Siemens – Predictive Maintenance

Author: Venkata Sandeep Kumar Reddy
Date: May 26, 2025

A machine learning system to detect equipment failure in advance using industrial sensor data.

Abstract

Predictive maintenance is a crucial aspect of modern smart manufacturing. This project aims to build a classification model that predicts machine failure using real-world sensor data inspired by Siemens' industrial equipment.

Business Objective

Minimize downtime and reduce maintenance costs by predicting potential failures before they happen.

Dataset

The dataset includes sensor readings such as air temperature, process temperature, rotational speed, torque, and tool wear. The target is binary — 1 for failure, 0 for normal operation.

Data Preprocessing

Categorical variables were encoded (Type: A/B/C), and numerical features were scaled. Train-test split was stratified to balance target classes.

Model Selection

We used two models:

1. Logistic Regression (baseline)

2. Random Forest (ensemble)

Evaluation Metrics

Accuracy, ROC AUC, Precision, Recall, F1-score were used. Due to small sample size, ROC AUC was preferred for judging ranking performance.

Results

Logistic Regression:

- Accuracy: 0.50

- ROC AUC: 1.00

Random Forest:

- Accuracy: 0.50

- ROC AUC: 1.00

Despite low accuracy due to limited test samples, ROC AUC showed perfect ranking — models correctly differentiated failure risk.

Interpretation

ROC AUC scores indicate the models can correctly assign risk scores. Accuracy is affected due to very small test size (only 2 samples).

Business Impact

A model like this can help Siemens reduce unplanned downtimes, improve maintenance scheduling, and extend equipment lifespan.

Limitations

Dataset was small and synthetic. In production, a larger dataset and more failure examples would improve generalization.

Future Improvements

Add time-series analysis, sensor drift correction, anomaly detection, and integrate with IIoT platforms.

Deployment

This model can be deployed as a real-time prediction service on the factory floor using Streamlit, Flask, or an API.

Conclusion

Predictive maintenance saves time, cost, and operational risks. This project demonstrates the foundation for scalable machine failure prediction systems.

Code Snippet

```
model = RandomForestClassifier()
model.fit(X_train, y_train)
y_pred = model.predict(X_test)
print(classification_report(y_test, y_pred))
roc_auc_score(y_test, model.predict_proba(X_test)[:, 1])
```