Advances in Reliability, Safety and Security

ESREL 2024
Collection of Extended Abstracts

Method To Evaluate Similarity Of Completion Equipment Using Deep Learning

Davi Souza[a], Eduardo Menezes[a], Isis Lins[a], Márcio Moura[a], Feliciano da Silva[b], Marcos Nóbrega[b]

[a]*Federal University of Pernambuco, Recife, Brazil*
[b]*Petrobras, Rio de Janeiro, Brazil*

Keywords: autoencoder, similarity, equipment development, reliability

In the O&G industry, the estimation of reliability is usually an intricate task due to several reasons. First, the long useful life of the equipment makes it difficult to obtain time-to-failure data. Second, the involved costs are often very expensive. Third, the only available data is commonly provided from similar components but exposed to different operational conditions. In this context, the importance of evaluating similarity between diverse equipment is fundamental for an appropriate reliability estimation, since data from multiple devices can be included as partially relevant information. The conventional methods for analyzing relevance use the arbitrary imposition of similarity weights by experts or the calculation of similarity measures using Euclidean distance. Otherwise, the use of artificial intelligence techniques can improve the similarity assessment, by making it independent from human intervention and capturing non-linear relationships with deep neural networks.

This is a very recent field of application, based on the previous use of deep learning for clustering tasks. For temporal series, some research has shown that computing similarity by autoencoders can improve the classification of cluster elements (Alqahtani et al., 2021). Indeed, the autoencoder is a special type of deep neural network, whose objective is to reconstruct at the output the same signal of the input (Li et al., 2023). It has an encoder layer, which makes the compression of information, and a decoder layer, which transforms the codified data to its original content. Among these two layers, the autoencoder includes a latent space, from which the codified information can be extracted. The main idea is to use the latent space data as a measure of similarity obtained through deep learning (Jo and Jun, 2022). The input time-series are reduced to lower dimensionality and the calculation of similarity distances is more meaningful. In this case, the proposed measure of similarity in the latent layer is the one from (Segaran, 2007):

$$s(z_0, z_1) = \frac{1}{d(z_0, z_1) + 1} \tag{1}$$

Where $d(z_0, z_1)$ is the Euclidean distance between the latent vectors. For the O&G industry, the employed time-series can be extracted from standard qualification tests (API, 2018). All the completion equipment must pass through ranges of pressure and temperature prior to its field application. Therefore, one can use this information to evaluate the similarity of O&G devices. The first step taken was to write a supervised autoencoder and compare it against results from an Ai4i2020 dataset (Jo and Jun, 2022). This dataset is used only to verify the code, which will need another validation when used with the O&G equipment. For the specific case, the supervised AE used ReLU as the activation function and a similarity score threshold of $\delta = 0.8$ for the comparison, obtaining the same accuracy (Jo and Jun, 2022). The combined loss convergence is shown in Figure 1.

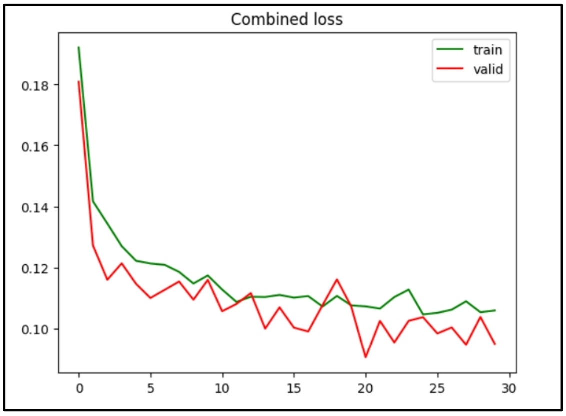

Fig. 1. AE combined loss from training and validation.

Another step taken was to use a DT generated data of bearings to evaluate the similarity between different types of equipment. This is done by first generating healthy and faulty data for one specific bearing that will be used to train the AE. Since theoretically, the fault frequencies of bearings are related to the equipment design (Randall, 2011), another kind of bearing can be run through the DT model (Souza et al, 2023). This new equipment data is then used as input for the trained AE, where the information from the latent space is taken and used to calculate the Euclidean distance between new and original data. This second step was used to verify the usability of the code in known working equipment.

Lastly, several qualification tests from an expandable packer are inputted into an autoencoder to obtain codified latent space information. After the training and validation process, data from similar equipment is passed to the same trained network, and the distance between the latent layer information is used to calculate the similarity measure. Comparisons against traditional distance methods and experts' elicitation will be presented for the proposed approach. The deep learning similarity is expected to attain a more refined similarity evaluation when compared to simple distance metrics, according to the assessment of specialized O&G staff.

Acknowledgements

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

References

Alqahtani, A., Ali, M., Xie, X., Jones, M. W. 2021. Deep time-series clustering: A review. Electronics 10(23). MDPI. https://doi.org/10.3390/electronics10233001

API. 2018. Recommended Practice on Subsea Equipment Qualification API RP 17Q.

Jo, H., Jun, C.H. 2022. A personalized classification model using similarity learning via supervised autoencoder. Applied Soft Computing, 131. https://doi.org/10.1016/j.asoc.2022.109773

Li, P., Pei, Y., Li, J. 2023. A comprehensive survey on design and application of autoencoder in deep learning. Applied Soft Computing 138, 110176. https://doi.org/https://doi.org/10.1016/j.asoc.2023.110176

Segaran, T. 2007. Programming Collective Intelligence: Building Smart Web 2.0 Applications. O'Reilly Media.

Randall, R.B., Antoni, J. 2011. Rolling element bearing diagnostics - a tutorial. Mech. Syst. Sig. Process. 25.

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. In: European Conference On Safety And Reliability (ESREL), Southampton.