PROJECT REPORT

on

MACHINE PREDICTION FAILURE USING MACHILE LEARNING

A Project Report submitted as part of the completion of an certification course of SkillDunia

Course Name: Artificial Intelligence Self Paced 25
Platform: SkillDunia

Submitted by:
Name: P. Mohana Chaitanya Teja
Email: 24311a05e7@cse.sreenidhi.edu.in

Course Duration: 11/2025 - 02/2026

This project is an original work completed as part of the above-mentioned course.

ABSTRACT

Machine failure prediction is a key use of machine learning in predictive maintenance. The project aims to forecast machine failures using sensor data collected from machines while they operate.

The dataset includes various sensor readings like temperature, air quality, current usage, rotational speed, pressure, and ultrasonic sensor values. These features help identify whether a machine is likely to fail.

In this project, we handle data preprocessing and exploratory analysis. Then, we train a Random Forest Classifier to predict machine failure. We assess the trained model using accuracy score, confusion matrix, and classification report.

The results indicate that machine learning can accurately foresee failures, which helps industries lower downtime and boost maintenance efficiency.

Table of Contents

INTRODUCTION

With the advancement of Industry 4.0, machines are increasingly equipped with sensors that continuously monitor their operational conditions. These sensors generate large volumes of data that can be analyzed to detect early signs of machine failure.

Traditional maintenance techniques often result in unexpected breakdowns and high maintenance costs. Machine learning enables predictive maintenance by identifying failure patterns from historical data.

This project aims to build a machine learning model that predicts machine failure using sensor data, allowing timely maintenance and reducing operational risks.

PROBLEM STATEMENT

Unexpected machine failures can lead to production loss, safety risks, and increased maintenance costs.

The problem addressed in this project is to predict whether a machine will fail or not based on multiple sensor readings collected during its operation.

The objective is to develop a supervised machine learning model that classifies machine states into failure (1) or non-failure (0) categories.

Objectives of the Project

- To analyze machine sensor data
- To preprocess and clean the dataset
- To build a machine learning classification model
- To predict machine failure using sensor readings
- To evaluate the model performance using appropriate metrics

Dataset Description

The dataset used in this project is related to machine failure prediction and contains sensor data collected from machines under different operating conditions. Each record in the dataset represents a machine state, and the target variable indicates whether a failure occurred.

Target Variable: fail

- 1 → Machine failure
- 0 → No failure

ATTRIBUTES DESCRIPTION

- **footfall**
  Number of people or objects passing near the machine
- **temp_mode**
  Temperature mode or setting of the machine
- **AQ**
  Air Quality Index near the machine
- **USS**
  Ultrasonic sensor readings
- **CS**
  Current sensor readings
- **VOC**
  Volatile Organic Compounds level
- **RP**
  Rotational position / RPM of machine parts
- **IP**
  Input pressure to the machine
- **Temperature**
  Operating temperature
- **fail**
  Machine failure indicator
  (1 = failure, 0 = no failure)

Methodology

The methodology followed in this project consists of the following steps:

1. Loading the dataset using Pandas
2. Checking for missing values
3. Separating input features and target variable
4. Splitting the dataset into training and testing sets
5. Training the Random Forest model
6. Evaluating the model performance

Algorithm Used: Random Forest

Random Forest is an ensemble learning algorithm that builds multiple decision trees and combines their predictions to improve accuracy.

It reduces overfitting and handles complex datasets effectively. Due to its robustness and high performance, Random Forest was chosen for this project.

Advantages:

- High accuracy
- Handles non-linear data
- Resistant to overfitting

Tools & Technologies Used

- Python
- Jupyter Notebook
- Pandas
- NumPy
- Scikit-learn
- Matplotlib
- Seaborn

Implementation Details

The dataset is loaded into the Jupyter Notebook using Pandas. Input features are separated from the target variable.

The data is split into training and testing sets. A Random Forest Classifier is trained using the training data.

The trained model predicts outcomes on test data and is evaluated using accuracy score, confusion matrix, and classification report.

Results

The Random Forest model achieved good accuracy in predicting machine failures.

The confusion matrix shows correct classification of both failure and non-failure cases. The classification report indicates balanced precision and recall values.

Conclusion

This project demonstrates the effectiveness of machine learning in predicting machine failures using sensor data.

The Random Forest model successfully identifies potential failures, helping industries improve maintenance planning and reduce downtime.