Predictive Maintenance using Machine Learning

(AI4I 2020 Industrial Dataset)

1. Introduction

Predictive maintenance helps industries reduce machine downtime by detecting early signs of failure using data. Instead of repairing machines after breakdowns, machine learning models can predict failures in advance, improving reliability and lowering costs.

This project uses the AI4I 2020 industrial dataset to build a machine learning model that identifies whether a machine is likely to fail based on sensor readings.

2. Dataset Overview

The dataset contains 10,000 records with features such as:

Air and process temperatures

- Rotational speed (rpm)
- Torque (Nm)
- Tool wear (min)
- Machine type (L, M, H)
- Five types of failure indicators
- The target variable is Machine Failure (0 = normal, 1 = failure).

3. Methodology

- Data Preprocessing
- Removed non-useful ID columns
- Converted machine type into numerical dummy variables
- Split dataset into 80% training and 20% testing
- Used class balancing to handle fewer failure samples

Model Used

A Random Forest classifier was selected because it handles non-linear relationships well, works robustly with industrial datasets, and provides feature importance insights.

- Training
- 300 decision trees
- Balanced class weights
- Random state = 42

4. Results

- Model Performance
- The Random Forest model achieved:
- Accuracy: 99.5%
- High precision and recall for failure class
- Only 3 misclassifications in the test set

Important Features

According to feature importance analysis:

1. Rotational speed (rpm)

2. Torque (Nm)

3. Power Failure (PWF)

4. Heat Dissipation Failure (HDF)

5. Tool wear (min)

These align with real-world causes of equipment breakdown such as high load, overheating, and mechanical stress.

5. Discussion

The model shows excellent reliability in predicting machine failures using sensor data. With 99.5% accuracy, it can support condition-based maintenance, reduce downtime, and improve safety in manufacturing systems.

The project also demonstrates how Industrial Engineering and Data Analytics combine effectively within Industry 4.0 systems.

6. Conclusion

This project successfully implemented a predictive maintenance model using machine learning. The results indicate that Random Forest is highly effective for identifying machine failures based on operational data.

The approach can help industries move from reactive to proactive maintenance, improving efficiency and reducing operational losses.