Weekly Progress Report

Name: Agamya David
Domain: Data Science and Machine Learning
Date of submission: 21 July 2025

Week Ending: 01

I. Overview:

This week, the primary focus was on selecting a project and becoming familiar with its objectives and datasets. Efforts were concentrated on understanding the problem of predictive maintenance and exploring the structure of the provided data files.

II. Achievements:

1. **Project Selection and Understanding:**
    - Selected the project: **"Predictive Maintenance: Turbofan Engine Failure Prediction"**.
    - Studied the project brief to understand the core objective: predicting the Remaining Useful Life (RUL) of turbofan engines.
    - Successfully set up a basic Python development environment.
2. **Dataset Familiarization:**
    - **Name of the project:** Predictive Maintenance: Turbofan Engine Failure Prediction
    - Located and reviewed the complete set of files for the project.
    - Understood the purpose of the different datasets (FD001 through FD004) and the roles of the train, test, and RUL files within each set.
3. **Initial Setup:**
    - Outlined a plan to begin analysis, starting with the FD001 dataset as it represents the simplest scenario.

III. Challenges:

1. **Data Complexity:**
    - Recognizing that the dataset is comprehensive, with multiple files representing different operational conditions and fault modes.
    - Understanding the meaning of the 26 columns of data for each engine requires careful reading of the documentation.
2. **Project Scoping:**
    - The initial challenge is to scope the project appropriately, starting with the simplest dataset (FD001) before tackling the more complex ones.

IV. Learning Resources:

1. **Project Documentation:**
    - Utilized the official readme.txt file associated with the turbofan dataset to understand the data dictionary and file structure.
    - Reviewed the high-level project description to grasp the overall goal.
2. **Technical Documentation:**
    - Briefly reviewed documentation for the Pandas library in preparation for loading and analyzing the data.

V. Next Week's Goals:

1. **Data Loading and Initial Exploration:**
    - Load the train_FD001.txt data into a Pandas DataFrame.
    - Perform initial Exploratory Data Analysis (EDA) by generating descriptive statistics (.describe()).
    - Visualize the sensor data for a few sample engines to identify any obvious trends as they approach failure.
2. **Begin Preprocessing Plan:**
    - Based on the initial exploration, start to plan the first steps of data preprocessing, such as calculating the RUL for the training set.

VI. Additional Comments:

This first week was productive in setting a clear direction for the project. I have a good understanding of the data I'll be working with and have a clear plan for the initial analysis to be performed next week.