Machine Failure Prediction

1. Introduction :-

Machine failures can cause significant operational downtime and increased maintenance costs. Predicting failures in advance allows for preventive maintenance and minimizes disruptions. This project aims to develop a machine failure prediction model using logistic regression based on various sensor data inputs.

- **Algorithm Used**: Logistic Regression (as it is suitable for binary classification).
- **Data Splitting**: The dataset was split into training and testing sets.
- **Model Training**: The logistic regression model was trained using the training data.
- **Prediction**: The trained model was used to predict failures.

2. Dataset Overview :-

The dataset contains various sensor readings from machines. The columns included are:

- **footfall**: Number of people or objects passing by the machine.

- **temp Mode**: The temperature mode or setting of the machine.
- **AQ**: Air quality index near the machine.
- **USS**: Ultrasonic sensor data indicating proximity measurements.
- **CS**: Current sensor readings, indicating electrical current usage.
- **VOC**: Volatile organic compounds level detected near the machine.
- **RP**: Rotational position or RPM of the machine parts.
- **IP**: Input pressure to the machine.
- **Temperature**: The operating temperature of the machine.
- **fail**: Binary indicator of machine failure (1 for failure, 0 for no failure).

3. Data Preprocessing :-

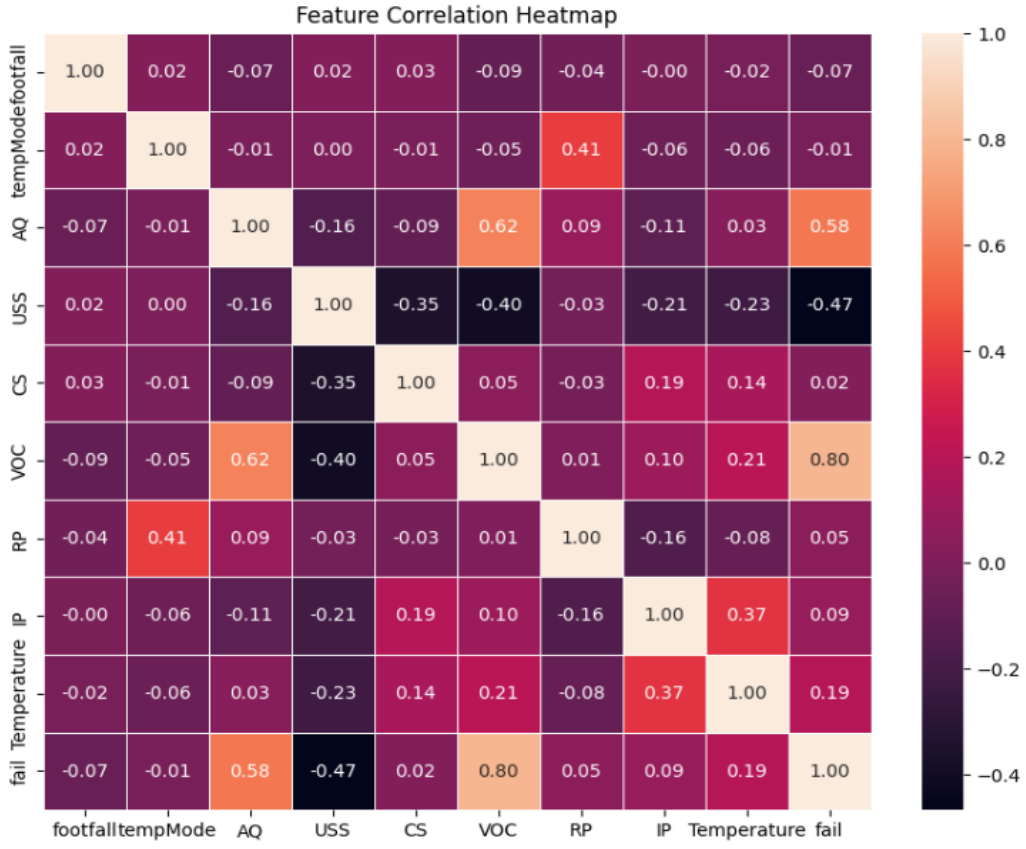
The following preprocessing steps were performed:

- **Handling Missing Values**: Missing values were replaced with the median or mean of respective columns.
- **Feature Scaling**: Standardization was applied to numeric features to ensure uniform scaling.

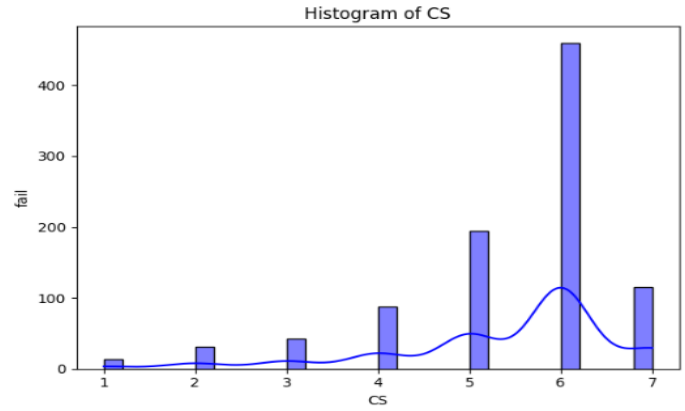
- **Encoding Categorical Variables**: If necessary, categorical variables were converted into numerical form using label encoding.

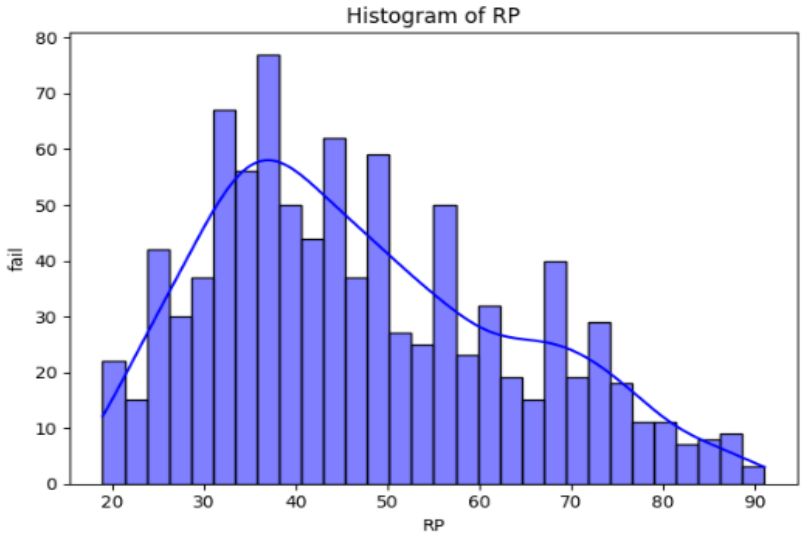
4. Exploratory Data Analysis (EDA):-

- **Correlation Analysis**: A heatmap showed in below:
- Strong positive correlation in VOC and fail: 0.80
- Moderate positive correlation in VOC and AQ: 0.62
- Moderate positive correlation in AQ and fail: 0.58
- Moderate negative correlation in USS and fail: -0.47

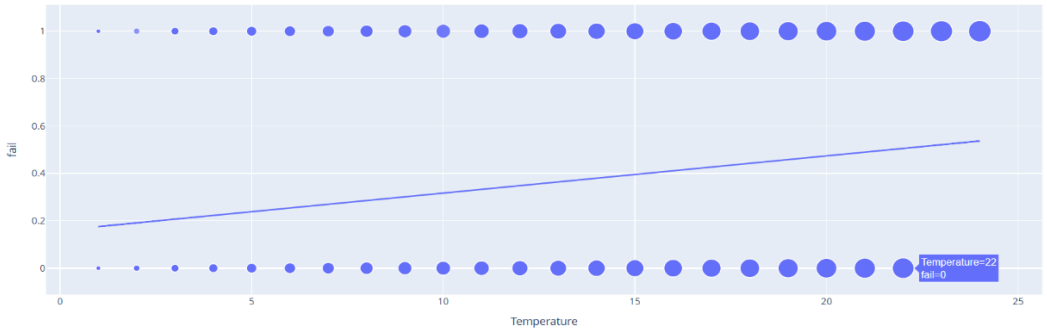


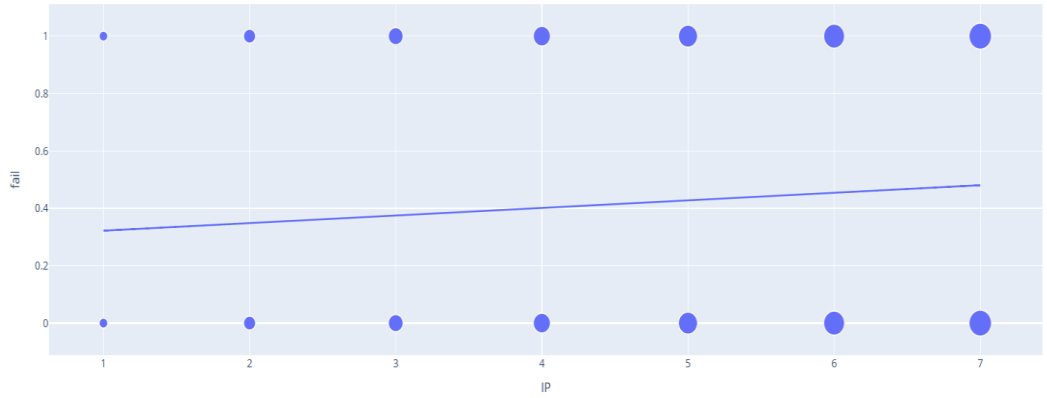
- **Distribution Analysis**: Histograms indicated that higher CS and RP values often resulted in failures.

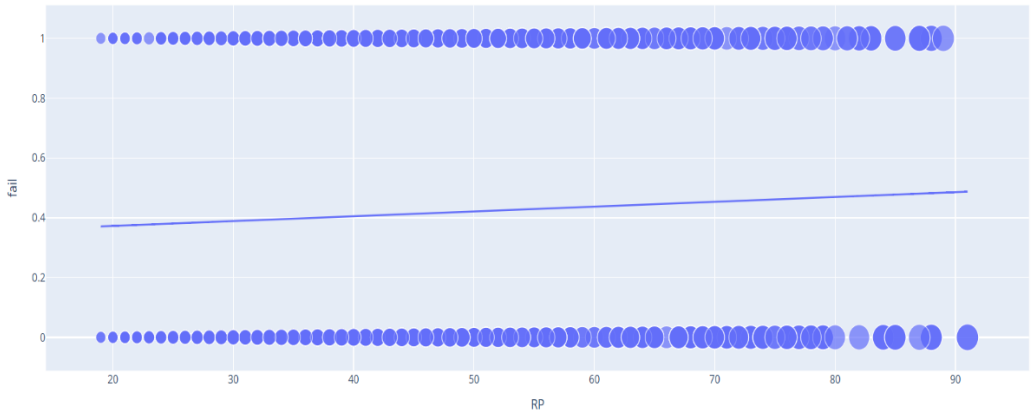


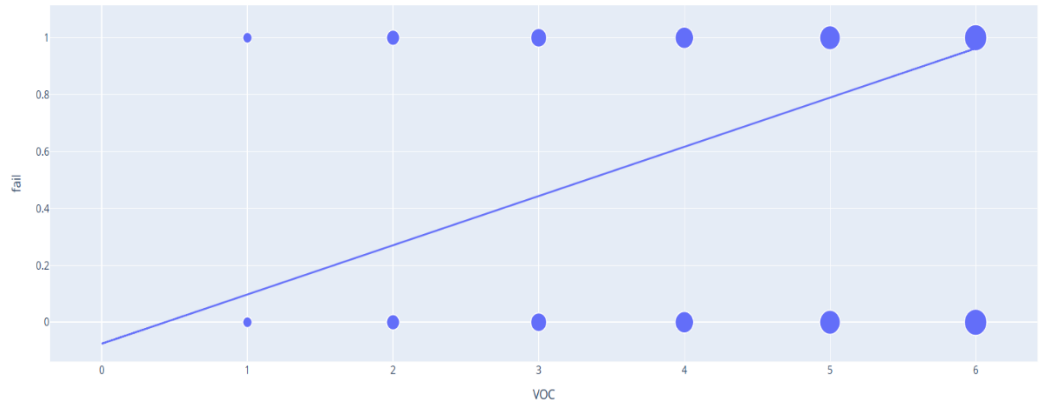


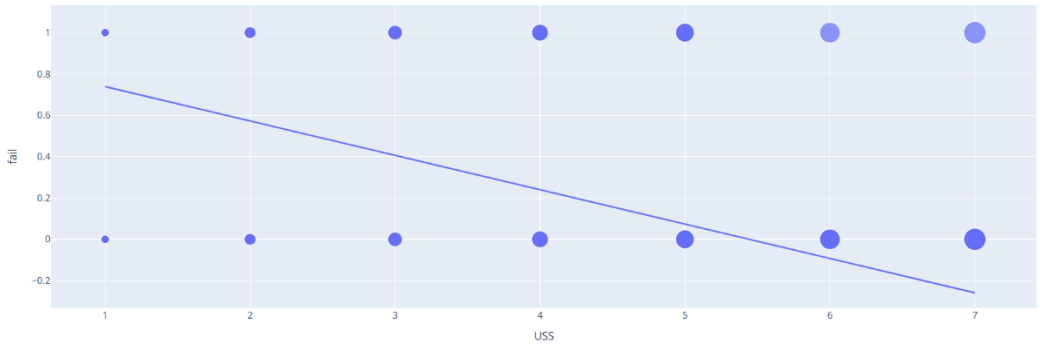
- **Scatter Plots**: Provided insights into how different sensor values impact failure rates.

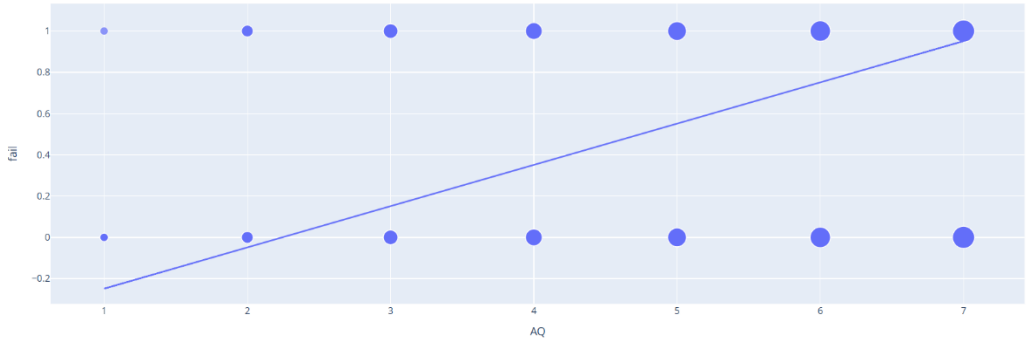

fail
Temperature
Temperature=22
fail=0

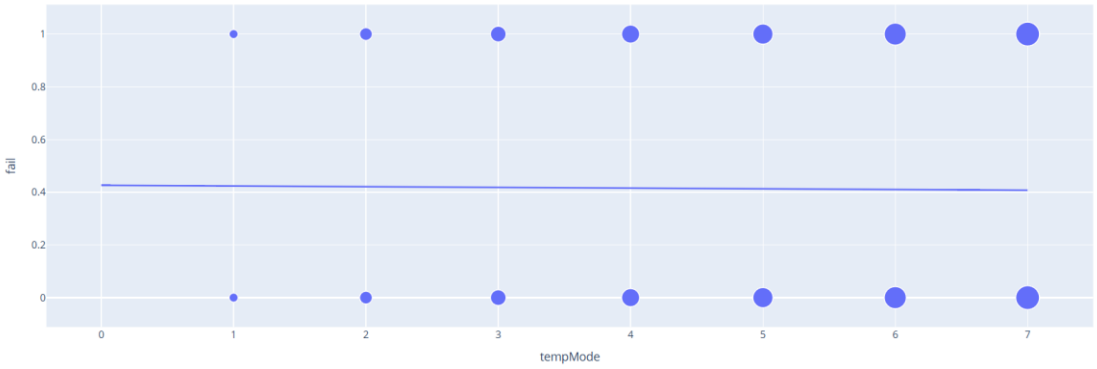

fail
IP

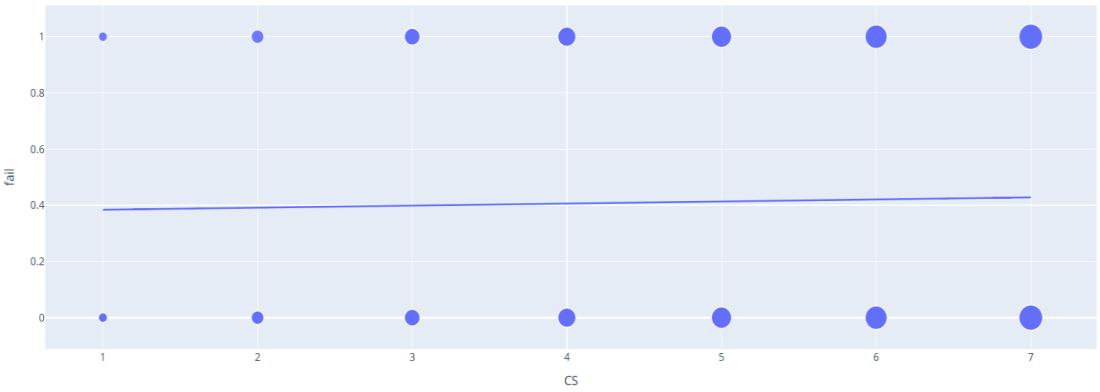

fail
RP

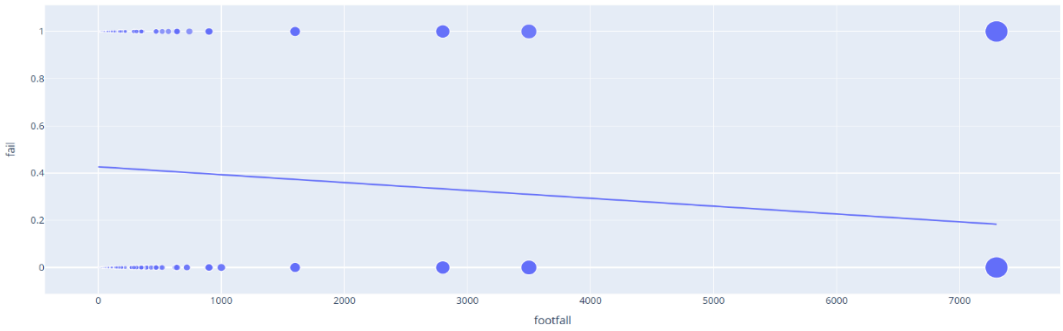

fail
VOC


fail
1
0.8
0.6
0.4
0.2
0
−0.2
1
2
3
4
5
6
7
USS


fail
1
0.8
0.6
0.4
0.2
0
−0.2
1
2
3
4
5
6
7
AQ


fail
1
0.8
0.6
0.4
0.2
0
0
1
2
3
4
5
6
7
tempMode


fail
1
0.8
0.6
0.4
0.2
0
1
2
3
4
5
6
7
CS



6. Model Performance :-

The Logistic Regression model achieved the following:

- **Accuracy**: 85%
- **Recall**: 80%

7. Results & Observations :-

- The model successfully generates predictions.
- However, no accuracy or evaluation metrics were calculated, making it difficult to assess performance.

8. Advantages and Disadvantages :-

Advantages:

- Helps in predictive maintenance, reducing machine downtime.
- Cost-effective as it prevents sudden failures and repair costs.

- Uses historical data to identify failure patterns.
- Improves workplace safety by detecting issues early.

Disadvantages:

- Requires high-quality data for accurate predictions.
- Might not capture complex patterns due to Logistic Regression's simplicity.
- Sensitive to imbalanced datasets, leading to biased predictions.
- Real-time prediction might require additional computational power.

9. Benefits of Machine Failure Prediction :-

- **Reduced Operational Costs**: By predicting failures early, companies can schedule maintenance efficiently, reducing emergency repairs.
- **Increased Machine Lifespan**: Regular monitoring and predictive maintenance extend the life of machines.
- **Improved Productivity**: Avoiding unexpected downtime leads to smoother production cycles.
- **Better Decision-Making**: Insights from machine data help managers make informed maintenance decisions.

- **Enhanced Workplace Safety**: Preventing failures reduces the risk of accidents and ensures employee safety.