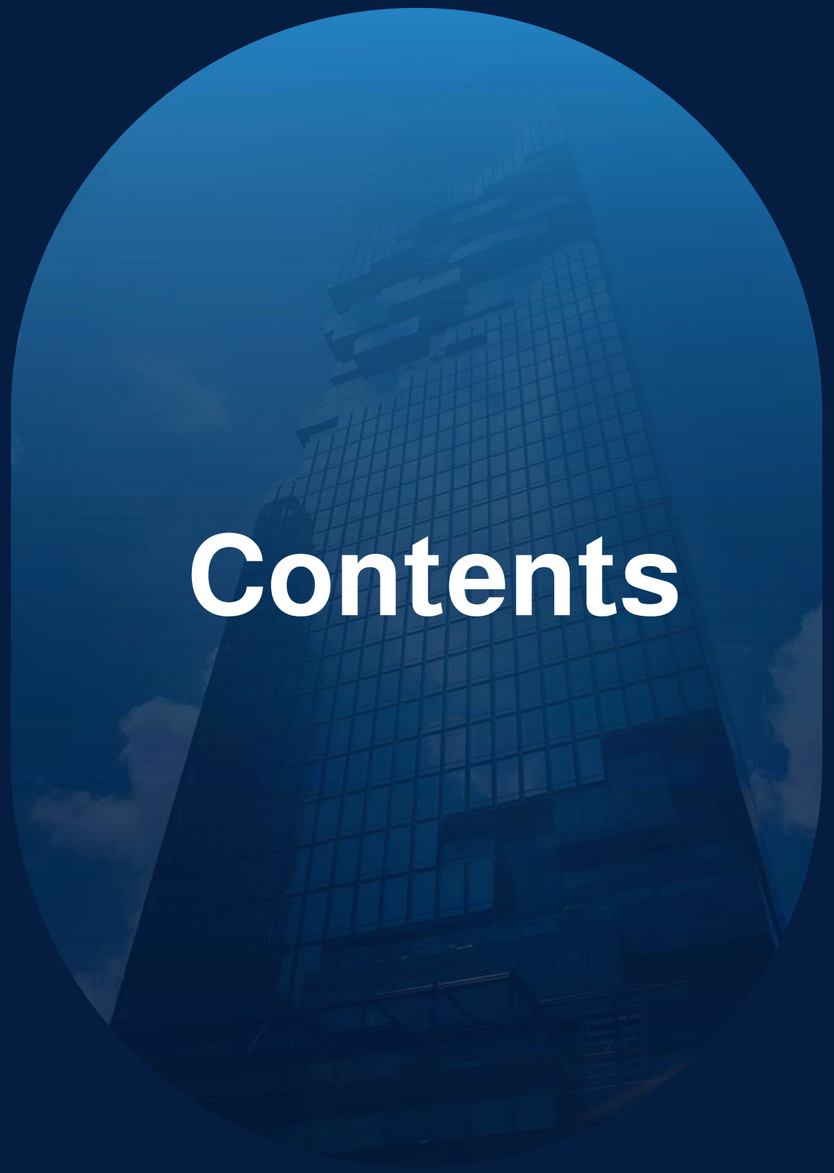
Ahmed Safwat

Predictive Maintenance for Vehicle Health

final project Digital Egypt Builders Initiative (DEBI)

supervised by Corelia



Contents

01 Introduction

This project aims to develop a smart system that predicts potential vehicle failures before they occur, using real-time sensor data and machine learning algorithms.

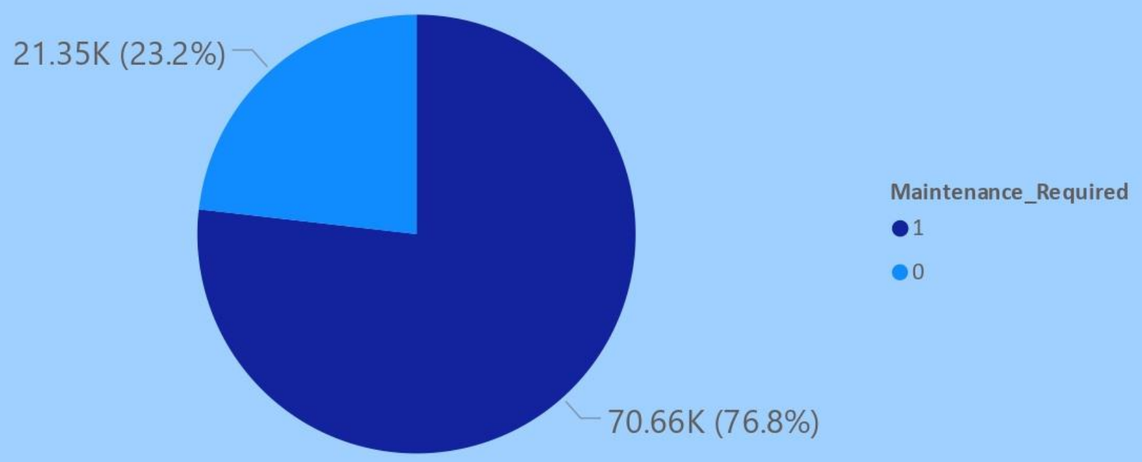
Instead of relying on traditional maintenance schedules, the system detects early signs of malfunction and alerts users for timely repairs.

By applying predictive analytics, the solution helps reduce unexpected breakdowns, lower maintenance costs, and improve overall vehicle safety and efficienc.

02 Dashboard

Dash Board

Maintenance Required



Size of data set

92.00K

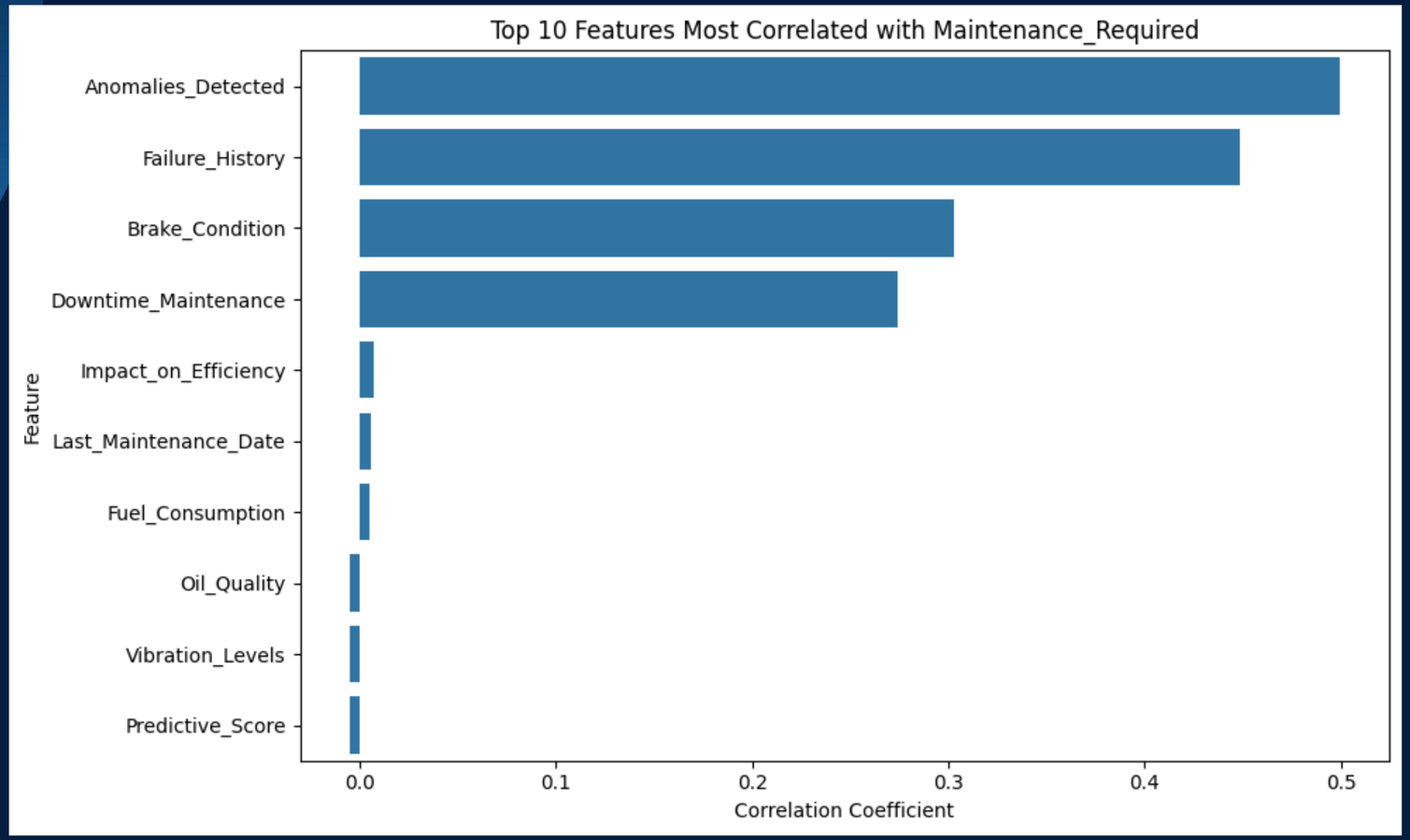
Anomalies_Detected
- 0
- 1

Failure_History
- 0
- 1

Brake_Condition
- Fair
- Good
- Poor

Model	Accuracy	F1 Score	Precision	Recall	Run time(m)
Random Forest	1.00	1.00	1.00	1.00	10.00
XGBoost	1.00	1.00	1.00	1.00	2.00

03 Data description



- Removed constant column: 'Engine_Temperature Engine_Temperature' = 120.
- Processed 'Last_Maintenance_Date' into day, month, year.
- Applied encoding to categorical features.
- Used non-linear models: Random Forest, XGBoost.
- Handled class imbalance with SMOTE / class weighting.
- Ongoing tuning & feature engineering to boost accuracy.

04 Data cleaning & prepocessing

#	Column	Non-Null Count	Dtype
0	Make_and_Model	92000 non-null	float64
1	Year_of_Manufacture	92000 non-null	float64
2	Vehicle_Type	92000 non-null	float64
3	Usage_Hours	92000 non-null	float64
4	Route_Info	92000 non-null	float64
5	Load_Capacity	92000 non-null	float64
6	Actual_Load	92000 non-null	float64
7	Maintenance_Type	92000 non-null	float64
8	Maintenance_Cost	92000 non-null	float64
9	Tire_Pressure	92000 non-null	float64
10	Fuel_Consumption	92000 non-null	float64
11	Battery_Status	92000 non-null	float64
12	Vibration_Levels	92000 non-null	float64
13	Oil_Quality	92000 non-null	float64
14	Brake_Condition	92000 non-null	float64
15	Failure_History	92000 non-null	float64
16	Anomalies_Detected	92000 non-null	float64
17	Predictive_Score	92000 non-null	float64
18	Maintenance_Required	92000 non-null	float64
19	Weather_Conditions	92000 non-null	float64
20	Road_Conditions	92000 non-null	float64
21	Delivery_Times	92000 non-null	float64
22	Downtime_Maintenance	92000 non-null	float64
23	Impact_on_Efficiency	92000 non-null	float64
24	Maintenance_Year	92000 non-null	float64
25	Maintenance_Month	92000 non-null	float64
26	Maintenance_Day	92000 non-null	float64

- Dropped unnecessary columns and those with constant values.
- Renamed columns for clarity and removed duplicate labels.
- Handled missing values and ensured data types were appropriate.
- Extracted day, month, and year from 'Last_Maintenance_Date'.
- Encoded categorical variables using Label Encoding.
- Checked for class imbalance in the target variable.
- Saved the cleaned data to a new CSV file.

05 Models comparison

Maintenance_Required	count
1.0	56566
0.0	56566

The XGBoost model is the most efficient because it takes short time

Model	Accuracy	F1 Score	Precision	Recall	Run time(m)
Random Forest	1.00	1.00	1.00	1.00	10.00
XGBoost	1.00	1.00	1.00	1.00	2.00

06

References

Project

Data

Model build

THANK YOU