Vibration Analysis

By Akshay Nair

With Help from Tim Gianitsos

Abstract:

This paper covers differences in accuracies of an artificial intelligence model that classifies different sets of sensor outputs (in this case, machinery shaft vibrations) collected by sensors into varying levels of weight on the shaft. The model uses a neural network to categorize the data, and there are varying accuracies depending on weights at the end of the shaft. This paper goes over the causes for these changes and patterns one can notice by going through the results.

Introduction:

Many factories have machines involving shafts, and most breakdowns in equipment are due to an excess or imbalance in weight on these shafts. When these breakdowns happen, most factory processes have to be shut down to fix it, which takes a lot of productivity out of the process. We thought of making a model that could predict the difference between weight imbalances on a shaft based on sensor readings of how much the shaft vibrated. More vibration meant the shaft had a larger weight imbalance. This research plans to help lower shutdown downs when factory machines are in need of repair by telling when mechanical shafts start to wear. This will allow factory owners to replace the shaft before serious damage is done to products and surrounding factory equipment.

Background:

Before starting the model, we researched whether this was an actual problem. We found that many factories that follow lean manufacturing practices, like automotive and clothing goods factories, have frequent shaft breakdowns that affect machine productivity. This showed that it is a real problem, and it meant that automotive factories, as one of the most in-demand industries, weren't able to meet demand as efficiently as they could be. We then researched potential datasets, coming across one for which similar analyses had been done. .

Dataset:

The dataset used in our research was a series of arrays containing electronic pulses returned by shaft-mounted sensors that measure vibration movement. Data is recorded at a sampling rate of 4096 values per second. Five different unbalance strengths are captured in the dataset, ranging from no imbalance to very high imbalance. These dataset components are labeled 0 to 4 for least no unbalance to very high unbalance, respectively, and D or E for training and testing, respectively (D is for development, E is for evaluation). There is enough data to get accurate predictions (11 gigabytes), which, to our convenience, have been pre processed into numpy arrays by the authors of the dataset.

Methodology/Models:

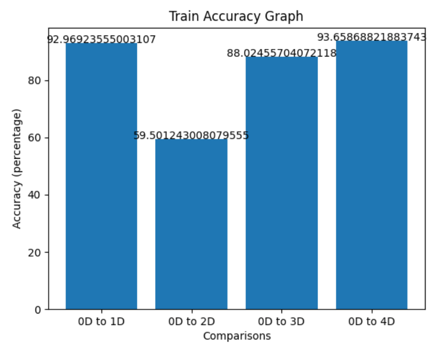
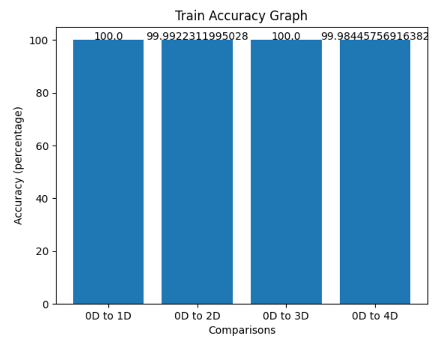
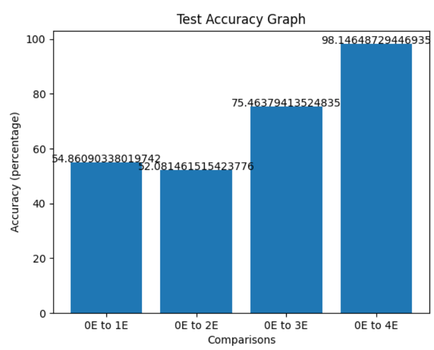
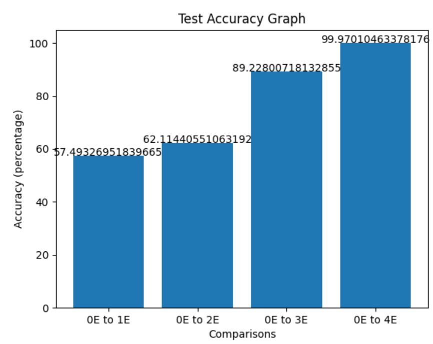
The initial model used the MLP Classifier (multi-layer perceptron classifier) to classify preprocessed sensor readings into 5 groups depending on the weight imbalance. The MLP classifier is a neural network that works by passing input data through multiple layers. The output layer classifies the output. The hidden layer(s) have neurons through which the data is passed through multiple times, training the model.

The final model uses a Random Forest Classifier to classify the data points into five categories depending on the weight imbalance (no imbalance returns a 0, most imbalance returns a 4). A Random Forest works by forming multiple uncorrelated decision trees to predict the outcome of one input, which increases the accuracy of the model. It can be thought of as a peer review assignment. Peer reviews are useful because the student being reviewed gets the feedback from multiple people when they are given the assignment. The Random Forest is useful because the data input gets a prediction of the output from multiple decision trees that are unaffected by the outcome of each other, allowing for a better, well-rounded prediction.

Results:

Out of the two models, The Random Forest performed better in most cases. Multiple dataset-to-dataset cases were predicted, as in the table below.

Comparison	Model	
	MLP Classifier	Random Forest Classifier
No weight imbalance to very small weight imbalance (0-1)	54.86090	57.49327
No weight imbalance to small weight imbalance (0-2)	52.08146	62.11441
No weight imbalance to large weight imbalance (0-3)	75.46379	89.22801
No weight imbalance to very large weight imbalance (0-4)	98.14649	99.97010

Type	MLP Classifier	Random Forest Classifier
Train Accuracy		
Test Accuracy		

The Random Forest model performed better in all four comparisons than the MLP Classifier. The Random Forest also took much less time to run (about 12 seconds) compared to the MLP Classifier (about four minutes).

Conclusion:

The main takeaway from this research and model was the difference between accuracies of different imbalance comparisons. The 0-4 comparison, or the no imbalance to very large imbalance comparison, was the most accurate with 99.97010%, and the 0-1 comparison, or the no imbalance to very small imbalance comparison, was the least accurate with just 57.49327%. This is because it is easier to tell the difference between two entities when there is a larger difference. For reference, it is easier to tell the difference between a wolf and a cat than a wolf and a husky. In this case, the 0 data input array was the wolf, the 1 data input array was the husky, and the 4 data input array was the cat. And predictably, the correlation between prediction accuracy and shaft weight imbalance was positive as the difference between imbalances increased.

Acknowledgements:

This research, model, and paper would have been impossible without my supportive and helpful mentor Tim Gianitsos, who helped me understand AI, terminals, and different models better than anybody. I also thank my parents who introduced me to artificial intelligence and helped me throughout the way.

References: (need to format to whatever protocols)

Adams, Helen. “The Foundation of Lean Manufacturing.” *Manufacturing Digital*, 25 Dec. 2021, https://manufacturingdigital.com/lean-manufacturing/foundation-lean-manufacturing.

Admin. “A Closer Look at Shaft Failures and What to Do about Them.” *Vissers Sales Corp.*, Vissers Sales Corp, 13 Oct. 2021, https://visserssales.com/a-closer-look-at-shaft-failures-and-what-to-do-about-them/.

Admin. “What Causes Centrifugal Pump Shafts to Snap or Break?” *Vissers Sales Corp.*, Vissers Sales Corp, 3 Aug. 2021, https://visserssales.com/what-causes-centrifugal-pump-shafts-to-snap-or-break/.

Jishnu. “Vibration Analysis on Rotating Shaft.” *Kaggle*, 23 Feb. 2022, https://www.kaggle.com/datasets/jishnukoliyadan/vibration-analysis-on-rotating-shaft.

Meltzer, Rachel. “What Is Random Forest? [Beginner's Guide + Examples].” *CareerFoundry*, 15 July 2021, https://careerfoundry.com/en/blog/data-analytics/what-is-random-forest/#:~:text=2.,group%20than%20they%20do%20alone.

Mey, Oliver, et al. “Machine Learning-Based Unbalance Detection of a Rotating Shaft Using Vibration Data.” *ArXiv.org*, 31 July 2020, https://arxiv.org/abs/2005.12742.

Sachs, Neville. “Failure Analysis of Machine Shafts.” *Efficient Plant*, 15 Jan. 2018, https://www.efficientplantmag.com/2012/07/failure-analysis-of-machine-shafts/.