# Weekly Progress Report

Name: Hemanth Yarra

Domain: Predictive Maintenance / Industrial Manufacturing

Date of submission: 11-06-2025

Week Ending: 02

## I. Overview:

The second week focused on building foundational models for both predictive maintenance and quality prediction problems. This involved completing data preprocessing, generating structured sequences, and training initial models. Evaluation frameworks were also put in place to assess model performance and prepare for optimization in the next phase.

## II. Achievements:

1. Turbofan Engine RUL Prediction - Completed feature scaling using Min-Max Normalization. - Constructed multivariate time-series sequences using a rolling window strategy. - Designed and trained a baseline LSTM model using TensorFlow/Keras on FD001 subset. - Implemented early stopping and dropout layers to prevent overfitting. - Evaluated model using RMSE, and visualized actual vs. predicted Remaining Useful Life (RUL).

2. Mining Process Quality Prediction - Addressed missing data using forward fill and KNN-based imputation. - Merged multiple sources and aligned time-series features with varying frequencies. - Built regression models using Random Forest and XGBoost. - Conducted feature importance analysis with SHAP to understand variable impact. - Measured performance using RMSE, MAE, and R² on test data.

## III. Challenges:

1. Temporal Alignment: - Aligning different time-sampled variables in the flotation dataset was complex and required extensive interpolation and resampling.

2. Model Instability: - LSTM model performance was unstable with small sequence sizes; tuning required for window length and batch size.

3. Feature Importance Misleading: - SHAP occasionally highlighted irrelevant features due to collinearity, prompting a re-check of preprocessing steps.

## IV. Lessons Learned:

- Gained practical knowledge in preparing time-series datasets for deep learning.

- Understood the importance of evaluation metrics in both regression and sequence modeling.

- Improved skills in TensorFlow/Keras for LSTM model development.

- Learned advanced imputation and time alignment techniques for real-world datasets.

- Gained experience in using SHAP for interpretability in industrial ML settings.

## V. Next Week's Goals:

1. Turbofan Engine Project

- Tune LSTM model architecture using hyperparameter search.

- Extend training to additional FD subsets (FD002, FD003).

- Begin experiments with attention-based sequence models.

2. Mining Process Project

- Refine feature engineering pipeline.

- Improve model robustness and reduce overfitting.

s- Explore deployment-ready pipeline using real-time data simulation..

## VI. Additional Comments:

Week 2 has shown meaningful progress with initial models and performance baselines established. Dealing with real-world industrial datasets continues to provide valuable learning and exposure to practical ML workflows. Optimizing and preparing for deployment will be the core focus of Week 3.