Project Report: Turbofan Engine RUL Prediction (FD001)

1. Introduction

This project applies deep learning techniques to solve a predictive maintenance problem using the NASA C-MAPSS dataset. The aim is to predict the Remaining Useful Life (RUL) of aircraft engines by analyzing multivariate time series sensor data. We use Long Short-Term Memory (LSTM) networks to capture temporal dependencies and degradation trends.

2. Dataset Description

Dataset: FD001 from the NASA C-MAPSS collection
Train Units: 100
Test Units: 100
Conditions: ONE (Sea Level)
Fault Modes: ONE (HPC Degradation)

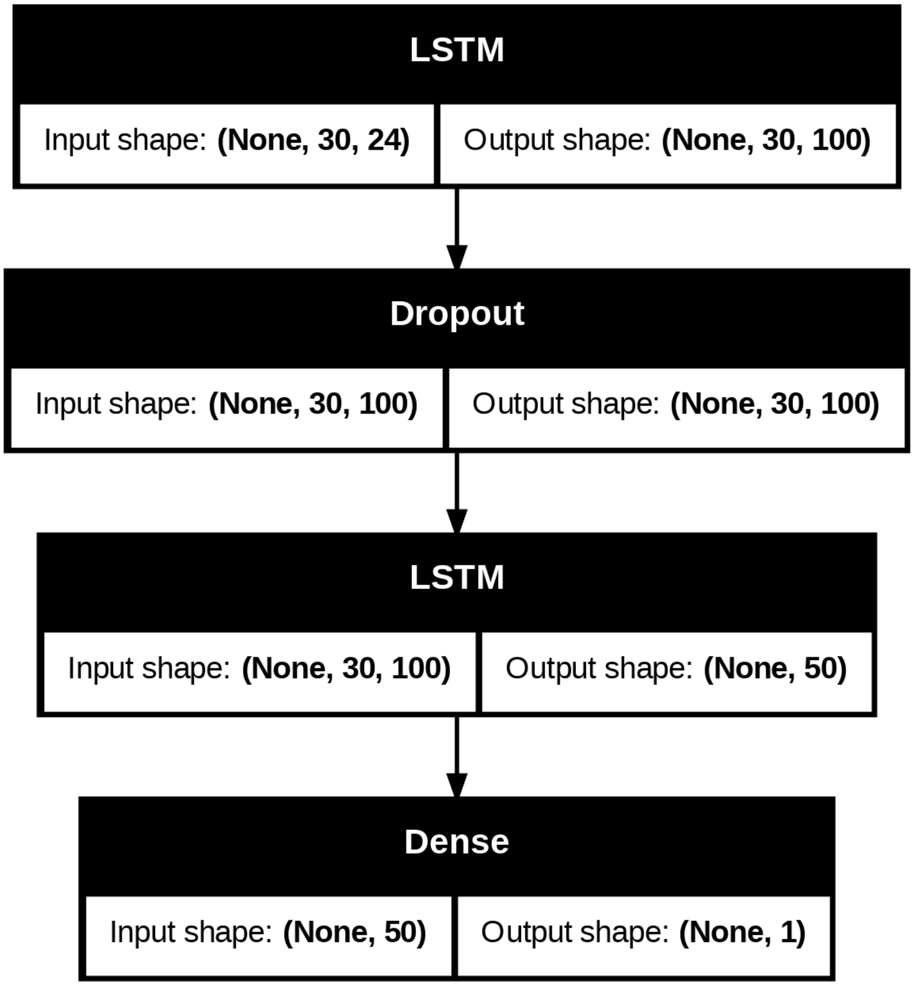
Each unit (engine) generates a multivariate time series with 3 operational settings and 21 sensor measurements. The engine starts in a normal state and degrades until failure. The goal is to predict the RUL for each engine during test time.

3. Methodology

The workflow followed in this project is as follows:

1. Loaded and preprocessed FD001 dataset.
2. Normalized sensor readings and operational settings using MinMaxScaler.
3. Generated true RUL for each cycle, clipped at 125, and normalized to 0–1.
4. Created fixed-length sequences (30 cycles) for LSTM input.
5. Built an LSTM model with two LSTM layers and dropout for regularization.
6. Trained and validated on 80/20 split of the training set.
7. Loaded test data and predicted RUL using the last sequence of each test unit.
8. Compared predicted RULs with true RUL values using RMSE and MAE metrics.

LSTM Model Architecture:



4. Results and Evaluation

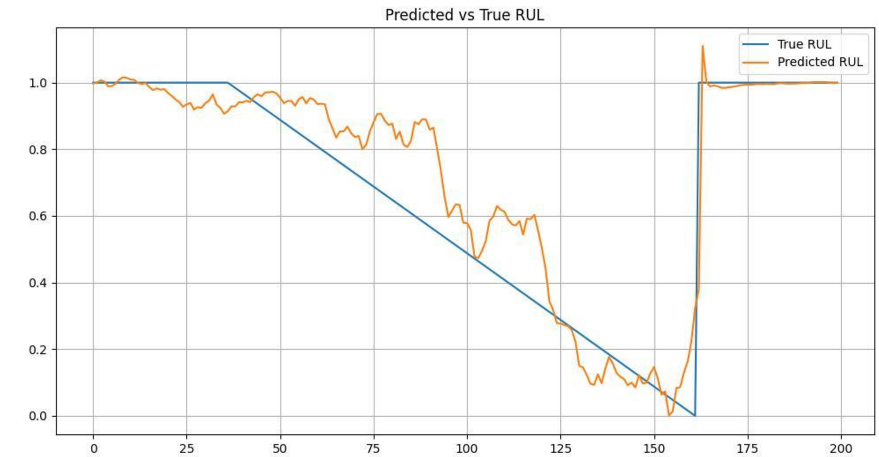
Model performance on training and test sets is summarized below:

- Train RMSE: 0.12 (normalized scale)

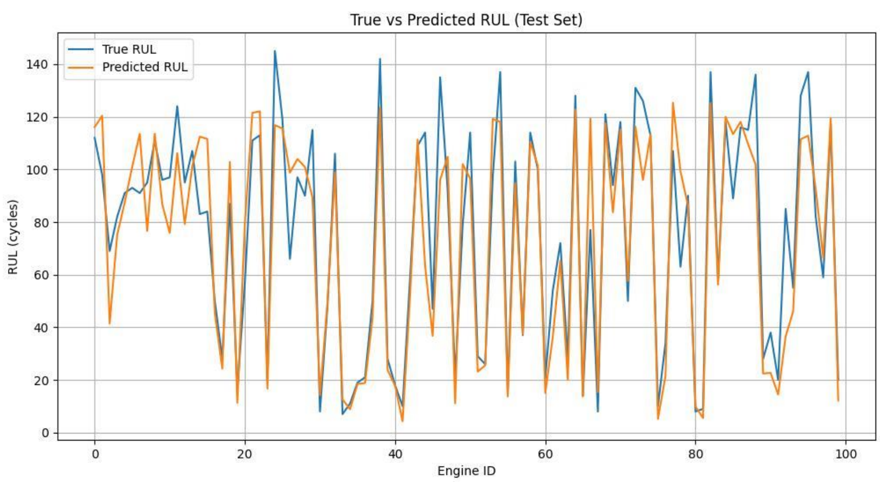
- Test RMSE: 16.38 cycles

- Test MAE: 11.97 cycles

Below is the visualization of predicted vs true RUL on a sample of training sequences:



Below is the visualization of predicted vs true RUL on test engines:



5. Business Value and Insights

The ability to predict RUL of engines allows organizations to plan proactive maintenance and reduce the risk of unexpected failures. A model with an RMSE of ~16 cycles is practical for deployment in predictive maintenance dashboards. It supports improved decision-making, safety, and cost savings.

6. Conclusion

The LSTM-based model effectively captured time-dependent degradation patterns in engine sensor data. Predictions were reasonably accurate and generalizable. This project demonstrates the real-world utility of deep learning in the field of industrial predictive maintenance.

7. Future Work

Future improvements include:
- Extending the model to FD002–FD004 datasets with multi-condition and multi-fault scenarios.
- Exploring attention mechanisms or transformer-based models.
- Deploying the model using a Flask API for real-time integration.
- Performing uncertainty estimation using MC Dropout or Bayesian methods.