A Data-driven Method Based on Deep Belief Networks for Backlash Error Prediction in Machining Centers

Zhe Li[1], Yi Wang[2] and Kesheng Wang[1]

[1] Department of Production and Quality Engineering, Norwegian University of Science and Technology, 7491 Trondheim, Norway
[2] School of Materials, University of Manchester, M13 9PL Manchester, UK

Abstract Backlash error occurs in a machining center may lead to a series of changes in the geometry of the components and subsequently deteriorate the overall performance of the equipment. Due to the uncertainty of mechanical wear between kinematic pairs, it is challenging to predict backlash error through physical models directly. An alternative method is to leverage data-driven models to map the degradation. This paper proposes a data-driven method for backlash error predication through Deep Belief Network (DBN). The proposed method focuses on the assessment of both current and future geometric errors for backlash error prediction and subsequent maintenance in machining centers. During the process of prognosis, a DBN via stacking Restricted Boltzmann Machines is constructed for backlash error prediction. Energy-based models enable DBN to mine information hidden behind highly coupled inputs, which makes DBN a feasible method for fault diagnosis and prognosis when the target condition is beyond the historical data. In the experiment, to confirm the effectiveness of deep learning for backlash error prediction, similar popular regression methods, including Support Vector Machine Regression and Back Propagation Neural Network, are employed to present a comprehensive comparison in both diagnosis and prognosis. The experimental results show that the performances of all these regression methods are acceptable in the diagnostic stage. In the prognostic stage, DBN demonstrates its superiority and significantly outperforms the other models for backlash error prediction in machining centers.
Keywords Data mining · Machining Centers · Data-driven Method · Deep Belief Network · Backlash Error

1. Introduction

In recent years, machining centers are widely utilized as vital manufacturing equipment in modern production. Their application can be found everywhere in modern production; from large mechanical parts to miniature medical equipment, from mould making to aerospace manufacturing and from daily repair tools to scientific research instruments, for their excellent orientation ability, high material removal rate, and lower production

---

[1]Corresponding author: Zhe Li
E-mail address: zhe.li@ntnu.no

cost (Zhang et al. 2013). The term "machining center" can be used to describe any Computer Numerical Control (CNC) milling and drilling machine that includes an automatic toolchanger and a table that clamps the workpiece in place. Geometric errors that occur in a machining centers are the errors on account of the inaccuracies built in during assembly and from the components used in the machine. The errors may be affected by many error sources (Lee and Yang 2013; Zhu et al. 2012; Stryczek 2016; Cheng et al. 2015). These error sources may cause a series of changes in the geometry of the components and present in the structural loop, including the spindle shaft, the ball screws, the bearings, the housing, the guideways and frame, and work-holding fixtures. The errors may not only cause significant quality and accuracy degradation but also fatal breakdown of machines, which can lead to serious economic loss (Aydın et al. 2015). Therefore, it is especially crucial to accurately detect the existence of geometric errors as early as possible and predict the error in a period of working time (Zhong et al. 2015; Siguenza-Guzman et al. 2015; Cheng et al. 2014). Schwenke et al. (2008) reviewed various technologies to evaluate the geometric errors of machines and their basic characteristics. As reported in that paper, backlash usually affect uncertainties in the measured parameters since they are not modelled adequately. The uncertainties may cause correlations or just erroneous estimations. Therefore, many companies choose the preventive maintenance for backlash error in machining centers, which means the error would be checked and eliminated from time to time following planned guidelines. However, this strategy is both costly and time-consuming.

During the last few decades, many researchers have studied methods to monitor, model and control backlash error for mechanical systems, and many diagnosis approaches have been proposed (Chen et al. 2016; Fines and Agah 2008; Liu et al. 2010; Prasanga et al. 2013; Slamani et al. 2012). Prasanga et al. (2013) proposed a method to compensate the backlash error through two parallel thrust wires without any encoder or force sensor at the end effector. Slamani et al. (2012) evaluated the backlash error of an industrial serial robot under various conditions using a laser interferometer measurement instrument and represent the relationship between the backlash error and the robot configuration with a polynomial model. All these methods require additional measurement equipment or system along with elaborate engineering and considerable domain expertise.

Nowadays, with the trend of smart manufacturing, companies are increasingly using sensors and wireless technologies to capture data at all stages of a product's life (Kusiak 2017). For this reason, "Big Data" has attracted not only researchers' but also manufacturers' attention along with the development of data-driven methods from various perspectives such as product lifecycle management (Li et al. 2015), manufacturing (Tao et al. 2017), and maintenance (Mosallam et al. 2016). Some machine learning and artificial intelligence approaches, such as neural networks, support vector machine and fuzzy logic systems, also have been applied in backlash error evaluation or prediction. Chen et al. (2016) compensated the backlash nonlinearity by a smooth backlash inverse with the help of parameter estimations and fuzzy logic system-based approximation for an active vibration isolation system. Fines and Agah (2008) applied artificial neural network for positioning error compensation in a machine tool, and proved the feasibility to calculate compensation values. Liu et al. (2010) employed back-propagation neural network to map the backlash error in a vertical machining center and compared the result with polynomial models.

However, most of these researches only focus on the diagnosis or evaluation of current or historical backlash error. It still lacks a method with high generalization for backlash error prediction, especially when the target condition is beyond the historical data. Therefore, this paper presents a data-driven method for backlash error detection and predication for machining centers based on Deep Belief Network (DBN). The research focuses on the assessment of both current and future geometric errors for backlash error compensation and subsequent maintenance in machining centers. During diagnosis stage, the missing prior data including the historical data and current backlash error will be interpreted, while the prognosis stage is responsible for the prediction of future backlash error based on the prior data provided by the former stage through a deep neural network.

The remaining part of this paper is organized as follows. Section 2 briefly introduces the sources and effects of backlash error in machining centers. Section 3 details the theory of deep learning, Restricted Boltzmann Machines (RBMs) and DBN. Section 4 proposes a novel hierarchical diagnosis and prognosis system (HDPS) for backlash error prediction with deep learning algorithm. Section 5 reports the experiment for backlash error detection and prediction during our research in detail and conforms the effectiveness and feasibility of the proposed method. Section 6 compares and discusses the performances of several intelligent regression approaches in backlash error prediction. Conclusions and future work are summarized in the last section of this paper.

2. Backlash Error in Machining centers

In order to perform an error mapping and subsequent compensation for backlash error, an understanding of the sources and effects of backlash error in machining centers are necessary. In mechanical engineering, backlash is a kind of nonlinear position dependent-error caused by the existence of clearance between two mechanical elements. It may occur in a rotational mechanical element as well as in a translational mechanical element (Kao et al. 1996). Normally, a machining center is equipped with various high-precision sensors such as gratings, rotary encoders, current sensors, temperature sensors, and linear scales to guarantee the accuracy through closed or half closed loop control. Many parameters such as machine temperature, geometric position of ball screw, torque and current, can be obtained directly from the control system. In machining centers, backlash error can be acquired through some extra methods, most of which are either time-costly such as laser interferometer or only yield the maximum value (Liu et al. 2010).

In a machining center, backlash error occurs when there exists a gap between the ball screw and spindle at the kinematic pair. As shown in Fig. 1, when the direction of motion reversed, the spindle will not move until the gap is taken up in the opposite direction. The distance that the ball screw travels before the table will move again is the geometric error caused by backlash, which is also called as backlash error. In general, all loosely connected elements in the driving mechanism may influence the backlash error of the system.

Backlash error varies at different axis positions and depends on the moving direction. In addition, the error affects the contouring accuracy and increases over time due to wear in the machining center, which means it is almost impossible to establish an accurate physical model for backlash error prediction. Therefore, it is significant and necessary to apply machine-learning approaches to monitor, model, and predict backlash error in mechanical systems to maintain the desired level of accuracy.

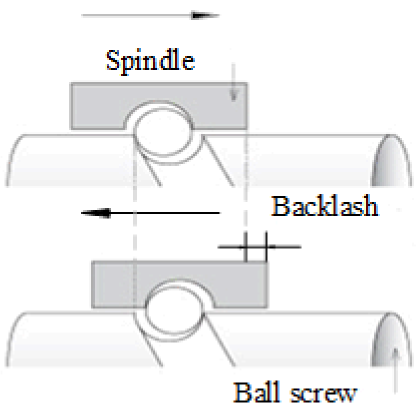

Fig. 1 Backlash error in machining centers

In our pioneer work (Wang et al. 2015), we applied back-propagation networks to map the relationship between backlash error and axis positons of the spindle. It proved the feasibility to apply machine-learning approaches to detect backlash error. However, that paper did not take nature wear into consideration, since conventional machine learning methods require elaborate feature extractors before applying learning models (Zhang et al. 2017), and it is impractical to fulfill this requirement without exact parameters from future condition in this case. Therefore, this paper focuses on the application of deep learning approach to map the future backlash error while the important parameters to map the backlash error is missing. In addition, a hierarchical diagnosis and prognosis system based on DBN is also constructed in this paper.

3. DBN

3.1 Deep Learning

As a branch of machine learning, deep learning is a series of algorithms which can be applied to model, approximate, or map high level abstractions in data. The essence of deep learning is about how to compute hierarchical features or representations from the objective data. The family of deep learning approaches have been growing increasingly richer, encompassing variety of neural networks with multiple processing layers, hierarchical probabilistic models, and kinds of unsupervised or supervised feature learning algorithms (Deng and Yu 2014). Schmidhuber (2015) summarized all relevant work about deep learning in neural networks and distinguished the Shallow and Deep Learners by the depth of their credit assignment paths.

As reported by LeCun et al. (2015), conventional machine-learning techniques were usually limited by their ability to process natural data in their raw form. For decades, to construct a machine learning system required elaborate engineering and considerable domain expertise to design a feature extractor that transformed the raw data into a suitable internal representation or feature vector from which the learning subsystem, often a classifier, could detect or classify patterns in the input. More recently, deep learning, as a latest research area of machine learning, has accelerated its application in fault diagnosis and prognosis (Gan and Wang 2016). The key advantage of deep learning is that the features are not designed by human engineers but learned from data itself through a generalized self-learning procedure. Various deep learning algorithms, such as deep belief networks (Hinton and Salakhutdinov 2006) and convolutional neural network (LeCun et al. 1998), have been applied successfully in the field of computer vision (Krizhevsky et al. 2012; Tompson et al. 2014; Zhou et al. 2010; Ribeiro et al. 2011), speech recognition (Hinton et al. 2012; Sainath et al. 2013; Seltzer et al. 2013), sentence modelling (Er et al. 2016) and proved the good performance in predicting the activity of potential drug molecules (Hermann et al. 2015), analysing particle accelerator data (Ciodaro et al. 2012), reconstructing brain circuits (Helmstaedter et al. 2013), and predicting the effects of mutations in non-coding DNA on gene expression and disease (Xiong et al. 2015). In this paper, the deep neural network is constructed via the stacked of RBMs, which also called as DBN. A DBN is a deep neural network composed of multiple layers of RBMs (Hinton et al. 2006).

3.2 RBMs

A Restricted Boltzmann Machine (RBM) is a special type of Markov Random Field (MRF), which consists of two layers, one with stochastic visible or observable units and the other with stochastic hidden units (Keyvanrad and Homayounpour 2014). RBMs can be represented as bipartite graphs as shown in Figure 2, where all visible units v are connected to all hidden units h, and there are no visible-visible or hidden-hidden connections (Yu and Deng 2011). w_{ij} represents the interaction term between visible unit v_i and hidden unit h_j, while vector **a** and **b** are bias terms for hidden units and visible units respectively.

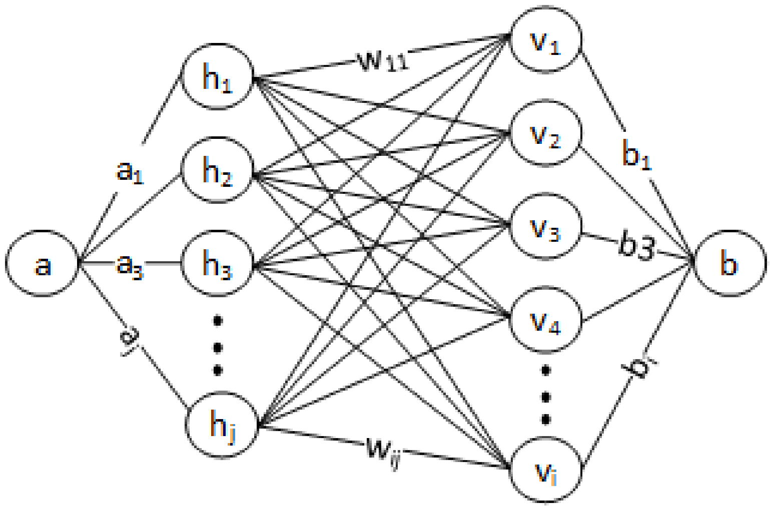

Fig. 2 Structure of RBMs

The energy of the joint configuration with bias in RBMs is defined as equation (1):

$$E(\boldsymbol{v},\boldsymbol{h};\boldsymbol{W},\boldsymbol{a},\boldsymbol{b}) = -\boldsymbol{v}^T\boldsymbol{W}\boldsymbol{h} - \boldsymbol{b}^T\boldsymbol{v} - \boldsymbol{a}^T\boldsymbol{h}$$

$$= -\sum_{i=1}^{I}\sum_{j=1}^{J} w_{ij}v_ih_j - \sum_{i=1}^{I} b_iv_i - \sum_{j=1}^{J} a_jh_j \qquad (1)$$

Where **W** is the concurrent weights between visible and hidden units. I and J represent the numbers of visible and hidden units. Then the joint probability distribution for all visible and hidden pairs can be defined as follows:

$$p(\boldsymbol{v},\boldsymbol{h};\boldsymbol{W},\boldsymbol{a},\boldsymbol{b}) = Z^{-1}e^{-E(\boldsymbol{v},\boldsymbol{h};\boldsymbol{W},\boldsymbol{a},\boldsymbol{b})} \qquad (2)$$

Where Z is the partition, which can be obtained by summing all possible pairs of visible and hidden units as equation (3):

$$Z = \sum_{\boldsymbol{v}}\sum_{\boldsymbol{h}} e^{-E(\boldsymbol{v},\boldsymbol{h};\boldsymbol{W},\boldsymbol{a},\boldsymbol{b})} \qquad (3$$

Then, the probability assigned from the network for the visible vector **v** can be obtained by marginalizing out the hidden vector:

$$p(\boldsymbol{v};\boldsymbol{W},\boldsymbol{a},\boldsymbol{b}) = \sum_{h} p(\boldsymbol{v},\boldsymbol{h};\boldsymbol{W},\boldsymbol{a},\boldsymbol{b}) = Z^{-1}\sum_{\boldsymbol{h}} e^{-E(\boldsymbol{v},\boldsymbol{h};\boldsymbol{W},\boldsymbol{a},\boldsymbol{b})} \qquad (4)$$

Due to the specific structure of RBM, there are no direct connection between hidden units. Therefore, all the visible and hidden units are conditionally independent (Hinton 2010), and the conditional probabilities can be efficiently calculated as equation (5) and (6):

$$p\left(h_j = 1 | \boldsymbol{v}; \boldsymbol{W}, \boldsymbol{a}, \boldsymbol{b}\right) = S\left(a_j + \sum_{i=1}^{I} v_i w_{ij}\right) \tag{5}$$

$$p(v_i = 1 | \boldsymbol{h}; \boldsymbol{W}, \boldsymbol{a}, \boldsymbol{b}) = S\left(b_i + \sum_{j=1}^{J} h_j w_{ij}\right) \tag{6}$$

Where $S(x)$ is the logistic sigmoid function $S(x) = 1/(1 + e^{-x})$

Then, the RBM model with binary units can be learned through negative log-likelihood gradients (Gan and Wang 2016). The derivative of the log probability of a training vector can be obtained as follows:

$$\frac{\partial \log p(\boldsymbol{v}; \boldsymbol{W}, \boldsymbol{a}, \boldsymbol{b})}{\partial w_{ij}} = \langle v_i h_j \rangle_{data} - \langle v_i h_j \rangle_{model} \tag{7}$$

$$\frac{\partial \log p(\boldsymbol{v}; \boldsymbol{W}, \boldsymbol{a}, \boldsymbol{b})}{\partial \boldsymbol{a}} = \langle h_j \rangle_{data} - \langle h_j \rangle_{model} \tag{8}$$

$$\frac{\partial \log p(\boldsymbol{v}; \boldsymbol{W}, \boldsymbol{a}, \boldsymbol{b})}{\partial \boldsymbol{b}} = \langle v_i \rangle_{data} - \langle v_i \rangle_{model} \tag{9}$$

Where the angle brackets denote the expectations with the distribution specified by the subscript that follows. Then the learning rule for all parameters can be obtained as follows:

$$\Delta w_{ij} = \mathcal{E}_w\left(\langle v_i h_j\rangle_{data} - \langle v_i h_j\rangle_{model}\right) \tag{10}$$

$$\Delta a_{ij} = \mathcal{E}_a\left(\langle h_j\rangle_{data} - \langle h_j\rangle_{model}\right) \tag{11}$$

$$\Delta b_{ij} = \mathcal{E}_b(\langle v_i\rangle_{data} - \langle v_i\rangle_{model}) \tag{12}$$

Where $\mathcal{E}_w$, $\mathcal{E}_a$ and $\mathcal{E}_b$ represent learning rate of weight, hidden bias and visible bias, respectively. According to the previously mentioned RBM property, an unbiased sample of $\langle\ .\rangle_{data}$with the respect to the data distribution can be easily obtained, while attaining an unbiased sample of $\langle\ .\rangle_{model}$ is intractable, since it can be done through starting from any random state of the visible units and performing sequential Gibbs sampling for a long time (Keyvanrad and Homayounpour 2014). Therefore, the Contrastive Divergence (CD) method (Hinton 2002) is applied to approximate the gradient objective function, where $\langle\ .\rangle_{model}$ is replaced by k iterations of Gibbs sampling. During Gibbs sampling, each iteration updates all hidden units according to equation (11), followed by updating of all visible unites through equation (12), as shown in Fig. 3. Although CD method is not a perfect gradient computation method, the results has been proved acceptable (Carreira-Perpinan and Hinton 2005).

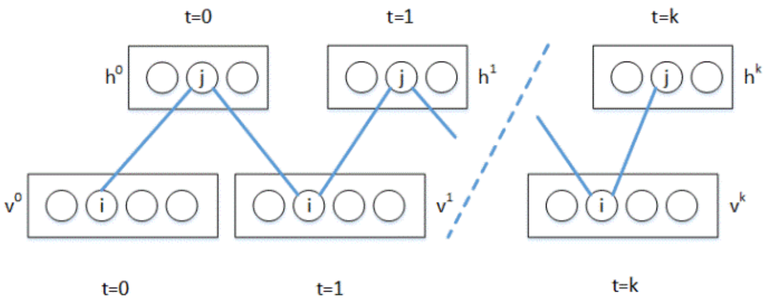

Fig. 3 Contrastive Divergence Training

3.3 DBN Construction for Fault Diagnosis and Prognosis

As mentioned above, A DBN can be constructed through stacking RBMs, each of which contains one visible layer and one hidden layer respectably. As one of the major technologies of deep learning, DBN has excelled in AI areas with notable success (Wang and Jiang 2017). The construction process of DBN is well described in (Hinton et al. 2012). Each RBM is pre-trained with their own training data by CD training algorithm, and its output serves as the training data for the next RBM layer. As shown in Fig. 4, the input layer and the first hidden layer h_1 construct the first RBM. Then the states of the binary hidden units of the first trained RBM is used to train the next hidden layer h_2, then hidden layer h_1 and hidden layer h_2 form the second RBM. These layer by layer unsupervised training method can effectively pre-train the DBN.

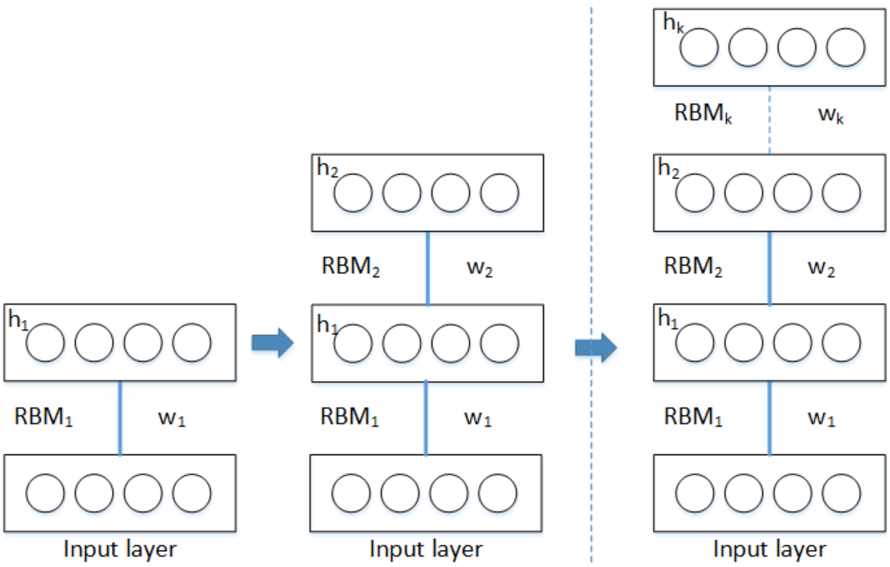

Fig. 4 Overall construction of DBN

To achieve high classification or regression performance for fault diagnosis and prognosis, a final decision layer with variables, which represent the desired outputs or labels, is added to the stacked RBM. The final structure of DBN for fault diagnosis and prognosis in shown in Fig. 5, which is composed of several successive RBM layers and a final decision layer for faults clustering, faulty component's identity, or evaluation of potential failures. Once the DBN is initialized, the back propagation (BP) algorithm, which is a supervised learning method and applied in Back-propagation Neural Network (BPNN), can be employed to adjust the weights.

Because Deep Belief Networks are based on RBMs, which are particular energy-based models, the learning process corresponds to modifying that energy function so that its shape has desirable properties (Bengio 2009). In a DBN, each RBM is trained to encode in its weight matrix a probability distribution that predicts the activity of the visible layer through the hidden layer. By stacking such models, and letting each layer predict the activity of the layer below, higher RBMs learn increasingly abstract representations of sensory inputs (O'Connor et al. 2013). Due to the layer-by-layer unsurprised learning process with fine-tuning procedure, DBN has the superiority to capture intrinsic characteristics and discover discriminative information about potential failures from massive data. Therefore, DBN can be leveraged to establish diagnosis and prognosis models with high generalization for backlash error detection and prediction.

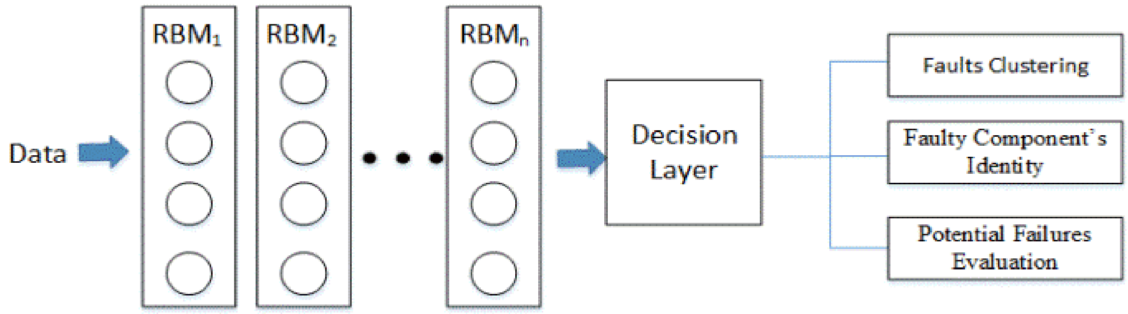

Fig. 5 DBN for fault diagnosis and prognosis

4. Hierarchical Diagnosis and Prognosis System

As mentioned above, since it is intractable to establish physical model for backlash error, and the backlash error varies with different axis positions, most company chooses to check and eliminate backlash according to the planned guidelines, which is both costly and time-consuming. Therefore, it still lacks method to predict the backlash error instead of detecting from time to time. In this section, a novel HDPS is proposed for backlash error detection and predication in machining centers based on DBN models. The purpose of the system is to make maintenance decision based on the result of faults diagnosis and prognosis. With the help of HDPS, the maintenance team can prevent occurrence and development of failures effectively, ensure the safety of equipment and personnel, and reduce economic loss caused by failures. It can use fault diagnosis, performance assessment of degrading level, fault prognosis models to reach near-zero-breakdown performance and improve productivity for a company.

As shown in Fig. 6, HDPS proposed in this paper can be divided into three layers, which are Data Acquisition Layer, Diagnosis Layer and Prognosis Layer. In Data Acquisition Layer, the data, including all the information we may require during the diagnosis and prognosis, is collected from machining centers. This is the first step to achieve diagnosis and prognosis based on data mining for machine centers. The first task of this layer is to select and collect suitable parameters which can represent the current working condition or install additional sensors for this purpose. Then, the backlash error at current axis position can be calculated form geometric measurement information through backlash error interpretation method, which will be introduced in next section. Then, all the measured parameters and obtained backlash error are stored in the data warehouse.

As mentioned before, since backlash error varies at different axis positions and the measurement is both costly and time-consuming, it is almost impracticable to collect the backlash error at all positions. Therefore, in Data Acquisition Layer, only one or several axis positions' backlash error may be acquired, and the Diagnosis Layer is responsible to fill up the others. At first, training samples require to be selected from the data warehouse and divided into groups for training and testing the diagnosis model. Several data-driven models like BPNN, Support Vector Machine Regression (SVMR) and DBN, can be applied as diagnosis model here according to the user. Once the diagnosis model is trained by historical data, it can be used to detect current backlash error in all position and fill up all missing backlash errors back to the data warehouse.

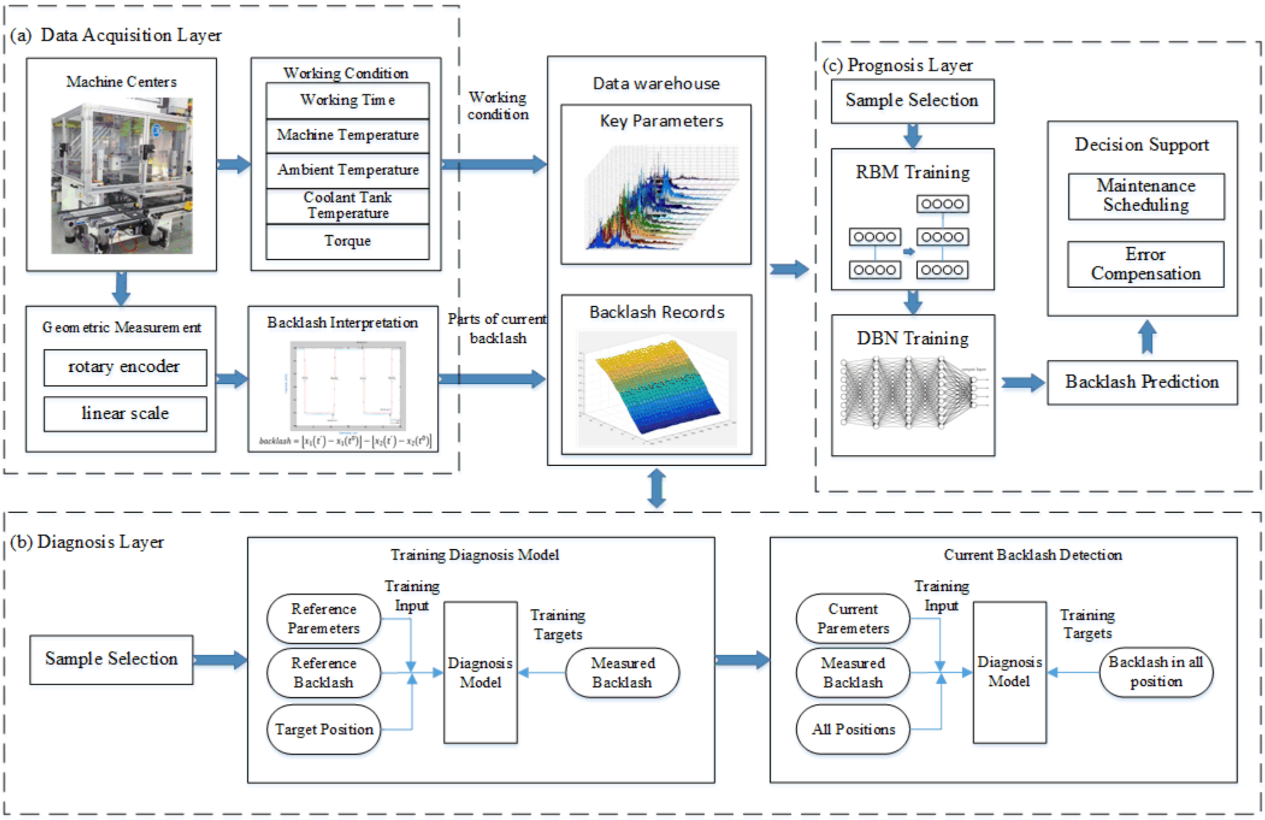

Fig. 6 Structure of HDPS

In Prognosis Layer, the historical data will be selected from the data warehouse to train the prognosis model. Then the data can be used to pre-train the RBMs in an unsupervised method as we discussed in section 3. The DBN model can be constructed by stacking these RBMs and a final decision layer, which may adjust the weight of the network according to the target values. Once the prognosis is trained, it can be applied to predict the backlash error in the future through the current working condition. In next section, the experiment for backlash error detection and prediction during our research is introduced in detail to illustrate how the HDPN and deep learning approach can work for backlash error detection and prediction in machining centers. And the results conform the effectiveness and feasibility of the proposed method.

5. Backlash Error Detection and Prediction Experiment

5.1 Experiment Set Up

In order to measure the backlash error in a machining center, at least two geometric measurements are required: displacement of the ball screw and the linear position of the spindle. In our experiment, the displacement of the ball screw is measured through a rotary encoder in the motor. It records the rotation angle and converts the signal into linear displacement. The position of the spindle is acquired by a calibrated linear scale. Figure 7 shows the setup of the measurement system, in which the linear scale records the direct position of the spindle x_1, and rotary encoder measures the displacement of the ball screw (the indirect position of the spindle x_2).

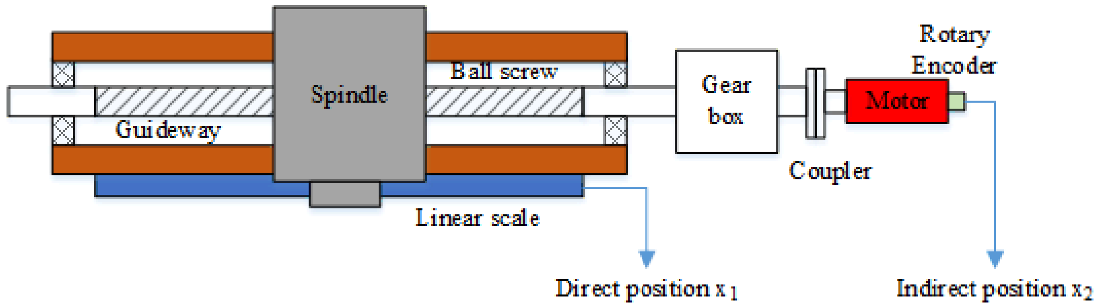

Fig. 7 Setup of the measurement system

5.2 Backlash Error Interpretation

According to the definition, the geometric error caused by backlash error can be interpreted as the difference between the displacements of the spindle and ball screw when the direction is reversed. Simply described as the following equations:

$$B[p] = \int_{t^b}^{t^*} f_1(t)\,dt - \int_{t^b}^{t^*} f_2(t)\,dt \tag{13}$$

$$p = x_1(t^b), \tag{14}$$

Where:

$f_1(t)$ is the velocity of the spindle with time, recorded by linear scale.

$f_2(t)$ is the velocity of the ball screw with time, recorded by rotary encoder.

$x_1(t)$ is the direct position of the spindle with time, recorded in linear scale.

p is the position on linear scale, where backlash occurs.

$B[p]$ is the geometric error caused by backlash at position p.

t^b is the time when backlash occurs.

t^* is the time when backlash disappears (the gap is taken occurs)

However, due to the uncertainty of backlash, it is almost impossible to capture the exact time when the gap is filled, which means t^* in the equation (1) is unavailable or with low precision. Therefore, during the measurement, we extended the measurement distance and let backlash occurred three times in one sample to eliminate the measurement error. As shown in figure 8, three blocks of geometric error, caused by backlash, can be recognized during one sampling process. The time interval between each sampling units is two milliseconds.

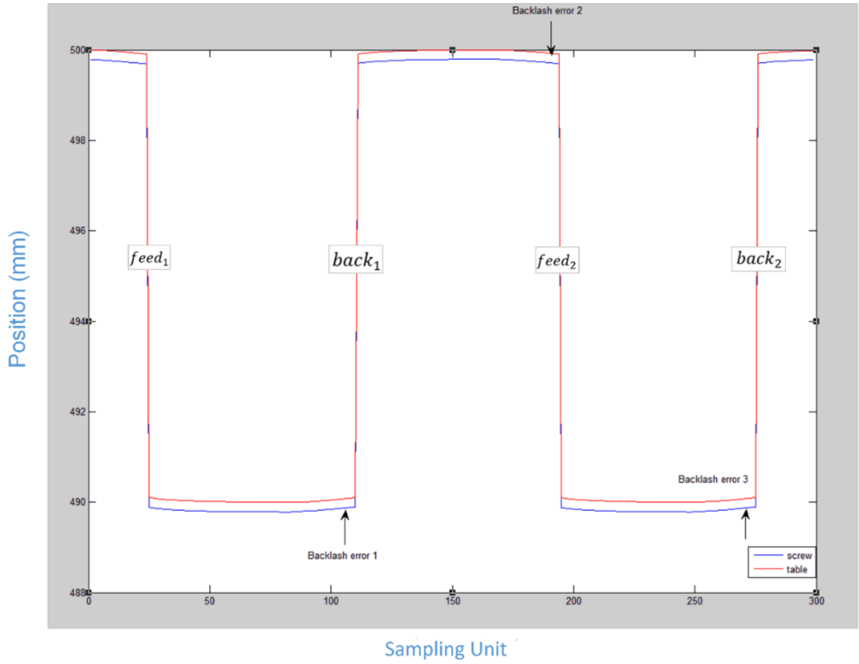

Fig. 8 Three blocks of backlash in one sample

The position of spindle and ball screw can be decomposed into the amount of feed and return, the initial position, and backlash error, as shown in the following equation:

$$x_1(t') = x_1(t^0) + \int_{t^0}^{t'} feed(t)\,dt + \int_{t^0}^{t'} back(t)\,dt \ + \sum backlash \tag{15}$$

$$x_2(t') = x_2(t^0) + \int_{t^0}^{t'} feed(t)\,dt + \int_{t^0}^{t'} back(t)\,dt \tag{16}$$

Where:

$x_2(t)$ is the position of ball screw with time, recorded in rotary encoder.

t^0 The initial time of the measurement.

t' The end time of the measurement.

$feed(t)$ The feed rate at t, which is equal to zero if not under feed movement.
$back(t)$ The return rate at t, which is equal to zero if not under return movement.

Since the backlash errors occur in very close positions, we consider the difference of the value is approach to zero. Therefore, the backlash error can be calculated according to equation (17):

$$backlash = [x_1(t') - x_1(t^0)] - [x_2(t') - x_2(t^0)] \tag{17}$$

5.3 Data Acquisition

To investigate the actual performance of proposed system, the experiment was carried out on a Pietro Carnaghi AC 16 TM vertical machining center with a collection of 25 weeks' data, from the 7th week to 31st since the last maintenance. The machining center is placed in a plant with some manufacturing tasks every day. During the data collection period, data was collected through a very rigorous testing procedure after daily work to ensure all the data collected is under a similar condition. The collected parameters are shown in Table 1.

During the experiment, we divided the moving distance into 24 points with the interval of 20 mm from 1090 to 1550 mm according to the calibrated linear scale. Then, all the backlash errors with current parameters can be calculated through backlash error interpretation method, which was introduced in the Section 5.2. Fig. 9 shows the obtained backlash error only with working weeks and axis position (It is more than a 3-dimension issue, but according to empirical knowledge, it is intuitive to visualize the backlash error with working weeks and axis position).

Table 1 Parameters collected from a vertical machining center

Parameters	Meaning
x_1	Direct position measured by linear scale
x_2	Displacement recorded by rotary encoder
w	Number of weeks since the last maintenance
T_1	Temperature of the coolant tank
T_2	Temperature of the machining center

T_3	Ambient temperature
TRQ	Machining torque
t	Sampling units during the testing procedure
P	Axis position of the spindle

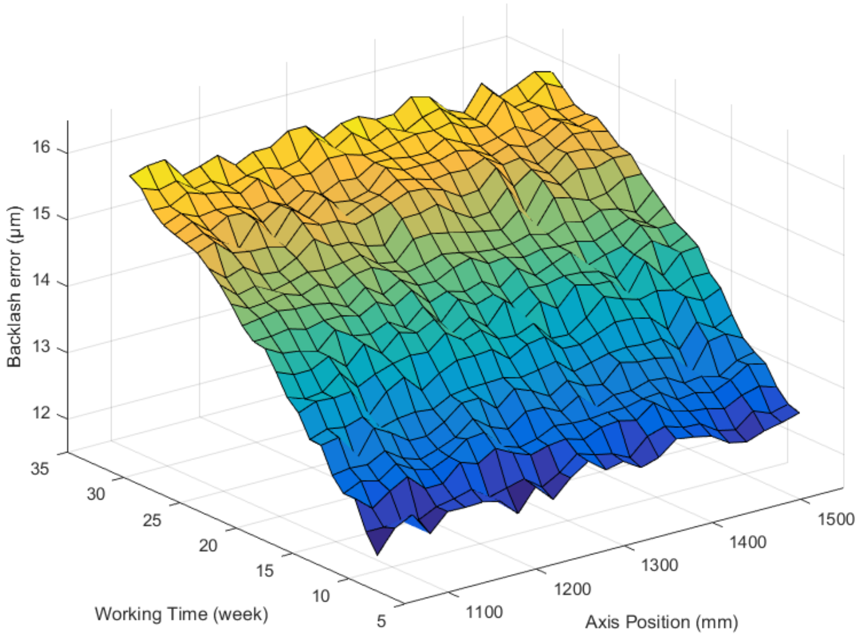

Fig. 9 Backlash error with working time and axis position

5.4 Diagnosis for Detection

As mentioned above, backlash errors in some positions may be missed in Data Acquisition Layer. So Diagnosis Layer requires to detect and fill up the current backlash error in all positions through the collected parameters. Actually, the essence of the diagnosis model is to deal with a nonlinear regression problem, in which all the variables are collected or mapped under the same working condition. The regression function F can be simply shown as equation (18):

$$b_t \approx F(T_1, T_2, T_3, t, TRQ, p_r, b_r, p_t) \quad (18)$$

Where b_t is the backlash error at target position, p_r represents the backlash error at reference position, which is obtained and interpreted in Data Acquisition Layer. p_t and b_r mean the target position and reference position, respectively.

In Diagnosis Layer, the diagnosis is first trained through the historical data, and responsible to map the backlash error in all positions through the data collected under current working condition. During the experiment, a selection of 1152 samples was applied to train the regression model. We employed three methods including BPNN, SVMR, and DBN as the regression model, and compared the results as shown in Table 2, where Maximum Error (ME), Mean Squared Error (MSE) and training time of each model are listed.

Table 2 Diagnosis result of BPNN, DBN, and SVMR

Model	BPNN	DBN	SVMR
Structure	50 nodes in the hidden layer	4 layers, 50 nodes in each	Gaussian Kernel Function
MSE (μm)	0.01148	0.01056	0.00964
ME (μm)	0.2758	0.2807	0.2812
Training time	Instantly	30 mins	Instantly

The result shows that all three methods has the capacity to deal with the diagnosis problem when all the relative parameters were obtained. In addition, both BPNN and SVMR can finish the training process instantly, and DBN is relatively time-consuming.

5.5 Prognosis for Backlash Error

As mentioned above, Prognosis Layer is responsible to predict future backlash error according to the historical data and current working condition of the machining center. Actually, it can also be considered as a regression problem, though all the parameters that can be used to represent the working condition are missing since we never know the exact values of these parameters in the future. Therefore, we applied the deep learning approach to predict the backlash error first, and then tested other methods to compare the effectiveness. Table 3 shows the input variables of the prognosis model.

Table 3 Inputs for Prognosis

Inputs	Meaning
w	Number of weeks since the last maintenance
T_1	Temperature of coolant tank
T_2	Temperature of machine
T_3	Ambient temperature
TRQ	Machine torque
P	Axis position
$backlash_{w-1,P-1}$	Backlash Error in $w-1$ weeks at $P-1$ position
$backlash_{w-2,P-2}$	Backlash Error in $w-2$ weeks at $P-2$ position

During the experiment, it is supposed that prognosis model is established at week 29th and employed to predict the backlash error in the future, which means the backlash errors in week 30th and 31st are beyond the training data. According to the author's empirical knowledge, this problem is common and challenging since in many situations, one may not have the data under faults. Some machines may run several years without any failures. However, it may exist potential faults that would occur one day. When they happen, they may cause terrible disasters in both economy and personal safety. This is the reason why it is also crucial to evaluate potential failures or degradations beyond the historical data. During the experiment, the applied DBN model is constructed through stacking four RBMs. The structure for the utilized DBN is shown in Table 4 along with the parameters of each Restricted Boltzmann Machine.

During the training process, samples are randomly divided into two groups, 80 percent for training and 20 percent for testing. Fig. 10 shows the training result, in which the best MSE from week 9th to week 29th is 0.012207 µm at the 14075th epochs.

Table 4 parameters for utilized DNB

Parameters	RBM_1	RBM_2	RBM_3	RBM_4

Type	Bernoulli	Bernoulli	Bernoulli	Bernoulli
Number of neurons	50	50	30	30
Learning rate	0.01	0.01	0.01	0.01
Number of epochs	500	500	300	300

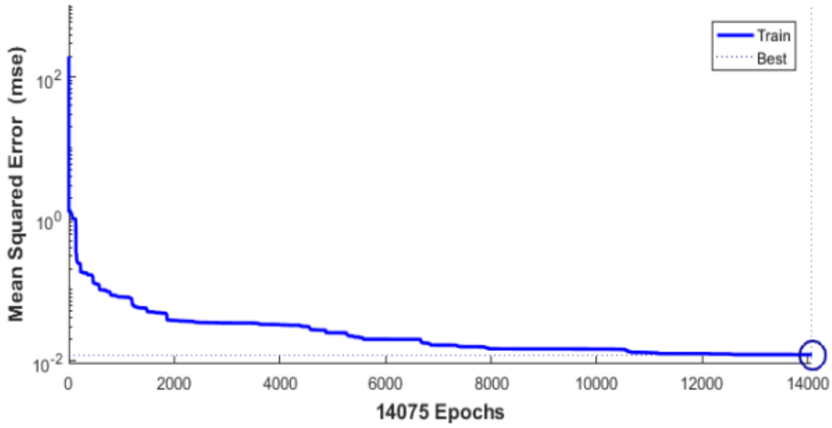

Fig. 10 Training result of DBN

Fig. 11 and Fig. 12 compared the actual backlash error and predicted backlash error in week 30th and 31st respectively. The MSE of backlash error prediction in week 30th and 31st is 0.0102 µm and 0.0142 µm with 0.1808 and 0.2275 as ME, respectively. In this case, the Maximum Permissible Error (MPE) for the backlash error is set as 16µm. In week 29th, we can predict that the backlash error in week 31st may exceed the MPE considering about the mean prediction error, which means the fault could be forecasted two weeks in advance. Then, subsequent maintenance should be arranged in week 30th to prevent the error from growing up and finally exceeding the MPS.

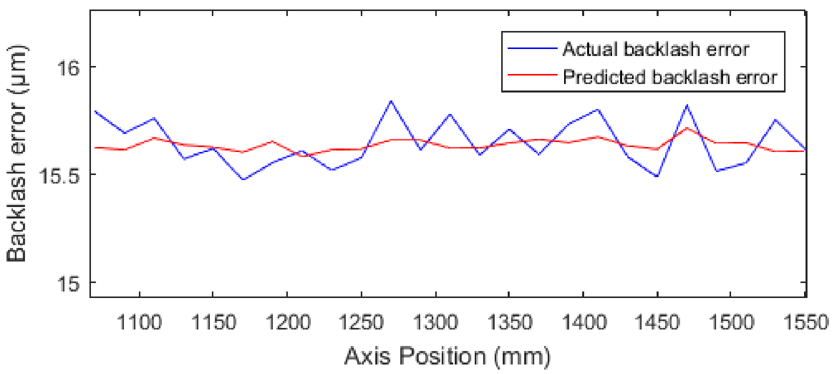

Fig. 11 Predicted backlash error in week 30th

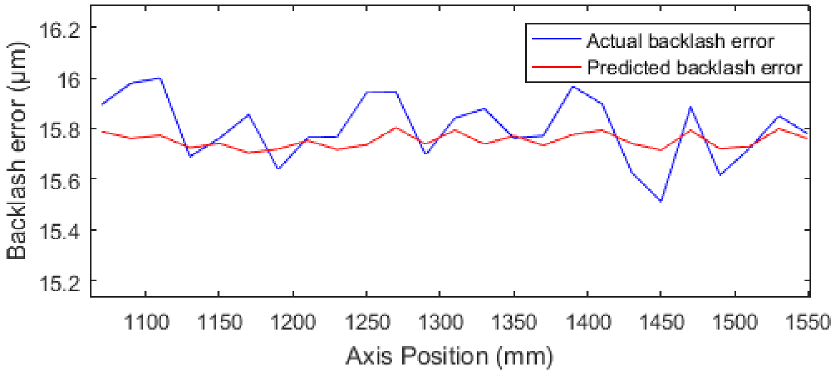

Fig. 12 Predicted backlash error in week 31st

6. Discussion

During the experiment, to confirm the effectiveness of deep learning for backlash error prediction, two other widely used regression methods, SVMR and BPNN, are employed for comparison. The BPNN applied in the experiment consisted of one hidden layer with 50 neurons. The Levenberg-Marquard optimization rule is applied as training algorithm for the BPNN. For the SVMR, Gaussian Function is applied as the kernel function. As mentioned above, all the training samples are randomly selected. In order to obtain more accurate and comprehensive evaluation results, we did this random selection for 20 times. The prediction results of each algorithm with their best performance over the 20 runs are summarized in Table 5.

Table 5 Prediction results of BPNN, DBN, and SVMR

Model	BPNN	DBN	SVMR
Structure	50 nodes in the hidden layer	4 layers, 50 * 50 * 30 * 30	Gaussian Kernel Function
Training MSE (μm)	0.020894	0.012207	0.0102
MSE in week 30th	0.2566	0.0102	1.2335
MSE in week 31st	1.6450	0.0142	2.8029

ME in week 30st	0.9904	0.1808	1.3669
ME in week 31st	1.6338	0.2275	1.8969
Training time	Instantly	40 mins	instantly

Through observation, DBN outperforms other two methods with better performance in both MSE and ME for backlash error prediction. The numerical result indicates that DBN can effectively deal with the backlash error prediction issue with better stability and accuracy compared with the conventional methods, when the target condition is beyond the training data. To be more specific, energy-based models enable DBN to mine information hidden behind highly coupled inputs, which makes DBN a suitable method to predict backlash error in the future through current condition. Although DBN requires more training time and higher computational complexity compared with conventional machine learning algorithms, the outstanding performance in prediction outweighs this shortage.

As to BPNN and SVMR, both of them have good performances in fault diagnosis if all the necessary parameters can be obtained. In addition, compared with DBN, they showed the advantage in learning speed during the experiment. However, it is also obvious that neither of them has the ability to predict future backlash error when the important information in target condition is missing. The experiment also demonstrated the necessity and feasibility of applying DBN as the prognosis model for backlash error prediction in the proposed HDPS.

7. Conclusion and future work

This paper presents a novel hierarchical diagnosis and prognosis system for backlash error detection and prediction in machining centers based on DBN. The method builds on unsupervised pre-training process through energy-based models, which can help to predict future backlash error through current condition. The proposed system mainly includes three layers, Data Acquisition Layer, Diagnosis Layer and Prognosis Layer. In Data Acquisition Layer, important parameters including geometric measurement information, temperature, and machine torque is obtained from machining centers. Diagnosis Layer is responsible to detect the current error in all position and fill up all missing backlash errors back to the data warehouse through diagnosis model.

In the Prognosis Layer, the backlash error in the future can be predicted with the help of Deep Learning approach.

The performance of the proposed method is evaluated during the experiment. To provide a comprehensive evidence for the effectiveness HDPS, two other intelligent algorithms, BPNN and SVMR, are also applied to replace the DBN as the prognosis model. The numerical results demonstrate the superiority to apply deep learning method for backlash error prediction, when the target condition is beyond the training data.

Future work may focus on the following two points:

1. Due to the limitation of data source, the authors only verify the effectiveness of HDPS for backlash error prediction in 25 weeks. Further research can be made when the observation time extends.
2. The proposed method may be utilized and tested for other failures or errors.

References

Aydın, İ., Karaköse, M., & Akın, E. (2015). Combined intelligent methods based on wireless sensor networks for condition monitoring and fault diagnosis. *Journal of Intelligent Manufacturing, 26*(4), 717-729.

Bengio, Y. (2009). Learning deep architectures for AI. *Foundations and trends® in Machine Learning, 2*(1), 1-127.

Carreira-Perpinan, M. A., & Hinton, G. On Contrastive Divergence Learning. In *AISTATS, 2005* (Vol. 10, pp. 33-40): Citeseer

Chen, C., Liu, Z., Zhang, Y., Chen, C. P., & Xie, S. (2016). Actuator Backlash Compensation and Accurate Parameter Estimation for Active Vibration Isolation System. *IEEE Transactions on Industrial Electronics, 63*(3), 1643-1654.

Cheng, Q., Zhao, H., Zhang, G., Gu, P., & Cai, L. (2014). An analytical approach for crucial geometric errors identification of multi-axis machine tool based on global sensitivity analysis. *The International Journal of Advanced Manufacturing Technology, 75*(1-4), 107-121.

Cheng, Q., Zhao, H., Zhao, Y., Sun, B., & Gu, P. (2015). Machining accuracy reliability analysis of multi-axis machine tool based on Monte Carlo simulation. *Journal of Intelligent Manufacturing*, doi:10.1007/s10845-015-1101-1.

Ciodaro, T., Deva, D., De Seixas, J., & Damazio, D. Online particle detection with neural networks based on topological calorimetry information. In *Journal of Physics: Conference Series, 2012* (Vol. 368, pp. 012030, Vol. 1): IOP Publishing

Deng, L., & Yu, D. (2014). Deep Learning: Methods and Applications. *Foundations and Trends® in Signal Processing, 7*(3–4), 197-387, doi:10.1561/2000000039.

Er, M. J., Zhang, Y., Wang, N., & Pratama, M. (2016). Attention pooling-based convolutional neural network for sentence modelling. *Information Sciences, 373*, 388-403.

Fines, J. M., & Agah, A. (2008). Machine tool positioning error compensation using artificial neural networks. *Engineering Applications of Artificial Intelligence, 21*(7), 1013-1026.

Gan, M., & Wang, C. (2016). Construction of hierarchical diagnosis network based on deep learning and its application in the fault pattern recognition of rolling element bearings. *Mechanical Systems and Signal Processing, 72*, 92-104.

Helmstaedter, M., Briggman, K. L., Turaga, S. C., Jain, V., Seung, H. S., & Denk, W. (2013). Connectomic reconstruction of the inner plexiform layer in the mouse retina. *Nature, 500*(7461), 168-174.

Hermann, M., Pentek, T., & Otto, B. (2015). Design principles for Industrie 4.0 scenarios: a literature review. *Technische Universität Dortmund, Dortmund*.

Hinton, G. (2010). A practical guide to training restricted Boltzmann machines. *Momentum, 9*(1), 926.

Hinton, G., Deng, L., Yu, D., Dahl, G. E., Mohamed, A.-r., Jaitly, N., Senior, A., Vanhoucke, V., Nguyen, P., & Sainath, T. N. (2012). Deep neural networks for acoustic modeling in speech recognition: The shared views of four research groups. *IEEE Signal Processing Magazine, 29*(6), 82-97.

Hinton, G. E. (2002). Training products of experts by minimizing contrastive divergence. *Neural computation, 14*(8), 1771-1800.

Hinton, G. E., Osindero, S., & Teh, Y.-W. (2006). A fast learning algorithm for deep belief nets. *Neural computation, 18*(7), 1527-1554.

Hinton, G. E., & Salakhutdinov, R. R. (2006). Reducing the dimensionality of data with neural networks. *Science, 313*(5786), 504-507.

Kao, J., Yeh, Z.-M., Tarng, Y., & Lin, Y. (1996). A study of backlash on the motion accuracy of CNC lathes. *International Journal of Machine Tools and Manufacture, 36*(5), 539-550.

Keyvanrad, M. A., & Homayounpour, M. M. (2014). A brief survey on deep belief networks and introducing a new object oriented MATLAB toolbox (DeeBNet V2. 1). *arXiv preprint arXiv:1408.3264*.

Krizhevsky, A., Sutskever, I., & Hinton, G. E. Imagenet classification with deep convolutional neural networks. In *Advances in neural information processing systems, 2012* (pp. 1097-1105)

Kusiak, A. (2017). Smart manufacturing must embrace big data. *Nature, 544*(7648), 23.

LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature, 521*(7553), 436-444.

LeCun, Y., Bottou, L., Bengio, Y., & Haffner, P. (1998). Gradient-based learning applied to document recognition. *Proceedings of the IEEE, 86*(11), 2278-2324.

Lee, K.-I., & Yang, S.-H. (2013). Measurement and verification of position-independent geometric errors of a five-axis machine tool using a double ball-bar. *International Journal of Machine Tools and Manufacture, 70*, 45-52.

Li, J., Tao, F., Cheng, Y., & Zhao, L. (2015). Big data in product lifecycle management. *The International Journal of Advanced Manufacturing Technology, 81*(1-4), 667-684.

Liu, H., Xue, X., & Tan, G. (2010). Backlash error measurement and compensation on the vertical machining center. *Engineering, 2*(06), 403.

Mosallam, A., Medjaher, K., & Zerhouni, N. (2016). Data-driven prognostic method based on Bayesian approaches for direct remaining useful life prediction. *Journal of Intelligent Manufacturing, 27*(5), 1037-1048.

O'Connor, P., Neil, D., Liu, S.-C., Delbruck, T., & Pfeiffer, M. (2013). Real-time classification and sensor fusion with a spiking deep belief network. *Frontiers in neuroscience, 7*.

Prasanga, D. K., Tanida, K., Mizoguchi, T., & Ohnishi, K. Evaluation of a backlash compensation method using two parallel thrust wires. In *Industrial Electronics (ISIE), 2013 IEEE International Symposium on, 2013* (pp. 1-6): IEEE

Ribeiro, B., Gonçalves, I., Santos, S., & Kovacec, A. Deep learning networks for off-line handwritten signature recognition. In *Iberoamerican Congress on Pattern Recognition, 2011* (pp. 523-532): Springer

Sainath, T. N., Mohamed, A.-r., Kingsbury, B., & Ramabhadran, B. Deep convolutional neural networks for LVCSR. In *2013 IEEE International Conference on Acoustics, Speech and Signal Processing, 2013* (pp. 8614-8618): IEEE

Schmidhuber, J. (2015). Deep learning in neural networks: An overview. *Neural networks, 61*, 85-117.

Schwenke, H., Knapp, W., Haitjema, H., Weckenmann, A., Schmitt, R., & Delbressine, F. (2008). Geometric error measurement and compensation of machines—an update. *CIRP Annals-Manufacturing Technology, 57*(2), 660-675.

Seltzer, M. L., Yu, D., & Wang, Y. An investigation of deep neural networks for noise robust speech recognition. In *2013 IEEE International Conference on Acoustics, Speech and Signal Processing, 2013* (pp. 7398-7402): IEEE

Siguenza-Guzman, L., Saquicela, V., Avila-Ordóñez, E., Vandewalle, J., & Cattrysse, D. (2015). Literature review of data mining applications in academic libraries. *The Journal of Academic Librarianship, 41*(4), 499-510.

Slamani, M., Nubiola, A., & Bonev, I. A. (2012). Modeling and assessment of the backlash error of an industrial robot. *Robotica, 30*(07), 1167-1175.

Stryczek, R. (2016). A metaheuristic for fast machining error compensation. *Journal of Intelligent Manufacturing, 27*(6), 1209-1220.

Tao, F., Cheng, J., Qi, Q., Zhang, M., Zhang, H., & Sui, F. (2017). Digital twin-driven product design, manufacturing and service with big data. [journal article]. *The International Journal of Advanced Manufacturing Technology*, doi:10.1007/s00170-017-0233-1.

Tompson, J. J., Jain, A., LeCun, Y., & Bregler, C. Joint training of a convolutional network and a graphical model for human pose estimation. In *Advances in neural information processing systems, 2014* (pp. 1799-1807)

Wang, C., & Jiang, P. (2017). Deep neural networks based order completion time prediction by using real-time job shop RFID data. [journal article]. *Journal of Intelligent Manufacturing*, doi:10.1007/s10845-017-1325-3.

Wang, K.-S., Li, Z., Braaten, J., & Yu, Q. (2015). Interpretation and compensation of backlash error data in machine centers for intelligent predictive maintenance using ANNs. *Advances in Manufacturing, 3*(2), 97-104.

Xiong, H. Y., Alipanahi, B., Lee, L. J., Bretschneider, H., Merico, D., Yuen, R. K., Hua, Y., Gueroussov, S., Najafabadi, H. S., & Hughes, T. R. (2015). The human splicing code reveals new insights into the genetic determinants of disease. *Science, 347*(6218), 1254806.

Yu, D., & Deng, L. (2011). Deep learning and its applications to signal and information processing [exploratory dsp]. *IEEE Signal Processing Magazine, 28*(1), 145-154.

Zhang, Y., Er, M. J., Zhao, R., & Pratama, M. (2017). Multiview convolutional neural networks for multidocument extractive summarization. *IEEE transactions on cybernetics, 47*(10), 3230-3242.

Zhang, Y., Yang, J., & Zhang, K. (2013). Geometric error measurement and compensation for the rotary table of five-axis machine tool with double ballbar. *The International Journal of Advanced Manufacturing Technology, 65*(1-4), 275-281.

Zhong, G., Wang, C., Yang, S., Zheng, E., & Ge, Y. (2015). Position geometric error modeling, identification and compensation for large 5-axis machining center prototype. *International Journal of Machine Tools and Manufacture, 89*, 142-150.

Zhou, S., Chen, Q., & Wang, X. Discriminative Deep Belief networks for image classification. In *2010 IEEE International Conference on Image Processing, 2010* (pp. 1561-1564): IEEE

Zhu, S., Ding, G., Qin, S., Lei, J., Zhuang, L., & Yan, K. (2012). Integrated geometric error modeling, identification and compensation of CNC machine tools. *International Journal of Machine Tools and Manufacture, 52*(1), 24-29.