

# Project Overview and Brief Litreature Review

Aaryadev Ghosalkar

September 22, 2023

## 1 Metro System Failure Prediction

### Overview

This project aims to develop a predictive model for detecting and predicting failures in a metro transportation system. The project leverages machine learning and data analysis techniques to enhance the system's reliability and safety as well as providing a framework to deploy these models into production.

### Literature Review

The initial phase of data extraction, along with the comprehensive delineation of the task and the methodology employed for data acquisition, encompassing the data format obtained from all sensors installed within the metro system, was undertaken by the work conducted by [1]. In their work, [1] expounded upon the operational principles of diverse sensors and articulated the central objective of predicting metro system failures. Moreover, [1] 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.

Another comparison was done by [2] which explores contemporary approaches to address this issue. It underscores the need for Deep Learning (DL) due to the development of advanced sensors generating substantial data, making traditional Machine Learning (ML) methods inadequate for modern Predictive Maintenance (PdM) challenges. Common methods discussed including Artificial Neural Networks (ANN) for binary classification of system failure, Convolutional Neural Networks (CNN) for transforming 1D time series data into images and using image classification for failure prediction. Recurrent Neural Networks (RNN), particularly Long Short-Term Memory (LSTM) networks, suitable for both regression and classification in PdM.

The AzureML team at Microsoft [3] introduced data preparation techniques for Predictive Maintenance (PdM). They addressed three key PdM tasks: binary

classification, multiclass classification, and regression, the latter also referred to as Remaining Useful Life (RUL) prediction.

The application of Convolutional Neural Networks (CNN) to time series data, typically used for images, requires a method for converting time series into image-like representations. Reference [4] introduced the concept of GAF (Gramian Angular Field) feature transformation, which facilitates the conversion of one-dimensional time series data into images. Leveraging this transformation, they achieved an impressive accuracy rate of 93% on the SF103 dataset.

## 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] 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.
- [3] AzureML team (Microsoft). (2015, April) Predictive maintenance: Step 1 of 3 - data preparation and feature engineering. Accessed: September 22, 2023. [Online]. Available: <https://gallery.azure.ai/Experiment/Predictive-Maintenance-Step-1-of-3-data-preparation-and-feature-engineering-2>
- [4] W. Silva, “Cnn-pdm: A convolutional neural network framework for assets predictive maintenance,” 2019. [Online]. Available: <https://api.semanticscholar.org/CorpusID:260505498>