

Assessment and Evaluation of an Unsupervised Machine Learning Model for Automotive and Industrial NVH Applications

Abdul Haq Azeem Paracha¹, Johannes Blickensdorff², Dr. David Scott Johnson³

¹ Technische Universität Ilmenau, 98693 Ilmenau, Germany, Email: abdulhaq.ah@gmail.com

² Schaeffler Technologies AG & Co. KG, Germany, Email: blickjha@schaeffler.com

³ Fraunhofer IDMT, 98693 Ilmenau, Germany, Email: david.scott.johnson@idmt.fraunhofer.de

Abstract

Rapid changes in the global industry like the emergence of electric vehicles and high-resolution data have posed new challenges for NVH engineers.

Current analysis techniques involve an interdisciplinary knowledge of structural dynamics, signal processing and psychoacoustics but most notably they require experienced professionals to analyse and assess the ever-expanding amount of acquired industrial NVH data.

Concurrently recent advances in machine learning show data driven model inference of feature representations - without human intervention. Unsupervised data driven methods have the potential to support NVH teams to focus on actual solutions by reducing manual efforts for pre-processing, classification and assessment of measurement and simulation-based data.

Introduction

Machine learning algorithms have shown significant success in image and speech recognition applications, but the learning and recognition performance depends also greatly on the choice and selection of data features [1].

They provide key methods for inductive inference through which specific statistical phenomena are generalized for new samples. Figure 1 shows a stem overview of methods, for the present work an unsupervised autoencoder (AE) approach was chosen [2].

The here so-called “classical approaches” require either an assessment of time- and frequency data by human experts (to confirm and validate NVH properties of specimen with respect to hardware design, root causes and simulation results) or they implement well-known signal processing techniques in automated test stands or other devices.

For the latter we challenged an unsupervised deep neural network model (DNN) based on AE to detect anomalies in industrial NVH data and compared findings to classical annotated automation data for the present use case of electric drive units (EDU) [3].

Related Work

The current industrial approaches reduce signal properties to a set of elementary acoustical metrics which represent NVH-target characteristics of specimen and have to work in quite robust production environments [4].

In comparison to a “classical” feature extraction (FE)

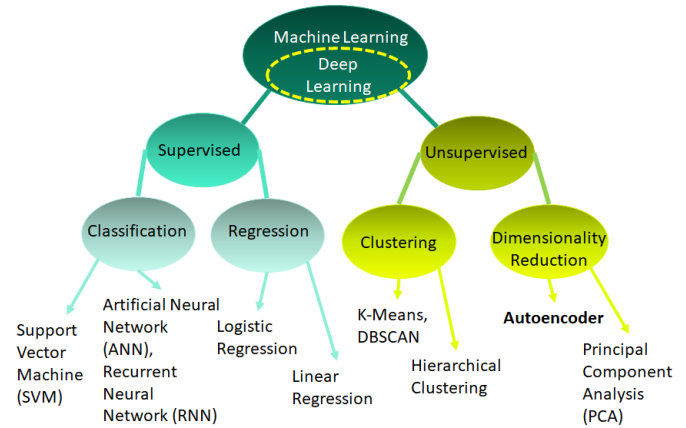


Figure 1: Types of machine learning algorithms

like Short-Time Fourier Transform (STFT) and order-based analyses, DNN provide an alternative approach of extraction, learning and whole system representation. We focused on a comparison of STFT and order features by an AE model having the same architecture and hyperparameter configuration, benchmarking the effectiveness of DNN results to the annotations made by classical FE (expert-based methods).

Dataset

Industrial multichannel data sets usually include airborne- and structure-borne noise data in conjunction with reference values and metadata and have application specific sampling rates from mid to higher ranges (that also can differ in-between channels).

The used experimental data set basket represents a virtual reject rate and consists of 77 samples in total. All samples were originally classified by the same automated EoL fault diagnosis setup that implements classical NVH quality control targets. Seventy samples were originally identified as “accepted” (normal) and seven samples in the basket were annotated as “refused”.

Each EoL measurement contains triaxial and single axial acceleration time data, acquired by sensors at the test bench and from automated tactile measurements on the specimen surface during the test. Reference channels are torque and rotational speed with same sampling rates.

Methodology

For the evaluation of the AE multilayer perceptron (MLP) model on the EoL dataset, usually three steps

are required: pre-processing, training and analysis. The dataset is divided into three subsets: Training, validation and test data. A most significant step in the evaluation of AE is to train the model only on normal data. Therefore it was ensured to keep training and validation data separate from the refused specimen data. Training and validation data include all normal specimen while the test data include also refused specimen. Loss or cost function used in AE is mean squared error (MSE) which was minimized between input and output to reconstruct the data sample in latent space by adding a penalty (reduced dimension layer in the encoder) on the model. MSE are commonly used in AE applications because their efficiency and simplicity make them a good choice as a starting point [5]. The mean absolute error (MAE) was calculated between original input data and predicted output data to find the reconstruction error. As performance metric the area under curve (AUC) of precision and recall curves was used. Furthermore, based on the distribution of reconstruction errors, the overlap of normal and refused specimen data was separated by a threshold that was defined by the intersection point of precision and recall curves. Finally, the confusion matrix was used to visualize the true positives.

Tool Chain

The model is implemented in Tensorflow using Keras API and STFT features are extracted directly in Python pipelines while order features had to be calculated in Matlab. To process the provided industrial datasets with Python the originally supplied ASAM-ATFX transport files were converted to ASAM-MDF4 format [6].

Experimental Setup

One data channel is used along with a rotational speed channel in the preprocessing step, order and STFT features were extracted. The procedure of anomaly detection demands test data or new data to be predicted by the model, including both: Accepted and refused specimen. Therefore 63 out of 70 normal samples were used for training and validation. Each of the seven normal and seven refused samples were used in the test subset. A relatively low number of samples indicates a deliberate effort to evaluate the performance of the model in less specimen R&D data conditions. After feature extraction, all data were scaled through min-max normalisation to keep them in a certain range for the objective function to work properly. Architecture of the AE is primarily inspired from [7], but modified with a number of preliminary experiments with different layer- and hyperparameter configurations. The model used for training uses rectified linear activation function (ReLU) as activation that has 10 fully connected layers. The first layer of the encoder starts with 256 neural units and progressively decreases to 16 units in the latent space. The decoder starts with 16 units and increases to 256 units in the last layer. The optimization algorithm used is Adam and learning rate is 0.0001 [8].

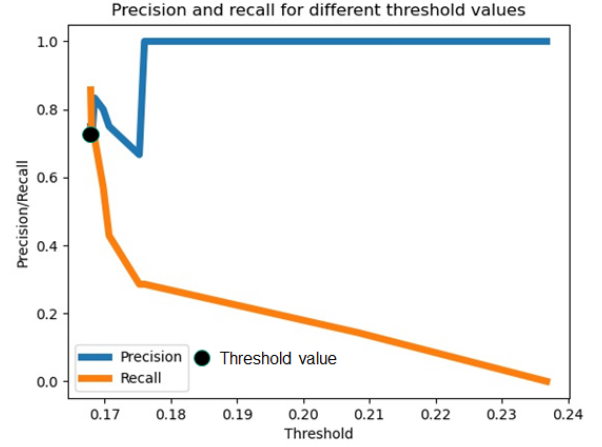


Figure 2: Precision and recall curves with respect to threshold of experiment A, the black point marks the threshold value

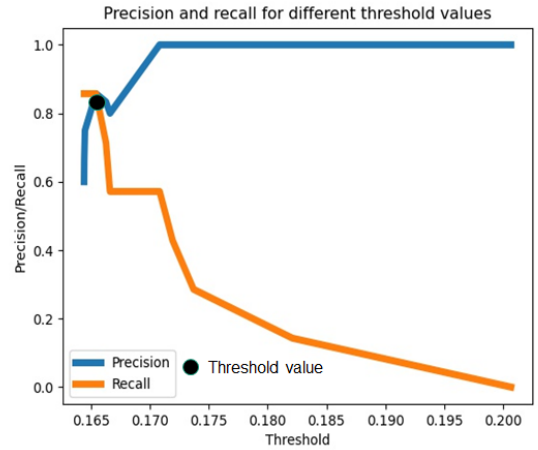


Figure 3: Precision and recall curves with respect to threshold of experiment B, the black point marks the threshold value

Results and Discussion

There were two experiments conducted with the same model parameters, experiment A (STFT) and experiment B (orders features). Training and validation loss of the model for both experiments is 0.04, threshold of experiment A according to precision and recall vs. threshold curve is 0.169 as shown in Figure 2.

Threshold for experiment B is 0.166 as shown in Figure 3. AUC of precision-recall curves are 0.828 (A) and 0.878 (B) as shown in Figure 4. Results from experiment B (order features) show a better efficiency in comparison to experiment A. More detailed analysis of the experiment A can be seen in the confusion matrix of Figure 5 where the true class contains annotations identified by an NVH expert through classical analysis and predicted class is the model output. The predicted "refused" label indicates true positives. As shown in Figure. 6 all seven refused samples were identified correctly based on the threshold value in Figure 2 but three normal samples were also erroneously categorized as refused. In experiment B seven out of six samples were identified correctly

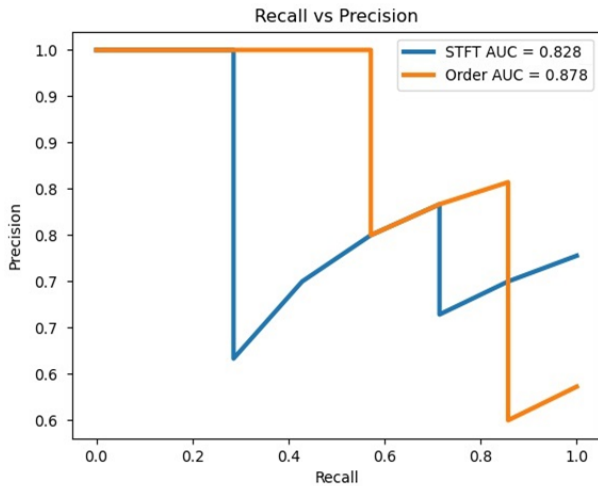


Figure 4: Performance of experiment A and B through AUC precision vs recall curve

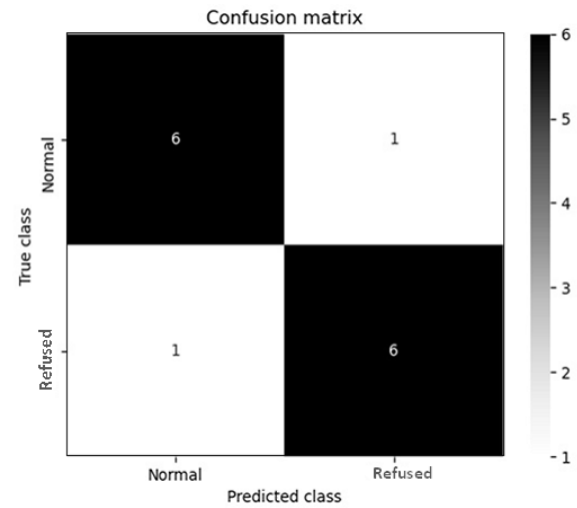


Figure 6: Confusion matrix experiment B

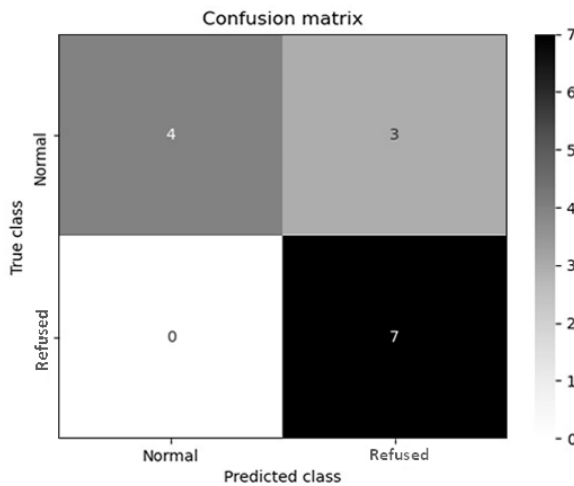


Figure 5: Confusion matrix experiment A

while one normal sample was identified as refused and one identified vice versa.

Conclusion

The best comparative accuracy achieved is 0.878 in experiment B based on order features. A (so called) classical NVH approach requires in depth analyses of simulated, normal and fault-based data using different signal processing techniques and (expensive) expert tools. Unsupervised autoencoders show good results in accuracy compared to classical approaches and “just” require normal data to find noticeable specimen. Presuming consistency and a certain amount of diversity in the underlying data sets, such data driven methods may help to reduce time consuming data examination tasks for NVH engineers significantly in the (very) near future.

Acknowledgment

This research is conducted in collaboration of Schaeffler AG and Fraunhofer IDMT. Special thanks go to Schaeff-

ler AG for providing datasets and NVH support.

References

- [1] Bengio, Yoshua et al. “Representation learning: A review and new perspectives”. *IEEE transactions on pattern analysis and machine intelligence* 35. 8(2013): 1798–1828.
- [2] Rättsch, Gunnar. “A brief introduction into machine learning”. Friedrich Miescher Laboratory of the Max Planck Society. (2004).
- [3] Koizumi, Yuma et al. “Unsupervised detection of anomalous sound based on deep learning and the neyman–pearson lemma”. *IEEE/ACM Transactions on Audio, Speech, and Language Processing* 27. 1(2018): 212–224.
- [4] Knieper, Johannes et al. “Methoden für die akustische Analyse und Bewertung von E-Achs-Getrieben”. Deutsche Gesellschaft für Akustik eV: DEGA. (2020).
- [5] Shao, Haidong, et al. “A novel deep autoencoder feature learning method for rotating machinery fault diagnosis.” *Mechanical Systems and Signal Processing* 95 (2017): 187-204.
- [6] “ASAM MDF.” Association for Standardization of Automation and Measuring Instrument, www.asam.net/standards/detail/mdf/.
- [7] Ribeiro, Alexandrine, et al. “Deep Dense and Convolutional Autoencoders for Unsupervised Anomaly Detection in Machine Condition Sounds.” *arXiv preprint arXiv:2006.10417* (2020).
- [8] Paracha, Abdul Haq Azeem. “Assessment and Application of Machine Learning Methods to NVH Related Component Data in the Automotive R&D Environment.” Technische Universität Ilmenau, 2021.