Predictive Maintenance Model

Data Science Capstone
Sarah Jones

Introduction

A predictive maintenance model takes in historical maintenance data, usage patterns/frequency, and possible sensor reading and predicts when machinery or manufacturing equipment is likely to fail. This can help with maintenance cost and production downtime/efficiency.

Explanation

Collecting data

- Historical maintenance records - repair logs, failure incidents, and maintenance schedules
- Sensor data - real-time or past readings (temperature, pressure, vibrations)
- Usage patterns - operational data like workload, runtime hours, and operating conditions

Modeling approach

This will be chosen based on the data obtained and what we want the final product to look like. A classification model would be used for a binary prediction output, regression model for estimating time-to-failure output, or potentially a hybrid if multiple data types are involved.

Why

This model is useful and important to have especially for industries that rely heavily on machinery and equipment. Using data-driven predictions can transition a company from reactive or scheduled maintenance to a more proactive approach. Being more proactive helps lower repair costs by predicting exactly when it requires replacement instead of replacing a part frequently which could lead to unnecessary costs or fixing equipment right when it breaks which is expensive and inefficient.

These types of models can also help reduce downtime in production and enhance production efficiency. Unplanned equipment failures lead to costly production stops. A predictive maintenance model forecasts failures before they occur, allowing maintenance teams to intervene proactively, reducing unscheduled downtime. Continuous, uninterrupted operation of machines improves overall equipment effectiveness and productivity. Predictive models help ensure production schedules are met without unexpected disruptions.

Steps to completion and possible continuation of project

This can be broken down into 9 steps:

1. Define problem and objectives of the model
    a. Identify equipment
    b. Define what types of failures we want to predict (mechanical, overheating, etc)

    c. Specify objectives of the model
        i. How long the equipment will operate before failing or predict whether a failure will occur within a certain timeframe
2. Collect the data
3. Preprocess the data
    a. Data cleaning
    b. Feature engineering
    c. Label creation (defining failures)
    d. Normalization/Scaling (scaling features for consistency)
4. Data Analysis
    a. Visualize data trends and distributions
    b. Analyze correlations between features and failures
    c. Investigate outliers or anomalies in data
5. Choose modeling approach
    a. Classification
    b. Regression models
    c. Hybrid models
6. Model Training
    a. Split data
    b. Handle imbalance
    c. Hyperparameter tuning
    d. Evaluation metrics
7. Validate the model
    a. Cross-validation to ensure the model generalizes well
    b. Test on unseen data the evaluate robustness
8. Deploy and monitor
    a. Deploy the model (can be done on own system and not business databases or software) in the form of a dashboard
        i. Integrate into maintenance management system
        ii. Set up pipelines for real-time inference
    b. Monitor model performance
        i. Regularly evaluate metrics and retrain the model as needed
        ii. Adapt the model to changes in equipment
9. Iterate and improve
    a. Gather feedback and adjust if needed
    b. Incorporate new data sources (upgrades, external factors, etc)
    c. Refine feature engineering and model algorithms based on performance