Proceedings of the 33rd European Safety and Reliability Conference (ESREL 2023)
Edited by Mário P. Brito, Terje Aven, Piero Baraldi, Marko Čepin and Enrico Zio

doi: 10.3850/978-981-18-8071-1_P391-cd

Deep Learning Models Applied to Intelligent Diagnosis of Rotating Machines

Caique Emanuel da Silva Nunes

Center of Risk Analysis, Reliability and Environmental Modeling, Department of Production Engineering, Universidade Federal de Pernambuco, Brazil. E-mail: caique.nunes@ufpe.br

Thales Henrique Castro de Barros

Department of Eletronic and Systems, Universidade Federal de Pernambuco, Brazil. E-mail: thales.henrique@ufpe.br

Isis Didier Lins

Center of Risk Analysis, Reliability and Environmental Modeling, Department of Production Engineering, Universidade Federal de Pernambuco, Brazil. E-mail: isis.lins@ufpe.br

Márcio José das Chagas Moura

Center of Risk Analysis, Reliability and Environmental Modeling, Department of Production Engineering, Universidade Federal de Pernambuco, Brazil. E-mail: márcio.cmoura@ufpe.br

AI algorithms can help detect anomalies and identify failure modes under specific conditions, making them valuable tools in maintenance management. However, there is no consensus on which of them is the most effective because each author builds a different architecture based on the main deep learning models, changing functions, parameters and normalizations, and with different databases, making a fair comparison between the models impossible. To address this issue, this work proposes a brief review using the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) methodology to identify works in the literature that perform intelligent diagnostics on available datasets of rotating machines using deep learning algorithms. After this review, this work also presents new results from the use of models, such as multilayer perception (MLP), auto-encoder (AE), convolutional neural network (CNN) and recurrent neural network (RNN), making direct comparisons of the result obtained with the outcomes found in the literature after the review. To support the discussions about the results, confusion matrix, accuracy and losses' graphs were generated for all combinations between models and input types applied.

Keywords: Intelligent diagnosis, reliability, prognostic health management, deep learning, rotating machines.

1. Introduction

Prognostics and Health Management (PHM) is an approach used in engineering and manufacturing to monitor the health of machines, predict when they will need maintenance or repair, and optimize their performance. The main goal of it is to reduce downtime, maintenance costs, and safety risks, by enabling proactive maintenance and repair actions based on the actual condition of the machine, rather than on a fixed schedule or reactive response to failures. PHM can be applied to a wide range of complex systems, such as aircraft, trains, manufacturing lines, and rotating machines (Fink et al. 2020; Heng et al. 2009).

Pumps, motors, turbines, and generators are examples of rotating machines that are essential for many industrial processes. These machines are subject to various stresses, including, mechanical, thermal, and electrical stresses, which can lead to wear, damage, and failure over time. Accurately diagnosing and predicting the health of rotating machines is an important requirement for ensuring their reliable operation and preventing catastrophic failures that can result in safety hazards, production losses, and environmental impacts (Rezaeianjouybari and Shang 2020; Heng et al. 2009).

Thus, accurate and reliable intelligent diagnosis can significantly enhance the effectiveness of PHM and improve the overall reliability of rotating machines. It involves the use of machine or deep learning algorithms to analyze the data collected from sensors and other sources, identify patterns, and make predictions about the health of the machine (Rezaeianjouybari and Shang 2020).

In the literature, there are some intelligent diagnosis models that have been developed and applied in practice. However, each author designs a specific architecture based on the main deep learning models and modifies functions, parameters, and normalizations according to their problem. Moreover, they also use different databases that may have different characteristics and quality. This variation in approaches makes it difficult to compare the effectiveness of different models in detecting anomalies and failures and evaluate their performance objectively.

From this scenario, the present work performed a brief search to select the main datasets related to failures in rotating machines referenced in the literature and in the public domain. These datasets were then utilized to test various deep-learning models with a standardized architecture. By doing so, the study aimed to analyze the impact of parameters and functions on failure detection and determine which models performed better through a fair comparison.

2. Literature Review

As already mentioned, before applying the proposed methods to process different datasets, it was necessary to perform a brief review to identify the main methods and datasets used by other works. This previous step is important because it contributes to avoiding the use of an irrelevant dataset at the same time which allows understanding of the state of the art in this context.

To perform the search in literature, a protocol based on a simplification of the methodology Preferred Reporting Items for Systematic Reviews and Meta-Analysis (PRISMA) Page et al. (2021) was used. The first step in the implementation of this tool was the definition of databases used to investigate the state of art in literature. Normally, two or three databases are used Page et al. (2021) but given the brief nature of the intended review and the focus on locating the most influential works of recent years, which normally can be easily found in more than one database, was enough to choose a single database: the Scopus. After this definition, the next step, maybe the most important, was the choice of the keywords Page et al. (2021): "deep learning", "bearing fault diagnosis", "rotating machinery" and "intelligent diagnosis". To further narrow the results some filters also were applied: only peer-reviewed articles published and written in English were considered.

The implementation of this protocol resulted in the construction of a network that supports the identification of the most influential works, besides allowing the clustering of these works in subnetworks, improving the understanding of the scopes of their contributions. Fig. 1 shows the first network generated from the search in literature.

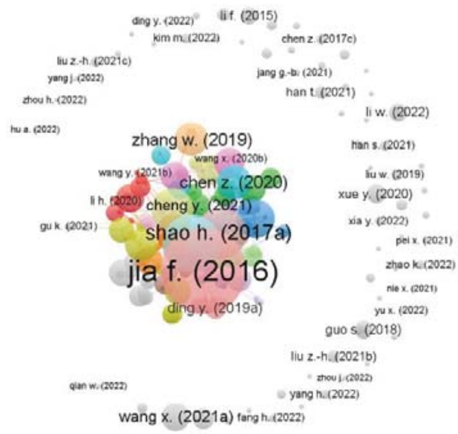

Fig. 1. Network built from citation analysis between the work found in literature search. Each color is assigned to a subcluster while articles not correlated to any other are in gray.

From Fig. 1, was possible separate the works in two different categories: those works that compose the central cluster and those that are not related to any other. This second group usually represents works totally or partially unrelated to the topic and can be filtered to keep the focus on the main tools used. The application of this second layer of filtering results in the final layout used in the analysis, shown in Fig. 2.

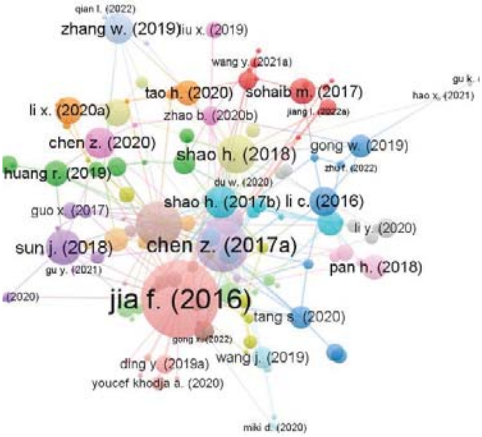

Fig. 2. Central cluster obtained after the application of the second layer of filtering.

Without deep changes after the second filtering, Jia et al. (2016) remain as the most relevant work found. This occurs because this work is one of the first to present deep neural networks as a tool to process massive datasets. Due to this, it is a work commonly cited by others in the construction of the problem. Chen and Li (2017) and Shao et al. (2017b) also present great relevance, being one of the first to apply an auto-encoder to process the datasets. Posteriorly, Shao et al. (2018) and Sun et al. (2017) continue to propose an implementation of methods based on auto-encoders, the first exploring different particularities in activation functions to optimize the results obtained while the second evaluating the impacts of data compression on results.

Another work worthy of mention is Zhang et al. (2019) which proposes a residual learning algorithm to improve network training. A critical bottleneck in data processing is due to the use of larger data sequences. Moreover, despite Li et al. (2020) not appear among the top 10 most cited, this work also presents a high impact which can be associated with the interesting proposal presented in the work: a novel deep learning method for rotating machinery fault diagnosis, allowing the creation of artificial samples for algorithm training. Moreover, this work also compares Gaussian noise, masking noise, signal translation, amplitude shifting and time stretching. The main dataset used is the Case Western Reserve University (CWRU) and to process it, different methods were used and compared, such as Simple Neural Network (NN), Convolutional Neural Network (CNN), and Recurrent Neural Network (RNN). Besides the CWRU dataset, other datasets also were used as in Yang et al. (2021) propose the utilization of an auto-encoder model, aiming to learn important features from limited raw vibration signals, to build a framework, performing tests using the Southeast University Gearbox Dataset (SEU) dataset.

Another model applied to CWRU and especially based on a CNN is presented in Zhao et al. (2020). To validate the proposed model the author performs a direct comparison between this model performance and other algorithms with and without a pre-processing of the dataset. Despite Tao (2020) proposing a solution to a dataset not used in this work, the author implements a pre-processing step based in the short-time Fourier transform (STFT) also used in the present work.

Other works as Tang et al. (2020), that makes a comprehensive comparison between several techniques, or Cheng et al. (2021), which combines a convolutional neural network with a local binary convolution layer to improve performance, are fundamental to a complete understanding about the topic. To reaffirm the CNN as an important processing tool, Cao et al. (2022) present an unsupervised domain-share CNN aiming to simultaneously extract the domain-invariant features from the source domain and the target domain. Lastly, another application of autoencoders was presented by Wu et al. (2021) that evaluated the performance of the processing with models which kept a dependence between the fault diagnosis and labelled data, as CNN implementations.

One of the most extensive works found is Zhao et al. (2020) which explores other datasets such as JNU Bearing Dataset, XJTU-SY Bearing Dataset, UoC Gear Fault Dataset and PU Bearing Dataset. Zhao et al (2020) also use different framework models applied to these datasets as Auto-Encoder and CNN. The variety of methods and datasets used to set this work as one of the most important references for the present study.

3. Methodology

As noted, PHM has increasingly relied on deep learning as a critical tool. This is due to deep learning algorithms have shown remarkable success in absorbing complex patterns in the sensor data and providing insights into the underlying causes of equipment failures. However, it is not simple to say which deep learning model is the most suitable or to make any other type of comparison when using different parameters, functions, and databases.

Therefore, the objective of this work is to be able to equate deep learning models so algorithms that are more persuasive and appropriate for specific situations can be built. The four main deep learning models chosen for testing with this purpose were: Multilayer Perceptron (MLP), Convolutional Neural Network (CNN), Recurrent Neural Network (RNN) and Autoencoder (AE). In general, the architecture of each model will be kept the same for tests with more than one database. A concise summary of the models and the way the tests were conducted will be explained as follow.

Before anything else, the data was preprocessed in 5 different ways. The first form of preprocessing considers the time domain, which means the data were used without any transformation. In the second, the data are adapted to the frequency domain using the Fast Fourier Transform - FFT. In the third form of preprocessing, the Fourier transform is not applied for the complete duration of the signal but considering small segments or frames using the Short-time Fourier Transform (STFT) and leaving the data in the Time-Frequency domain. Continuous Wavelet Transform – CWT was used to preprocess the data in a fourth way and let them in the Wavelet domain. And finally, each unprocessed data was reshaped into smaller images in a preprocess called slicing image.

After being preprocessed, each of the five ways of inputting the data was tested in the models under study. Initially, MLP was tested because it is a basic machine learning model consisting of multiple layers of the perceptron. The model is a supervised process and is trained through three phases, including generating an output, comparing it with the actual classification of data, and adjusting model weights using backpropagation (Hashem Samadi et al. 2023).

Training an MLP also involves dividing the dataset into smaller batches, enabling more efficient computation and frequent weight updates. The process is repeated multiple times (epochs) until the model improves and approaches the perfect result. Tests with the MLP for this research were conducted varying the number of layers, epochs, and batch size.

Similarly, the CNN model specializes in pattern recognition, was tested (Géron, 2019). In summary, CNN operates by utilizing a filter (also known as a kernel) to scan the input data in search of a specific pattern. The filter passes over a small section of pixels, generating a convolution each time, which allows for the identification of the desired pattern. In mathematical terms, convolution refers to an operation that generates an output by modifying the shape of one object through another object.

There exist certain mechanisms that facilitate the tuning of the CNN model. Besides varying the number of epochs and batch size, the experiments with CNN for this research were carried out with variations in the kernel size and the number of classification, convolutional, and max pooling layers. Pooling is another adjustment mechanism that seeks to eliminate noise and make only meaningful information pass through the neural network.

This research made experiments also with a RNN model. RNNs are neural networks that incorporate cycles/loops to enable past information to persist within the network and influence current information. Essentially, this type of network can be seen as numerous replicas of the same neural network. However, each copy transmits a different message to the following network (Olah 2015).

The batch size and the number of epochs were the only characteristics that varied for the tests with the RNN models.

Finally, tests were carried out with AE models. This kind of deep learning model conducts two operations on the data: encoding and decoding. The encoding phase compresses the input data, thereby reducing its dimensionality to a single dimension. Conversely, the decoding phase aims to reconstruct the original input data by decompressing it. Throughout the reduction and decompression process, the goal is to retain as much significant information as possible while minimizing error and noise. As an unsupervised learning model, it is frequently employed as the foundation for generative models (San Martin et al. 2019). Experiments with the variation of kernel size and epochs parameters were made for this research.

4. Dataset description

In view of this, six datasets related to the rotating machine were selected to this research. The two main selection criteria were: being publicly available and being commonly cited in the literature. More than one database was chosen to encompass diverse data types, thereby producing impartial and authentic comparative outcomes. A brief description of the databases under study is given below.

The JNU datasets (JNU, 2019), provided by Jiangnan University, consist of three vibration datasets of bearings with different rotational speeds, and the data were collected at 50 kHz. They contain a healthy (healthy) state and three failure modes which include inner ring failure, outer ring failure, and rolling/rolling element failure. Therefore, the total number of classes was equal to twelve according to different working conditions.

The datasets from the CWRU (Case Western Reserve University. Bearing Data Center Websitesite 2023) were sourced from the Bearing Data Center at the university. The data of vibration were obtained from normal bearings, bearings afflicted with single-point defects (single-point drive), and fan end defects (fan end defects) caused in the laboratory. The experiments were conducted under four varying engine loads (engine rotation speed in rpm).

The University of Connecticut (UoC) (Cao 2018) gear failure datasets were provided by the university itself. The data were classified into nine categories (one healthy state and eight failure states) to test performance.

The SEU ("Mechanical-Datasets" 2018) gearbox datasets were provided by Southeast University. They contain two subsets of data, including a bearing dataset and a gearing dataset, which were both acquired in the simulator.

The Machinery Failure Prevention Technology (MFPT) (Bechhoefer 2012) dataset was provided by the Society for Machinery Failure Prevention Technology. They consist of three sets of bearing data that have been classified into 15 categories (one health state and 14 failure states) according to different loads (which influence the mode of operation).

And lastly, the Paderborn University (PU) (Universitat Paderborn 2016) datasets were provided by the Paderborn University Bearing Data Center. They consist of 32 sets of current and vibration signals. The motor current signal from an electromechanical drive system is used for bearing diagnostics.

5. Results

As mentioned before, the data from the six databases under study were preprocessed using time domain, FFT, STFT, CWT, and slicing. Then, one hundred tests were executed to each dataset. Each model (MLP, CNN, RNN, and AE) went through five test categories, and for each category associated with a preprocessing type, five variations in specific parameters (such as the number of layers, epochs, batch size, etc.) were tested five times.

The metric chosen to evaluate the performance of each model given the variations in parameters and database was accuracy. It is obtained by dividing the number of correct predictions by the total number of predictions and represents the ability of a trained model to predict results for new data. The accuracy of the models will be analysed first individually for each database and then in a more general way.

For the JNU database, the model that showed the highest accuracy, among all the one hundred variations tested, was the MLP. One point that attention should draw to how good the results referring to the preprocessing done in FFT and STFT are in relation to the results generated from other forms of preprocessing, even for different models. In Fig. 3 we can compare each type of preprocessing through the highest accuracies generated using them in MLP. Another interesting aspect is the architecture that presents the best result is in this scenario (MLP with data imputed in the frequency domain) and has a variation of up to 3% of accuracy with different function parameters (number of layers, epochs, and batch size).

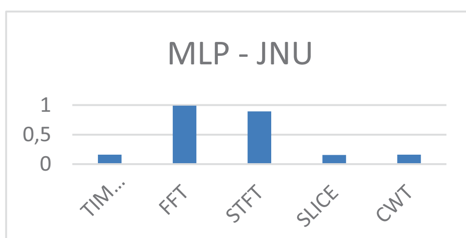

Fig. 3. Analyzing the preprocess methods on MLP in JNU

It is impressive how well the RNN model fits with the CWRU database, regardless of the type of pre-processing as illustrated in Fig. 4. For comparison purposes, the highest accuracy values for the RNN were always close to 99% while for the MLP the maximum accuracy was 87%. For the other models, the pre-processing using STFT presents the best results.

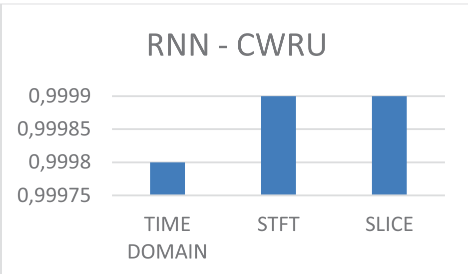

Fig. 4. Analyzing the preprocess methods on RNN in CWRU.

The UoC is a more complex database, where each parameter makes a difference in the accuracy of the models. It is highlighted how much the pre-processing was able to change the result of the models as shown in Fig. 5 and Fig. 6 respectively. Keeping the same scenario of parameters, in the MLP the accuracy was 92% for pre-processed data in the frequency domain to 11% in the time domain. In the RNN, the accuracy was 92% for pre-processed data in the time domain and 12% in the time-frequency domain in the same scenario of parameters. In other words, there is no constancy with the resulting values of accuracy of the models to assert which type of architecture best fits the UoC dataset. And this may be linked to the complexity of the data.

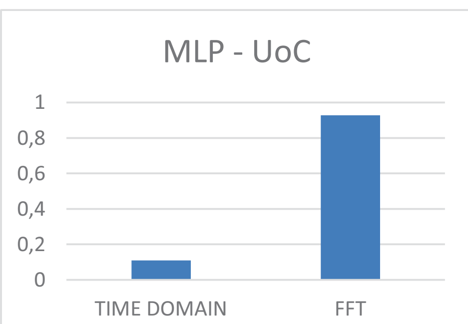

Fig. 5. Analyzing the preprocess methods on MLP in UoC.

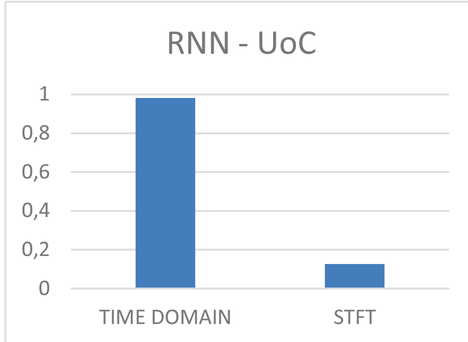

Fig. 6. Analyzing the preprocess methods on RNN in UoC.

The SEU database also showed low accuracy results in general. The lowest accuracy observed was only 4% for preprocessed data in the STFT domain when utilizing RNN, and even lower at 2% for preprocessed data in the time domain or when using slice images with MLP or CNN. SEU can also be considered a complex database.

Differently, the MFPT database presented a consistent average and high-accuracy results for most models. Its worst model was the AE but still, the accuracy values were on average 29% regardless of the type of pre-processing.

The last analyzed database was the PU. Most of the models presented considerably satisfactory results. The lowest results among the one hundred tests performed with the different types of variations came from the RNN models, reaching an accuracy of 17%.

Observing the databases under study as one, a phenomenon is noticed. The accuracy of the models is adjusted as some parameters of basic functions of deep learning models are changed, such as the number of layers or the size of the batch, for example. However, even changing these parameters, the difference in the accuracy value of each of the architectures with different parameters does not reach 10%. In other words, changing the parameters of basic functions of deep learning models is not a sufficient factor to make one model better than another.

However, the processing time is different (see Table 1). Models with more layers or larger batches tend to take longer to train and test data, consuming more computational resources. This can be a limiting factor for choosing the most appropriate model for each problem. Therefore, in addition to accuracy, processing time must also be considered as an evaluation criterion for deep learning models.

Table 1. Analyzing the time on MLP model in MFPT

	MFPT - MLP			
	1	2	3	4
Preprocess type	Feature extraction	Feature extraction	FFT	FFT
Number of layers	4	6	6	4
Batch	128	64	64	64
Epochs	150	100	100	100
Test Accuracy	66.83%	68.23%	75.28%	66.42%
Time used	1 min	12min	15 min	3 min

Furthermore, it is evident how much the test results reinforce the importance that adequate data pre-processing influences the accuracy of a model. Therefore, seeking to have minimum knowledge about the data and how they were extracted to choose a pre-processing method is a fundamental strategy that affects the performance of a model and determines how assertive it will be in making predictions.

In general, the RNN and CNN models, as well as the data that were pre-processed using FFT and STFT, were the ones that presented the highest accuracy values among the four models tested in the six databases under study. A likely explanation for this is that the worked failure data, mostly obtained from accelerated wear experiments, make sense in sequence, and must have a pattern before failure.

6. Conclusion

In conclusion, this research presented a study on the application of deep learning models for intelligent diagnosis of failures in rotating machines, which is a key component of Prognostics and Health Management (PHM). PHM is a valuable approach for ensuring the reliability and availability of machines by monitoring their health condition and predicting their remaining useful life. To achieve this, intelligent diagnosis models use data from sensors and other sources to detect and classify faults in the machines. However, there is no consensus on the best deep learning model or data preprocessing method for this task, as different authors use different architectures, parameters, and functions in their models.

Therefore, it was conducted a comparative analysis of four deep learning models (MLP, CNN, RNN, and AE) using a standardized architecture and six main databases related to failures in rotating machines. It also evaluated the impact of different data preprocessing methods and 5 types of input data (such as time domain, frequency domain, time-frequency domain, wavelet domain, and slicing images) on the performance of the models.

The results showed that RNN and CNN models outperformed the other models in terms of accuracy and that data preprocessed using FFT and STFT yielded better results than data. These findings can provide useful insights and guidance for researchers and practitioners who want to design effective deep learning models for intelligent diagnosis of failures in rotating machines as part of PHM.

References

Bechhoefer, E. (2012). "Fault Data Sets." 2012. https://www.mfpt.org/fault-data-sets/.

Cao, H., Shao, H., Zhong, X., Deng, Q., Yang, X., & Xuan, J. (2022). Unsupervised domain-share CNN for machine fault transfer diagnosis from steady speeds to time-varying speeds. Journal of Manufacturing Systems, 62, 186-198.

Cao, P., Zhang, S., & Tang, J. (2018). "Gear Fault Data. Figshare. Dataset." 2018. https://figshare.com/articles/dataset/Gear_Fault_Data/6127874/1.

Case Western Reserve University. Bearing Data Center Websitesite. "CWRU PAGE." (202. https://engineering.case.edu/bearingdatacenter/download-data-file.

Chen, Z., Mauricio, A., Li, W., & Gryllias, K. (2020). A deep learning method for bearing fault diagnosis based on cyclic spectral coherence and convolutional neural networks. Mechanical Systems and Signal Processing, 140, 106683.

Chen, Z., & Li, W. (2017). Multisensor feature fusion for bearing fault diagnosis using sparse autoencoder and deep belief network. IEEE Transactions on Instrumentation and Measurement, 66(7), 1693-1702.

Cheng, Y., Lin, M., Wu, J., Zhu, H., & Shao, X. (2021). Intelligent fault diagnosis of rotating machinery based on continuous wavelet transform-local binary convolutional neural network. Knowledge-Based Systems, 216, 106796.

Fink, O., Wang, Q., Svensén, M. Dersin, P., Wan Jui Lee, W. J., & Ducoffe, M. (2020). "Potential, Challenges and Future Directions for Deep

Learning in Prognostics and Health Management Applications." Engineering Applications of Artificial Intelligence 92 (June). https://doi.org/10.1016/j.engappai.2020.103678.

Samadi, H. Ghobadian, S. B., Nosrati, M., & Rezaei. M. (2023). "Investigation of Factors Affecting Performance of a Downdraft Fixed Bed Gasifier Using Optimized MLP Neural Networks Approach." Fuel 333 (February). https://doi.org/10.1016/j.fuel.2022.126249.

Aiwina, H., Zhang ,S., Tan, A.C.C. , & Mathew, J.. (2009). "Rotating Machinery Prognostics: State of the Art, Challenges and Opportunities." Mechanical Systems and Signal Processing. https://doi.org/10.1016/j.ymssp.2008.06.009.

Jia, F., Lei, Y., Lin, J., Zhou, X., & Lu, N. (2016). Deep neural networks: A promising tool for fault characteristic mining and intelligent diagnosis of rotating machinery with massive data. Mechanical systems and signal processing, 72, 303-315.

Li, X., Zhang, W., Ding, Q., & Sun, J. Q. (2020). Intelligent rotating machinery fault diagnosis based on deep learning using data augmentation. Journal of Intelligent Manufacturing, 31, 433-452.

"Mechanical-Datasets." (2018). https://github.com/cathysiyu/Mechanical-datasets.

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., et al (2021). The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ, 71, doi: 10.1136/bmj.n71

Christopher, O. (2015). "Understanding LSTM Networks." August 27, 2015. https://colah.github.io/posts/2015-08-Understanding-LSTMs/.

Rezaeianjouybari, Behnoush, & Shang, Y. (2020). "Deep Learning for Prognostics and Health Management: State of the Art, Challenges, and Opportunities." Measurement: Journal of the International Measurement Confederation 163 (October). https://doi.org/10.1016/j.measurement.2020.107929.

San Martin, G., Droguett, E. L., Meruane, V. & Moura, M. C. (2019). "Deep Variational Auto-Encoders: A Promising Tool for Dimensionality Reduction and Ball Bearing Elements Fault Diagnosis." Structural Health Monitoring 18 (4): 1092–1128. https://doi.org/10.1177/1475921718788299.

Shao, H., Jiang, H., Zhao, H., & Wang, F. (2017). A novel deep autoencoder feature learning method for rotating machinery fault diagnosis. Mechanical Systems and Signal Processing, 95, 187-204.

Shao, H., Jiang, H., Lin, Y., & Li, X. (2018). A novel method for intelligent fault diagnosis of rolling bearings using ensemble deep auto-encoders. Mechanical Systems and Signal Processing, 102, 278-297.

Sun, J., Yan, C., & Wen, J. (2017). Intelligent bearing fault diagnosis method combining compressed data acquisition and deep learning. IEEE Transactions on Instrumentation and Measurement, 67(1), 185-195.

Tang, S., Yuan, S., & Zhu, Y. (2019). Deep learning-based intelligent fault diagnosis methods toward rotating machinery. Ieee Access, 8, 9335-9346.

Tao, H., Wang, P., Chen, Y., Stojanovic, V., & Yang, H. (2020). An unsupervised fault diagnosis method for rolling bearing using STFT and generative neural networks. Journal of the Franklin Institute, 357(11), 7286-7307.

Universitat Paderborn. (2016). "Bearing DataCenter." 2016. https://mb.uni-paderborn.de/kat/forschung/datacenter/bearing-datacenter/.

Wu, X., Zhang, Y., Cheng, C., & Peng, Z. (2021). A hybrid classification autoencoder for semi-supervised fault diagnosis in rotating machinery. Mechanical Systems and Signal Processing, 149, 107327.

Yang, D., Karimi, H. R., & Sun, K. (2021). Residual wide-kernel deep convolutional auto-encoder for intelligent rotating machinery fault diagnosis with limited samples. Neural Networks, 141, 133-144.

Zhang, W., Li, X., & Ding, Q. (2019). Deep residual learning-based fault diagnosis method for rotating machinery. ISA transactions, 95, 295-305.

Zhao, B., Zhang, X., Li, H., & Yang, Z. (2020). Intelligent fault diagnosis of rolling bearings based on normalized CNN considering data imbalance and variable working conditions. Knowledge-Based Systems, 199, 105971.