Journal of Manufacturing Systems 43 (2017) 248–256

Contents lists available at ScienceDirect

Journal of Manufacturing Systems

journal homepage: www.elsevier.com/locate/jmansys

Technical Paper

Multi-bearing remaining useful life collaborative prediction: A deep learning approach

Lei Ren[a,b,*], Jin Cui[a,b], Yaqiang Sun[a,b], Xuejun Cheng[a,b]

[a] School of Automation Science and Electrical Engineering, Beihang University, Beijing, China
[b] Engineering Research Center of Complex Product Advanced Manufacturing System, Ministry of Education, China

ARTICLE INFO

Article history:
Received 30 September 2016
Received in revised form 22 January 2017
Accepted 21 February 2017
Available online 6 March 2017

Keywords:
Cyber manufacturing
Bearing health analysis
Remaining useful life prediction
Data analytics
Deep learning
Industrial big data

ABSTRACT

Rolling bearing health analysis and remaining useful life prediction have become an increasingly crucial research area that can promote reliability and efficiency in the modern manufacturing industry. Internet-of-Things and cyber manufacturing techniques make it convenient to collect large volumes of sensor data that can provide powerful support for efficient data analytics such as deep learning. The combination of a massive amount of available data and advanced machine learning models brings new opportunities for bearing remaining useful life prediction. This paper proposes an integrated deep learning approach for multi-bearing remaining useful life collaborative prediction by combining both time domain features and frequency domain features. The method can extract high-quality degradation patterns of rolling bearing from vibration signals. Regarding features extracted from bearing vibration signals, in addition to three conventional time domain features, a novel frequency domain feature is adopted in the proposed method as well. Based on the extracted features, the deep neural network model is introduced to predict the remaining useful life of rolling bearing. We evaluate the performance of the proposed method on a real dataset and compare it with several commonly used shallow prediction methods Numerical experiment results show the effectiveness and superiority of the proposed approach.

1. Introduction

Rapid development of Internet-of-Things as well as cyber manufacturing [1–6] techniques are changing modern manufacturing industry dramatically. A massive amount of industrial data are generated from sensors in cyber manufacturing environment, providing new possibilities for further improvement of reliability and efficiency for manufacturing industry. In addition, advances in AI (Artificial Intelligence) oriented big data analytics are creating new research opportunities for large-scale manufacturing data processing and analysis.

As an indispensable element of modern factory machines, rolling bearings play a critical role in industrial manufacturing systems, especially where rotating machinery and equipments serve as the essential components. Bearing faults are usually considered as one of the most frequent causes of mechanical failures [7,8]. Bearing reliability has a crucial impact on dependability, durability and efficiency of the equipments in manufacturing industry. In recent years, research related to bearing degradation process analysis and service time prediction has become an increasingly important area [9].

The health status of a rolling bearing is influenced by a wide variety of factors, i.e., running load, operating temperature, lubrication, installation, corrosion, material defects, and operation mode, during the whole service life of the bearing. Each of these factors has a unique effect on bearing health status, thus making bearing remaining useful life prediction a rather complex problem [10]. In fact, multi-bearing collaborative analysis introduces more challenging issues in remaining useful life prediction, which aims at predicting remaining useful life of a bearing by considering a group of bearings with the same type under similar operating conditions. Indeed, there often exist differences between respective degradation patterns in a group of bearings, so multi-bearing remaining useful life collaborative prediction still faces great challenges.

Current existing bearing remaining useful life prediction approaches can be generally classified into three categories, model-based method, knowledge-based method, and data-driven method [11]. Model-based method usually requires constructing control equation to describe the operation principles [12]. Precision of a model-based method depends on the accuracy of the established model. Since it is difficult to describe the complex process

* Corresponding author at: School of Automation Science and Electrical Engineering, Beihang University, Beijing, China.
E-mail address: renlei@buaa.edu.cn (L. Ren).

http://dx.doi.org/10.1016/j.jmsy.2017.02.013
0278-6125/

of bearing degradation clearly and comprehensively, mechanism model for bearing remaining useful life prediction is usually hard to construct and has drawbacks from the perspective of prediction accuracy. Knowledge-based method makes prediction based on expert systems which take advantages of empirical knowledge. Typically, knowledge-based methods are good at qualitative evaluation rather than quantitative prediction [13]. Therefore, knowledge-based methods are more suitable to make qualitative predictions but have limitations in forecasting bearing remaining useful life with high accuracy. Data-driven method, featured data mining and machine learning, need not establish complex control equations at the initial analytical stage. In addition, data-driven methods are able to provide quantitative prediction results or probabilistic distributions of predictive variables. Building appropriate learning models for machine learning algorithms plays an essential role in data-driven approaches [14]. Both quantity and quality of available data have an effect on the prediction precision of a data-driven method.

In recent years, data-driven methods have become increasingly popular by leveraging the advantages of data analytics methodology [15]. The development and application of deep learning approach [16] have a tremendous impact on big data analysis technique research. Deep learning approach features high precision [17] and big data processing capability [18], allowing the decrease of modeling complexity [19]. Therefore, deep learning based data-driven approach has the potential to provide new opportunities for multi-bearing remaining useful life prediction. In addition, the development of high-performance computing also supports the implementation of complex machine learning algorithms to process the large-scale data efficiently.

To address the issues of multi-bearing remaining useful life prediction, this paper proposes an integrated deep learning approach based on collaborative analysis of monitoring data from multi-bearing vibration signals by combining both time domain features and frequency domain features. The feature parameters system consists of three features of time domain and one feature of frequency domain. The time domain features include root mean square (RMS), crest factor and kurtosis. All these three features derive from classical time domain features for bearing vibration signals analysis. The frequency domain feature is a newly defined feature, named as frequency spectrum partition summation (FSPS), which is represented as a six-dimensional vector in this paper. The features extracted from the bearing vibration signals almost cover a whole process of bearing degradation. The frequency domain parameter is sensitive to the earlier stage and the later stage, while the time domain features have advantages in representing the middle stage of bearing degradation process. The fully connected deep neural network is adopted in the proposed multi-bearing remaining useful life prediction model. And the deep neural network related parameters are determined according to a series of grid search experiments.

In order to validate the effectiveness of the proposed method, numerical experiments are implemented on a real dataset, which is provided by AS2M department of FEMTO-ST Institute, for performance comparison with the gradient boosting decision tree (GBDT) method, the support vector machine (SVM) method, BP neural network, Gaussian regression method and Bayesian Ridge method. Experimental results show the effectiveness and superiority of the proposed approach.

The remaining paper is organized as follows. Section 2 investigates the related work. Section 3 presents descriptions of the problem and the proposed methodology. Numerical experiment details and result analysis are given in Section 4. Section 5 concludes the paper.

2. Related work

In general, solutions for rolling bearing remaining useful life prediction could be classified into three categories, model-based prediction approach, knowledge-based prediction technique, and data-driven prediction method.

System mechanism model is the foundation of model-based prediction method. System state equations or control equations for rolling bearing degradation process are established by analyzing the complex relationships among bearing health status and the main affect factors. Based on the state/control equations and the current state of bearing, which can be used for initialization parameters extraction, the remaining useful life of rolling bearing is derived and calculated. Considering the noise in bearing state monitoring data, filtering algorithms like Kalman filter [20] and particle filter [21] are commonly adopted for signal preprocessing in order to improve the performance of the model-based prediction method. Theoretically, model-based prediction method is able to reflect the nature and the law of a system adequately. The established system model is expected to fully describe the mechanism and characteristics of bearing degradation process, thus model-based method has the potential to produce a reasonable forecasting result with high accuracy. However, since there exists various affect factors for bearing degradation process and some effect mechanisms are unclear or unpredictable, so it is usually quite difficult or even impossible to establish a precise and reliable mechanism model for bearing remaining useful life prediction problem. It is for this reason that model-based prediction method can only be used in very limited application fields.

By leveraging the advantages of accumulated technical experience or knowledge on the relevant issue, knowledge-based method makes prediction or judgment without the need of precise system mechanism model. Accumulation of domain knowledge and reasonable judgment of application situation are the crucial aspects of the knowledge-based prediction method. Expert system [22] and fuzzy logic [23] are commonly used classic knowledge-based decision-making techniques. Knowledge and experience are fully used to support the decision-making process. Knowledge-based prediction method is widely adopted in various application fields, especially in some qualitative decision making scenarios. However, knowledge-based method has limitations in solving high precise quantitative prediction problems. Thus, knowledge-based prediction approaches may have difficulty in providing satisfactory forecasting results of bearing remaining useful life.

Data-driven prediction method learns the latent association relationships among bearing degradation process and its affect factors automatically from the sensor monitoring data of rolling bearing by utilizing machine learning algorithms or other intelligent data analysis techniques. Then the well-trained data-driven prediction model can be used for bearing remaining useful life prediction. The forecasting precise depends on the quality and quantity of the available training data as well as the effectiveness of the learning algorithm. Data-driven prediction method has advantages in both modeling aspect and quantitative prediction aspect, but it requires the support of numerous high-quality available learning data as well as effective data analysis techniques. Rapid development and widely application of Internet-of-Things and cyber manufacturing techniques as well as advances in intelligent data analysis and high-performance computing techniques create new opportunities for research of data-driven prediction method and promote its development as well.

As the selected signal features and learning model varies, there exists various data-driven prediction models for rolling bearing remaining useful life forecasting problem. However, only if features of high representational capability and suitable high efficient learning model are integrated properly will the prediction method be able to make precise predictions.

The features that extracted from bearing vibration signal and used for bearing remaining useful life prediction can be grouped into three categories, time domain feature, frequency domain feature, and time-frequency domain feature. In which, commonly used time domain features including mean, variance, RMS, crest factor and kurtosis [24]. Time domain characteristic indexes are able to reflect the general tendency of rolling bearing degradation process in an intuitive way but insensitive to small changes. Besides, noise signals may influence the prediction result seriously as time domain features are sensitive to noise. Frequency spectrum variance and frequency spectrum RMS are two commonly used frequency domain features in bearing vibration signal analysis. Frequency domain features are suitable for stationary signal processing and widely used in bearing failure diagnosis research. In addition, some studies also apply time-frequency domain features to bearing vibration signal analysis and bearing remaining useful life prediction [25]. Typical time-frequency domain feature extraction techniques including wavelet packet transform [26] and empirical mode decomposition. Time-frequency domain features are capable of representing weak signal characteristics and suitable for nonlinear signal processing. However, these features have drawbacks in information redundant and usually need complex mathematical derivation for feature extraction. Therefore, a set of reasonable and efficient vibration signal characteristics indexes is crucial to the prediction result of bearing remaining useful life. The selection of signal features needs a comprehensive consideration of degradation pattern representational ability, information redundancy, and complexity of signal feature.

Classic learning models account for the great majority of learning models or algorithms applied in bearing remaining useful life prediction. Among them, neural network model plays an important role and makes excellent predictions in some scenarios because of the outstanding capability for the neural network to fitting a nonlinear system [27]. However, from another perspective, how to avoid over-fitting problem [28] and improve forecasting effect are two knotty problems for using neural network model. With the development of data analysis techniques and advances in efficient intelligent learning model research, the deep learning method is becoming popular gradually [16]. Deep learning method makes up the shortcomings and deficiencies of neural network model, especially in terms of error propagation. Deep learning method changes the error propagation mechanism of the neural network, preventing the error diffusion layer by layer [29], which greatly improves the prediction accuracy.

3. Problem and methodology

3.1. Multi-bearing remaining useful life collaborative prediction

Rolling bearing remaining useful life prediction means predicting the remaining useful life of rolling bearing according to condition monitoring data acquired in their operating process. Commonly used condition monitoring data of rolling bearing include vibration signal, acceleration signal, and temperature signal. In this paper, vibration signal is applied in rolling bearing remaining useful life forecasting and health evaluation analysis.

Existing studies on rolling bearing remaining useful life prediction mainly focus on single bearing analysis. However, performance decline features of different bearings in same operating condition show some similarity. All the performance decline features and delicate relations among them are implied in the vibration data. Machine learning algorithms provide an effective tool for observation data analysis.

The multi-bearing remaining useful life collaborative prediction refers to the problem of remaining useful life prediction for multi bearings based on the condition monitoring data acquired in same operating condition. Multi-bearing remaining useful life collaborative prediction aimed at forecasting remaining useful life of a bearing according to the available monitoring data of the bearing itself as well as monitoring data of other bearings of the same type and operating conditions.

3.2. Feature extraction

3.2.1. Time domain feature

Time domain signal features are effective to reflect the running condition and the fault information of rolling bearing. Time domain features only depend on the probability density functions of signal amplitude and are sensitive to the faults and defects of rolling bearing. As a result of balance between information representation ability and information dimensionality, three classical time domain features, root mean square X_{RMS}, crest factor X_{Crest}, and kurtosis $X_{Kurtosis}$, are used in this paper. Formulas for these three features are presented as follows:

$$X_{RMS} = \sqrt{\frac{\sum_{i=1}^{n} (x(i))^2}{n}} \tag{1}$$

$$X_{Crest} = \frac{max\left\{x(i)\right\}}{X_{RMS}} \tag{2}$$

$$X_{Kurtosis} = \frac{\sum_{i=1}^{n} (x(i) - \bar{x})^4}{n(X_{RMS})^4} \tag{3}$$

where $x(i)$ is a series of vibration signal and n refers to the number of vibration signal data points.

3.2.2. Frequency domain feature

Rolling bearing is of different health condition when in different degradation stage. The degradation features change over time throughout the whole life-cycle of rolling bearing. However, time domain features are not sensitive to signal frequency. Therefore, time domain features are not sufficient enough to reflect the bearing degradation process. In order to extract more comprehensive health condition information of rolling bearing from the monitoring vibration data, a new frequency domain feature named Frequency Spectrum Partition Summation (FSPS) is defined.

Given a series of vibration signal $x(i)$ for $i = 1, 2, \ldots, n$, $s(j)$ refers to its frequency spectrum which obtained through Fourier transformation and $j = 1, 2, \ldots, m$. Then, the FSPS index $X_{FSPS}(k)$ can be calculated as follows:

$$X_{FSPS}(k) = \sum_{j=\frac{K+m(k-1)}{K}}^{\frac{mk}{K}} s(j) \tag{4}$$

where $k = 1, 2, \ldots, K$. It should be note that the new defined frequency domain feature $X_{FSPS}(k)$ is a one-dimensional vector which be consist of K elements. K is an empirical parameter and is generally determined by the concrete domain problem.

3.3. Deep neural network

Deep neural network refers to a neural network that has two or more hidden processing layers between the input and output layers [16]. Deep neural network is a deep learning model developed based on the basic neural network. Compared with shallow neural network with a single hidden layer, deep neural network has the potential to solve more complex learning problems. A deep neural network consists of an input layer and an output layer separated by two or more hidden layers. A reasonable number of the hidden units can lead to higher prediction accuracy. Weights and bias values, as well as activation functions are primary effect factors of the performance of a deep neural network. A deep neural network model can be trained discriminatively by using the standard backpropagation algorithm. To combat the overfitting problem, regularization methods can be applied during the training process.

Currently, with the rapid development and successful application of deep learning techniques, many open source deep learning frameworks and libraries are developed to facilitate related research. Among them, Keras is one of the deep learning libraries. Keras is a python written modularity framework developed to enabling fast deep learning experimentation. Modularity and extensibility makes Keras a powerful tool for rapid deep learning experiment implementation.

3.4. Performance index

Two commonly used performance indicators, the mean absolute error X_{MAE} and the root mean square error X_{RMSE}, are employed to evaluate the performance of the proposed model. Mathematical descriptions of the indexes are given as follows:

$$X_{MAE} = \frac{\sum_{i=1}^{n} |f_i - \hat{f}_i}{|} n \tag{5}$$

$$X_{RMSE} = \sqrt{\frac{\sum_{i=1}^{n} \left(f_i - \hat{f}_i\right)^2}{n}} \tag{6}$$

where f_i is the real remaining useful life and $\hat{f}_i$ is the predicted remaining useful life.

4. Experiment and analysis

4.1. Data description

In order to validate the effectiveness of the newly defined signal frequency domain feature and the deep architecture collaborative prediction model, the multi-bearing remaining useful life collaborative prediction model was applied to a dataset, provided by AS2M department of FEMTO-ST Institute, as a numerical experiment. The vibration data of rolling bearing were generated from run-to-failure tests performed on the PRONOSTIA experimental platform [30]. Overview illustration of the test rig is shown in Fig. 1.

The rolling bearing was installed on the experimental platform and run-to-failure experiment was performed under constant load condition and constant rotating speed. For the bearings used in this paper, the rotating speed was maintained at 1650 rpm and the payload weight is 4200 N. Two vibration sensors, placed on vertical axis and horizontal axis separately, are adopted to acquire the bearing vibration data with a frequency of 25.6 kHz. And the measure is acquired at a frequency equal to 100 Hz. The bearing is regarded as failure as long as the amplitude of vibration signal overpassed 20 g. A dataset with four groups of bearing testing vibration signals was used in the following multi-bearing remaining useful life collaborative prediction experiment. Fig. 2 depicted a vibration raw signal acquired from an experiment.

4.2. Experiment implementation

4.2.1. Identify the failure time of testing rolling bearings

The very first step of the numerical experiment is data preprocessing. As a bearing is considered as failure if its vibration signal amplitude exceeds 20 g, the failure time of testing rolling bearing are determined by examining both vertical vibration signal and horizontal vibration signal. For simplicity, dimensionless variable time step is adopted instead of accurate time with its unit. Measure frequency of PRONOSTIA experimental platform equal to 100 Hz, thus, 0.1 s is defined as one time step in the numerical experiment. The remaining useful life of a rolling bearing is defined as the available service time form the present moment to the moment it failed. For example, for a bearing at sampling point t_1, it failed at sampling point t_2, then remaining useful life of the bearing is calculated by $t_2 - t_1$.

4.2.2. Extract both time domain and frequency domain features from the observed vibration signal data

High quality signal features are able to represent the useful information of the dataset more intensively and to filter the unrelated information effectively. As described in previous section, three time domain features, root mean square, crest factor, and kurtosis as well as a frequency domain feature, FSPS index, are employed to extract degradation information from the vibration data acquired from run-to-failure experiment. Considering the individualization characteristic of multi-bearing remaining useful life collaborative prediction problem, the empirical parameter K for frequency domain feature FSPS is set as 6. Figs. 3 and 4 show the RMS curve and the crest factor curve of one testing bearing, respectively.

4.2.3. Feature data normalization processing and data splitting

Different features have different ranges of values. To eliminate the negative effect caused by different ranges of values, min-max normalization method was applied to transform the range of values as [0,1] for all extracted features. We adopted the nine-dimensional vector to represent the features of bearing vibration signal, including three time domain features and one frequency domain feature that represented as a six-dimensional vector. As bearing vibration signal data in the dataset was collected from two directions, vertical direction and horizontal direction, so the dimensionality of the model input data is 18. Besides, the corresponding normalized remaining useful life is used as the label for each data record.

In addition, to perform the multi-bearing remaining useful life collaborative prediction numerical experiment, the preprocessed vibration dataset is separated as training data and testing data. A certain percentage of data points are selected as testing data randomly and the rest are used as training data. In our experiment, six testing datasets, account for different percentages of the whole dataset (5%, 10%, 15%, 20%, 25%, 30%), are set for comparison.

4.2.4. Deep neural network model construction and training

In this paper, we constructed a fully connected deep neural network for multi-bearing remaining useful life prediction. The whole experiment is developed and implemented based on the Keras deep learning framework. Input layer of the deep neural network model is fully connected to the first hidden layer, and by the same rule, the second hidden layer is fully connected to the first hidden layer and so on up to the output layer. Number of layers and amount of neurons for each layer are determined according to a series of grid search experiments. The deep neural network is consisting of 8 hidden layers with different amount of neurons (300, 200, 150, 100, 80, 50, 30, 1). Keras framework provides various probabilistic distribution based parameter initialization methods. In this paper, uniform distribution is employed for weight parameter initialization. Because of their excellent performance in the numerical experiments, ReLU function is applied as activation function for middle layers and activation function for the last layer is a sigmoid function. The sigmoid activation function matches the normalized remaining useful life values well since the output of a sigmoid function ranges from 0 to 1. As for the training setting initialization, mean squared error and RMSprop algorithm are selected as the optimization objective and the optimizer, respectively. After

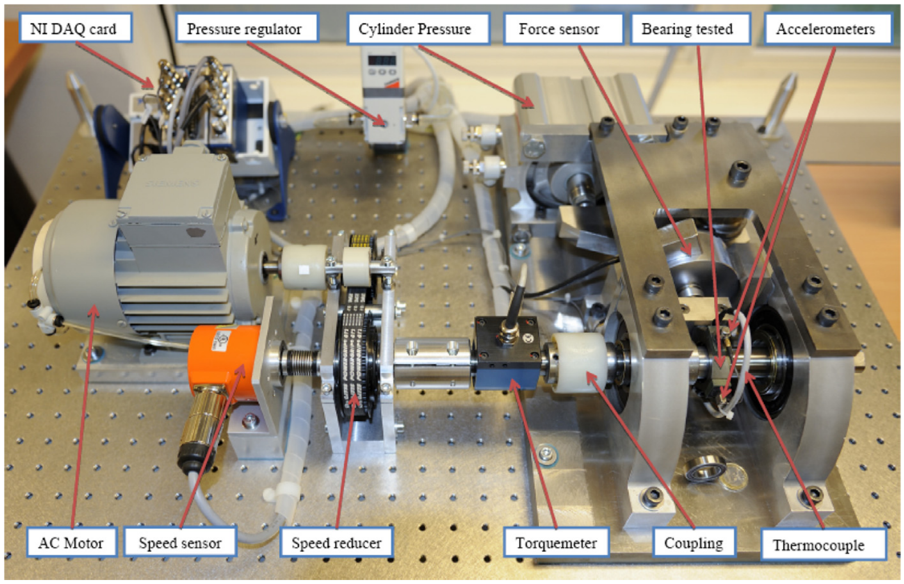

Fig. 1. Overview illustration of PRONOSTIA experimental platform [30].

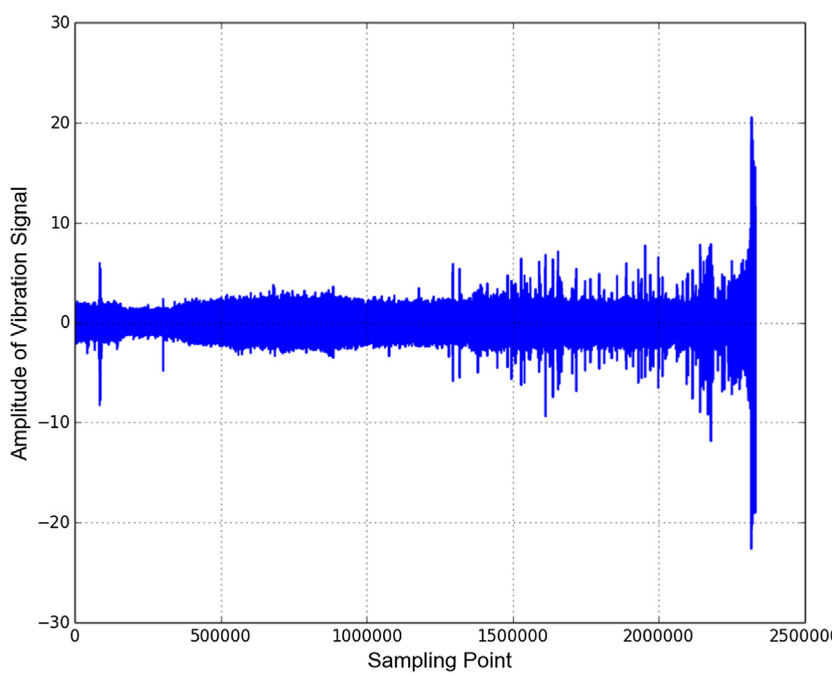

Fig. 2. Vibration raw signal gathered during a run-to-failure experiment.

parameter setting, the deep neural network model is trained based on the training data selected in the previous step. The dropout technique is employed in the Keras deep learning library to overcome the overfitting problem. Dropout is a technique where randomly selected neurons are ignored during training. This means that their contribution to the activation of downstream neurons is temporally removed on the forward pass and any weight updates are not applied to the neuron on the backward pass. The effect is that the network is capable of better generalization and is less likely to overfit the training data.

The metrics evaluation program and deep learning part are developed in python with the use of Keras deep learning library to speed up the experiment implementation. All tests have been performed on a PC equipped with Intel Xeon 5520 2GHZ with 8GB RAM and NVIDIA Quadro 2000 with 1GB VRAM.

4.2.5. Multi-bearing remaining useful life prediction and comparison among different predicting models

The well trained deep neural network is applied for multi-bearing remaining useful life prediction. We compared the performance of the deep neural network model with the GBDT method, the SVM method, BP neural network, Gaussian regression method and Bayesian Ridge method. Classical indicators, the mean absolute error (MAE) and the root mean square error (RMSE) are adopted to evaluate the performance of different prediction models. As the testing dataset is selected randomly according to a

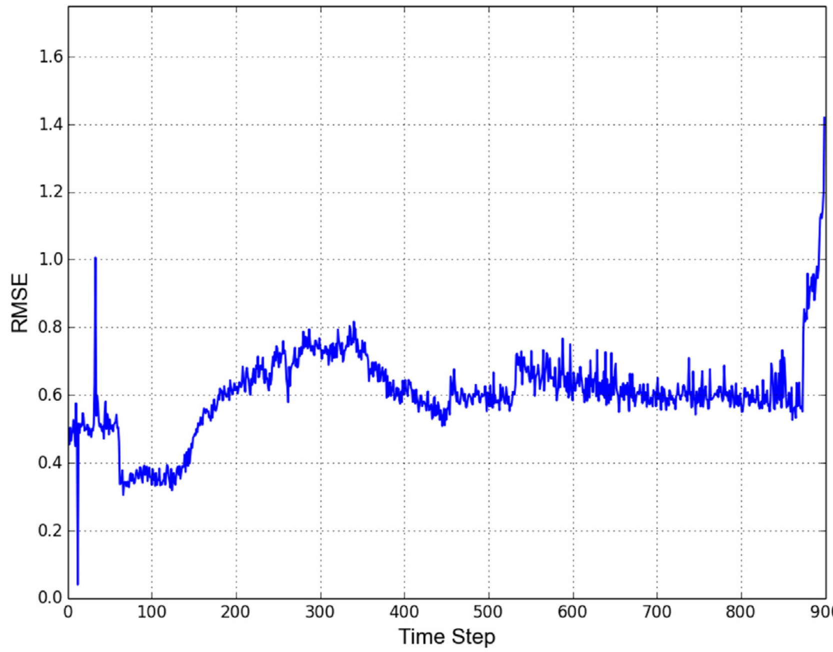

Fig. 3. RMS curve of a testing bearing vibration signal data.

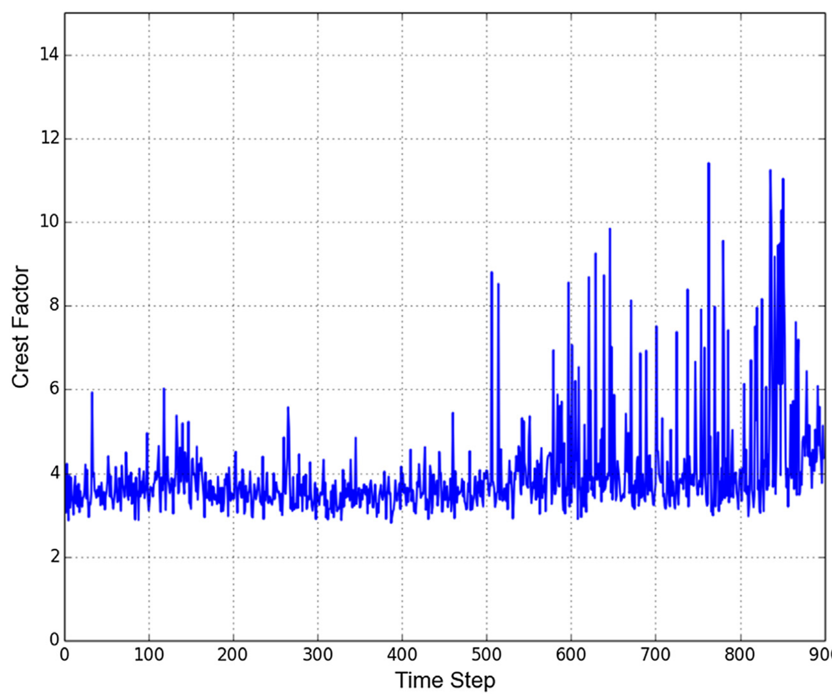

Fig. 4. Crest factor curve of a testing bearing vibration signal data.

specified percentage, we run each prediction model for 20 times and then got the average prediction result for comparison.

4.3. Results and analysis

Fig. 5 presents one of the typical prediction outputs for four bearings remaining useful life, which was marked as the black dotted line in the figure. Meanwhile, the red line for comparison represents the observed remaining useful life of the testing bearings. There are four groups of curves in Fig. 5 and each group corresponding to a rolling bearing. In Fig. 5, it is shown that the predicted remaining useful life has similar pattern with the real observed result. The predicted remaining useful life also matches well with the observed data, especially for bearing 3.

We compared the prediction performance of deep neural network with five commonly used high quality forecasting models. Evaluation metrics, MAE and RMSE, are adopted to assess the effectiveness of different models quantitatively. Figs. 6 and 7 show the MAE curves and RMSE curves for prediction output of different forecasting models, respectively. Horizontal axis in Figs. 6 and 7 represents the percentage of training data while vertical axis represents the metric value. As the bearing remaining useful life are

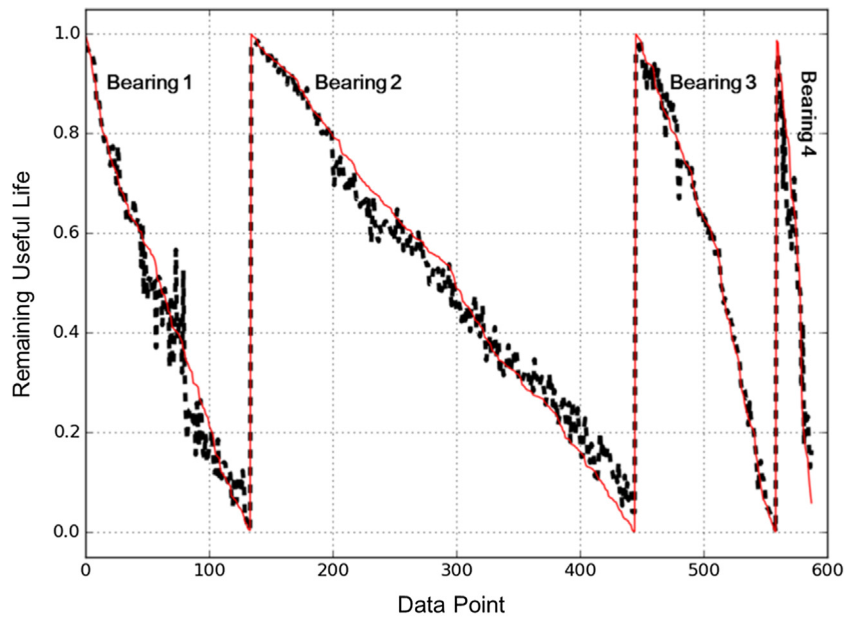

Fig. 5. Typical forecasting output of deep neural network for four-bearing remaining useful life collaborative prediction problem.

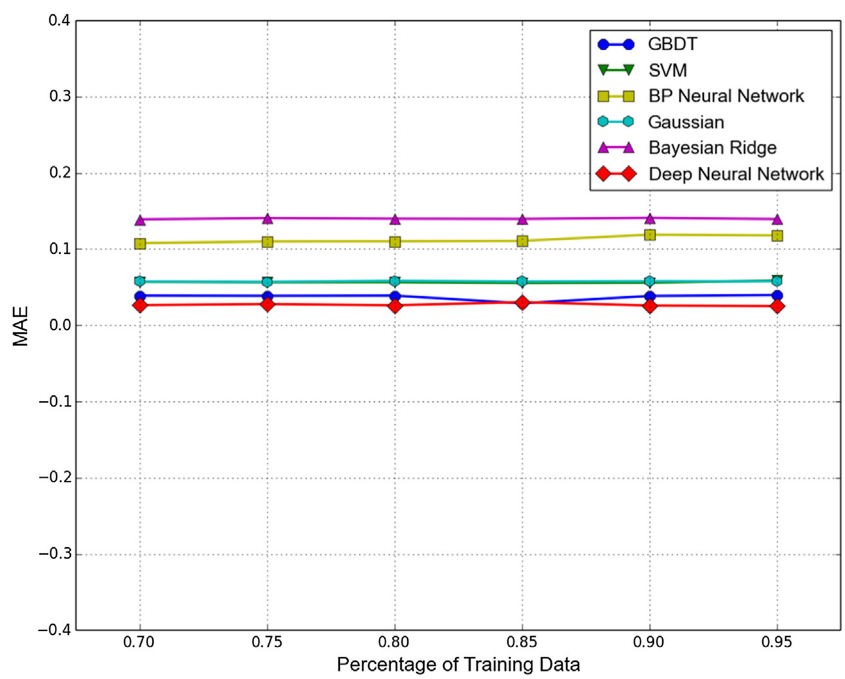

Fig. 6. MAE curves for prediction output of different forecasting models.

normalized into a range of [0,1], so even a small change of the evaluation metric value imply a remarkable difference of prediction performance.

As shown in Fig. 6, deep neural network model has the minimum MAE value compared with other models and the average MAE value for deep neural network on datasets of different percentages is 0.0270. SVM method and Gaussian regression model almost have same result on MAE indicator. Bayesian Ridge method has the biggest MAE with an average value of 0.1399.

As can be seen in Fig. 7, performance ranking order of the six prediction models is deep neural network, GBDT method, SVM model, Gaussian regression method, BP neural network and Bayesian Ridge method from the perspective of RMSE. Compared with MAE, RMSE of the prediction models are larger. For deep neural network, the average RMSE on different testing datasets is 0.0414. And this metric for Bayesian Ridge model even reaches 0.1710. The effectiveness of deep neural network for multi-bearing remaining useful life prediction is promising.

5. Conclusion

The requirement of high reliability and high efficiency in modern manufacturing industry makes bearing health analysis and remaining useful life prediction become an increasingly crucial research area. Advances in cyber manufacturing as well as AI oriented big data analytics provide new opportunities for bearing

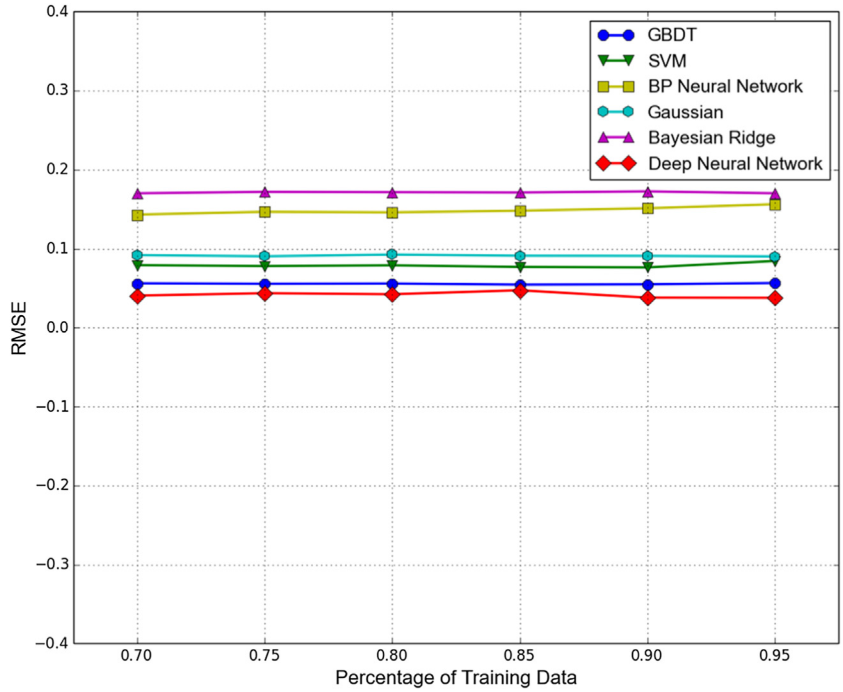

Fig. 7. RMSE curves for prediction output of different forecasting models.

health analysis and remaining useful life prediction. This paper proposes an integrated deep learning approach for multi-bearing remaining useful life collaborative prediction. The method can extract high quality degradation patterns of rolling bearing from its vibration signals. The extracted bearing vibration signal features include three features of time domain and one proposed six-dimensional frequency domain feature. We have adopted the deep neural network model for bearing degradation pattern learning and remaining useful life prediction. Based on the Keras deep learning rapid experiment implementation framework, we evaluate the performance of the proposed method on the dataset provided by FEMTO-ST Institute and compare it with the GBDT method, the SVM method, BP neural network, Gaussian regression method and Bayesian Ridge method. Numerical experiment results show the effectiveness and superiority of the proposed approach from both MAE perspective and RMSE perspective.

For future work, we plan to investigate other deep learning algorithms in solving multi-bearing remaining useful life prediction problems. In addition, it also would be interesting to work on multi-bearing remaining useful life collaborative prediction under alterative working condition.

Acknowledgements

The research is supported by the NSFC (National Science Foundation of China) Projects (No. 61572057) in China, the National High-Tech Research and Development Plan of China under Grant No. 2015AA042101.

References

[1] Li BH, Zhang L, Ren L, Chai X, Tao F, Luo Y, et al. Further discussion on cloud manufacturing. Comput Integr Manuf Syst 2011;17(3):449–57.
[2] Wu D, Thames L, Rosen DW, Schaefer D, et al. Towards a cloud-Based design and manufacturing paradigm: looking backward, looking forward. ASME 2012 international design engineering technical conferences & computers and information in engineering conference 2012:315–28.
[3] Wu D, Rosen DW, Wang L, Schaefer D, et al. Cloud-based design and manufacturing: a new paradigm in digital manufacturing and design innovation. Comput Aided Des 2015;59:1–14.
[4] Ren L, Zhang L, Tao F, Zhao C, Chai X, Zhao X, et al. Cloud manufacturing: from concept to practice. Enterp Inf Syst 2014;9(2):186–209.
[5] Ren L, Zhang L, Wang L, Tao F, Chai X, et al. Cloud manufacturing: key characteristics and applications. Int J Comput Integr Manuf 2014:1–15.
[6] Ren L, Cui J, Li N, Wu Q, Ma C, Teng D, et al. Cloud-based intelligent user interface for cloud manufacturing: model, technology, and application. J Manuf Sci Eng 2015;137(4):040910.
[7] Boškoski P, Gašperin M, Petelin D. Bearing fault prognostics based on signal complexity and Gaussian process models. In: Prognostics and health management; 2012. p. 1–8.
[8] Rasmussen C, Williams C. Gaussian processes for machine learning. MIT Press; 2004.
[9] Dong S, Luo T. Bearing degradation process prediction based on the PCA and optimized LS-SVM model. Measurement 2013;46(9):3143–52.
[10] Maio FD, Tsui KL, Zio E. Combining relevance vector machines and exponential regression for bearing residual life estimation. Mech Syst Signal Process 2012;31(8):405–27.
[11] Zarei J, Tajeddini MA, Karimi HR. Vibration analysis for bearing fault detection and classification using an intelligent filter. Mechatronics 2014;24(2):151–7.
[12] Zhao M, Jin X, Zhang Z, Li B, et al. Fault diagnosis of rolling element bearings via discriminative subspace learning: visualization and classification. Expert Syst Appl 2013;41(7):3391–401.
[13] Bellini A, Filippetti F, Tassoni C, Capolino GA, et al. Advances in diagnostic techniques for induction machines. IEEE Trans Ind Electron 2008;55(12):4109–26.
[14] Mallat S. Wavelet zoom – a wavelet tour of signal processing. Second Edition – VI Academic Press; 1998.
[15] Aydin I, Karakose M, Akin E. A new method for early fault detection and diagnosis of broken rotor bars. Energy Convers Manage 2011;52(4):1790–9.
[16] Schmidhuber J. Deep Learning in neural networks: an overview. Neural Netw 2014;61:85–117.
[17] Sutskever I, Martens J, Dahl G, Hinton G, et al. On the importance of initialization and momentum in deep learning. International conference on machine learning 2013.
[18] Deng L, Yu D. Deep learning: methods and applications. Found Trends Signal Process 2013;7(3):197–387.
[19] Deng L, dynamic A. feature-based approach to the interface between phonology and phonetics for speech modeling and recognition. Speech Commun 1998;24(4):299–323.
[20] Welch G, Bishop G. An Introduction to the Kalman Filter. University of North Carolina at Chapel Hill; 1995.
[21] Arulampalam MS, Maskell S, Gordon N, Clapp T, et al. A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking. IEEE Trans Signal Process 2002;50(2):174–88.
[22] Lauritzen SL, Spiegelhalter DJ. Local computations with probabilities on graphical structures andtheir application to expert systems. J R Stat Soc 1988;50(157):157–224.

[23] Camci F. Process monitoring, diagnostics and prognostics using support vector machines and hidden Markov models. Detroit: Graduate School of Wanye State University; 2005.
[24] Hong S, Zhou Z, Lu C, F.Camci B. Process monitoring, diagnostics and prognostics using support vector machines and hidden Markov models. Detroit: Graduate School of Wanye State University; 2005.
[24] Hong S, Zhou Z, Lu C, Wang B, Zhao T, et al. Bearing remaining life prediction using Gaussian process regression with composite kernel functions. J Vibroeng 2015;17(2):695–704.
[25] Gebraeel N, Lawley M, Liu R, et al. Residual life predictions from vibration-based degradation signals: a neural network approach. IEEE Trans Ind Electron 2004;51(3):694–700.
[26] Yao L, Feng L. Fault diagnosis and fault tolerant control for non-Gaussian time-delayed singular stochastic distribution systems. Int J Control Autom Syst 2016;14(2):435–42.
[27] Devlin J, Zbib R, Huang Z, Lamar T, Schwartz RM, Makhoul J, et al. Fast and robust neural network joint models for statistical machine translation. Meet Assoc Comput Linguist 2014:1370–80.
[28] Matthews WJ, Terhune DB, van Rijn H, Meck W, et al. Subjective duration as a signature of coding efficiency: emerging links among stimulus repetition, predictive coding, and cortical GABA levels. Timing Time Percepti Rev 2014;1(1):1–11.
[29] Gupta S, Agrawal A, Gopalakrishnan K, Narayanan P. Deep learning with limited numerical precision. The 32nd International Conference on Machine Learning 2015:1737–46.
[30] Nectoux P, Gouriveau R, Medjaher K, et al. PRONOSTIA: an experimental platform for bearings accelerated life test. In: IEEE international conference on prognostics and health management. 2012.