

Failure prediction for APU's on a Metro System

Aaryadev Ghosalkar

October 10, 2023

Abstract

Predictive maintenance is vital for ensuring the reliable operation of metro transportation systems, enhancing passenger safety, and minimizing service disruptions. In this research, we focus on the challenging task of predicting failures on the Air Processing Unit (APU) on a metro system at least two hours in advance. An APU is a component on the metro responsible for compressing and storing compressed air for use on various functions on the metro such as the suspension. Our study encompasses a comprehensive comparison of various machine learning models, including Support Vector Machines (SVM), Convolutional Neural Networks (CNN), Long Short-Term Memory (LSTM) networks, and transformer-based models. These models are evaluated for their performance in predicting impending failures using various metrics such as accuracy, precision, recall and F1-Score for classification and Mean Squared Error (MSE) for Regression, providing valuable insights into their applicability in this context.

Additionally, we introduce a practical demonstration of a web API designed to showcase how the model would be deployed in a real world situation. This API serves as a testament to the feasibility of deploying predictive maintenance models in production settings.

It is important to note that while our research offers valuable insights into model performance and deployment readiness, we refrain from proposing a final production model. This limitation arises from the complex domain knowledge required for selecting an optimal model and the to keep the paper more general since the requirements of various transportation companies need not be the same. Nevertheless, our findings contribute to the ongoing efforts to bolster metro system reliability and safety through proactive failure prediction.

Keywords: Metro system, predictive maintenance, failure prediction, machine learning, deep learning, model deployment.

1 Literature Review

The initial phase of data extraction was undertaken by the work conducted by Veloso et al [1]. In their work they expounded upon the operational principles of diverse sensors and articulated the central objective of predicting metro system failures. Moreover, they explained the precise criteria stipulated by the railway company for the assessment of the predictive model, thus serving as an instrumental starting point for the subsequent research endeavor.

The findings of Azure ML Team at Microsoft [2], show that there exist two primary approaches to tackle our problem: binary classification or failure prediction, where the objective is to predict whether the metro system will fail within the next two hours (yes or no), and regression, which in the context of our research pertains to predicting the number of cycles or amount of time before the metro system experiences a failure, or needs to be repaired. Also known as Remaining Useful life (RUL) Prediction. This paper will undertake a comparative analysis of both these approaches.

Davari et al [3] have conducted a survey on data-driven or Machine Learning (ML) and Deep Learning (DL) approaches to PdM. In their work they have compared both approaches to we will expand on this by providing more insight into which approaches are viable in the context for metro systems.

1.1 Failure Prediction

Numerous traditional ML algorithms are applied in the realm of failure prediction. As detailed in the works of Chaudhuri et al [4], where Support Vector Machines (SVM) and various SVM variants were employed to classify vehicles into three distinct risk categories: Immediate, Long-term, and Medium-term. This approach is simple and the results are easy to understand.

Convolution Neural Networks (CNN) can be applied to time series data, typically used for images, requires a method for converting time series into image-like representations using GAF (Gramian Angular Field) feature transformation, which facilitates the conversion of one-dimensional time series data into images. Leveraging this transformation for PdM Silva reported an accuracy rate of 93% on the SF103 dataset [5].

Long Short-Term Memory (LSTM) networks have gained prominence for failure prediction. Nguyen and Medjaher employed LSTM to predict failures by quantifying the probability of system failure within a specified time window [6]. This approach holds promise for our problem, as it not only offers the potential for interpretability but also grants flexibility to the company in defining the severity thresholds for failure prediction.

1.2 Remaining Useful Life Estimation

An alternative perspective on the problem involves estimating the Remaining Useful Life (RUL), which offers a versatile approach. This approach allows for the setting of thresholds to define significant risk levels based on the transportation companies needs this is usually done by monitoring specific variables that serve as reliable indicators of system deterioration however this approach requires domain knowledge about the features that can be used to predict the degradation of the system, due to this limitation we will be calculating the time to failure, the number of hours before failure based on the known failures.

Prior to 2011, a significant portion of research in Remaining Useful Life (RUL) prediction predominantly relied on statistical methods. Si et al.'s study exemplifies the utilization of various statistical techniques that were commonly employed, including linear regression and probabilistic models such as Markov models, in the context of RUL prediction [7].

In 2016 with the rise of Deep Learning, Wang et al. compared a more data driven approach such as Deep Neural Networks (DNN) and Shallow Neural Networks (SNN) in predicting lubricant pressure for wind turbines within a farm [8]. It's worth noting that selecting the appropriate variable for monitoring often requires domain expertise. As mentioned earlier we will take a slightly different approach in our study thus circumventing the need for domain-specific knowledge in variable selection.

Several models originally designed for failure prediction have found applicability in RUL estimation. For instance, Boujamza and Elhaq explored the utilization of an attention-based LSTM model for RUL estimation using the C-MAPSS Dataset [9]. This adaptation suggests the possibility of exploring other attention-based models, such as transformers, for similar tasks in RUL estimation, expanding the scope of potential approaches for our research.

2 Background

The dataset used in this study was provided by a company and contains sensor readings from the APU on metro trains. The data consists of sensor readings collected at a frequency of 1 Hz (1 reading per second) from 20 sensors within the APU. True labels for the observed events are also included in the dataset. The following features have been extracted from the data for analysis:

Table 1: Summary Statistics and Data Types of Sensor Readings

Column	Data Type	Mean	Median	Min	Max
timestamp	DateTime	NA	NA	NA	NA
TP2	Float	1.0103	-0.008	-0.03	10.806
TP3	Float	8.9730	8.99	0.938	10.38
H1	Float	7.9479	8.86	-0.034	10.368
DV_Pressure	Float	-0.0192	-0.026	-0.036	8.11
Reservoirs	Float	1.5892	1.606	1.35	1.726
Oil_Temperature	Float	66.2864	67.2	20.8	80.175
Flowmeter	Float	20.2065	19.0123	18.8347	37.0083
Motor_Current	Float	2.1342	3.665	-0.01	9.3375
COMP	Bool	0.8829	1	0	1
DV_Electric	Bool	0.1171	0	0	1
TOWERS	Bool	0.9413	1	0	1
MPG	Bool	0.8829	1	0	1
LPS	Bool	0.0079	0	0	1
Pressure_Switch	Bool	0	0	0	0
Oil_Level	Bool	0.0031	0	0	0
Caudal_Impulses	Bool	-7.7623	0	0	1
gpsLat	Float	37.0052	-8.6583	-8.6941	41.2401
gpsLon	Float	9.9405	41.1855	0	216
gpsSpeed	Float	0.8984	0	0	1
gpsSignal	Float	NA	1	0	1

It is crucial to acknowledge that these values are approximations and are solely intended to provide a broad overview of the data distribution. Furthermore, it is essential to emphasize that the dataset is does not contain any missing values. Additional information about the data is provided by the original authors [veloso et al\[1\]](#)

3 Data preprocessing

In the data preprocessing stage, three distinct copies of the dataset are being created, each tailored for a different Machine Learning (ML) task:

1. **Binary Classification:** The primary objective here is to detect system failures. The dataset is adapted to classify instances as either "failure" or "non-failure," framing it as a binary classification problem.
2. **Multiclass Classification:** In this task, the goal is to categorize the data into one of three states: normal operating conditions, pre-failure (indicating a condition two hours before failure), and failure. This transforms the dataset into a multiclass classification problem with three distinct classes.
3. **Regression:** This approach focuses on predicting the number of hours remaining before a system failure occurs. The dataset is prepared to facilitate regression analysis, enabling the prediction of the continuous variable representing the time until failure.

It's important to consider that additional data preprocessing steps during the model training phase to address specific challenges like converting data to time series format and dealing with imbalanced or rare failures and scaling the data. These steps can be incorporated as needed during the model development and training process to optimize performance and ensure robust results.

References

- [1] B. Veloso, R. P. Ribeiro, J. Gama, and P. M. Pereira, “The metropt dataset for predictive maintenance,” *Scientific Data*, vol. 9, no. 1, p. 764, Dec 2022. [Online]. Available: <https://doi.org/10.1038/s41597-022-01877-3>
- [2] AzureML team (Microsoft). (2015, April) Predictive maintenance: Step 1 of 3 - data preparation and feature engineering. Accessed: 23/09/2023. [Online]. Available: <https://gallery.azure.ai/Experiment/Predictive-Maintenance-Step-1-of-3-data-preparation-and-feature-engineering-2>
- [3] N. Davari, B. Veloso, G. Costa, P. Pereira, R. Ribeiro, and J. Gama, “A survey on data-driven predictive maintenance for the railway industry,” *Sensors*, vol. 21, p. 5739, 09 2021.
- [4] A. Chaudhuri, “Predictive maintenance for industrial iot of vehicle fleets using hierarchical modified fuzzy support vector machine,” *arXiv preprint arXiv:1806.09612*, 2018.
- [5] W. Silva, “Cnn-pdm: A convolutional neural network framework for assets predictive maintenance,” 2019. [Online]. Available: <https://api.semanticscholar.org/CorpusID:260505498>
- [6] K. T. Nguyen and K. Medjaher, “A new dynamic predictive maintenance framework using deep learning for failure prognostics,” *Reliability Engineering & System Safety*, vol. 188, pp. 251–262, 2019. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0951832018311050>
- [7] X.-S. Si, W. Wang, C.-H. Hu, and D.-H. Zhou, “Remaining useful life estimation – a review on the statistical data driven approaches,” *European Journal of Operational Research*, vol. 213, no. 1, pp. 1–14, 2011. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0377221710007903>
- [8] L. Wang, Z. Zhang, H. Long, J. Xu, and R. Liu, “Wind turbine gearbox failure identification with deep neural networks,” *IEEE Transactions on Industrial Informatics*, vol. 13, no. 3, pp. 1360–1368, 2016.
- [9] A. Boujamza and S. Lissane Elhaq, “Attention-based lstm for remaining useful life estimation of aircraft engines,” *IFAC-PapersOnLine*, vol. 55, no. 12, pp. 450–455, 2022, 14th IFAC Workshop on Adaptive and Learning Control Systems ALCOS 2022. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S2405896322007546>