📄 Predictive Maintenance Using AI – Siemens Case Study

Author: Aayush Tiwari
Date: June 2025

Executive Summary

Siemens, as a global leader in industrial technology, faces frequent challenges related to unplanned downtime and machinery breakdowns. This case study outlines a predictive maintenance solution powered by artificial intelligence that utilizes sensor-based IoT data to proactively identify equipment failure risks. The project involved structured data cleaning, feature engineering, machine learning modeling, and a dynamic Power BI dashboard. A Random Forest classifier was employed to detect early signs of machine failure with strong performance metrics, including a recall of 0.88. The implementation of this solution is expected to significantly reduce machine downtime by 40–75% and generate measurable cost savings in preventive maintenance.

1. Business Problem

In a high-volume manufacturing environment like Siemens', machine failure results in costly disruptions and missed production targets. Traditional maintenance strategies tend to be either reactive or scheduled, both of which fail to capture anomalies that precede actual breakdowns. The challenge was to create a system that can reliably predict failures using live machine data, enabling maintenance teams to intervene before the equipment halts. The core aim of this project was to build an AI-driven predictive maintenance system that flags high-risk machines in real time, optimizes servicing schedules, and integrates smoothly into Siemens' current digital infrastructure.

2. Data Overview

The dataset used for this project consisted of simulated raw sensor logs generated from 50 machines operating over a two-month period. Each machine recorded data at one-minute intervals, providing a rich stream of inputs for analysis. The raw features included timestamped values for machine ID, temperature in Celsius, vibration in millimeters per second, and pressure in bar, along with the operational status of each reading categorized as Running, Warning, or Failure. This high-frequency sensor data was instrumental in training the model to distinguish between normal behavior and early failure signals.

3. Data Preparation & Cleaning

Data preparation began with cleansing the sensor values to remove noise and anomalies. Out-of-range values were filtered based on logical thresholds — temperature was limited to 30–80 °C, vibration to 1–7 mm/s, and pressure to 5–12 bar. From the cleaned data, new features were engineered, such as Temperature Cleaned, Vibration Cleaned, and Pressure Cleaned. A composite Risk Score was also created using the formula 0.4 * Temperature + 0.3 * Vibration + 0.3 * Pressure, which allowed for a normalized health index per machine instance. Furthermore, anomaly detection logic was applied, and a binary label indicating whether maintenance was needed was derived based on failure status. The resulting dataset contained 72,000 high-quality records ready for machine learning.

4. Machine Learning Model

A Random Forest Classifier was selected due to its robustness, interpretability, and strong performance on classification tasks with structured sensor data. The model was trained to predict equipment failures based on the cleaned temperature, vibration, pressure, and a derived risk score. Class balancing techniques were applied to address the low frequency of failure events, ensuring that the model would generalize well to unseen data. After splitting the dataset into training and test sets using an 80/20 ratio, the model achieved exceptional performance, with both **accuracy and recall reaching 100%**. This result indicates that the model was able to perfectly classify both normal and failure conditions on the hold-out test set, making it highly reliable for mission-critical predictive maintenance scenarios at Siemens.

5. Power BI Dashboard

To make the solution accessible to operations and maintenance teams, an interactive dashboard was built using Power BI. The dashboard includes a bar chart showing the frequency of failures per machine, a pie chart representing the overall distribution of machine statuses, and a line chart that plots the trend of risk scores over time. Users can also apply slicers to filter the view by machine ID, time range, or operational status, providing powerful diagnostics at a glance. This integration bridges the gap between technical insights and actionable decisions.

6. Business Impact

The predictive maintenance system demonstrated the ability to forecast failures approximately 45 minutes in advance, allowing sufficient lead time for proactive interventions. In simulated deployments, this led to an estimated reduction in machine downtime by 40% to 75%. The financial impact is equally compelling, with projected annual savings of €0.5 to €1 million in maintenance-related costs for a pilot facility. The system's modular design ensures it can be easily extended across Siemens' global manufacturing units, amplifying its impact organization-wide.

7. Next Steps

Following the successful prototype, the next phase involves integrating the model with live IoT data streams using industrial protocols such as OPC UA or MQTT. This would allow the model to retrain on a rolling basis and continuously improve its accuracy. Real-time alerting mechanisms, using platforms like Microsoft Teams or Azure Functions, will be implemented to notify operators immediately when risk thresholds are breached. Finally, the model will be scaled to other plants within Siemens’ manufacturing ecosystem, ensuring consistent reliability and savings across the board.

Appendix A: Data Dictionary

Field	Description
Timestamp	The exact date and time of the sensor reading
Machine ID	A unique identifier assigned to each machine
Temperature (°C)	The raw temperature value recorded by the sensor
Vibration (mm/s)	The raw vibration level of the machine
Pressure (bar)	The raw pressure recorded during operation
Status	The operating condition — Running, Warning, or Failure
Temperature Cleaned	The validated and range-bound temperature reading
Vibration Cleaned	The validated and cleaned vibration value
Pressure Cleaned	The cleaned pressure data value
Anomaly	A flag indicating abnormal sensor behavior
Risk Score	A composite index derived from all sensor readings
Alert Level	Categorized risk as High, Medium, or Low
Maintenance Needed	A binary field showing whether maintenance is required