Predictive Maintenance for CNC Milling Machines Using XGBoost

Abstract

Unplanned downtime in industrial CNC milling machines leads to significant productivity losses and maintenance costs. Traditional reactive maintenance strategies fail to anticipate failures before they occur. This project presents an AI-driven predictive maintenance system that utilizes machine sensor data and an XGBoost classifier to detect early failure risks. The solution integrates cloud-based model training, real-time diagnostics, and an interactive web dashboard, offering a scalable and practical maintenance support tool for modern manufacturing environments.

1. INTRODUCTION

CNC milling machines operate under high mechanical and thermal stress, making them susceptible to failures such as tool wear, overheating, and torque overload. These failures often occur unexpectedly, resulting in costly downtime.
The objective of this project is to transition from reactive maintenance to predictive maintenance by leveraging machine learning techniques. By continuously analyzing sensor data, the system predicts failure risks before they occur, enabling proactive intervention

2. Problem Statement

Industrial CNC machines frequently suffer from:

- Sudden mechanical failures
- Thermal overload and excessive tool wear
- Inefficient scheduled maintenance

These issues lead to:

- Production interruptions
- Increased operational costs
- Reduced equipment lifespan

The challenge is to design a **data-driven system** that can identify early warning signs of failure and present them in a form usable by operators and maintenance engineers.

3. Methodology

3.1 Data Source

The project uses the **AI4I 2020 Predictive Maintenance Dataset**, which provides realistic industrial sensor telemetry, including:

- Air Temperature
- Process Temperature
- Rotational Speed
- Torque
- Tool Wear

These parameters closely resemble real CNC machine operating conditions.

3.2 Machine Learning Model

An **XGBoost Classifier** was selected due to:

- High performance on tabular sensor data
- Robust handling of non-linear relationships
- Ensemble learning capability using decision trees

The model evaluates sensor thresholds and outputs a **failure risk score** for each machine state.

3.3 Cloud Training Environment

Model training was performed using **AWS SageMaker**, ensuring scalability and professional deployment standards.

Training Details:

- Training Job ID: sagemaker-xgboost-2025-12-19-08-50-24-690
- Training Time: ~3 minutes
- Model Type: Ensemble of hundreds of decision trees

Each tree contributes a vote toward the final failure prediction, improving accuracy and robustness.

3.4 System Deployment

To ensure usability, results were deployed through an interactive dashboard developed using **Streamlit**.

Dashboard Features:

- Real-time failure risk visualization
- Interactive Plotly charts with hover diagnostics
- Custom CSS for enhanced readability
- Downloadable diagnostic reports

4. Step-by-Step Technical Implementation

Step 1: Data Preparation

Sensor data was cleaned, normalized, and analyzed to identify correlations between operating conditions and machine failure events.

Step 2: Model Training

The XGBoost model was trained using AWS SageMaker's managed infrastructure. The ensemble approach allows multiple decision trees to evaluate different sensor thresholds simultaneously.

Example logic:

Torque > 50 Nm → Increased failure probability
Tool Wear > threshold → Elevated risk score

Step 3: Streamlit Application Development

A professional UI was developed to:

- Display risk scores using red-gradient visualization
- Highlight critical sensor triggers
- Explain AI decision logic in a "How the AI Works" section

This makes the system understandable even for non-AI experts.

The trained XGBoost predictive maintenance model was deployed through an interactive web dashboard developed using Streamlit. The primary objective of the deployment phase was to translate complex machine learning outputs into an intuitive and operator-friendly interface suitable for industrial use. The dashboard allows users to upload CNC sensor datasets in CSV or Parquet format and immediately receive failure risk predictions.

The interface, shown in Figure 1, includes a dedicated section explaining how the AI model works using decision tree logic. This explainability component ensures transparency by describing how multiple trees evaluate sensor thresholds such as torque, temperature, and tool wear to determine machine health. Custom CSS styling was applied to enhance readability and ensure high contrast visibility in industrial environments.
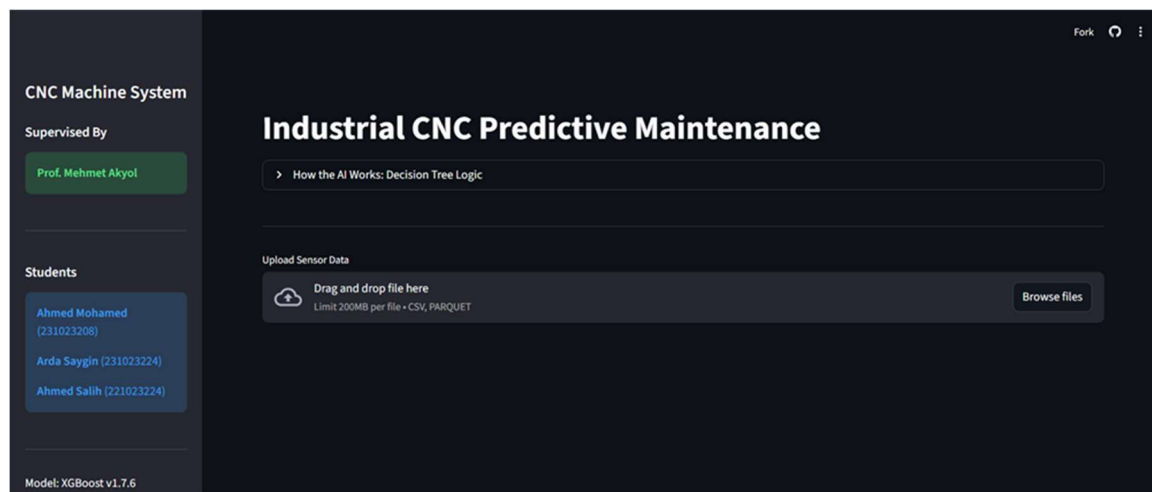

Figure 1.Streamlit-based predictive maintenance dashboard interface developed for CNC milling machines.

5. Results and Discussion

The deployed system successfully identified critical machine states from the input sensor data. The predictive model classified machine conditions into healthy and critical categories using a risk score generated by the ensemble of decision trees. The results demonstrate the model's effectiveness in early failure detection.

Figure 2 presents an interactive visualization of failure risk scores plotted against air temperature. Critical machine states are highlighted in red, while healthy states are shown in green. The hover-based diagnostics allow users to inspect hidden telemetry values, enabling precise identification of the sensor parameters responsible for increased failure risk.

Figure 2: Interactive failure risk visualization showing healthy and critical machine states.

In addition to visual analysis, the system provides a numerical diagnostic report, shown in Figure 3. This table presents raw sensor measurements alongside predicted failure states, enabling maintenance engineers to perform data-driven decision-making. The report can be exported for documentation and long-term maintenance planning.

Figure 3: Detailed telemetry report displaying numerical sensor values and predicted machine states.

7. Conclusion

This project delivers a complete **end-to-end predictive maintenance pipeline**, starting from sensor data analysis and cloud-based model training to real-time visualization and diagnostics.

By integrating **AWS SageMaker**, **XGBoost**, and **Streamlit**, the system provides a scalable and industry-ready solution for modern CNC maintenance strategies.

8. Future Work

Future improvements include:

- Integration with live CNC machine sensor streams (IoT)
- Expansion to other machine types
- Automated maintenance scheduling based on risk trends