger load as resilience indicators, allowing us to assess how resilient the studied system is relative to a perturbation by considering crisis management procedures.

Xu et al. [10] showed their method's effectiveness by examining the link between inherent traits and system resilience using historical data. Using data mining and machine learning techniques, they developed an advanced business analytical system to proactively predict potential disruptions and assist the operational team in improving organizational resilience. Nair et al. [11] paper developed a large-scale, data-driven ensemble forecasting system to generate forecasts using a two-stage random forest model to increase the prediction accuracy of train recovery times. The ensemble, tested on the Deutsche Bahn, performed overall better than constituent models, giving high-fidelity forecasts. To estimate conflict-free running times and dwell times, Kecman and Goverde [12] presents several data-driven approaches and discovered that decision trees and linear regression models were outperformed by random forests, providing the most precise running time predictions. Neural Network models, such as convolutional neural networks (CNN) and artificial neural networks (ANN), are among the most widely used data-driven models. Recurrent neural networks proposed as a solution to forecasting problems by Connor et al. [13] offered more accurate time-series prediction abilities than traditional neural networks. As with other neural network techniques, it was claimed that the input configuration is essential to successful prediction performance when using recurrent neural network designs. Another pioneering study, in 1996, was by Martinelli and Teng [14] who developed a neural network model to optimize solutions for the train formation problem. The study shows how the neural network model effectively identified the limitations of the conventional model and its objective functions. Yaghini et al. [15] presented an artificial neural network model to predict train delay of passenger trains.

Various architecture strategies and input approaches were tested, and decision trees and multinomial logistic regression models were used to evaluate the quality. The findings demonstrate that the delay prediction model has high accuracy and requires little training, making it a valuable tool for railway operators. Oneto et al. [16] developed a neural network train delay prediction model for extensive rail networks. The model employs machine learning and statistical techniques to use big data analytic methodologies and data processing technologies through its framework. Li et al. [17] considers the arrival routes of predicted trains and route conflicts with forward trains at multi-line stations when developing a train arrival delay prediction model. While Wen et al. [18] used a Long Short-Term Memory (LSTM) train delay prediction model to examine the correlation between different railway system features and train arrival delays. The model outperformed the random forest model and the artificial neural network model in comparison to accuracy tests of prediction accuracy.

A similar type of assignment has been undertaken in the present research. The present study aims to use available data and convert actual delays as learning features to analyze their effect on the robustness of the prediction model. This research is a part of an EU-funded project "FuTuRe". The flagship project FP6 – FutuRe ((GA 101101962) [19] under Europe's Rail Joint Undertaking) aims at providing new innovative technical requirements, methods, solutions, developments and services based on the latest leading-edge technologies to make regional rail cost-efficient while meeting safety standards and improving the reliability, availability and capacity of the railway system. The work presented here is related to the FutuRe project area Regional Rail Customer Services, focusing on customer service and aiming to develop highly accurate multimodal passenger information on-board and/or at stations for passenger and freight management.

2. Methodology

The data was acquired from Bane NOR for the actual arrival and scheduled arrival of different trains running on the Oslo (OSL)-Trondheim(TND) line. This data contained the records from 1 January 2021 to 28 February 2023. The stations covered in this study for the prediction modeling of train delay are Oslo S, Lillestrøm, Gardermoen, Hamar, Brumunddal, Moelv, Lillehammer, Ringebu, Vinstra, Kvam, Otta, Dovre, Dombås, Hjerkinn, Kongsvoll, Oppdal, Berkåk, Støren, Heimdal, Trondheim S. These stations are in order when a train travels from Oslo S station to Trondheim S station. However, the data includes trains running in both directions.

The Bane NOR delay estimate model uses an event-driven methodology. The model is created upon a system of linear equations, with an iterative process and multi-step forecasts, to predict train delays. The estimates are generated based on either GPS information from the train every 10 seconds, information from sensor trigger points, or manual information registration by the train dispatcher or the train driver. The following stations are automatically updated when an estimate at one station is updated, including the run time between stations and dwell time at stations. The data also included negative delays. These are the early departure times at non-boarding stations or technical halts.

The machine learning models based on neural networks were used on the data in the present research to suggest improvements in the Bane NOR model. Different prediction algorithms, i.e., recurrent neural network (RNN), gated recurrent unit (GRU), and long short-term memory (LSTM), were used. The prediction results are compared to look at the insights of the train delays in the given period. Departure delays at a station are predicted based on departure delays from the five previous stations. These propagated delays are called lags in the prediction model.

3. Analysis and results

The details of prediction models created using RNN, GRU, and LSTM are summarized as losses against epochs in Figure 1. The figure shows that LSTM is the best method as the losses converge faster. However, the magnitude of losses is the smallest in the RNN.

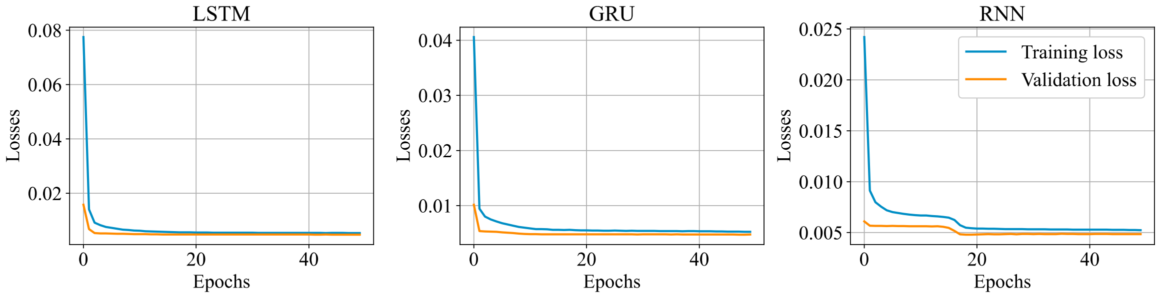

Figure 1. Prediction loss with the progression of training

The Pearson correlation coefficient (R^2) matrix acquired in the present study is shown in Table 1. R^2 value nearer to 1 represents a better fit and vice versa. It can be seen that unnecessarily adding features to the prediction model leads to lower prediction accuracy. This is evident from the decreasing R^2 values by adding lags (previous delays). However, the results in the present study are shown for the most conservative approach (five lags).

Table 1. Summary of performance metrics for LSTM models

Lags	**RNN**	**LSTM**	**GRU**
0	0.8057	0.8107	0.8100
2	0.7559	0.7791	0.7839
5	0.7617	0.7870	0.7854

The performance matrices in the form of mean absolute error (MAE, in minutes) and root mean square error (RMSE, in minutes) of the present study are shown against the studied literature in Table 2.

Table 2. Overview of literature using MAE and RMSE to evaluate the performance of their proposed model

Source	Model	Type	MAE (min)	RMSE (min)
	Data-driven	all operational trains	-	5.52
	Data-driven	all urban services	-	8.43
	Data-driven	all long-distance trains	-	13.08
	RandomForest	operational train > 3min	1.708	2.863
	RandomForest	operational train <= 3min	0.687	1.681
	ANN	operational train > 3min	1.857	3.051
	ANN	operational train <=3min	0.788	3.051
	XGBoost	operational train > 3min	1.724	-
	XGBoost	operational train <= 3min	0.723	-
	RandomForest	operational train	0.82	1.18
	CNN	operational train	0.98	1.47
	LSTM	operational train	0.66	0.78
	RandomForest	high-speed railway network	1.513	2.4
	RandomForest	high-speed railway network, route-related variables	1.223	2.196
	DELM	high-speed railway network	2.389	3.842
	DELM	high-speed railway network, route-related variables	1.591	2.82
	RNN	one long-distance train	0.7632	1.0182
	LSTM	one long-distance train	0.7277	0.9626
	GRU	one long-distance train	0.7294	0.9663

The predicted delays from the recurrent neural network (RNN), gated recurrent unit (GRU), and long short-term memory (LSTM) are compared with the original (Bane NOR) delays in Figure 2. Due to the scale of plotting, the plot is a bit gibberish. However, the limits of the delays can be easily seen, which shows that the predicted values are in a smaller range than the actual delays.

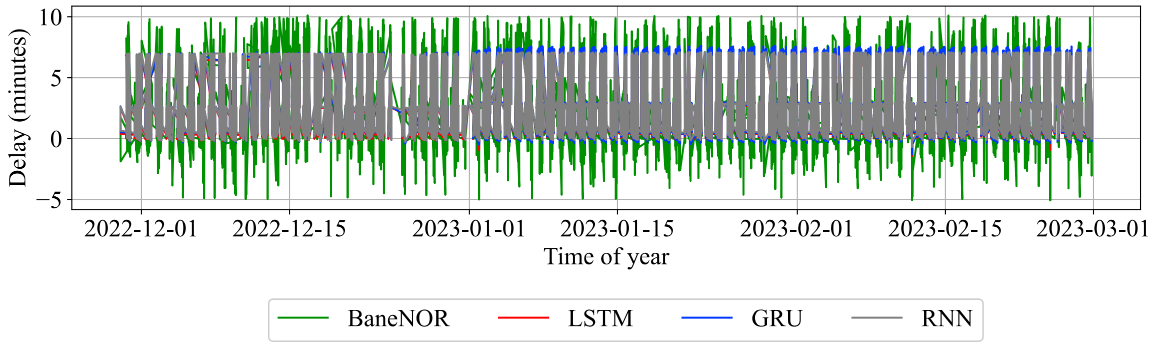

Figure 2. Comparison of predicted delays and actual delays

The delay averaged over the days of the week is shown in Figure 3. The trains get their highest delays on the Friday, aligning with the tendency of people to travel more on the weekend.

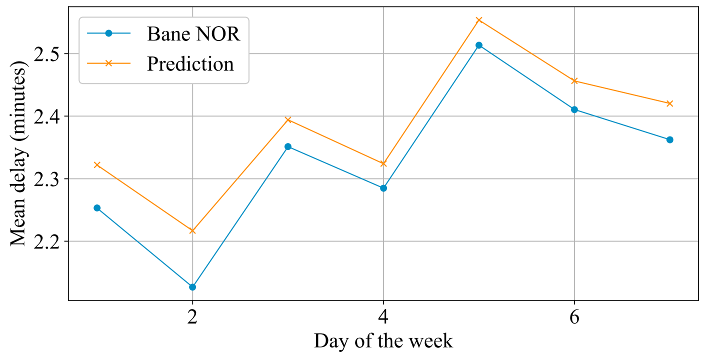

Figure 3. Delay averaged over the days of the week (1 is Monday and 7 is Sunday on the x-axis)

Conclusions

The major conclusion that can be drawn from the study and the literature studied in the present study are:

- A heavy feature engineering can sometimes harm the prediction model's viability. This generally leads to longer prediction time as well.
- The delay patterns align with customer behavior, as evident in the literature. However, the predicted values have a smaller range than the actual delay values.

Acknowledgements

We are delighted to acknowledge the help of Bane NOR for providing the delay data for this study. This research activity is part of the FP6 FutuRe project, partially funded by the European Commission through Europe's Rail Joint Undertaking under the Horizon Europe Program with the grant agreement number 101101962.

Data Availability Statement

The data used in the present research is made available on https://github.com/pranjalm/Future_Europes_rail.

References

[1] Nitin Gupta, Shashank Mujumdar, Hima Patel, Satoshi Masuda, Naveen Panwar, Sambaran Bandyopadhyay, Sameep Mehta, Shanmukha Guttula, Shazia Afzal, Ruhi Sharma Mittal, and Vitobha Munigala. Data quality for machine learning tasks. In Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery & Data Mining, KDD '21, page 4040–4041, New York, NY, USA, 2021. Association for Computing Machinery. ISBN 9781450383325. doi:10.1145/3447548.3470817.

[2] Yair Wand and Richard Y. Wang. Anchoring data quality dimensions in ontological foundations. Commun. ACM, 39(11):86–95, nov 1996. ISSN 0001-0782. doi:10.1145/240455.240479.

[3] Heather Maguire. Book review: Data quality: concepts, methodologies and techniques by c. batini and m. scannapieco. International Journal of Information Quality, 1(4):444–450, 2007. doi:10.1504/IJIQ.2007.016717.

[4] Rob MP Goverde and Ingo A. Hansen. Tnv-prepare: Analysis of dutch railway operations based on train detection data. Computers in Railways, 7:779–788, 2000. URL https://www.witpress.com/Secure/elibrary/papers/CR00/CR00075FU.pdf.

[5] Carla Conte and Anita Schobel. Identifying dependencies among delays. Proceedings of IAROR, 2007, 2007. doi:10.53846/goediss-3402.

[6] Nils Olsson and Hans Haugland. Influencing factors on train punctuality—results from some norwegian studies. Transport policy, 11:387–397, 2004. doi:10.1016/j.tranpol.2004.07.001.

[7] Anette Ostbo Sorensen, Johannes Bjelland, Heidi Bull-Berg, Andreas Dypvik Landmark, Muhammad Mohsin Akhtar, and Nils OE Olsson. Use of mobile phone data for analysis of number of train travellers. Journal of Rail Transport Planning & Management, 8(2):123–144, 2018. doi:10.1016/j.jrtpm.2018.06.002.

[8] Jiateng Yin, Xianliang Ren, Ronghui Liu, Tao Tang, and Shuai Su. Quantitative analysis for resilience-based urban rail systems: A hybrid knowledge-based and data-driven approach. Reliability Engineering & System Safety, 219:108183, 2022. doi:10.1016/j.ress.2021.108183.

[9] Kpotissan Adjetey-Bahun, Babiga Birregah, Eric Chatelet, and Jean-Luc Planchet. A model to quantify the resilience of mass railway transportation systems. Reliability Engineering & System Safety, 153:1–14, 2016. doi:10.1016/j.ress.2016.03.015.

[10] Donna Xu, Ivor W Tsang, Eng K Chew, Cosimo Siclari, and Varun Kaul. A data-analytics approach for enterprise resilience. IEEE Intelligent Systems, 34(3):6–18, 2019. doi:10.1109/MIS.2019.2918092.

[11] Rahul Nair, Thanh Lam Hoang, Marco Laumanns, Bei Chen, Randall Cogill, Jacint Szabo, and Thomas Walter. An ensemble prediction model for train delays. Transportation Research Part C: Emerging Technologies, 104:196–209, 2019. doi:10.1016/j.trc.2019.04.026.

[12] Pavle Kecman and Rob MP Goverde. Predictive modelling of running and dwell times in railway traffic. Public Transport, 7(3):295–319, 2015.doi:10.1007/s12469-015-0106-7.

[13] Jerome T Connor, R Douglas Martin, and Les E Atlas. Recurrent neural networks and robust time series prediction. IEEE transactions on neural networks, 5(2):240–254, 1994. doi:10.1109/72.279188.

[14] David R Martinelli and Hualiang Teng. Optimization of railway operations using neural networks. Transportation Research Part C: Emerging Technologies, 4(1):33–49, 1996. doi:10.1016/0968-090X(95)00019-F.

[15] Masoud Yaghini, Mohammad M Khoshraftar, and Masoud Seyedabadi. Railway passenger train delay prediction via neural network model. Journal of advanced transportation, 47(3):355–368, 2013. doi:10.1002/atr.193.

[16] Luca Oneto, Emanuele Fumeo, Giorgio Clerico, Renzo Canepa, Federico Papa, Carlo Dambra, Nadia Mazzino, and Davide Anguita. Train delay prediction systems: a big data analytics perspective. Big data research, 11:54–64, 2018. doi:10.1016/j.bdr.2017.05.002.

[17] Zhongcan Li, Ping Huang, Chao Wen, Xi Jiang, and Filipe Rodrigues. Prediction of train arrival delays considering route conflicts at multi-line stations. Transportation Research Part C: Emerging Technologies, 138: 103606, 2022. doi:10.1016/j.trc.2022.103606.

[18] Chao Wen, Ping Huang, Zhongcan Li, and Weiwei Mou. A predictive model of train delays on a railway line. Journal of Forecasting, 39:470–488, 2019. doi:10.1002/for.2639.

[19] Europe's Rail. Fp6 future homepage, 2022. URL https://projects.rail-research.europa.eu/eurail-fp6/.

[20] ZhongCan Li, Chao Wen, Rui Hu, Chuanlin Xu, Ping Huang, and Xi Jiang. Near-term train delay prediction in the dutch railways network. International Journal of Rail Transportation, 9(6):520–539, 2021. doi:10.1080/23248378.2020.1843194.

[21] Jianqing Wu, Yihui Wang, Bo Du, Qiang Wu, Yanlong Zhai, Jun Shen, Luping Zhou, Chen Cai, Wei Wei, and Qingguo Zhou. The bounds of improvements toward real-time forecast of multi-scenario train delays. IEEE Transactions on Intelligent Transportation Systems, 23(3):2445–2456, 2022. doi:10.1109/TITS.2021.3099031.