Advances in Reliability, Safety and Security

ESREL 2024
Collection of Extended Abstracts

Deep Learning For Fault Classification Using Digital Twin Centrifugal Pump

Davi Souza, Eduardo Menezes, Isis Lins, Márcio Moura

Universidade Federal de Pernambuco, Recife, Brazil

Keywords: autoencoder, simulation data, bearing fault, reliability

Rolling bearing are a common component in rotating systems, especially ones with centrifugal pumps used in Oil & Gas refineries. Its failure can cause massive damage or even stop production, and therefore its safety operation and reliability is very important. There are many ways to perform fault diagnosis on systems with rotating machinery, in specific rolling bearings (Haidong et al., 2018) using artificial intelligence models such as artificial and convolution neural network (ANN / CNN), support vector machine (SVM), autoencoders (AE) with deep or ensemble or even sparse methodologies.

All of these approaches work with big enough data sets so the algorithm can be trained with, using a major part of the set, and also have some remaining data to validate the procedure. For the particular case of rolling bearings, there is enough datasets available for some configurations. Another technology that's been used is the virtual representation of a physical model, called Digital Twin – DT (Rathore et al., 2021). By using a DT one can generate a dataset with faulty operations for a specific case like a O&G Naphtha Refinery, with the goal of evaluating reliability of the centrifugal pumps for the compound (Elbar et al., 2023).

In the present work the goal is to use a previous validated centrifugal pump DT (Souza et al., 2023) to generate data that can be used in a further analysis for fault detection and classification. The methodology accounts initially for a bearing fault mechanism via time series that after a Fast Fourier Transformation (FFT) will show in the frequency domain some characteristic frequencies that could be associated with either roller, inner race or outer race faults.

With the validated DT it is possible to generate fault datasets by varying specific parameters. A way to generate fault modes that will resemble a bearing fault is to add disturbances in specific frequency ranges around values that can be theoretically evaluated, using the bearing information (Hou et al., 2023). This was done and also a small random variation was applied to specific parameters that are related to the bearing and also impeller to replicate a physical degradation of the parts.

The dataset was generated composed of 70% healthy generated data that contains frequencies only on the rotor 34 Hz frequency and sub harmonics, and 10 % of the other 3 frequencies related to roller, inner race and outer race faults. The model was assembled in Simulink® within Matlab® coding. After the generation of the dataset, an Auto Encoder was used with 90% of the samples for training and the remaining for validation. An Adam solver was used in the AE. Some random samples from the original dataset were used in the AE and reconstructed. A random sample of this reconstructed data indicates a healthy condition and can be used for comparison with fault data where the different areas would indicate frequency values that can be traced to a specific fault, see Figure 1.

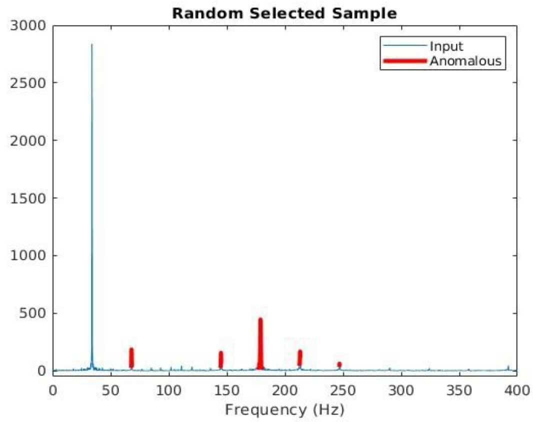

Fig. 1. Random sample anomaly detection using one AE.

Since the AE already captured the fault sample without categorizing it, a deep auto-encoder is used as a tool for this diagnostic (Haidong et al., 2018). For this case, a second AE will be used inside the hidden feature from the first AE and a third AE is used inside the hidden feature from the second AE. Also, different activation functions are used due to instability for some cases. For each tested activation function is calculated an acceptance threshold that will be used both as a criterion to use or not such function, and in the calculation of the weights. The weights are used in the predicted DAE as a combination for the ensemble. Another option is to use the generated data and use a Convolution Neural Network (CNN) for fault classification, which is well known for its accuracy in many tests (Sobie et al., 2018). The CNN is used here as complementary of the DEA formulation for performance comparison.

The result of the ensemble procedure is then used for the feature extraction and fault identification using trained samples. The validation is done with the test data set using all the different DAE combinations. After that the accuracy of each combination is evaluated and compared against each other in order to choose the rout within activation function, weights and DAEs. This rout is then used for fault diagnosis and applied to a random fault sample to see if the procedure identification corresponds to the correct fault. The fault classification from the CNN achieves more than 96% accuracy for the generated data and the confusion map shows no misclassification.

Acknowledgements

The authors wish to thank CNPq for support in this research.

References

Elbar, D., Elgazali, A., Haider, F., Genbril, M., Esmaael, R., Althalb, H., Abogtisha, A., Arciete, G. 2023. Failure Analysis of Multi stages Centrifugal Naphtha Pump. International Research Journal.

Haidong, S., Hongkai, J., Ying, L., Xingqiu, L. 2018. A novel method for intelligent fault diagnosis of rolling bearings using ensemble deep auto-encoders. Mechanical Systems and Signal Processing 102.

Hou, W. , Zhang, C., Jiang, Y., Cai, K., Wang, Y., Li, N. 2023. A new bearing fault diagnosis method via simulation data driving transfer learning without target fault data. Measurement.

Rathore, M. M., Shah, S. A., Shukla, D., Bentaft, E., Bakiras, S. 2021. The role of ai, machine learning, and big data in digital twinning: A systematic literature review, challenges, and opportunities. IEEE Access.

Sobie, C., Freitas, C., Nicolai, M. 2018. Simulation-driven machine learning: bearing fault classification. Mechanical Systems and Signal Processing.

Souza, D. B. Menezes, E. J. N., Lins, I. D., Moura, M. J. C. 2023. Design and validation of a digital twin for fault detection at O&G facility. ESREL.