

# Automatic Extraction of a Health Indicator from Vibrational Data by Sparse Autoencoders

Zhe Yang<sup>1</sup>, Piero Baraldi<sup>1\*</sup>, Enrico Zio<sup>1,2,3,4</sup>

<sup>1</sup>Energy Department, Politecnico di Milano, Via La Masa 34, 20156 Milan, Italy

<sup>2</sup>MINES ParisTech, PSL Research University, CRC, Sophia Antipolis, France

<sup>3</sup>Aramis Srl, Via pergolesi 5, Milano, Italy

<sup>4</sup>Eminent Scholar, Department of Nuclear Engineering, College of Engineering, Kyung Hee University  
Republic of Korea

e-mail: zhe.yang@polimi.it, piero.baraldi@polimi.it, enrico.zio@polimi.it

**Abstract**—We present a method for automatically extracting a health indicator of an industrial component from a set of signals measured during operation. Differently from traditional feature extraction and selection methods, which are labor-intensive and based on expert knowledge, the method proposed is automatic and completely unsupervised. Run-to-failure data collected during the component life are fed to a Sparse AutoEncoder (SAE), and the various features extracted from the hidden layer are evaluated to identify those providing the most accurate quantification of the component degradation. The method is applied to a synthetic and a bearing vibration dataset. The results show that the developed SAE-based method is able to automatically extract an efficient health indicator.

*Keywords*-feature extraction; health indicator; sparse autoencoder; deep learning; vibration data

## I. INTRODUCTION

The identification of the degradation state of an industrial component is fundamental for the development of condition-based maintenance. Conventional methods for the definition of health indicators from measured signals are typically based on handcrafted feature design and selection processes, which require the input of experts and the use of computationally intensive trial and error approaches [1-3].

Specifically, for identifying the degradation state of bearings from the monitoring signals, features are extracted from time-domain, frequency-domain, time-frequency domain or by using other advanced signal processing techniques [4, 5]. A novel Sparse Wavelet Reconstruction Residual (SWRR) feature, based on the use of Wavelet Packet Transform (WPT) and sparse representation theory, has been defined for rolling bearing diagnostics in [6]. Features extracted using wavelet and other time-frequency transforms have been used for the identification of the bearing degradation state in [7] and [8]. Greedy search methods are applied to the feature selection problem in [9]. A wrapper feature selection approach, which integrates a Binary Differential Evolution algorithm with the K-Nearest Neighbor classifiers, has been used for assessing bearing degradation from measured vibrational signals in [10]. In [11], a degradation indicator of slewing bearing was obtained

by applying the Principal Component Analysis (PCA) to fuse features extracted using statistical theory, wavelet and EMD from the vibration signal. A monotonic feature selection method based on rank mutual information was proposed in [12] to select degradation indicator of bearing. Features extracted using EMD were used for bearing performance degradation assessment in [13].

The above-mentioned methods highly rely on experts to design handcrafted features. However, the recent development of new and cheap types of sensors and storage devices has resulted in an exponential growth of the available monitoring data [14], which renders traditional feature extraction and selection methods difficult to apply. In this work, we present a method derived from deep learning, for the automatic definition of a health index from high-dimensional data.

Deep learning aims at extracting hierarchical representations from input data by building deep neural networks with multiple layers of non-linear transformations [15]. It has been successfully applied to various areas such as computer vision [16], automatic speech recognition [17] and natural language processing [18]. One of the most popular deep learning architecture is AutoEncoder (AE). AE is a neural network with symmetrical structure, composed of an “encoder” and a “decoder” network. The multilayer “encoder” network transforms the high-dimensional data into low-dimensional features and the “decoder” network recovers the data from the features [16]. Recently, in [19-21] AEs are used as a pre-training method to build supervised Deep Neural Networks (DNNs) for fault diagnostics, without exploiting their capability of extracting health indicators. In [22], an Extreme Learning Machine (ELM)-AE method is proposed to automatically extract degradation features; although the results obtained in the specific application are satisfactory, the authors do not explain the conditions in which the extracted features are health indicators.

In this work, we propose a novel method for the extraction of a health indicator when no information is available on the component degradation state, based on the use of SAE. Since degradation can be considered as a “monotonic” process, a SAE is trained using a run-to-failure trajectory of the monitoring data, and the most monotonic of

the extracted features, which is identified by performing the Mann-Kendall test, is selected as a health indicator.

## II. SPARSE AUTOENCODER AND THE PROPOSED FEATURE EXTRACTION METHOD

### A. Sparse Autoencoder

An AE is a neural network composed of an encoder and a decoder, trained to replicate its input as output [16] (Fig. 1b). Let  $\mathbf{X} = \{\mathbf{x}_i, i = 1, \dots, N\}$  be an unlabeled dataset containing a run-to-failure degradation trajectory for which the measurements of  $K$  signals at  $N$  time instants are available. The encoder aims at mapping the observation vector  $\mathbf{x}_i$  to a hidden representation,  $\mathbf{z}_i \in \mathbb{R}^{K_F}$ , whereas the decoder reconstructs the input data from the extracted features  $\mathbf{z}_i$ . A SAE is a variant of an AE, which encourages the extraction of discriminative features by adding a regularization term to the cost function [23], used to set the internal weights, to constrain hidden neurons to be inactive most of the times, i.e. by requiring the sparsity of the AE.

SAEs are typically organized in multilayered ‘deep’ architectures (Fig. 1), whose training process is based on the following two steps: a layer-wise pretraining and a fine-tuning. Pretraining consists of learning a stack of basic AEs, each one having only a single hidden layer of features. For example, the first basic AE (bottom in Fig. 1a) is trained to transform input  $\mathbf{x}_i$  to feature  $\mathbf{z}_i^{(1)} = [z_1^{(1)}(i), z_2^{(1)}(i), \dots, z_{K_1}^{(1)}(i)]$ :

$$\mathbf{z}_i^{(1)} = f(\mathbf{W}_1 \mathbf{x}_i + \mathbf{b}_1) \quad (1)$$

where  $f$ ,  $\mathbf{W}_1$ ,  $\mathbf{b}_1$  are the encoder activation function, weight matrix and bias vector, respectively. Once the features  $\mathbf{z}_i^{(1)}$  have been obtained, they are used as ‘data’ for training the next basic SAE which transforms  $\mathbf{z}_i^{(1)}$  to the feature vector  $\mathbf{z}_i^{(2)}$ . The procedure is repeated until the most inner basic SAE is pretrained. After pretraining, basic SAEs are ‘stacked’ to create a deep autoencoder, which is, then, fine-tuned using the backpropagation algorithm [24].

The training of a generic basic SAE aims at minimizing the cost function:

$$C = R_{error} + \beta R_{sparse} + \lambda R_{L2} \quad (2)$$

where  $R_{error}$ ,  $R_{sparse}$  and  $R_{L2}$  are the reconstruction error, the sparsity regularization and the  $L_2$  regularization terms, respectively,  $\beta$  and  $\lambda$  are coefficients indicating the relative importance of the terms in the cost function. The reconstruction error,  $R_{error}$ , quantifies the similarity of the reconstruction  $\hat{\mathbf{x}}_i$  provided by the decoder with the input vector  $\mathbf{x}_i$ :

$$R_{error} = \frac{1}{N} \sum_{i=1}^N \|\mathbf{x}_i - \hat{\mathbf{x}}_i\|^2 \quad (3)$$

It has been shown that the extraction of discriminative features  $z_j^{(1)}(i), j = 1, 2, \dots, K_1$ , is favored when the hidden neurons are constrained to be inactive most of the times, i.e. by requiring the sparsity of the AE [25]. To this aim, a sparse regularization term  $R_{sparse}$  is added to the cost function (2). Let  $\hat{p}_j$  be the average activation of the  $j$ -th neuron of the SAE hidden layer evaluated over all the training patterns  $\mathbf{x}_i, i = 1, \dots, N$ :

$$\hat{p}_j = \frac{1}{N} \sum_{i=1}^N z_j^{(1)}(i), j = 1, \dots, K_1 \quad (4)$$

where  $z_j^{(1)}(i)$  is the  $j$ -th element of the extracted feature vector  $\mathbf{z}_i^{(1)}, j = 1, 2, \dots, K_1$ . Since the sparsity condition requires that activation of the hidden neurons is mostly zero, most SAEs have small values of  $\hat{p}_j$ , e.g. 0.05, for all its neurons.  $R_{sparse}$  is, then, calculated using the Kullback-Leibler (KL) divergence function to measure the sparsity conditions, i.e. whether  $\hat{p}_j$  is close to a desired sparsity proportion  $p$ :

$$\begin{aligned} R_{sparse} &= \sum_{j=1}^{K_1} KL(p \parallel \hat{p}_j) \\ &= \sum_{j=1}^{K_1} \left[ p \log \frac{p}{\hat{p}_j} + (1-p) \log \frac{1-p}{1-\hat{p}_j} \right] \end{aligned} \quad (5)$$

Therefore, the KL function is zero when all  $\hat{p}_j$  are equal to  $p$  and increases when they diverge.

The regularization term,  $R_{L2}$ , is introduced in (2) to prevent that  $R_{sparse}$  becomes small during the training process, due to the increase of the weight values [26]; it is defined by:

$$R_{L2} = \frac{1}{2} \|\mathbf{W}\|^2 \quad (6)$$

where  $\mathbf{W}$  is the SAE weight matrix.

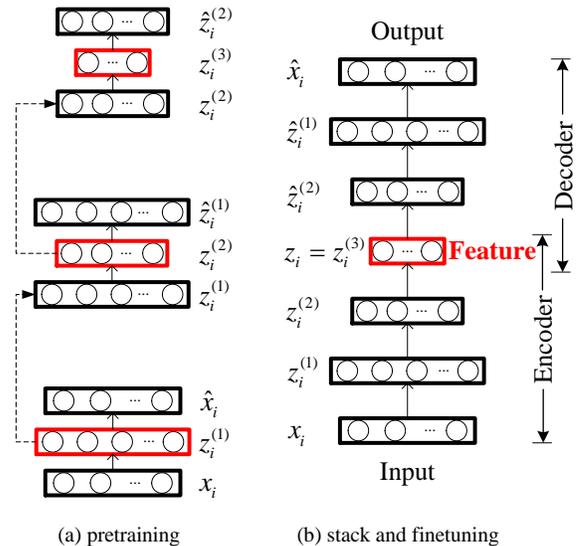


Figure 1. Layer-wise training of SAE

### B. The Proposed Method for Automatic Extraction of a Health Indicator

The signals measured in an industrial system typically contain information on different characteristics such as operational, environmental, degradation and failure conditions of its components. In this work, we assume to have available signal measurements collected during at least one run-to-failure degradation trajectory and that the component degradation is somehow influencing the measured signals, i.e. signal values are functions of various factors, among which there is the component degradation level.

The SAE is trained by using a run-to-failure trajectory of monitored data. The features extracted from the most inner of the SAE layers are, then, automatically evaluated to assess their capability of describing the component degradation. Being the degradation process irreversible, we select as health indicator the most monotonic of the extracted features, which is identified by performing the Mann-Kendall (M-K) monotonicity test [27] and computing the coefficient  $\rho_{MK}$ :

$$\rho_{MK} = \begin{cases} \frac{V-1}{\sigma} & \text{if } V > 0 \\ 0 & \text{if } V = 0 \\ \frac{V+1}{\sigma} & \text{if } V < 0 \end{cases} \quad (7)$$

where  $V = \sum_{i=1}^{N-1} \sum_{j=i+1}^N \text{sgn}(F(j) - F(i))$ ,  $\sigma = \sqrt{N(N-1)(2N+5)/18}$ ,  $N$  is the number of training patterns and  $\text{sgn}(\cdot)$  is the sign function: the larger is the absolute value of  $\rho_{MK}$ , the more monotonic is the feature.

### III. SYNTHETIC CASE STUDY

We consider a component characterized by a linear degradation process described by the health indicator  $D(t)$

$$D(t) = \alpha t, t = 1, 2, \dots, 200 \quad (8)$$

where  $\alpha = 0.025$  is the intrinsic degradation rate.

In order to mimic the complexity of an industrial real case, we assume that  $D(t)$  is not directly measurable, whereas there are 10 signals measured by sensors whose behaviours are influenced by the component degradation, the operational and environmental conditions, and a process noise. The operational and environmental conditions are represented by the factor:

$$C(t) = \sin\left(\frac{1}{25}\pi t\right) + w_c, w_c \sim N(\mu = 0, \sigma = 0.2) \quad (9)$$

where the periodic behavior describes the trend of the seasonal effects and the noise  $w_c$ , distributed according to a gaussian probability density function  $N(\mu = 0, \sigma = 0.2)$ , describes the stochasticity of the environmental changes. The process noise is represented by:

$$w(t) \sim N(0, 0.1) \quad (10)$$

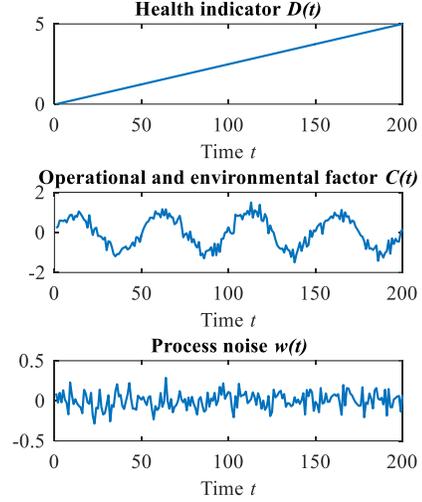


Figure 2. Evolution of the three factors used for generating the 10 signals of the synthetic dataset

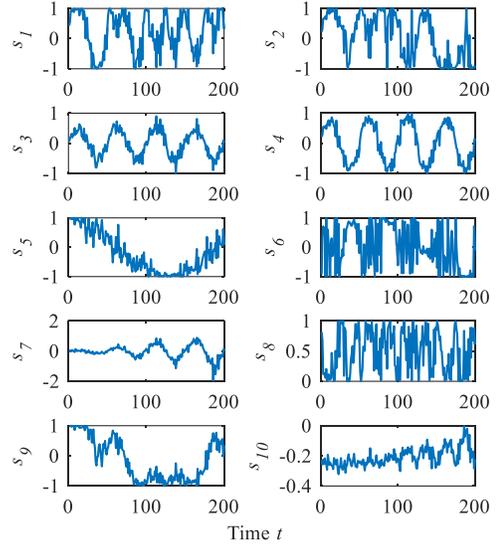


Figure 3. Time evolution of the 10 simulated signals during the considered run-to-failure trajectory

Fig. 2 shows the simulated evolution of the health indicator,  $D(t)$ , the operational and environmental factor,  $C(t)$ , and the process noise,  $w(t)$ , during a degradation trajectory.

The 10 measured signals  $s_i(t), i = 1, 2, \dots, 10$  are, then, obtained by combining at least 2 of the 3 factors  $D(t)$ ,  $C(t)$  and  $w(t)$  by using (11).

The simulated signal evolutions during the run-to-failure trajectory generated by  $D(t)$ ,  $C(t)$  and  $w(t)$  are shown in Fig. 3. The 10 signals  $s_1(t), s_2(t), \dots, s_{10}(t)$  are fed to a 4-hidden-layer SAE, formed by an encoder with layers of size 10-100-50-20-3 and a symmetric decoder. SAE hyper-parameters are set following a trial and error process with the objective of maximizing the monotonicity of the most monotonic feature extracted by the 3 nodes  $z_1^{(4)}, z_2^{(4)}, z_3^{(4)}$  of

the most inner hidden layer. The identified best performing hyperparameters are:  $p = 7e - 2$ ,  $\beta = 1$ ,  $\lambda = 5e - 8$  and sigmoid activation function. Fig. 4 shows the three extracted features  $z_1^{(4)}(t)$ ,  $z_2^{(4)}(t)$  and  $z_3^{(4)}(t)$ . Notice that  $z_3^{(4)}(t)$  is characterized by the largest M-K coefficient ( $|\rho_{MK}| = 15.7612$ ) and shows a monotonic behavior. Therefore,  $z_3^{(4)}(t)$  can be considered a satisfactory health indicator extracted from the data.

$$\begin{aligned}
s_1(t) &= \sin(1/2 D(t) + 2C(t)) \\
s_2(t) &= \sin(D(t) + \tan(C(t))) \\
s_3(t) &= \tan(1/2 C(t) - 2/3 w(t)) \\
s_4(t) &= \cos^3(w(t))\sin(C(t)) \\
s_5(t) &= \cos(D(t) - 3\tan(w(t))) \\
s_6(t) &= \sin(D(t) + w(t) - \tan^2(C(t) + 1)) \\
s_7(t) &= \tan(1/6 D(t)C_t) + 1/2 w(t) \\
s_8(t) &= \cos^2(7/10 D(t) + 2C(t)) \\
s_9(t) &= \cos(D(t) + 2 \tan(C(t)) \\
&\quad - \tan(1/2 w(t))) \\
s_{10}(t) &= \tan(1/4 w(t)) \\
&\quad - 1/4\cos(/3 D(t)\sin(C(t)))
\end{aligned} \tag{11}$$

It is also interesting to notice that  $z_2^{(4)}(t)$  with  $t \in [1,50]$  and  $z_1^{(4)}(t)$  with  $t \in [51,200]$  show a trend correlated to the operational and environmental factor  $C(t)$ .

#### IV. REAL CASE STUDY

We consider an experimental dataset taken from the Prognostics Data Repository of NASA [28] containing run-to-failure bearing vibration data collected from a bearing test rig, provided by the NSF I/UCR Center for Intelligent Maintenance Systems (IMS), University of Cincinnati. During this experiment, the rotation speed is kept constant at 2000 rpm, the vibration signal acceleration along the x-axis is acquired at 20KHz with a snapshot of 20,480 samples collected every 10 min. More details on the dataset and on the experiment design can be found in [29].

The proposed method for health indicator extraction is applied to the run-to-failure trajectory of the third bearing collected in the first experiment. The bearing fails after 35 days of operations during which 2156 snapshots of data have been collected.

Given the high dimensionality of the raw data (20480 values collected in each one of the 2156 snapshots originate  $20480 \cdot 2156 = 44,154,880$  patterns), the dataset has been pretreated by applying the Morlet 6 continuous wavelet transform [30], and by considering as features the time average of the 1333 frequency amplitudes obtained. This data preprocessing has allowed reducing the dimensionality of each snapshot from 20,480 raw acceleration values to 1333 average frequency amplitudes.

A 4-hidden-layer SAE, formed by an encoder with layers of size 1333-1500-700-200-5 and a symmetric decoder, has been developed for the feature extraction task. The identified best performing hyperparameters are:  $p = 1e - 1$ ,  $\beta = 1$ ,  $\lambda = 1e - 5$  and sigmoid activation function. Table I shows the coefficients  $|\rho_{MK}|$  of the 5 extracted features obtained by applying the M-K test. Among the extracted features, the most monotonic feature is shown in Fig. 5, characterized by a clear monotonically increasing degradation trend from the 2000<sup>th</sup> snapshot. Therefore, it can be considered a satisfactory health indicator of the bearing degradation.

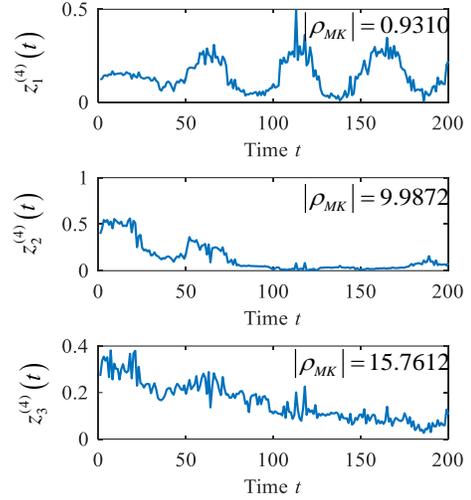


Figure 4. Features extracted from synthetic data

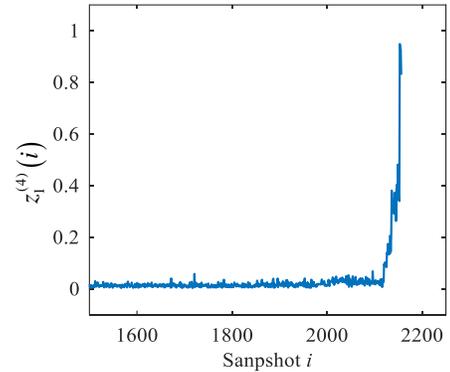


Figure 5. Trends of the extracted health indicator,  $z_1^{(4)}(i)$ , in the last part of the degradation trajectory

TABLE I. M-K COEFFICIENT  $|\rho_{MK}|$  OF THE EXTRACTED FEATURES

Feature	$ \rho_{MK} $
$z_1^{(4)}(i)$	27.7
$z_2^{(4)}(i)$	21.75
$z_3^{(4)}(i)$	9.62
$z_4^{(4)}(i)$	13.63
$z_5^{(4)}(i)$	11.18

## V. CONCLUSION

We have developed a method for the automatic extraction of a health indicator from monitoring data collected from a degrading component. The proposed method is based on 1) the application of a sparse autoencoder to run-to-failure data; 2) the application of the Mann-Kendall test for the selection of the most inner hidden layer of the SAE. The method has been applied to a synthetic case study and to real bearing vibrational data. The obtained results show that the proposed method can automatically extract satisfactory health indicators from different types of signals, without requiring the use of expert knowledge for feature extraction. The method can be very useful in PHM practice because it relieves the analyst from the difficult task of building a health indicator by expert analysis.

## ACKNOWLEDGMENT

Zhe Yang gratefully acknowledges the financial support from the China Scholarship Council (No. 201506280015). The participation of Enrico Zio in this research is partially supported by the China NSFC under grant number 71231001.

## REFERENCES

- [1] S. Das. Filters, wrappers and a boosting-based hybrid for feature selection. in *Icml*. 2001.
- [2] J. Zhao, X. Zhang, H. Li, X. Ni, and Z. Du. Health indicator selection and health assessment of rolling element bearing. in *Modelling, Identification and Control (ICMIC), 2017 9th International Conference on*. 2017. IEEE.
- [3] L. Liao. "Discovering prognostic features using genetic programming in remaining useful life prediction." *IEEE Transactions on Industrial Electronics*, 2014. 61(5): pp. 2464-2472.
- [4] R. B. Randall and J. Antoni, "Rolling element bearing diagnostics—A tutorial." *Mechanical systems and signal processing*, 2011. 25(2): pp. 485-520.
- [5] A. Rai and S. Upadhyay, "A review on signal processing techniques utilized in the fault diagnosis of rolling element bearings." *Tribology International*, 2016. 96: pp. 289-306.
- [6] A. Mosallam, K. Medjaher, and N. Zerhouni, "Data-driven prognostic method based on Bayesian approaches for direct remaining useful life prediction." *Journal of Intelligent Manufacturing*, 2016. 27(5): pp. 1037-1048.
- [7] R. Yan, R. X. Gao, and X. Chen, "Wavelets for fault diagnosis of rotary machines: A review with applications." *Signal processing*, 2014. 96: pp. 1-15.
- [8] Z. Feng, M. Liang, and F. Chu, "Recent advances in time–frequency analysis methods for machinery fault diagnosis: A review with application examples." *Mechanical Systems and Signal Processing*, 2013. 38(1): pp. 165-205.
- [9] T. W. Rauber, F. de Assis Boldt, and F. M. Varejão, "Heterogeneous feature models and feature selection applied to bearing fault diagnosis." *IEEE Transactions on Industrial Electronics*, 2015. 62(1): pp. 637-646.
- [10] P. Baraldi, F. Cannarile, F. Di Maio, and E. Zio, "Hierarchical k-nearest neighbours classification and binary differential evolution for fault diagnostics of automotive bearings operating under variable conditions." *Engineering Applications of Artificial Intelligence*, 2016. 56: pp. 1-13.
- [11] C. Lu, J. Chen, R. Hong, Y. Feng, and Y. Li, "Degradation trend estimation of slewing bearing based on LSSVM model." *Mechanical Systems and Signal Processing*, 2016. 76: pp. 353-366.
- [12] G. Niu, F. Qian, and B.-K. Choi, "Bearing life prognosis based on monotonic feature selection and similarity modeling." *Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science*, 2016. 230(18): pp. 3183-3193.
- [13] A. Rai and S. Upadhyay, "Bearing performance degradation assessment based on a combination of empirical mode decomposition and k-medoids clustering." *Mechanical Systems and Signal Processing*, 2017. 93: pp. 16-29.
- [14] J. Tang, S. Alelyani, and H. Liu, "Feature selection for classification: A review." *Data Classification: Algorithms and Applications*, 2014: pp. 37.
- [15] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning." *Nature*, 2015. 521(7553): pp. 436-444.
- [16] G. E. Hinton and R. R. Salakhutdinov, "Reducing the dimensionality of data with neural networks." *science*, 2006. 313(5786): pp. 504-507.
- [17] L. Deng, M. L. Seltzer, D. Yu, A. Acero, A.-r. Mohamed, and G. Hinton. Binary coding of speech spectrograms using a deep auto-encoder. in *Eleventh Annual Conference of the International Speech Communication Association*. 2010.
- [18] R. Collobert, J. Weston, L. Bottou, M. Karlen, K. Kavukcuoglu, and P. Kuksa, "Natural language processing (almost) from scratch." *Journal of Machine Learning Research*, 2011. 12(Aug): pp. 2493-2537.
- [19] F. Jia, Y. Lei, J. Lin, X. Zhou, and N. Lu, "Deep neural networks: A promising tool for fault characteristic mining and intelligent diagnosis of rotating machinery with massive data." *Mechanical Systems and Signal Processing*, 2016. 72: pp. 303-315.
- [20] C. Lu, Z.-Y. Wang, W.-L. Qin, and J. Ma, "Fault diagnosis of rotary machinery components using a stacked denoising autoencoder-based health state identification." *Signal Processing*, 2017. 130: pp. 377-388.
- [21] W. Sun, S. Shao, R. Zhao, R. Yan, X. Zhang, and X. Chen, "A sparse auto-encoder-based deep neural network approach for induction motor faults classification." *Measurement*, 2016. 89: pp. 171-178.
- [22] Y. Hu, T. Palmé, and O. Fink. Deep health indicator extraction: A method based on auto-encoders and extreme learning machines. in *Annual Conference of the Prognostics and Health Management Society*. 2016.
- [23] Q. V. Le. Building high-level features using large scale unsupervised learning. in *Acoustics, Speech and Signal Processing (ICASSP), 2013 IEEE International Conference on*. 2013. IEEE.
- [24] D. E. Rumelhart, G. E. Hinton, and R. J. Williams, "Learning representations by back-propagating errors." *nature*, 1986. 323(6088): pp. 533.
- [25] N. Andrew, "Sparse autoencoder." *CS294A Lecture notes*, 2011. 72.
- [26] B. A. Olshausen and D. J. Field, "Sparse coding with an overcomplete basis set: A strategy employed by V1?" *Vision research*, 1997. 37(23): pp. 3311-3325.
- [27] T. Pohlert, "Non-parametric trend tests and change-point detection." *CC BY-ND*, 2016. 4.
- [28] A. Saxena, K. Goebel, C. Larrosa, and F. Chang, "CFRP Composites dataset, NASA Ames Prognostics Data Repository". 2015.
- [29] H. Qiu, J. Lee, J. Lin, and G. Yu, "Wavelet filter-based weak signature detection method and its application on rolling element bearing prognostics." *Journal of sound and vibration*, 2006. 289(4): pp. 1066-1090.
- [30] C. Torrence and G. P. Compo, "A practical guide to wavelet analysis." *Bulletin of the American Meteorological society*, 1998. 79(1): pp. 61-78.