Weekly Progress Report

Name: Agamya David

Domain: Data Science and Machine Learning

Date of submission: 12 August 2025

Week Ending: 04

I. Tasks Completed & Milestones Achieved

Following the significant progress made in Week 3, this week was dedicated to refining the models and finalizing the experimental code. The primary goal was to ensure the robustness of the previous week's results and better understand the model's performance.

1. **Model Refinement and Verification:**
    - Spent time making minor adjustments to the LSTM model to see if performance could be improved further.
    - After several iterations, it was confirmed that the results achieved in Week 3 (especially the strong performance on the FD003 dataset) were the most effective. The decision was made to stick with that version of the model as the final one.
2. **Code Finalization and Cleanup:**
    - Cleaned and organized all four Jupyter Notebooks, ensuring a clear, sequential flow from data loading to model evaluation.
    - Renamed the notebooks to a descriptive format (e.g., Turbofan_Analysis_FD001_(Base_Case).ipynb) to clearly indicate the purpose of each file.
    - Added comments and Markdown explanations throughout the notebooks to make the code and methodology easy to understand and reproduce.

II. Challenges and Hurdles

1. **Diminishing Returns on Tuning:**
    - **Challenge:** While some minor improvements were made to the LSTM model, it became clear that significant further gains would require a much larger model and substantially more training time, which was not feasible.

- **Solution:** I focused on solidifying the existing results rather than pursuing marginal improvements. This involved documenting the most effective model architecture and parameters found during Week 3's successful experiments.

2. **Ensuring Code Reproducibility:**
    - **Challenge:** Making sure that the notebooks would run correctly from start to finish without errors and that all dependencies were documented.
    - **Solution:** I created a requirements.txt file listing all necessary Python libraries. I also restarted and ran each notebook from a clean kernel to confirm that all outputs could be reproduced reliably.

III. Lessons Learned

1. **Knowing When to Finalize a Model:** This week provided a practical lesson in understanding the point of diminishing returns. I learned that while it's possible to keep tuning a model indefinitely, it's important to recognize when a model is performing well enough and to focus on finalizing and documenting that successful result.
2. **The Importance of a Solid Baseline:** The strong performance of the Random Forest model in Week 3 served as an excellent benchmark. It provided a clear target for the LSTM and helped in making the final decision to stick with the LSTM, especially given its superior performance on the more complex datasets.
3. **Clean Code is a Key Deliverable:** The process of cleaning and commenting on the notebooks reinforced that the final code itself is a critical part of the project. A well-organized and documented notebook is essential for others (and my future self) to understand and build upon the work.