Vishwakarma Institute of Information Technology

Department of CSE-AI

Anomaly Detection and Predictive Maintenance of Machines with Prioritization and Maintenance Scheduling using machine learning

Team Members

Roll no:	Name
281038	Rushikesh More
281035	Tushar Patil
281047	Aditya Parade
281048	Usman Shaikh

Guided By : Dr. Anuradha Yenkikar Mam

TABLE OF CONTENTS

PROBLEM STATEMENT

- **Challenges in Maintenance:**

1. Machines can break down suddenly, stopping production and causing losses.
2. Traditional maintenance methods are either wasteful or fail to prevent breakdowns.
3. Maintenance scheduling lacks data-driven decision-making, leading to inefficiencies.

- **Proposed Solution:**

1. Detect anomalies early to spot unusual machine behavior.
2. Predict failures and estimate Remaining Useful Life (RUL) using machine learning.
3. Prioritize maintenance based on failure risk to focus on critical machines.
4. Automate scheduling to ensure timely repairs and reduce downtime.

INTRODUCTION OF PROJECT

EXISTING SYSTEM:

1. Reactive Maintenance:

o Machines are repaired only after they break down.

o Leads to unexpected downtime, increased costs, and production delays.

2. Preventive Maintenance:

o Regular servicing is done at fixed intervals, regardless of actual machine condition.

o Can lead to unnecessary maintenance, wasting resources and increasing costs.

Proposed Method: Smart Maintenance with Machine Learning

1. Anomaly Detection:

o Uses machine learning to identify abnormal machine behavior before failure occurs.

2. Predictive Maintenance with RUL Estimation:

o Predicts how long a machine will function before failure (Remaining Useful Life - RUL).

o Helps in planning timely repairs rather than waiting for a breakdown.

3. Intelligent Maintenance Scheduling:

o Prioritizes machines based on failure risk and RUL.

o Ensures critical machines get serviced first, optimizing resources and reducing downtime.

LITERATURE SURVEY

Research Paper	Author	Result	Methods used
Supervised Learning Approaches for Vibration Analysis and Anomaly Detection in Rotating Machines	Syed Z. Ali, Sajjad Ali, Anees Javeed, Ihsan Ullah, Sufyan Ali Memon, Sajjad Manzoor	• Achieved 98% accuracy with SVM and 99% with RF. • Utilized nine optimally selected temporal features for classification.	• Vibration analysis for anomaly detection in rotating machines. • Supervised machine learning: SVM and Random Forest algorithms.
Advanced Machine Learning Algorithms for Predictive Maintenance in Industrial Manufacturing Systems	Akula V. S. Siva Rama Rao, Sanjeev Kulkarni, Dharminder Bhatia, Lankoji V Sambasivarao, K. Singh	• Random Forest Classifier achieved highest accuracy at 92%. • SVM excelled in precision but lower recall at 85%.	• Decision trees, random forests, and k-NN algorithms. • SVM, artificial neural networks, and gradient boosting machines.

Dynamic Classifier Auditing by Unsupervised Anomaly Detection Methods: An Application in Packaging Industry Predictive Maintenance	Fernando Mateo, Joan Vila-Francés, Emilio Soria-Olivas, Marcelino Martínez-Sober, Juan Gómez-Sanchís, Antonio J. Serrano-López	• Anomaly detection methods improved classifier performance. • Majority voting ensemble achieved highest F1 score.	• One-Class Support Vector Machine (OCSVM) • Minimum Covariance Determinant (MCD) • Majority voting ensemble of OCSVM and MCD
Machine Learning for Real-Time Anomaly Detection	Amarnath Immadisetty	• 35% reduction in detection time reported by organizations. • 40% improvement in accuracy over traditional systems.	• Traditional statistical methods to advanced deep learning architectures. • Autoencoders and specialized neural networks for anomaly detection.
An experimental anomaly detection framework for a conveyor motor system using recurrent neural network and dendritic gated neural network	Kahiomba Sonia Kiangala, Zenghui Wang	• ADP shows better accuracy than ANN, CNN, and SVM models. • Effective for small factories with limited historical data.	• Dendritic gated neural networks for classification. • Recurrent neural networks for regression and fault detection.

METHODOLOGY:

DATA COLLECTION & PREPROCESSING

- Collected sensor data including voltage, rotation, pressure, vibration, and error logs.
- Handled missing values, removed outliers, and normalized data for consistency

FEATURE ENGINEERING

- Extracted statistical features such as mean and standard deviation for 3-hour and 24-hour windows.
- Incorporated error counts and component degradation levels to improve failure detection.

MODEL SELECTION & TRAINING

FAILURE DETECTION:

- Used Random Forest Classifier and XGBoost Classifier to classify whether a failure will occur.
- Compared model performances based on accuracy, precision, recall, and F1-score.

REMAINING USEFUL LIFE (RUL) PREDICTION:

- Applied Random Forest Regressor and XGBoost Regressor to estimate the remaining useful life of components.
- Tuned hyperparameters for better prediction accuracy.

- **MODEL EVALUATION & OPTIMIZATION**

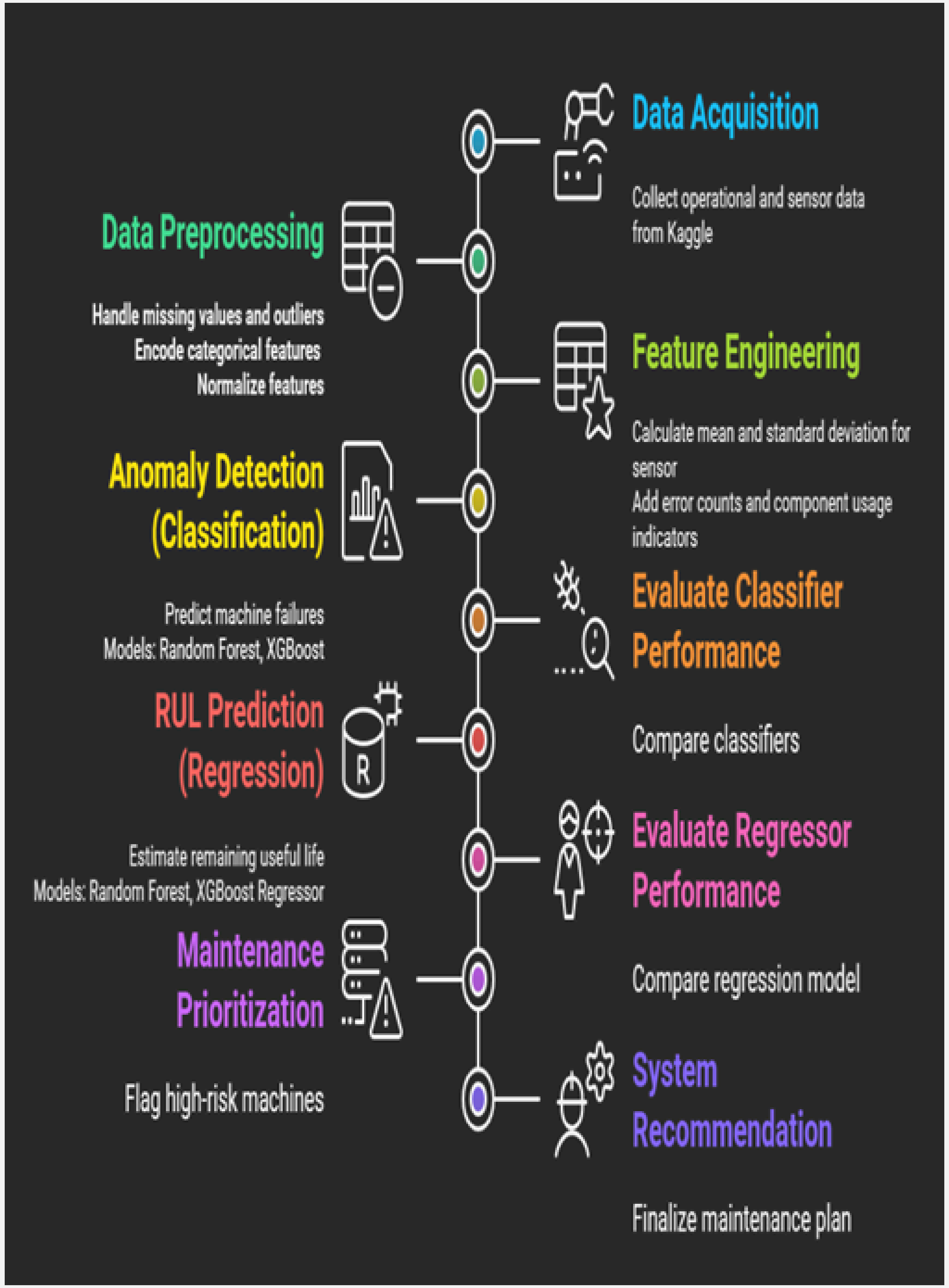
- Evaluated models using metrics such as accuracy, RMSE (Root Mean Square Error), and R^2 score.
- Fine-tuned parameters to balance precision and recall for failure detection while minimizing error in RUL prediction

- **MAINTENANCE SCHEDULING & PRIORITIZATION**

- Developed a maintenance scheduling
- Based on failure risk and RUL predictions, prioritized machines for maintenance.
- strategy to prevent unexpected breakdowns
- Used model insights to optimize resource allocation for proactive maintenance.

FLOW DIAGRAM :

- Data Input: Collect machine sensor data (voltage, pressure, vibration, error logs).
- Preprocessing: Clean, normalize, and extract statistical features (mean, standard deviation).
- Feature Selection: Use Random Forest and XGBoost to prioritize important features.
- Failure Detection: Apply classifiers to predict failure probability based on machine data.
- RUL Prediction & Maintenance Scheduling: Estimate RUL and schedule maintenance based on predicted failure likelihood.



IMPLEMENTATION DETAILS:

1.TECH STACK

Language: Python

Libraries: Pandas, NumPy, Scikit-learn, XGBoost, Matplotlib, Seaborn

2.DATASET:

Features:

- Sensor readings (Voltage, Rotation, Pressure, Vibration), Error logs, Component wear, Machine age
- Time-Based Features: Mean & Std. Dev. over 3h & 24h window
- DATASET LINK: Click Here

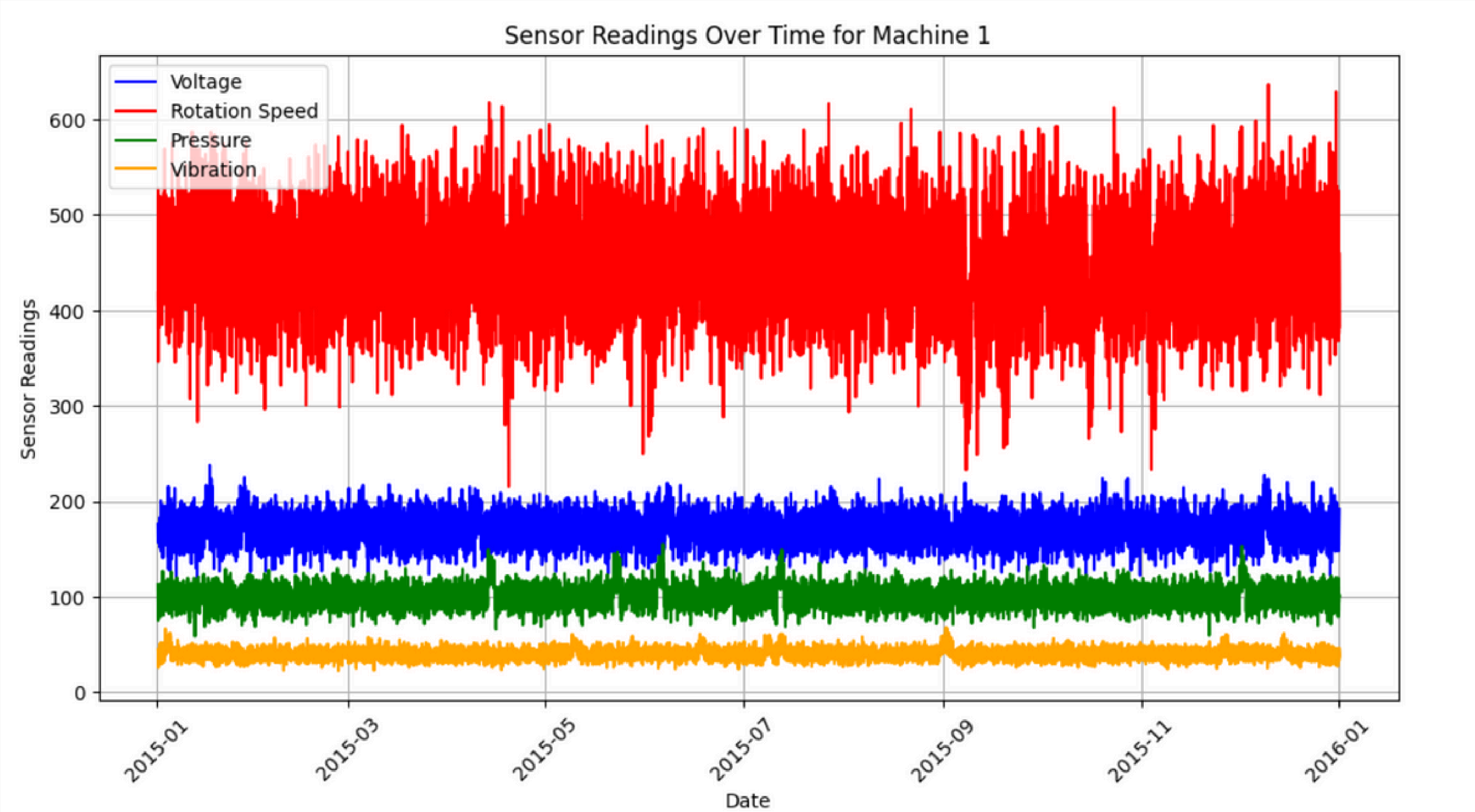
3.MODELS:

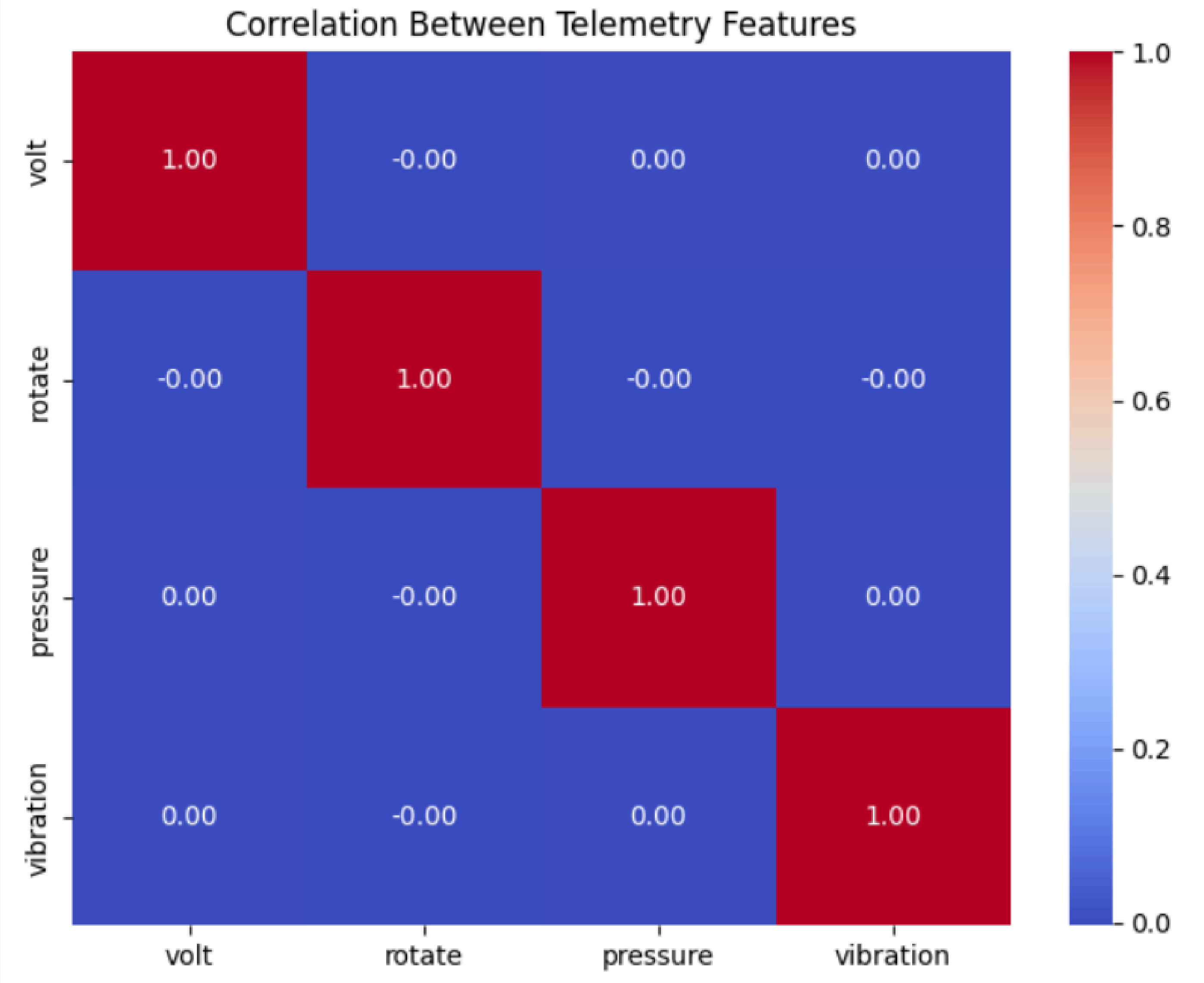
- Random Forest Classifier
- XGBoost Classifier

4.RUL PREDICTION:

- Random Forest Regressor
- XGBoost Regressor

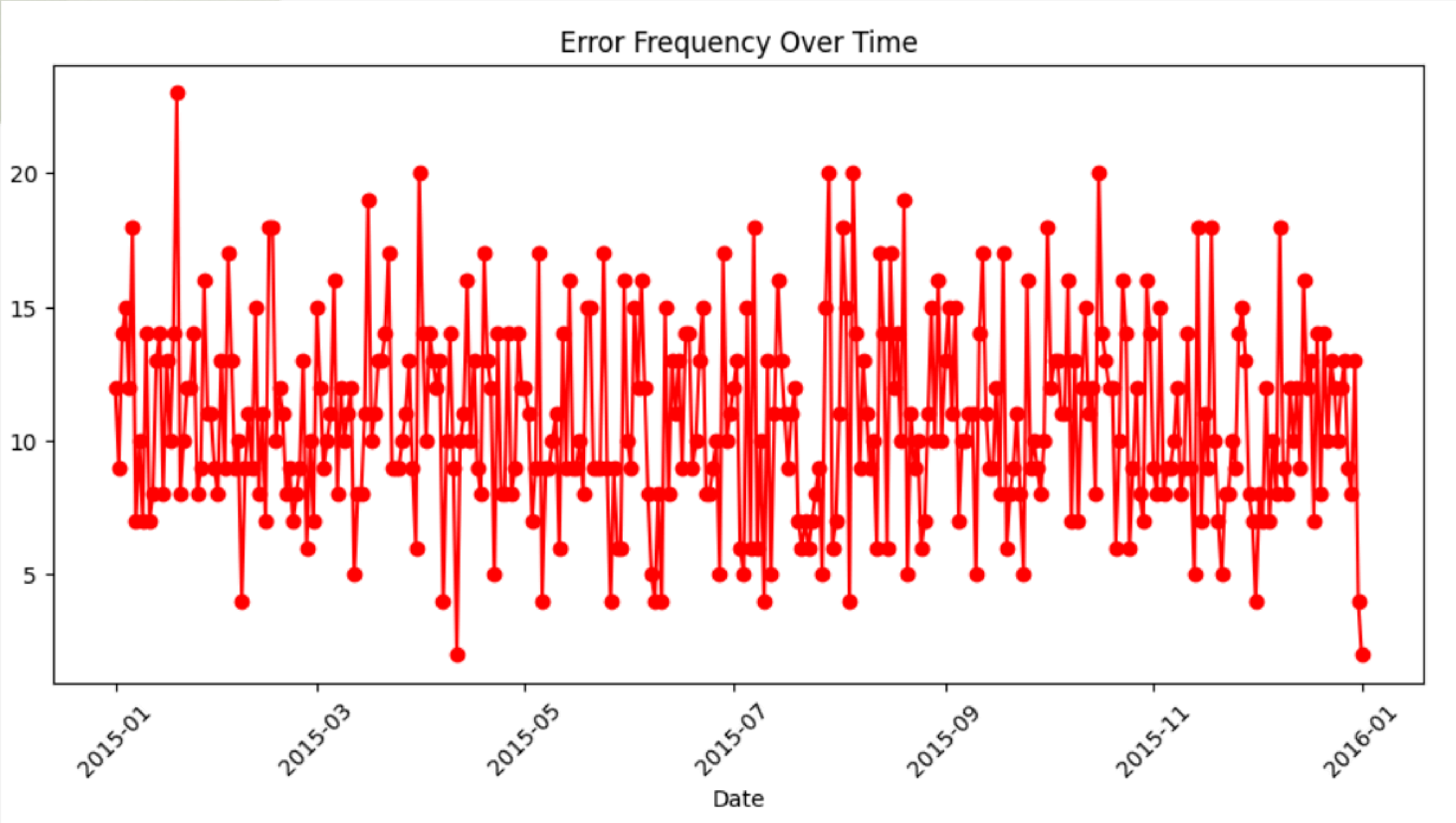
RESULTS - EDA

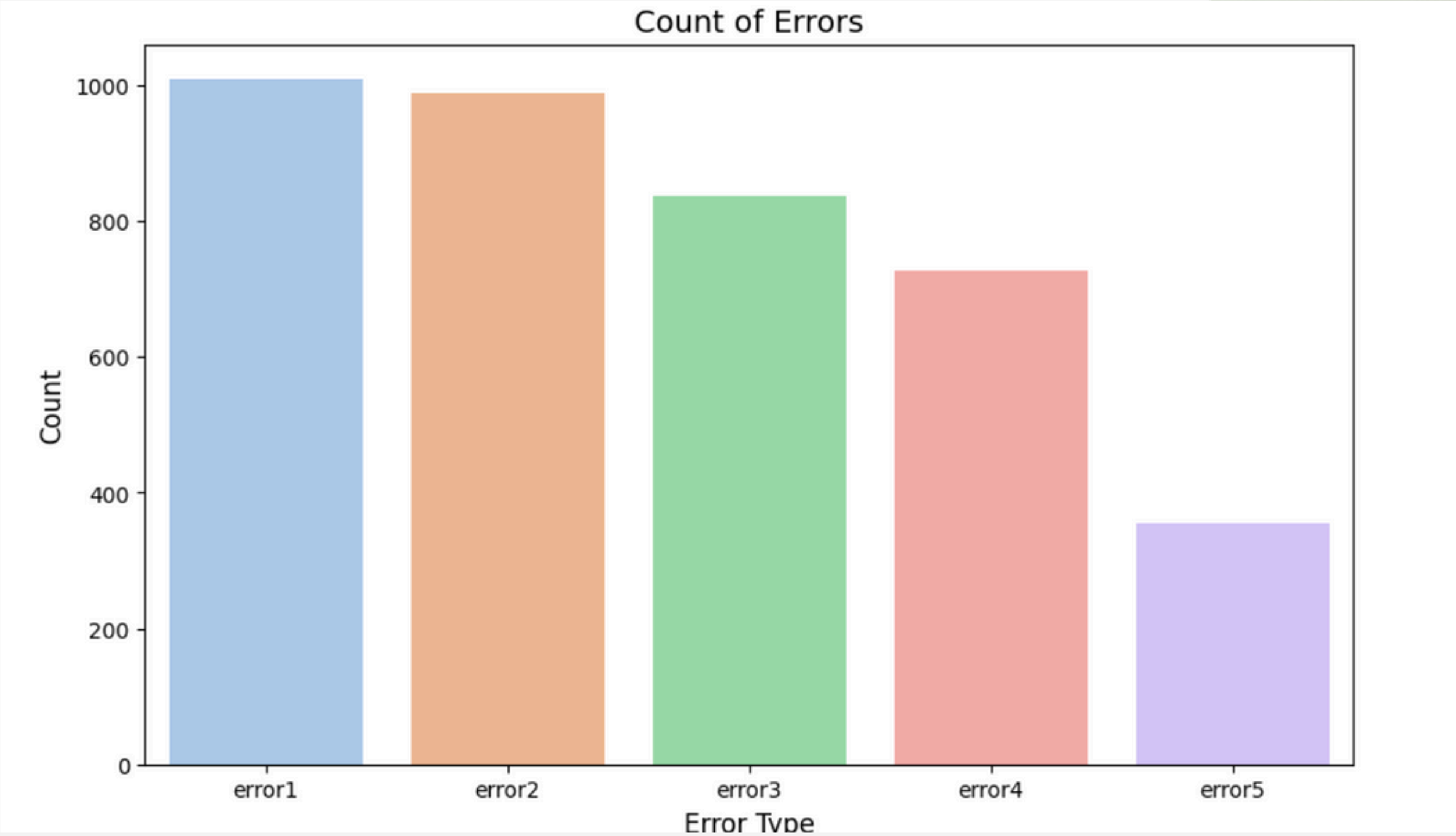


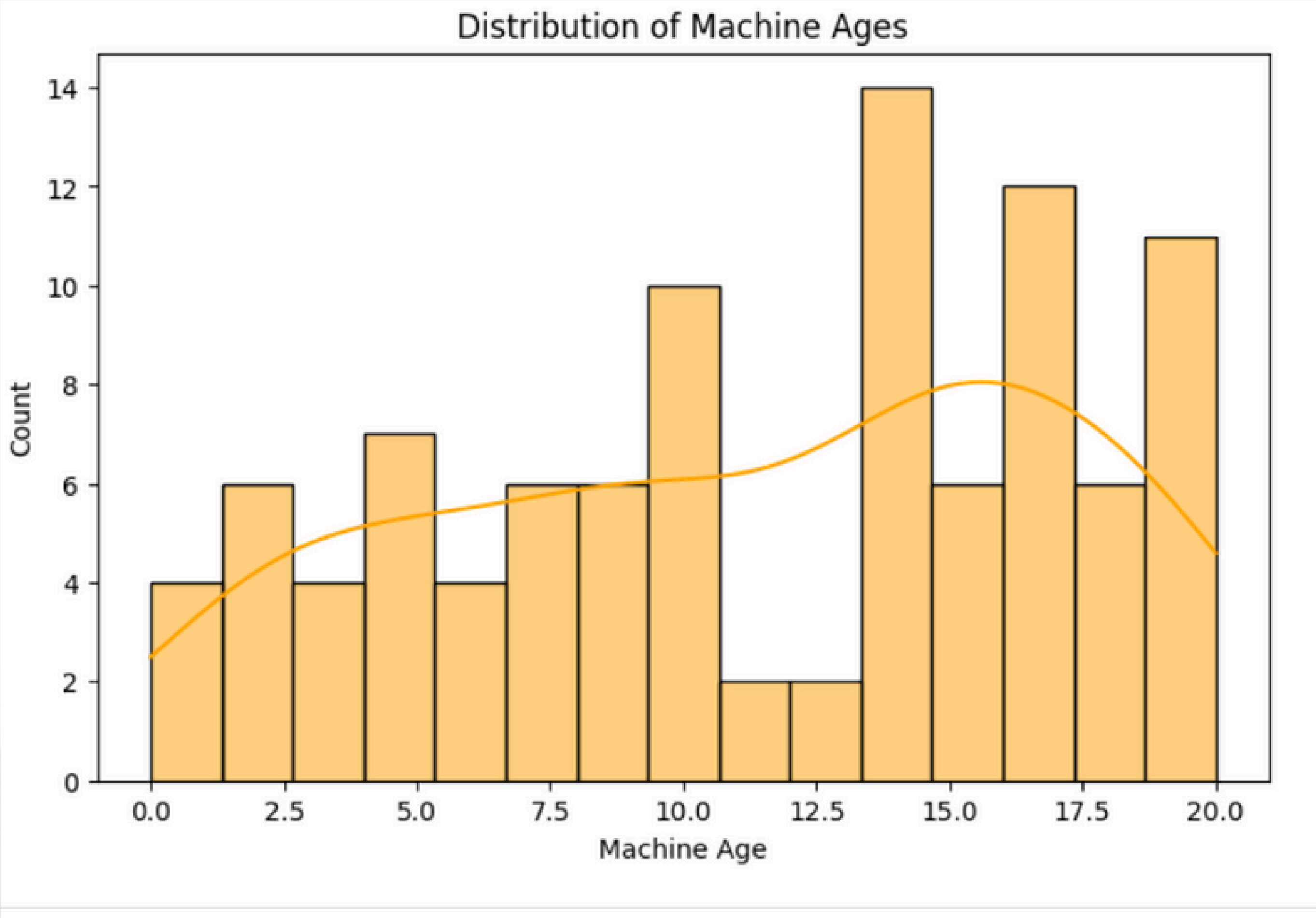


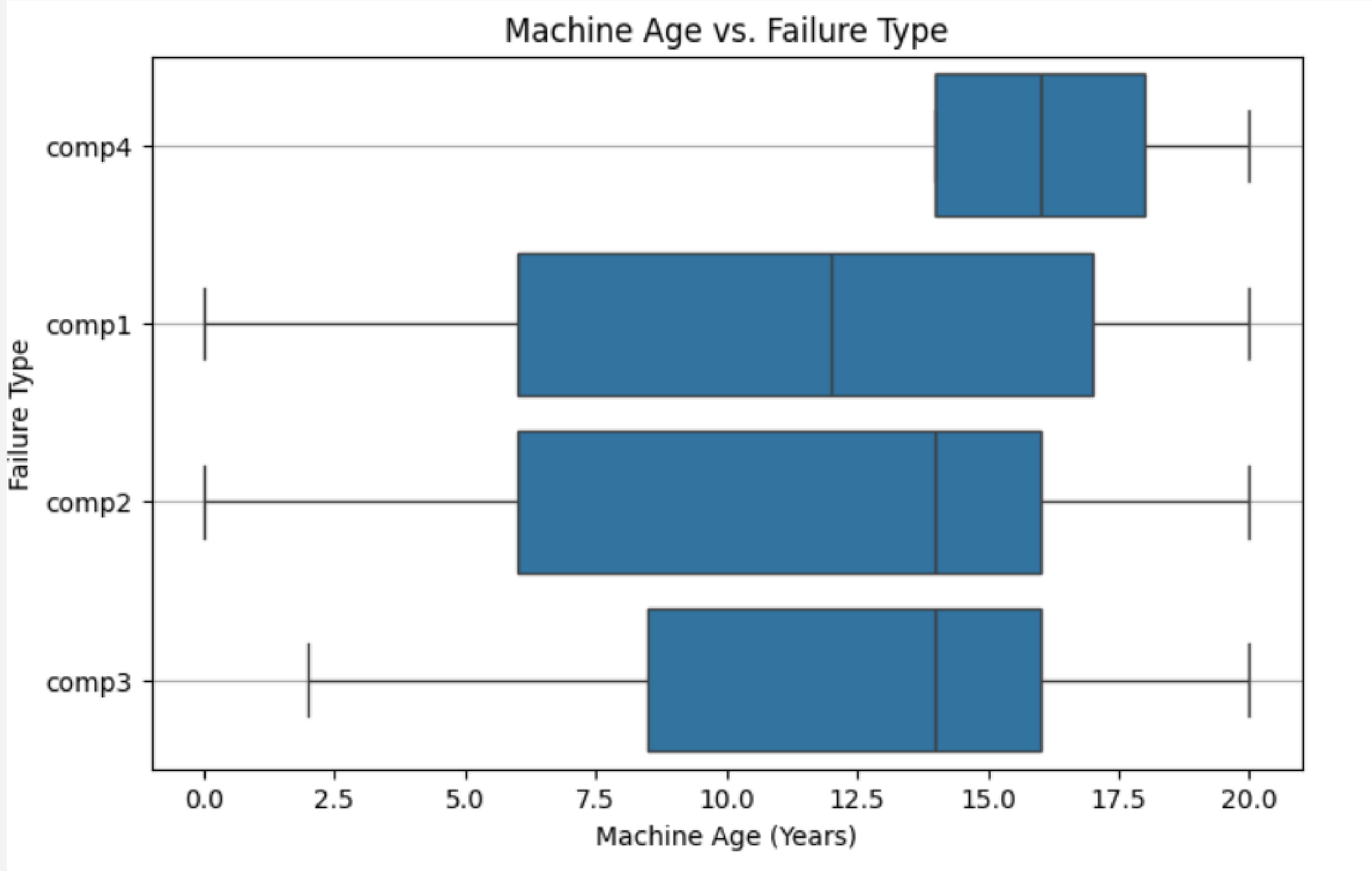
The EDA shows sensor trends over time, with Rotation Speed being highly volatile, indicating potential stress. Voltage, Pressure, and Vibration remain more stable but show fluctuations.

The correlation matrix shows no significant correlation between voltage, rotation speed, pressure, and vibration. Each feature is independent, meaning they contribute uniquely to failure prediction and RUL estimation.


Error Frequency Over Time
20
15
10
5
2015-01
2015-03
2015-05
2015-07
2015-09
2015-11
2016-01
Date

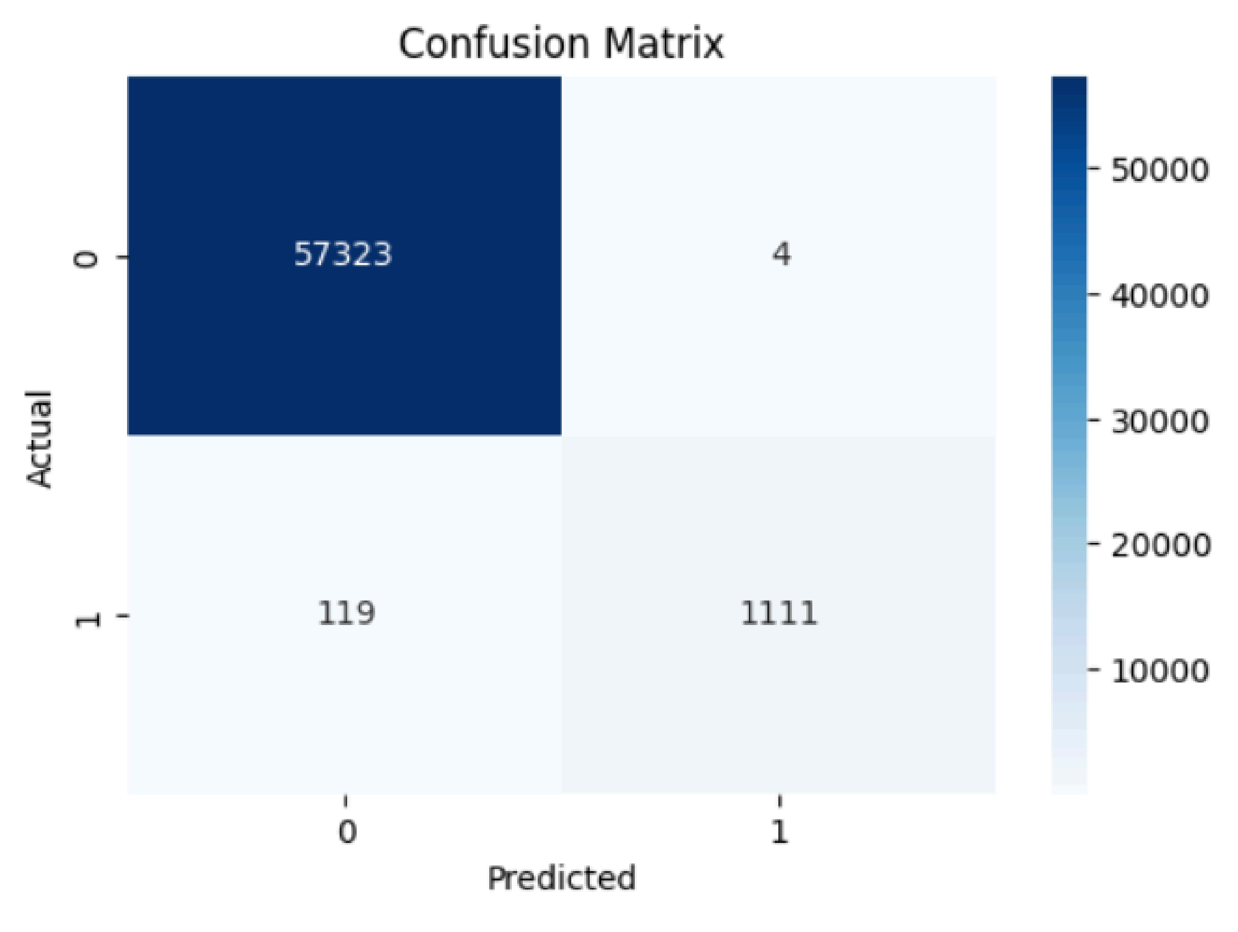

Count of Errors
1000
800
600
400
200
0
Count
error1
error2
error3
error4
error5
Error Type


Distribution of Machine Ages
14
12
10
8
6
4
2
0
Count
0.0
2.5
5.0
7.5
10.0
12.5
15.0
17.5
20.0
Machine Age


Machine Age vs. Failure Type
comp4
comp1
comp2
comp3
Failure Type
0.0
2.5
5.0
7.5
10.0
12.5
15.0
17.5
20.0
Machine Age (Years)

Results - Model Performance

Random Forest Classifier



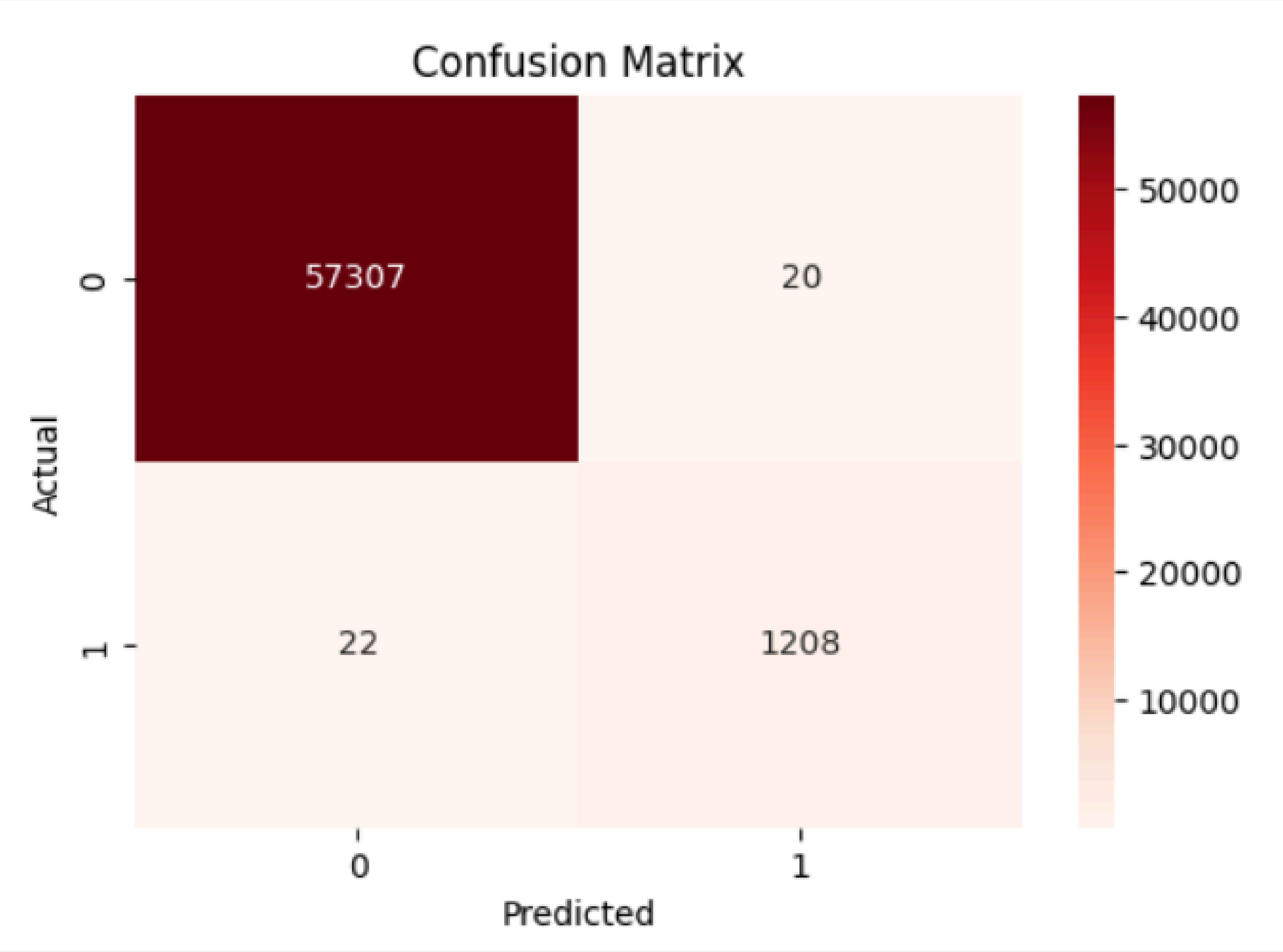

```
Accuracy: 0.9978994825554588
99.78994825554588
```

```
Classification Report:
              precision    recall  f1-score   support

           0       1.00      1.00      1.00     57327
           1       1.00      0.90      0.95      1230

    accuracy                           1.00     58557
   macro avg       1.00      0.95      0.97     58557
weighted avg       1.00      1.00      1.00     58557
```

Xgboost Classifier




```
Accuracy: 0.9992827501408884
99.92827501408884
```

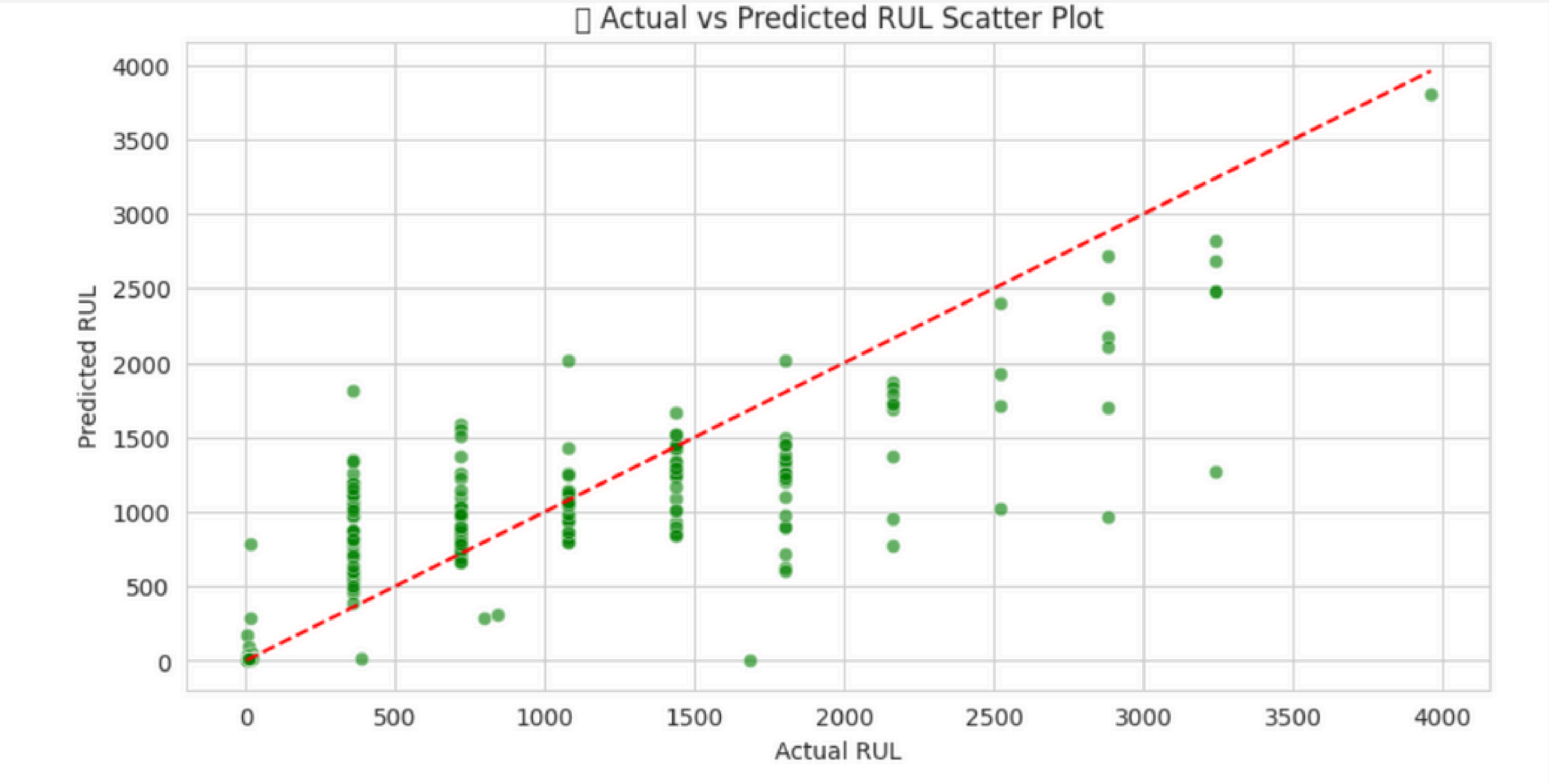
```
Classification Report:
              precision    recall  f1-score   support

           0       1.00      1.00      1.00     57327
           1       0.98      0.98      0.98      1230

    accuracy                           1.00     58557
   macro avg       0.99      0.99      0.99     58557
weighted avg       1.00      1.00      1.00     58557
```

RANDOM FOREST REGRESSOR FOR RUL PREDICTION

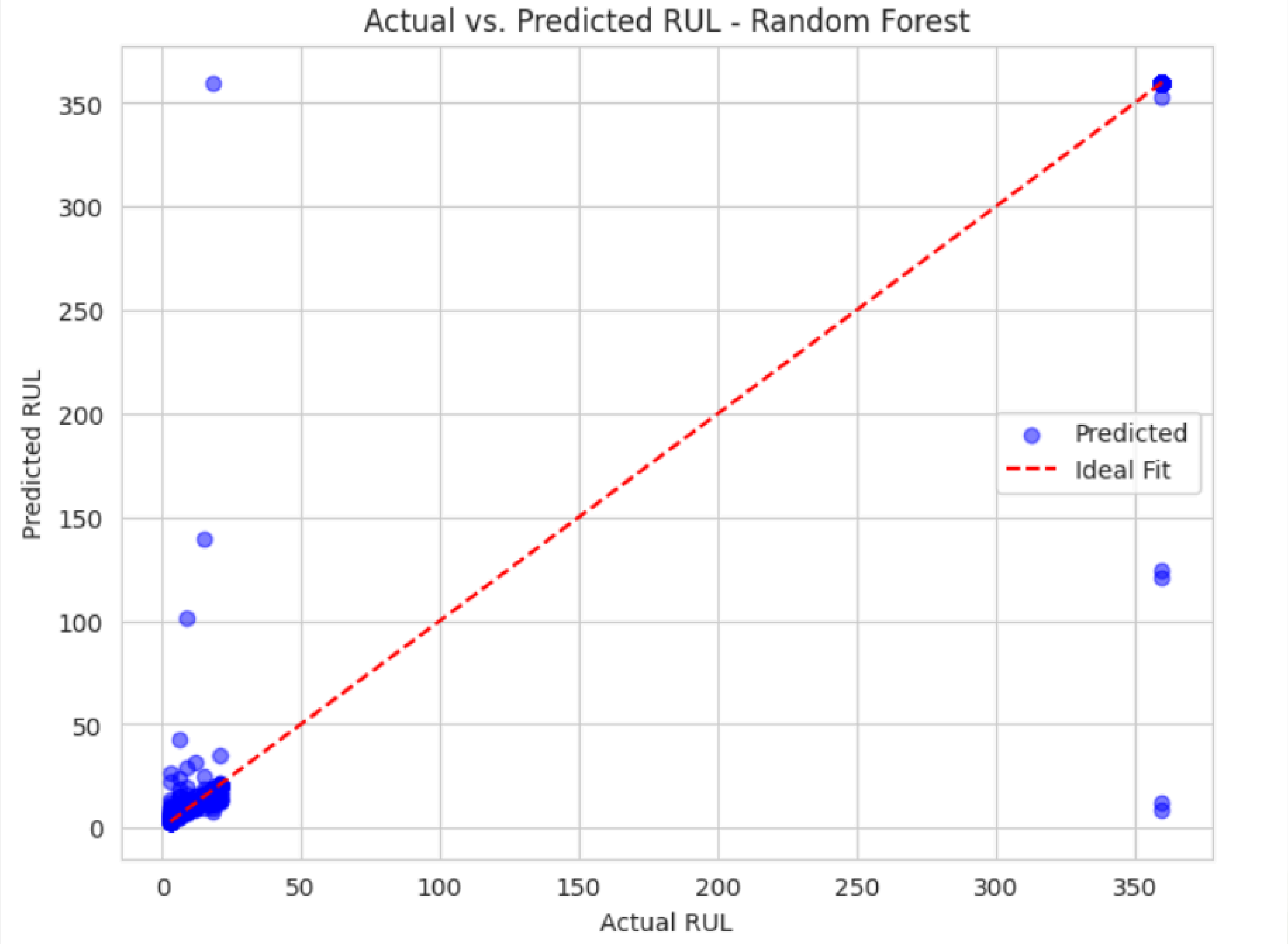
- 📌 Mean Absolute Error (MAE): 64.51
- 📌 Mean Squared Error (MSE): 47387.70
- 📌 Root Mean Squared Error (RMSE): 217.69
- 📌 R^2 Score: 0.8246



Initial result

After Winsorizing at 85% seems to remove extreme outliers without losing valuable data

- 📌 Mean Absolute Error (MAE): 2.20
- 📌 Mean Squared Error (MSE): 424.20
- 📌 Root Mean Squared Error (RMSE): 20.60
- 📌 R^2 Score: 0.9733



FINAL RESULT

Xgboost regressor for RUL prediction

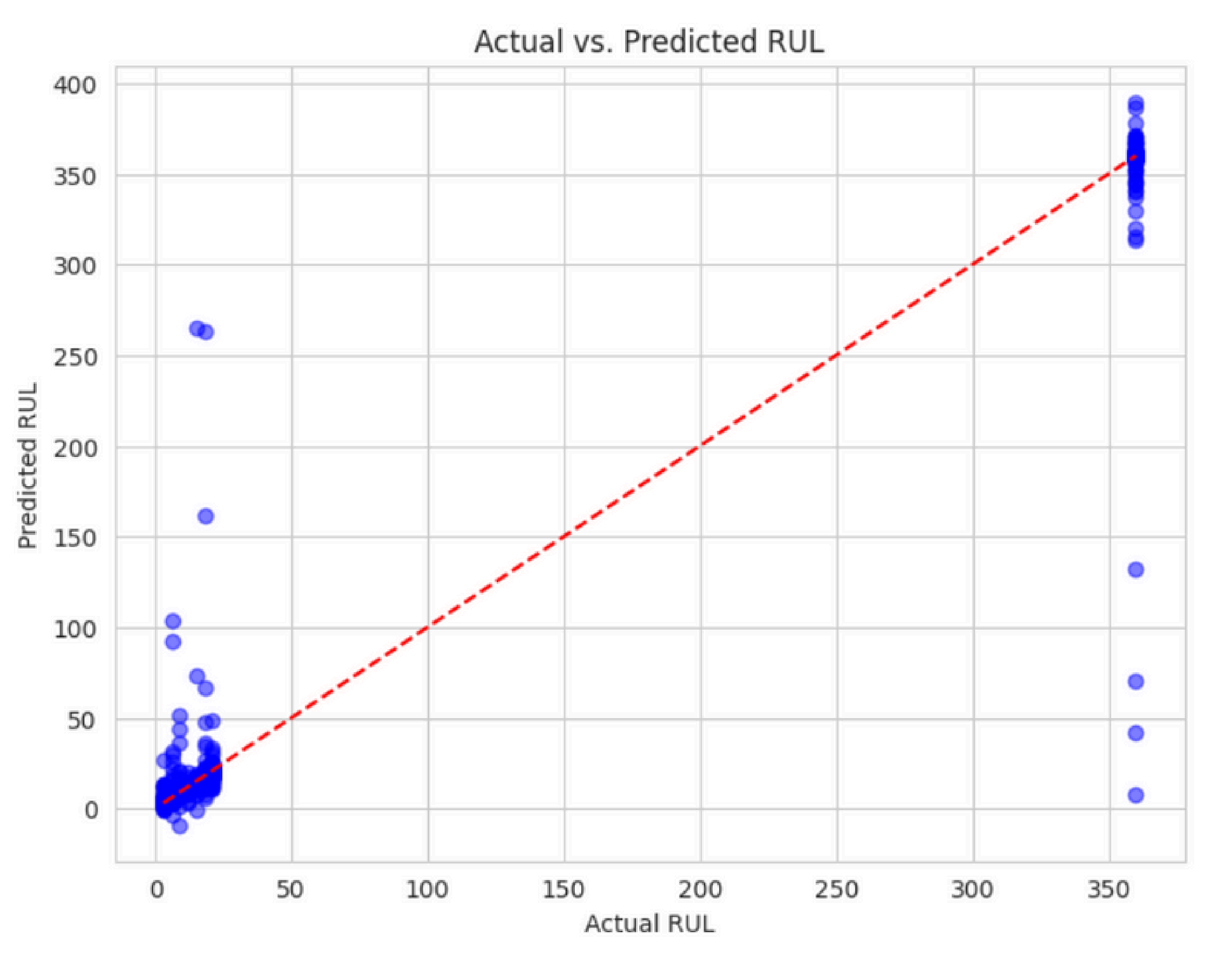
📌 Mean Absolute Error (MAE): 5.05
📌 Mean Squared Error (MSE): 597.39
📌 Root Mean Squared Error (RMSE): 24.44
📌 R^2 Score: 0.9624

Initial Result

After finding best parameter

📌 Mean Absolute Error (MAE): 3.85
📌 Mean Squared Error (MSE): 464.56
📌 Root Mean Squared Error (RMSE): 21.55
📌 R^2 Score: 0.9708

Final Result



UNIQUE CONTRIBUTION IN PROJECT

- Intelligent Maintenance Scheduling: Our approach optimizes maintenance actions based on Remaining Useful Life (RUL) predictions, ensuring proactive intervention before failures occur.

- Prioritization of High-Risk Machines: We introduce a risk-based prioritization mechanism, focusing on machines with the shortest RUL or highest failure probability, minimizing downtime and maintenance costs.

- Hybrid Approach: By integrating failure detection with predictive maintenance (future failure estimation), our model enhances reliability and operational efficiency.

Conclusion

Our project successfully integrates failure detection and predictive maintenance to enhance machine reliability. By leveraging RUL-based maintenance scheduling and risk-based prioritization, we optimize downtime reduction and cost efficiency. The hybrid approach improves failure prediction accuracy, enabling timely interventions.

Future Scope

- Integration with IoT for real-time data collection and cloud-based monitoring.

- Automated Maintenance Workflow using AI-driven decision-making.

- Scalability to handle complex industrial setups with diverse machine types.

REFERENCES

[1] S. Z. Ali, S. H. Ali, A. Javeed, I. Ullah, S. A. Memon, and S. Manzoor, “Supervised Learning Approaches for Vibration Analysis and Anomaly Detection in Rotating Machines,” 2024 3rd International Conference on Emerging Trends in Electrical, Control, and Telecommunication Engineering (ETECTE), pp. 1–6, Nov. 2024,
doi: https://doi.org/10.1109/etecte63967.2024.10824020.

[2] D. Akula., D. S. Kulkarni, S. Kaur, L. V. Sambasivarao, and K. S. Singh, “Advanced Machine Learning Algorithms for Predictive Maintenance in Industrial Manufacturing Systems,” South Eastern European Journal of Public Health, pp. 716–723, Nov. 2024,
doi: https://doi.org/10.70135/seejph.vi.2057.

[3] F. Mateo, J. Vila-Francés, E. Soria-Olivas, M. Martínez-Sober, and Antonio-José Serrano-López, “Dynamic classifier auditing by unsupervised anomaly detection methods: an application in packaging industry predictive maintenance,” May 20, 2024.

[4] A. Immadisetty, “Machine Learning for Real-Time Anomaly Detection.”
Available: https://www.ijfmr.com/papers/2024/6/33087.pdf

[5] Kahiomba Sonia Kiangala and Z. Wang, “An experimental anomaly detection framework for a conveyor motor system using recurrent neural network and dendritic gated neural network,” IET Collaborative Intelligent Manufacturing, vol. 7, no. 1, Jan. 2025, doi: https://doi.org/10.1049/cim2.70017.

[6] Amaia Arregi, A. Barrutia, and Iñigo Bediaga, “End-to-End Methodology for Predictive Maintenance Based on Fingerprint Routines and Anomaly Detection for Machine Tool Rotary Components,” Journal of Manufacturing and Materials Processing, vol. 9, no. 1, pp. 12–12, Jan. 2025, doi: https://doi.org/10.3390/jmmp9010012.

[7] R. de, L. Radrigan, and A. S. Morales, “Enhancing Predictive Maintenance in Mining Mobile Machinery through a TinyML-enabled Hierarchical Inference Network,” Nov. 11, 2024.

[8] H. Lee, S. Lee, and S. Choi, “Perturbation-driven data augmentation for time series anomaly detection improvement in predictive maintenance,” 2021 IEEE International Conference on Big Data (Big Data), pp. 1367–1372, Dec. 2024, doi: https://doi.org/10.1109/bigdata62323.2024.10825569

[9] W. Doorsamy and Pitshou Bokoro, “An Investigation into Unsupervised Anomaly Detection for Data-Driven Predictive Maintenance,” pp. 1–12, Aug. 2024, doi: https://doi.org/10.1109/is61756.2024.10705190.

[10] T. Parmar, “Recent Advancements in Anomaly Detection for Manufacturing Systems,” International Scientific Journal of Engineering and Management, vol. 04, no. 01, pp. 1–7, Jan. 2025, doi: https://doi.org/10.55041/isjem01290


Thank You!