Predictive Maintenance Solution for Device Failure Using Gradient Boosting Machines with Synthetic Data

1. Introduction

In industrial environments, it is important to maintain operational efficiency and minimize downtime. Predictive maintenance, which leverages sensor data to predict failures, offers significant advantages over traditional preventive strategies by performing maintenance only when necessary. This project aims to develop machine learning models that predict device failure using synthetic datasets that reflect real-world predictive maintenance scenarios. The dataset contains 10,000 data points with 14 features and includes various operational parameters and product quality.

2. Methodology

The project will proceed through the following stages:

2.1. Data Collection and Preprocessing

Load the synthetic dataset containing 10,000 data points with 14 features.

Clean and preprocess the data by handling missing values, normalizing sensor readings, and encoding categorical variables.

Ensure no data leakage by carefully separating the target variable (failure or not) from the features.

2.2. Exploratory Data Analysis (EDA)

Perform EDA to understand data distributions, correlations, and feature importance.

Visualize data trends, relationships, and anomalies related to device failures.

2.3. Model Development with Gradient Boosting Machines (GBM)

Split the data into training and testing sets.

Develop a GBM model to predict device failures.

Optimize hyperparameters using cross-validation to enhance model performance.

2.4. Model Evaluation

Evaluate the GBM model using metrics such as accuracy, precision, recall, F1-score, and ROC-AUC.

Focus on minimizing both false positives (unnecessary maintenance) and false negatives (missed failures).

3. AI and ML Algorithm:

Gradient Boosting Machines (GBM)

GBM is chosen for its ability to handle complex patterns and relationships in the data. It builds models sequentially to correct errors made by previous models, making it highly effective for binary classification tasks like predicting device failures.

4. Problem Solving:

The project will address the following problems:

Accurate Prediction of Device Failures

Utilize historical sensor data to train the GBM model to identify patterns and signals indicative of impending device failures, enabling proactive maintenance actions.

Minimizing False Positives and False Negatives

Optimize the GBM model to balance false positives and false negatives, ensuring maintenance tasks are performed only when necessary and avoiding unnecessary interventions.

Cost Savings and Operational Efficiency

Implement predictive maintenance to reduce maintenance costs by avoiding routine maintenance tasks and minimizing unplanned downtime, thus enhancing operational efficiency and productivity.

Scalability and Adaptability

The GBM solution will be scalable to accommodate a growing fleet of devices and adaptable to evolving sensor technologies and maintenance practices.

5. Dataset Description

The dataset consists of 10,000 data points with the following features:

UID: Unique identifier ranging from 1 to 10000.

Productid: Consisting of a letter L, M, or H for low, medium, and high product quality variants, respectively, and a variant-specific serial number.

Air temperature [K]: Generated using a random walk process, normalized to a standard deviation of 2 K around 300 K.

Process temperature [K]: Generated using a random walk process, normalized to a standard deviation of 1 K, added to the air temperature plus 10 K.

Rotational speed [rpm]: Calculated from power of 2860 W, overlaid with normally distributed noise.

Torque [Nm]: Normally distributed around 40 Nm with a standard deviation of 10 Nm.

Tool wear [min]: Quality variants H/M/L add 5/3/2 minutes of tool wear to the used tool in the process.

Machine failure: Binary label indicating whether the machine has failed.

6. Conclusion

By the end of this project, we aim to deliver a robust and reliable predictive maintenance solution using Gradient Boosting Machines that enhances the maintenance strategy of the company, leading to significant cost savings and improved device uptime.