DRAFT

**1. Background & Motivation**

Modern logistics and manufacturing systems are increasingly dependent on automated machines and sensor-equipped infrastructure. Unexpected machine failure can lead to costly downtime, safety issues, and disrupted supply chains.

This motivates the development of data-driven strategies that can **predict potential failures in advance**, enabling smarter maintenance and system optimization.

**2. Project Objective**

The proposed project will aim to:

* Explore the use of time-series sensor data for early detection of equipment failure
* Investigate statistical and machine learning techniques for anomaly detection
* Build a basic yet effective simulation of a predictive maintenance pipeline

**3. Methodology (Preliminary Plan)**

| **Step** | **Planned Approach** |
| --- | --- |
| **Data Exploration** | Start with small synthetic or public datasets related to machine sensors |
| **Anomaly Detection** | Use rolling average and z-score to flag sensor drift |
| **Visualization** | Plot sensor patterns and failure risk zones |
| **Prediction Modeling** | Apply binary classification (XGBoost or Logistic Regression) |
| **Validation** | Evaluate detection logic using simulated failure windows |

**4. Data Sources (Tentative)**

| **Dataset** | **Description** | **Source** |
| --- | --- | --- |
| **CMAPSS Jet Engine Sensor Data** | Simulated time-series data of jet engines with sensor readings and failure behavior | NASA via Kaggle |
| **UCI Machine Learning Repository** | Alternative datasets for time-series and industrial processes | UCI.edu |
| **Self-simulated sensor data** | Small test cases to validate logic before scaling to CMAPSS | Generated via Python |

**5. Tools & Skills To Be Used**

* Python (pandas, numpy, matplotlib, seaborn)
* XGBoost, scikit-learn for ML modeling
* Google Colab for computation
* GitHub for versioning and documentation

**6. Expected Outcome**

* An interpretable anomaly detection pipeline
* Key insights about sensor drift and failure timing
* Visualizations to support findings
* (Optional) ML model for predicting anomaly risk

📌 **Note:** This proposal serves as a foundation. Actual methods and scope may adjust based on data availability and performance results.