🗎 Weekly Progress Report

Name: Abhishek Kumar Gupta
Domain: Predictive Maintenance using ML
Date of submission: 30/08/25

Week Ending: 04

I. Overview

In Week 4, Project 6 – *Turbofan Engine Remaining Useful Life (RUL) Prediction* was completed. The focus was on:

- Running final experiments across multiple machine learning and deep learning models.
- Performing comparative evaluation using RMSE, MAE, R^2, and NASA scoring function.
- Separating production-ready models (LSTM/GRU) vs research models (BiLSTM, CNN-LSTM, Attention, Transformer).
- Preparing deployment workflow using Streamlit for interactive predictions.
- Documenting the pipeline, final results, and future enhancement scope.

II. Achievements

◇ Model Implementations & Results (NASA FD001 Dataset)

Production Models (Deployment Candidates):

Model	Seq Len	Units	Dropout	RMSE	MAE	R^2	NASA	Notes
LSTM	40	128	0.3	12.59	9.12	0.91	9628	Strong, stable
GRU	50	128	0.3	11.47	8.44	0.92	6234	Best performer
LSTM	50	128	0.3	12.47	9.18	0.91	8375	Consistent
GRU	40	128	0.5	12.27	9.06	0.91	7908	Robust generalization

Research Models (Explored but not stable):

Model	RMSE	MAE	R^2	NASA	Notes
BiLSTM	14.22	10.00	0.88	15265	Slight overfitting
CNN-LSTM	13.91	9.82	0.89	13341	Test collapse
Attention-LSTM	20.63	17.19	0.76	19071	Poor stability
Transformer	16.85	12.88	0.84	25928	Inconsistent

☑ Conclusion: GRU with seq len=50, units=128, dropout=0.3, LR=0.0005 was the best overall model (RMSE 11.47, R^2 0.92, NASA 6234).

◇ Deployment Readiness

- Built a Streamlit dashboard for interactive predictions.
  - File upload option for new engine sensor data.
  - Model predicts RUL and visualizes degradation curves.
  - Batch mode for predicting RUL across multiple engines daily.
- Models saved as .h5 and scaler.pkl for reproducibility.
- Final pipeline supports both research experimentation and real-world deployment.

III. Challenges

- NASA Score Consistency: Some models (AttnLSTM, Transformer) showed poor generalization → excluded from deployment.
- Overfitting: CNN-LSTM & BiLSTM overfit; solved partially with dropout & tuning.
- Computational Load: Training deep models required careful LR/batch size tuning.

IV. Lessons Learned

- Sequence models (LSTM, GRU) capture temporal degradation patterns effectively.
- Baseline ML models (Linear Regression, Random Forest) offer interpretability but lack predictive power.
- Hybrid architectures (CNN-LSTM, Transformer) need larger datasets and regularization for stable performance.
- Model interpretability and visualizations (Predicted vs Actual RUL) are crucial for client confidence.

V. Final Outcome

- Delivered a production-ready predictive pipeline for turbofan RUL estimation.
- Identified GRU & LSTM as deployment candidates.
- Streamlit dashboard prepared for client-facing demonstration.
- Results documented with structured comparisons of production vs research models.

VI. Future Enhancements

- Extend training across all NASA datasets (FD001–FD004) to improve robustness.
- Implement explainable AI (SHAP/attention maps) for sensor contribution analysis.
- Add uncertainty estimation in predictions for risk-aware decision-making.
- Wrap trained models in a REST API (FastAPI/Flask) for enterprise integration.
- Explore federated learning for scalability across multiple sites.