Predictive Maintenance System - Technical Documentation

This document presents a professional overview of the **Predictive Maintenance System**. The system is designed to proactively predict equipment failures, optimize maintenance schedules, and reduce unplanned downtimes using state-of-the-art machine learning and data-driven approaches.

1. System Approach

The Predictive Maintenance System follows a structured machine learning pipeline: **Data Acquisition:** Supports multiple data sources including Microsoft Azure Predictive Maintenance Dataset, NASA Bearing/Battery datasets, generic CSV files, and synthetic data generation. **Preprocessing:** Handles missing values, standardizes datetime formats, encodes categorical variables, and constructs engineered features. **Feature Engineering:** Incorporates lag features, rolling statistics, and error count summaries across multiple time windows (3-hour and 24-hour) to capture temporal behavior of machine telemetry. **Modeling:** Applies anomaly detection and supervised learning algorithms to predict upcoming equipment failures. **Reporting & Insights:** Generates failure distribution reports, highlights high-risk machines, and recommends proactive maintenance actions.

2. Data Handling

The system demonstrates flexibility in handling heterogeneous datasets:
Azure Dataset: Integrates telemetry, error logs, maintenance schedules, and failure records into a unified dataset. **NASA Dataset:** Adapts to bearing or battery datasets by aligning telemetry parameters with the system schema. **CSV Imports:** Automatically detects date and numeric fields, mapping them into standardized telemetry variables. **Synthetic Data:** Enables system testing without external datasets by simulating telemetry, errors, and component replacements for multiple machines. The resulting dataset is enriched with both raw telemetry (voltage, rotation, pressure, vibration) and engineered variables that improve predictive performance.

3. Model Selection

The system adopts a hybrid modeling strategy:
Anomaly Detection: Implements Isolation Forest, Local Outlier Factor, One-Class SVM, and Elliptic Envelope to detect abnormal operating conditions. **Failure Prediction:** Utilizes a diverse set of supervised classifiers: Random Forest – robust baseline for feature importance analysis XGBoost – gradient boosting model for high predictive accuracy Logistic Regression – interpretable model for binary/multiclass failures Support Vector Machines (SVM) – effective for complex decision boundaries **Feature Importance:** Random Forest and XGBoost highlight critical predictors such as vibration, pressure, and machine age. Models are trained on stratified train-test splits with standardized feature scaling. Performance metrics include accuracy, precision, recall, F1-score, and feature importance rankings.

4. Key Insights

Analysis across datasets and models reveals several practical insights:
Equipment age and accumulated error events strongly influence failure likelihood. Vibration and pressure anomalies are leading indicators of component failures. Anomaly detection methods serve as an early-warning system, flagging machines at risk before supervised models confirm failure predictions. Maintenance reports identify high-risk machines and common failure modes, guiding efficient allocation of inspection and repair resources. Proactive strategies informed by predictions

can reduce downtime, lower maintenance costs, and extend equipment lifespan.

5. Conclusion

The **Predictive Maintenance System** provides a research-grade, extensible, and production-ready framework for industrial failure prediction. By combining anomaly detection, supervised machine learning, and automated reporting, the system empowers organizations to transition from reactive to proactive maintenance strategies, significantly improving operational efficiency and reliability.