Anomaly Detection using Machine learning

Heiner Castro, Zuhair Almansouri, Christopher Han, Gunkyo Park, Saptarshi Saha , Suyash Johari, Saarang Srinivasan, Vikhyat Jagini, Khyathi Bishop, Aayushi Akhouri

Introduction/ Background

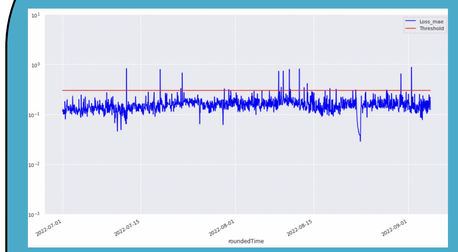
Project Background

Our team was tasked by Webee to understand if a machine is at an “on or off” state using the data provided as well find the duration of machine running time, the threshold of on/off speeds, and find the trends in data and detect anomalies and outliers.

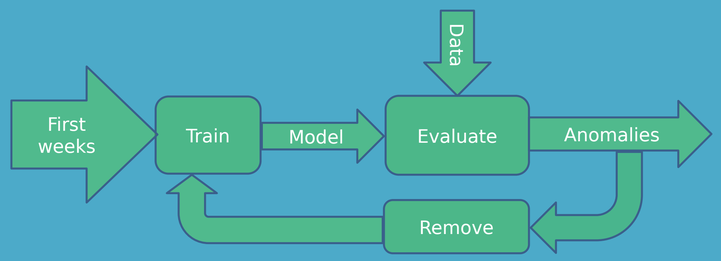
Company Background

Webee is known for monitoring the machine health by analyzing its behavior through sensor data. This includes vibration, temperature, and energy consumption.

Methodology

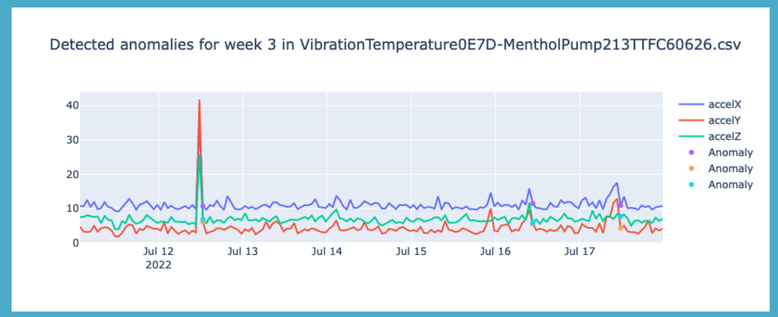


A. Anomaly Detection LSTM – Model learns to identify patterns and relationships in the data and uses this to detect anomalies.

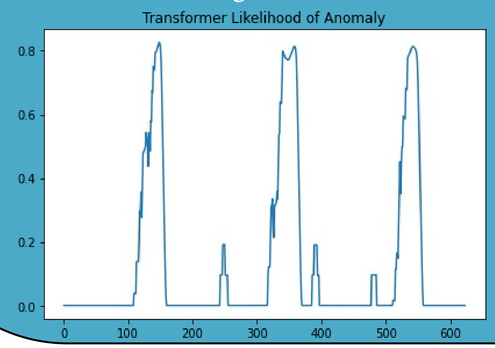


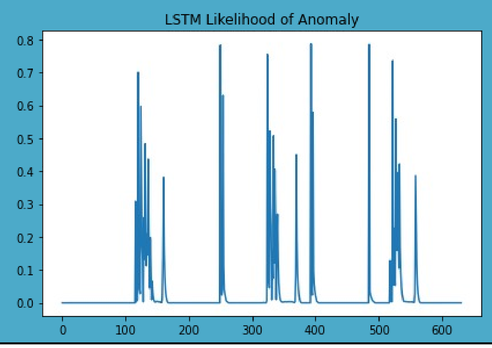
B. NN is trained with the first weeks of data. Then, evaluation take place with forward data. Next, the detected anomalies are removed from the original data, and the NN is trained again.

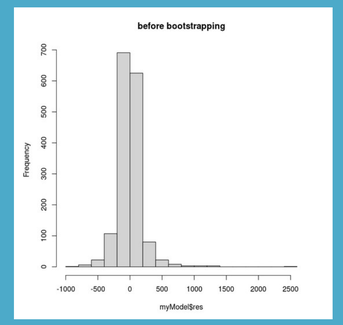
Conclusions



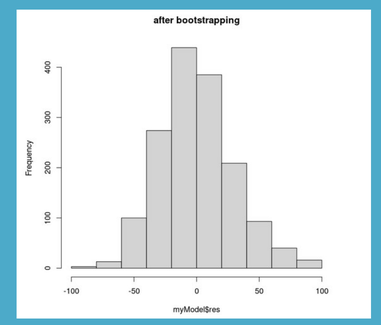
A. The first train is critical. The NN learns any anomaly in the first training.
B. The training can be updated every time a data stream is evaluated. Consequently, the evaluation tool can learn new trends.
C. LSTMs are more effective for short-term trends; transformers are more effective for long-term trends.



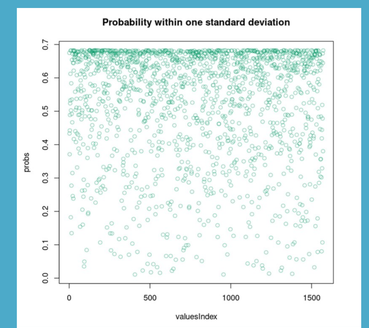




Original distribution of the data



Data distribution after being normalized



Data points' probabilities that indicate if they're outliers.

Future Goals

We can now have real-time updates regarding anomalous behavior in industrial machines to prevent breakdown.
The designed Neural Network is highly dependable on the type of scaling used on data. Future works should address this issue.
Future endeavors should study the inclusion of a preprocessing stage such as FFT (Fast Fourier Transform) analysis.

Acknowledgements

- Lucas Petralli and Dario Farfan from the Webee team
- Our CRP TA Matthew Choi
- The Data Mine Staff