Roller Bearing Fault and Wear Simultaneous Prognosis and Classification Using Artificial Neural Networks

ME 8813 Final Project Report

Enea Dushaj

Mechanical Engineering

Georgia Institute of Technology

Atlanta

5/2/2022

1 Introduction

Roller bearings are a fundamental building block to the smooth operation of modern day industrial machines. The maintenance of these critical components ensures the proper functioning of these machines to guarantee accuracy, safety, and reliability in any mechanical field, from automotive, to aerospace and manufacturing. The ability to predict the remaining useful life of a bearing can prove to be quite beneficial to industry and research applications. The ability to forecast the future behavior or performance of a bearing can allow for proper maintenance scheduling, and can help prevent potential premature faults, and increase cost savings by maximizing potential remaining useful life. Typically, Hidden Markov Models have been utilized for estimating remaining useful life [1] [2] [3], however they lack potential finer details that can be useful for prognosis. Rather than forecasting the continuous state of the bearing, a discrete bearing health state is predicted by hidden Markov models. This limits the information to be obtained about the health of the bearing, particularly if the user would like to maximize the remaining life of the bearing. One potential approach for extrapolating more detailed time-series data is through a recurrent neural network. Recurrent neural networks have been proven to perform very well at incorporating the importance of time in time series data, and create the potential for predicting future bearing performance metrics. Furthermore, various approaches to processing the vibration signals for data driven learning have been taken. Wavelet Packet Decomposition and statistical parameter extraction for measurements such as kurtosis, skewness, entropy, etc. have been used for feature extraction to be used in machine learning methods for bearing faliure prognosis [4]. Additionally, Recurrent Neural Networks have been used to predict the remaining useful life of turbomachinery [5] and bearings [4], however, this study aims to compare the effectiveness of another approach to signal feature extraction in conjunction with an ANN for bearing prognosis. Hjorth parameters are intended to be used in concurrence with time series data as input features in this study. Hjorth parameters are statistical time domain parameters related to the variance of a signal and the derivatives of the signal. Hjorth parameters have been shown to be able to detect potential faults in the bearing data set used[6] [7], however they have not yet been utilized in conjunction with a machine learning technique yet. The objective of this study is to create a method for predicting Hjorth parameters for a bearing, namely Activity, as a method for determining bearing health and remaining useful life of a roller bearing. An artificial neural network is the proposed machine learning model to be used. The propoesed ANN will be used to extrapolate the future activity of the bearing given some amount of its previous activity, and will be able to classify the impending failure of the bearing. The proposed work will create an algorithm that can predict future activity of the bearing, and can furthermore be used to accurately depict the remaining useful life of the bearing.

2 Methodology

This study makes use of bearing acceleration data provided by the Center for Intelligent Maintenance Systems (IMS), University of Cincinnati, published by the NASA Prognostics Center of Excellence. This data was generated in an experiment in which 4 bearings were installed on a shaft driven at a constant speed of 2000 RPM by an AC motor. A high sensitivity quartz accelerometer was installed on each bearing. Every 10 minutes, a 1 second interval of vibration acceleration data is gathered at approximately 20 kHz. The bearings are said to have a designed lifetime of approximately 100 million revolutions. All detected faults in the data occurred after this designed life time. Three tests were performed. In the first test, an inner race defect and roller element defect were detected in bearings 3 and 4 respectively. In the second test, an outer race defect was detected in bearing 1. In the third test, an outer race defect was detected in bearing 3. As stated previously, every 10 minutes, a 1 second interval of vibration acceleration data is gathered at approximately 20 kHz. This study makes use of each 1 second sample as discrete signals. From each signal, the Hjorth parameters are derived. The parameters are Activity, Mobility, and Complexity. It was demonstrated by Cavalaglio, et al. that Activity is a good measure of predicting bearing fault with this data set. Converting the time series data into Activity data tremendously condenses the amount of information necessary to process. This results in a tradeoff, however. With less data, the training of a neural network is less successful, however this drawback is overcome with sub-sampling, which will be discussed later on. Activity is simply the variance of the signal.

$$Activity = \sigma^2 \tag{1}$$

Where σ is the standard deviation of the one second acceleration signal of the bearing. A higher variance in the bearing acceleration indicates a large change in acceleration, which may indicate some defect. The Activity parameter of the bearings with detected faults approximately 150 hours before the end of data collection are shown in Figure 1. Additionally, a bearing in which no fault was detected at the end of testing is shown for comparison (Test 1 - Bearing 1).

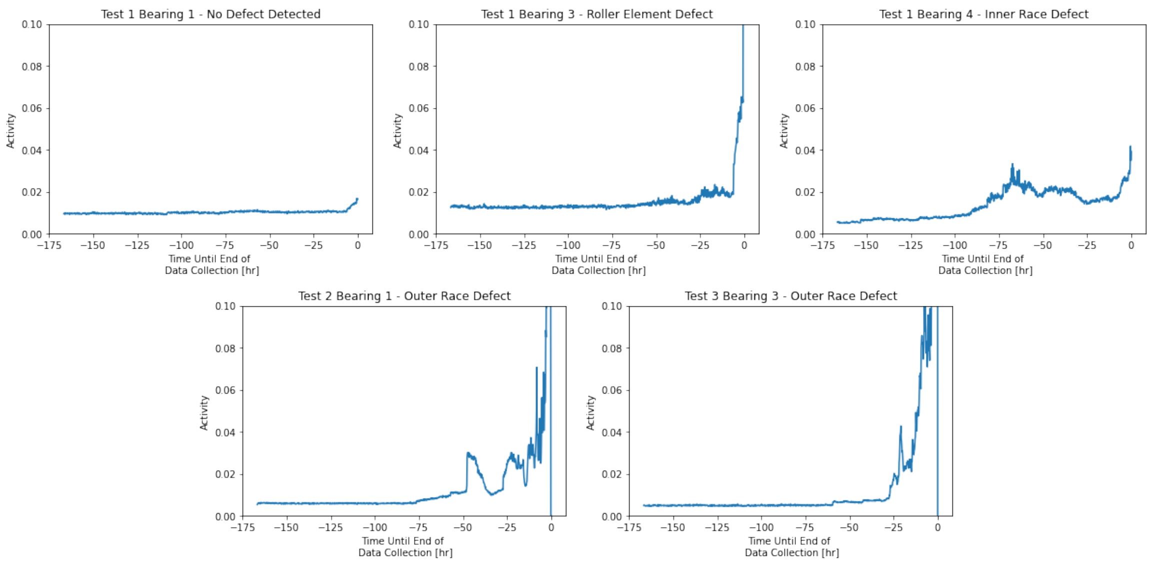


Figure 1: Activity Plot for each bearing with a defect detected, and one without a defect detected.

An example of how activity is calculated is shown in Figure 2. Each 1 second sample is condensed into an activity parameter (variance of the signal). Each point on the activity plot represents one second of acceleration data at 10 minute intervals.

An artificial neural network is utilized to simultaneously classify faults and predict the activity of the bearing in the future. The input and labeled data must be organized in a way that allows for this type of prognosis. Furthermore, in order to effectively train the model, more data than provided is required. Therefore, each activity profile of the bearing is split into multiple time series data points. For each activity profile, every n^{th} point is sampled into new time series data to effectively create n activity profiles our of a single bearing's activity profile. A sample of this increase in data is shown in Figure 3. This has the effect of increasing the available data for training and testing the model, but simultaneously lowers the resolution of the samples as the number of subsamples increases.

Next, each activity profile for each bearing is split into inputs and labels. A fraction, f of each time series

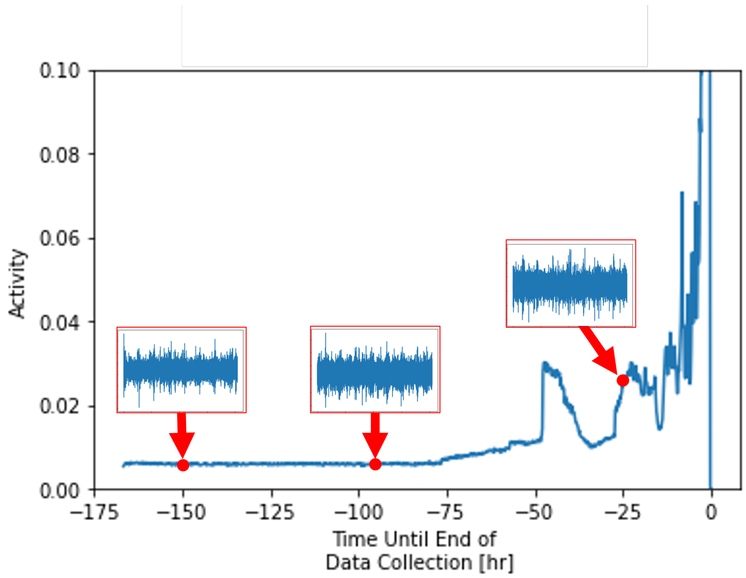


Figure 2: Each raw data signal has Hjorth parameters associated with it and is transformed into a single point.

is used as inputs into the model, and the remaining $1 - f\%$ of the data is used as a label to that training example. Additionally, bearing is classified into a corresponding failure mode. Appended to the end of each training example label is a vector classifying the failure of the bearing:

$$\begin{bmatrix} 1 & 0 & 0 & 0 \end{bmatrix} = No\ Failure\ Detected$$

$$\begin{bmatrix} 0 & 1 & 0 & 0 \end{bmatrix} = Outer\ Race\ Fault\ Detected$$

$$\begin{bmatrix} 0 & 0 & 1 & 0 \end{bmatrix} = Inner\ Race\ Fault\ Detected$$

$$\begin{bmatrix} 0 & 0 & 0 & 1 \end{bmatrix} = Roller\ Element\ Fault\ Detected$$

Finally, the model constructed is a 3 hidden layer artificial neural network consisting of b nodes in each layer. During preliminary model creation, it was determined that the number of hidden layers added after 3 layers provided diminishing returns for the training time required. To isolate the best parameters needed to construct a good fitting model that could classify and provide a prognosis of the performance of each bearing, a three-dimensional parametric sweep was done for the number of sub-samples for each bearing's

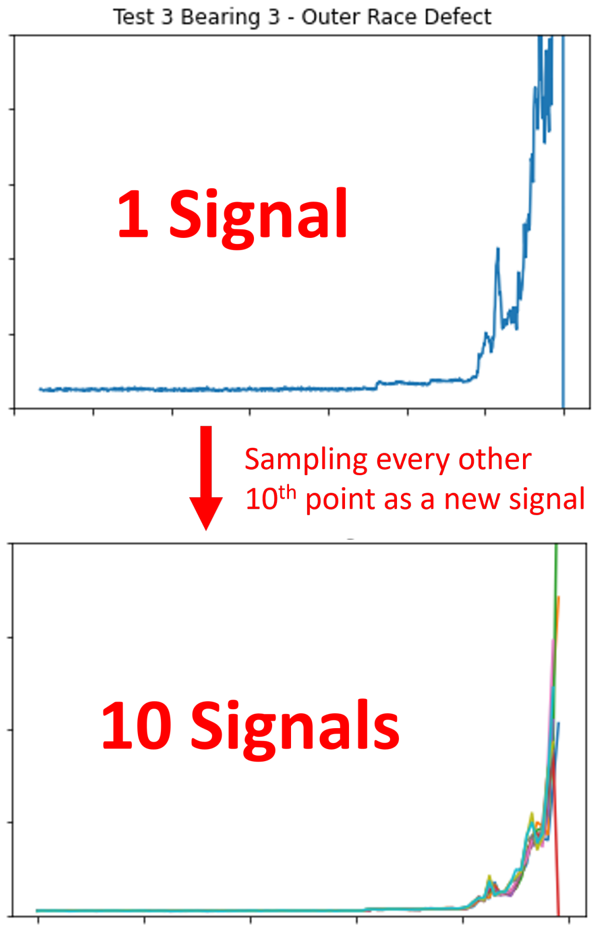


Figure 3: Each bearings activity profile is extended into n more samples - in this case 10.

activity, n, the fraction of each activity profile assigned as input features, f, as well as the number of nodes in each layer, b. The mean absolute error, cross validation mean absolute error, test data root mean square error, as well as the R^2 score for each model was recorded. The number of sub-samples, n, tested were 10, 25, and 40, the fractions of data used as input, f, were 10%, 25%, 50%, 75%, and 90%, and the number of nodes in each hidden layer , b, tested were 1000, 2000, and 3000. A total of 45 models were created testing different combinations of model and data preparation parameters, n, f, b. The parameter search with the given values was chosen since they provided a good balance between model accuracy and reasonable training time. Each model is trained for 100 epochs. Beyond 100 epochs, no models benefit from additional training. The data is normalized, excluding the fault classification, and the classification vector is appended to the end of each training and testing label. 70% of the total data was used for training, the remaining 30% was used for testing in every scenario.

3 Results

Through the parameter sweep described in the Methodology section, the best model determined was the ANN utilizing 40 subsets of pseudo bearing activity profiles, where 90% of the data was input data, the remaining 10% was label data, and the number of nodes in each hidden layer was 1000. This model gave yielded an R^2 score of 0.827. The parity plot of the prognosis predictions for this model and the confusion matrix for fault classification are shown in Figure 4. The model performs fairly well in classifying faults. Every inner race and roller element fault was predicted accurately. Every bearing that had no fault was correctly predicted to not have a fault, however not ever bearing with an outer race defect was predicted to have an outer race defect. Around 20% of bearings with outer race defects were predicted to have no faults. This model may prove relatively useful to a company performing prognosis for bearings on their machines, however false positive health indications of bearings that may have outer race faults could prove extremely costly, therefore this model must be improved to be successfully implemented in industry.

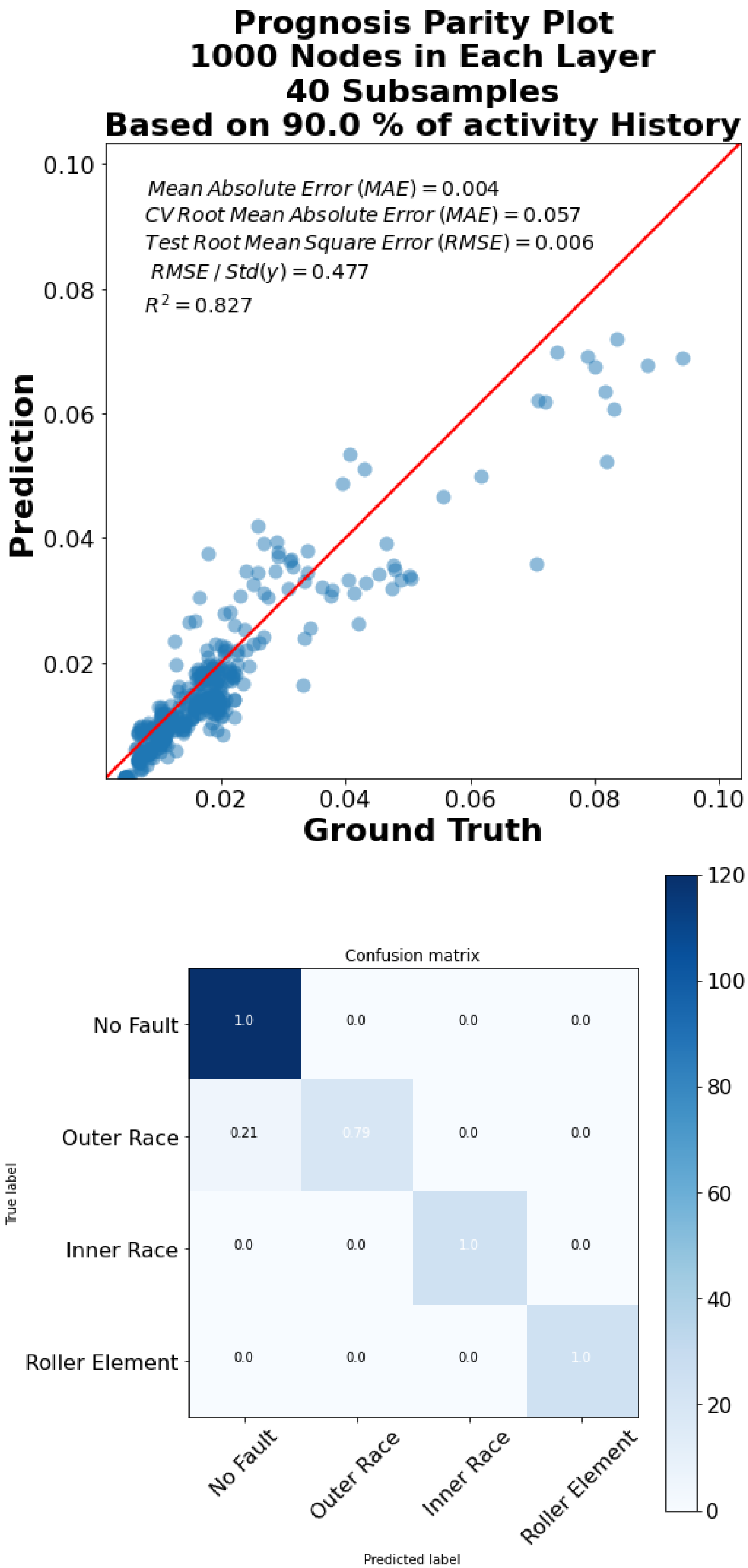


Figure 4: Best model for classification and prognosis given the methodology.

These results come with a trade-off however - This data is sub-sampled 40 times, and the prediction is based on 90% of the bearing history. This means that the inputs are few and far between, the model requires a lot of prior information to make a valuable prediction, and the model can only predict 2 time steps into the future, where each time step is approximately 8 hours into the future. Samples of these predictions are shown in Figure 5.

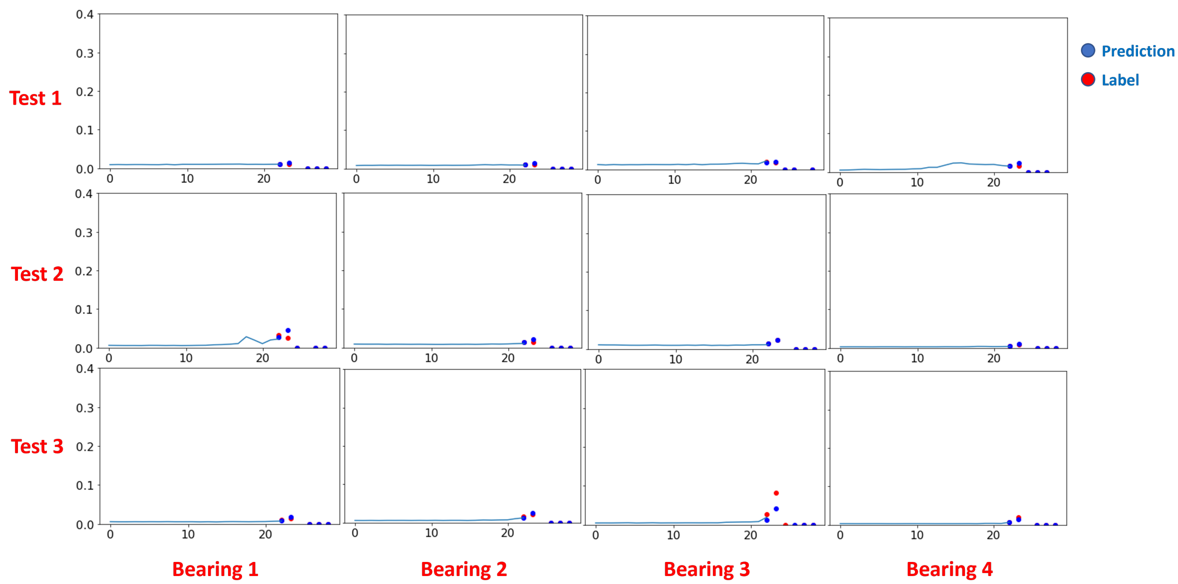


Figure 5: Sample prognostic predictions using the model with the highest determined R^2 value (0.872) in the parameter sweep.

Although this prediction may be more accurate, it may not be too useful for an organization since predictions can only be made a few hours in advance.

Prognostic predictions for a model that can make predictions far into the future with only minimal previous knowledge of the bearing activity are shown in Figure 6.

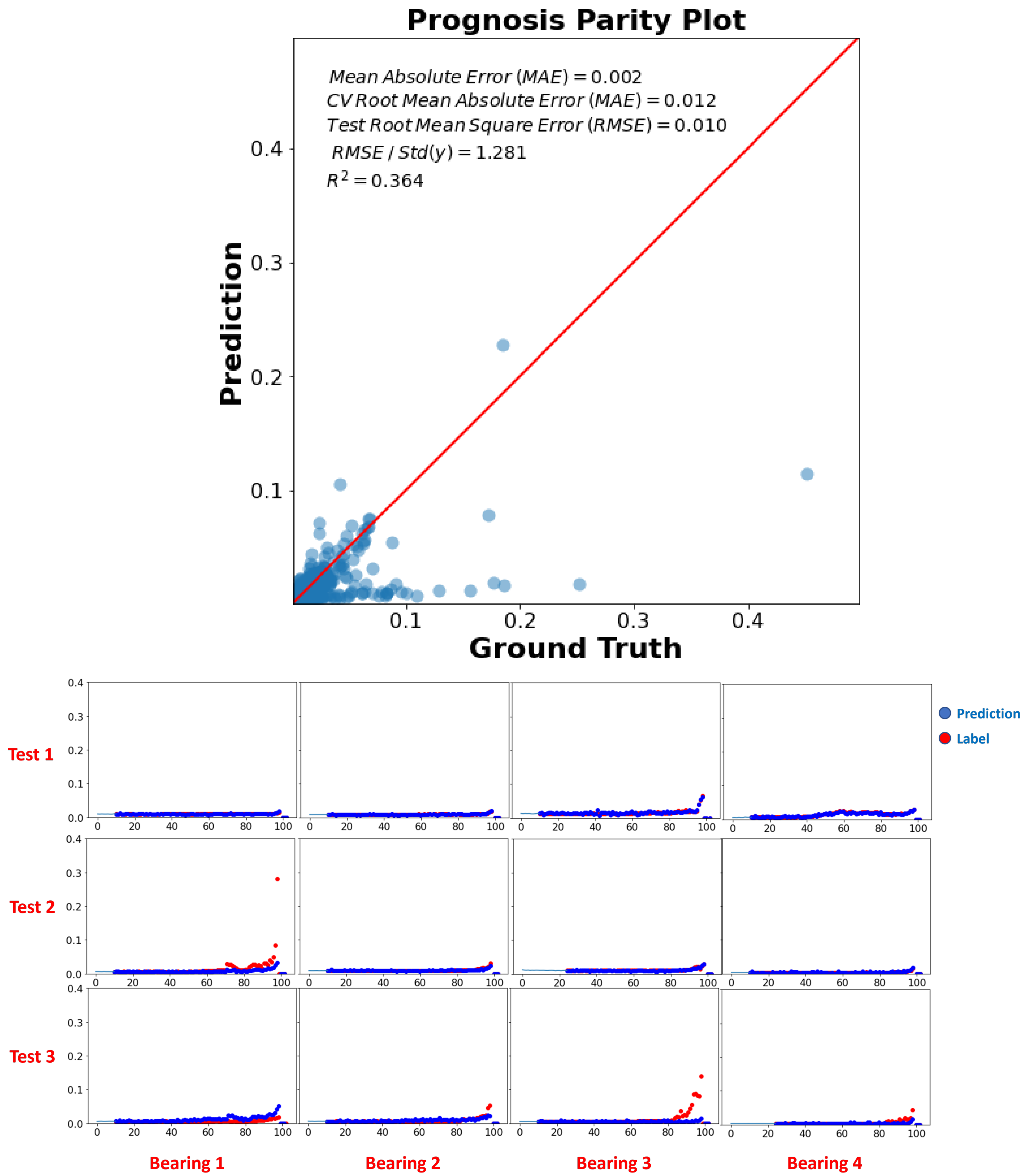


Figure 6: Sample prognostic predictions produced by a model with knowledge of 10% of prior data points.

While these models can produce predictions far into the future they may not be as accurate. Predictions for classifications are relatively inaccurate as well. Outer Race bearing faults are consistently mis-classified. The confusion matrix for this model is shown in Figure 7.

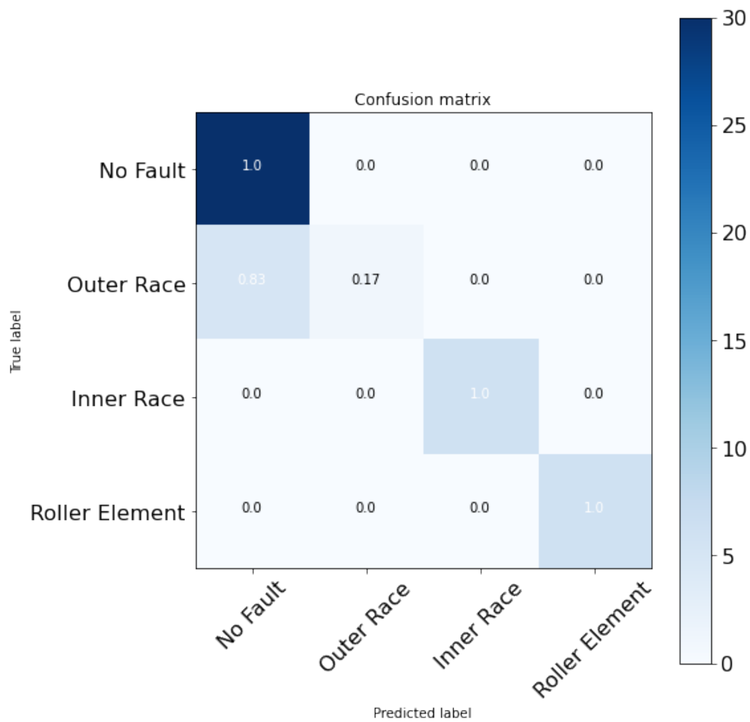


Figure 7: Confusion matrix for a model with knowledge of 10% of prior data points.

A typical set of predictions with 10 subsets, 1000 nodes in each hidden layer, and setting 75% of the time series to be training data is shown in Figure 8. While this data is less accurate, it provides higher resolution and a further outlook on the potential performance of a bearing. The parity plot for the prognosis predictions made by this model correlate directly with the under-predictions made for Bearings 2 and 3 in Test 3. The activity extrapolation prediction for Test 3 Bearing 2 and Bearing 3 are inaccurate. While this may prove discouraging for the model, Gousseau et al. [8] determined that Test 3 of the IMS dataset is inconsistent for diagnosing faults in Bearing 3, and that the data does not indicate an outer race fault in bearing 3. Generally, this is the reason data from test 3 is not utilized in literature. This may mean the model is actually performing well, and finding the inconsistencies of the data collection, and correctly predicting that there is no failure in Test 3 Bearing 3, as reported by the IMS.

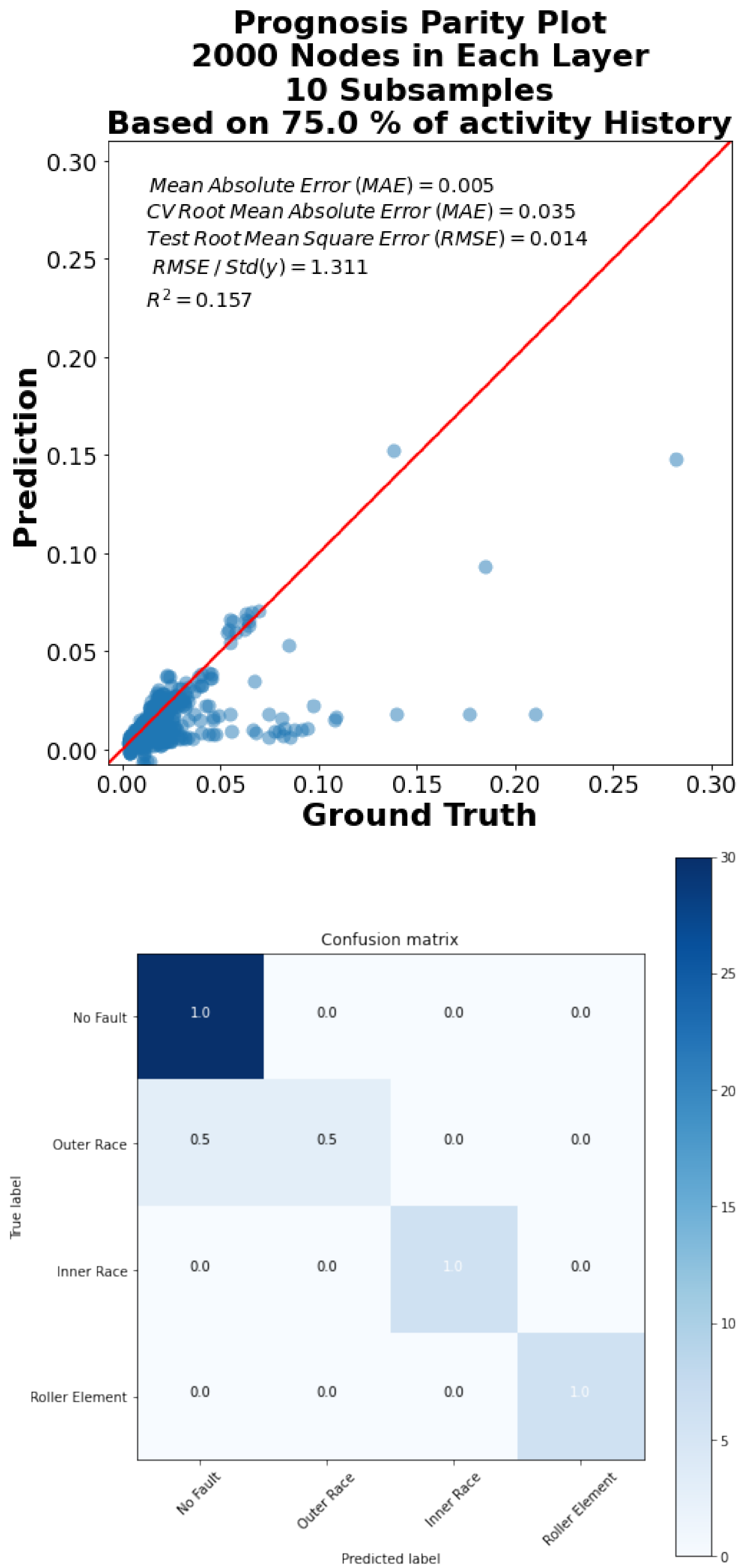
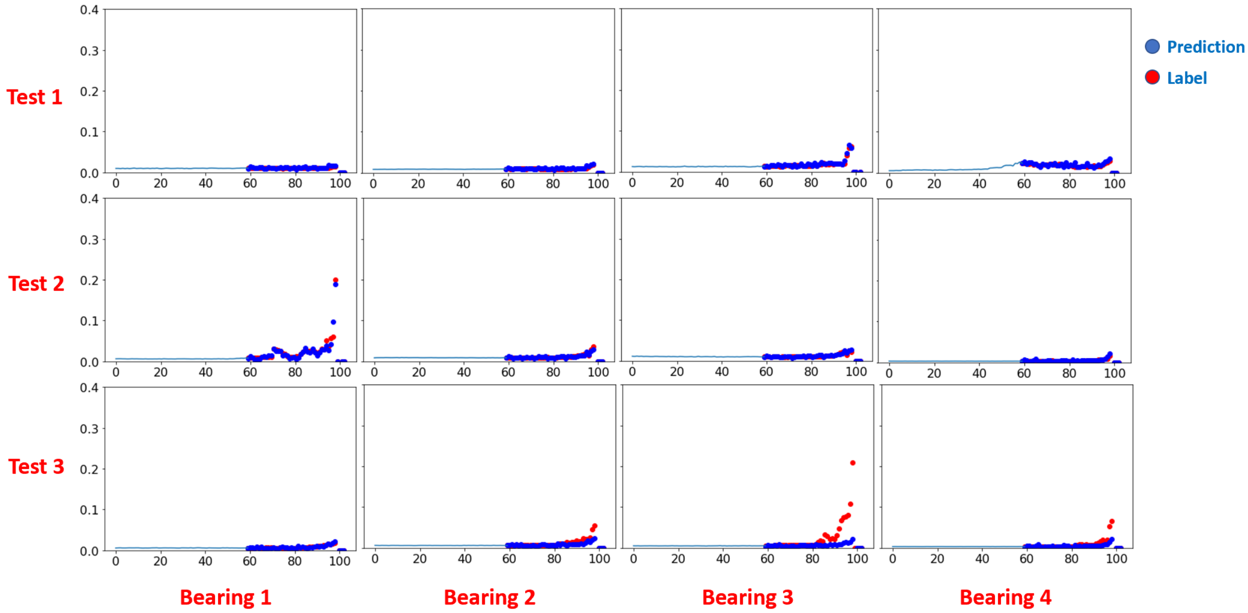


Figure 8: Sample prognostic prediction with 75% prior knowledge ($R^2 = 0.157$).



4 Summary and Future Work

Through this study, an artificial neural network capable of predicting bearing activity and classifying the predicted failure mode of the bearing was created. A parametric sweep of model and data processing parameters was performed to find the optimal performing model, and to determine the optimal amount of prior history needed for the model to make accurate predictions. The model did not perform well on bearing data from Test 3 of the bearing data set applied in this study, however literature suggests that Test 3 in this dataset provides inconsistent analysis and/or data, and therefore should not be used in model construction. Literature suggests that there is no indication of a bearing fault in any of the Test 3 bearings, and the model tends to align with that assessment of the data. The model can be used in predicting the statistical time-domain Hjorth parameter, activity, which is the variance of the time signal, well into the future, even with limited prior knowledge. To expand on this model framework, all 3 Hjorth parameters will be determined and fed into the model to create a multiple input model. Furthermore, the amount of cycles the bearing has been through will also be passed into the model as another feature. Some additional information appears to be necessary to properly diagnose faults. The model cannot seem to properly classify and provide a prognosis for the failure from Bearing 3 of test 3, and may require additional input information to do so. Additionally, the model can further be optimized by testing various architectures. This model is a 3 hidden-layer model with hundreds of nodes in each layer. Different architectures with more or less layers with different numbers of nodes in each layer can be explored. For simplicity, this exploration was excluded for this study, but may present large gains for model accuracy. Furthermore, techniques such as dropout and early stopping can be used for quicker model training and can lead to more accurate results and should be explored further. Finally, a method for training a recurrent neural network on multiple time series should be focused on, as recurrent neural networks may provide the best results for a model being trained on time series data such as this one. Currently, the main method for training recurrent neural networks is on a single time series - a method for training the network on multiple different time series would be beneficial especially for a study like this where many different time series can be produced from a single series. Furthermore, since the requirement for training neural networks is substantial amounts of data, another method for increasing the number of data points for training and testing would be to split the 1 second signals in half and considering each half of the signal as its own bearing. This concept is shown in Figure 9, where the blue and green line represent the activity profiles extracted from one bearings time series data. Ultimately, some more amount of data is needed. Finally, the method of increasing data in this study may lead to overfitting. The assumption that each sub-sample of the activity of a bearing can be considered an entirely new bearing is not entirely accurate since every signal is originating from the same source. A better method for extracting more samples from a limited data set is needed, since sub sampling is likely to lead to overfitting.

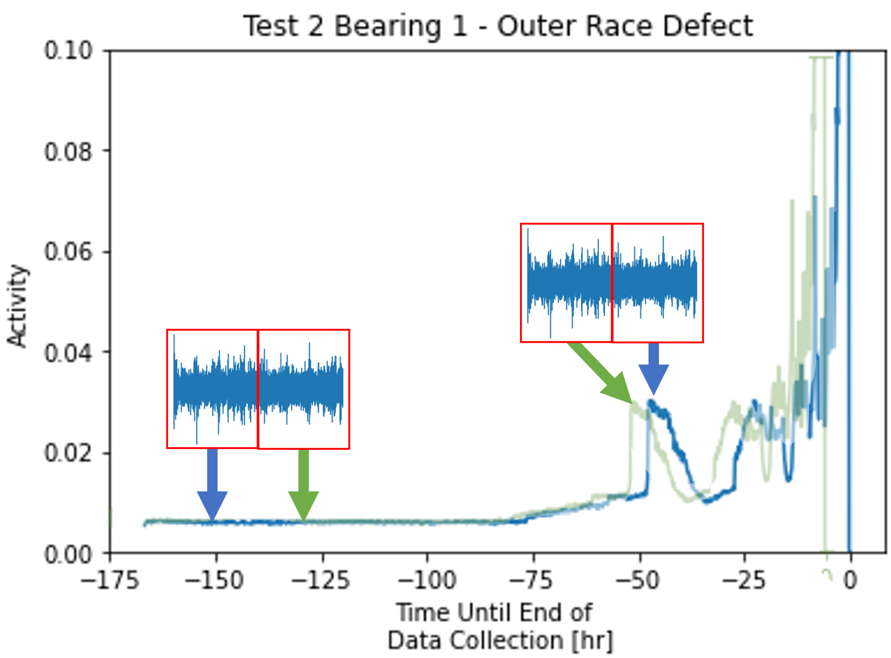


Figure 9: Another method for increasing the number of data points given only a few time series signals.

References

[1] Diego Alejandro Tobon-Mejia et al. "A Data-Driven Failure Prognostics Method Based on Mixture of Gaussians Hidden Markov Models". In: *IEEE TRANSACTIONS ON RELIABILITY* 61.2 (2012).

[2] Jianbo Yu. "Health Condition Monitoring of Machines Based on Hidden Markov Model and Contribution Analysis". In: *IEEE TRANSACTIONS ON INSTRUMENTATION AND MEASUREMENT* 61.8 (2012).

[3] Diego Tobon-Mejia et al. "Hidden Markov models for failure diagnostic and prognostic". In: *Prognostics and System Health Management Conference* (2011), pp. 1–6.

[4] Lianga Guo et al. "A recurrent neural network based health indicator for remaining useful life prediction of bearings". In: *Neurocomputing* (2017), pp. 98–109.

[5] Felix O. Heimes BAE Systems. "Recurrent Neural Networks for Remaining Useful Life Estimation". In: *2008 INTERNATIONAL CONFERENCE ON PROGNOSTICS AND HEALTH MANAGEMENT* (2008).

[6] Chhaya Grovera and Dr. Neelam Turk. "Rolling Element Bearing Fault Diagnosis using Empirical Mode Decomposition and Hjorth Parameters". In: *International Conference on Computational Intelligence and Data Science (ICCIDS 2019)* (2019).

[7] Cavalaglio Camargo Molano et al. "ANALYSIS OF NASA BEARING DATASET OF THE UNIVERSITY OF CINCINNATI BY MEANS OF HJORTH'S PARAMETERS". In: *DISMI, University of Modena and Reggio Emilia* (2018).

[8] F. Girardin W. Gousseau J. Antoni and J. Griffaton. "Analysis of the Rolling Element Bearing data set of the Center for Intelligent Maintenance Systems of the University of Cincinnati". In: (2016).