Predictive Maintenance in Manufacturing using IoT Sensor data

An Explainable Machine Learning Approach for Failure Prediction

Capstone Project Report

Nandini Garudacharpalya Amaresh (80305005)
Kavya Gottam (50982029)

Supervised and Guided by:
Prof. Dr. Humera Noor Minhas
Mr. Ebby George Varghese

University of Europe for Applied Sciences

ABSTRACT

MANUFACTURING INDUSTRIES FACE HIGH FINANCIAL LOSS DUE TO

UNEXPECTED MACHINE FAILURES

- Unplanned downtime affects
- Production schedules
- Maintenance cost
- Worker safety
- Traditional maintenance methods are inefficient
- Predictive Maintenance (PdM) uses IoT sensor data to

ANTICIPATE FAILURES

- Machine Learning enables early failure detection
- Explainable AI improves trust and adoption

INDUSTRY 4.0 & MAINTENANCE EVOLUTION

- Industry 4.0 enables data-driven manufacturing
- Machines operate under continuous mechanical stress
- Reactive maintenance causes unplanned downtime
- Preventive maintenance leads to over-maintenance
- Predictive maintenance anticipates failures using data
- Maintenance strategy directly impacts productivity
- Shift from reactive to predictive maintenance

PROBLEM STATEMENT

- Machine failures disrupt manufacturing operations
- Sensor data challenges:
- Sensor data is High dimensionality
- Nonlinear behavior
- Class imbalance Failure events are rare and highly imbalanced
- Existing ML models solutions lack transparency
- Maintenance decisions require explanation Requirement: reliable and interpretable solution

OBJECTIVES

- Develop an accurate and explainable predictive maintenance system using IoT sensor data and machine learning
- Analyze sensor data to understand machine behavior and failure patterns through exploratory data analysis
- Create domain-informed engineered features capturing mechanical stress, thermal imbalance, and tool wear
- Build and compare Logistic Regression, Random Forest, and XGBoost models under imbalanced data conditions
- Evaluate models using accuracy, precision, recall, F1-score, and ROC–AUC
- Apply SHAP-based explainability for global and individual prediction interpretation
- Deploy the best model as an interactive Streamlit dashboard for real-time failure prediction

PROJECT SCOPE

- Analysis of historical IoT sensor data
- Binary classification of machine failure
- Feature engineering using mechanical and thermal relationships
- Explainable AI using SHAP values
- Interactive dashboard-based deployment

METHODOLOGY

Dataset Description

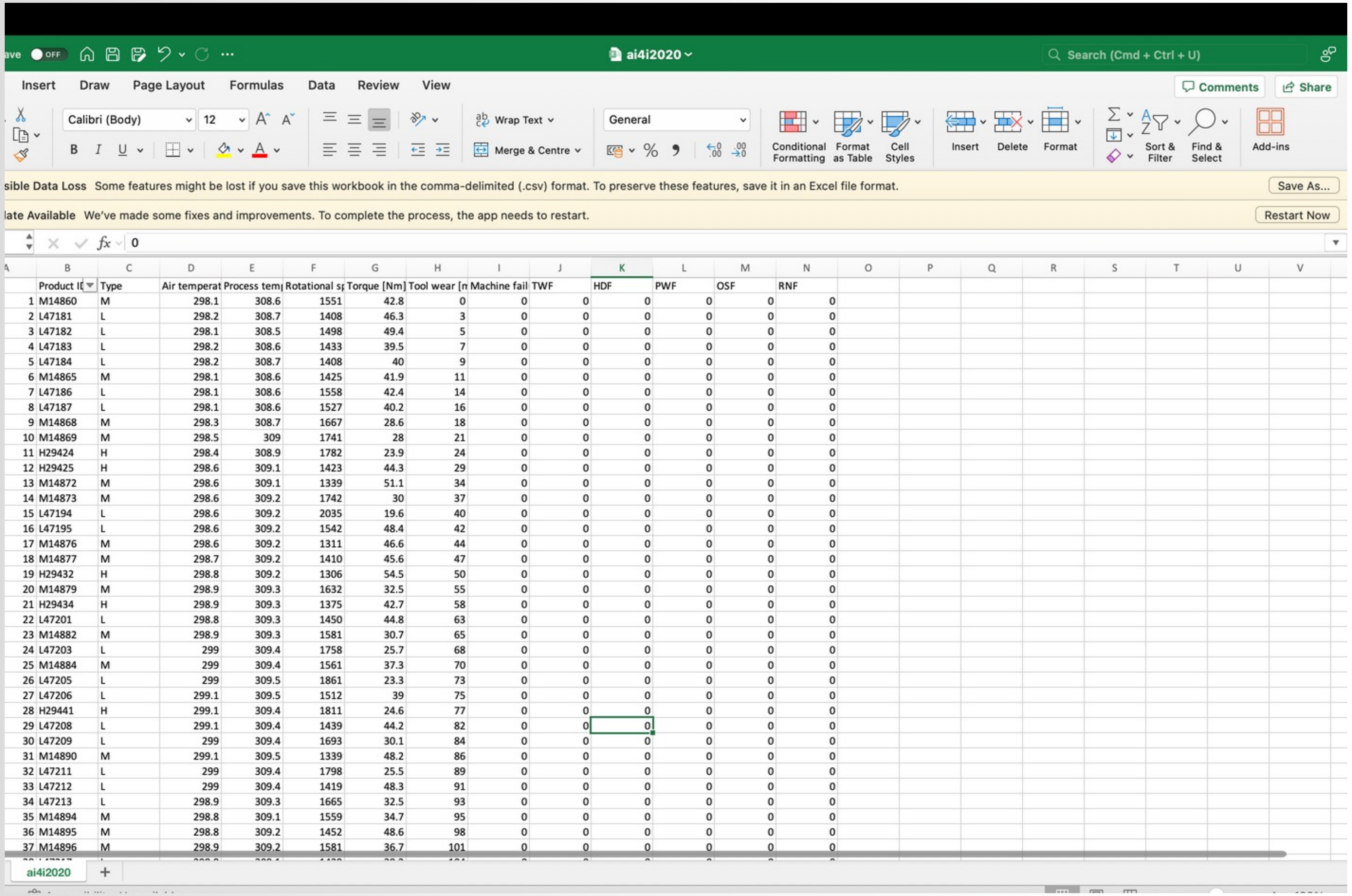
- Synthetic milling machine dataset
- 10,000 observations
- 14 sensor & operational features
- Binary target: Machine failure(0or 1)
- Product quality types: L, M, H

FAILURE MODES IN DATASET

Machine failure occurs due to:

- Tool Wear Failure (TWF)
- Heat Dissipation Failure (HDF)
- Power Failure (PWF)
- Overstrain Failure (OSF)
- Random Failure (RNF)

All combined into a single binary failure label



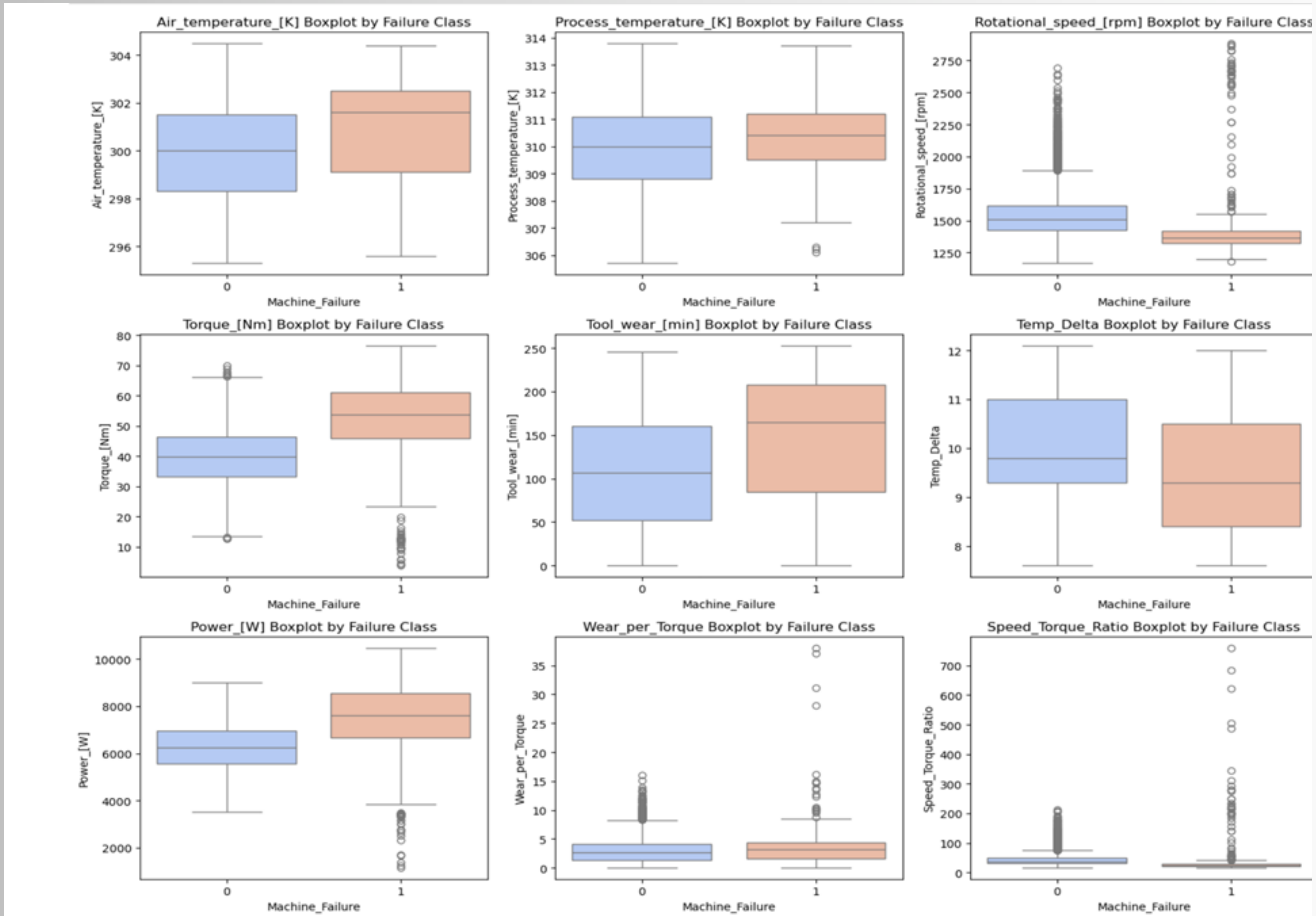
	Product ID	Type	Air temperat	Process tem	Rotational sp	Torque [Nm]	Tool wear [m	Machine fail	TWF	HDF	PWF	OSF	RNF
1	M14860	M	298.1	308.6	1551	42.8	0	0	0	0	0	0	0
2	L47181	L	298.2	308.7	1408	46.3	3	0	0	0	0	0	0
3	L47182	L	298.1	308.5	1498	49.4	5	0	0	0	0	0	0
4	L47183	L	298.2	308.6	1433	39.5	7	0	0	0	0	0	0
5	L47184	L	298.2	308.7	1408	40	9	0	0	0	0	0	0
6	M14865	M	298.1	308.6	1425	41.9	11	0	0	0	0	0	0
7	L47186	L	298.1	308.6	1558	42.4	14	0	0	0	0	0	0
8	L47187	L	298.1	308.6	1527	40.2	16	0	0	0	0	0	0
9	M14868	M	298.3	308.7	1667	28.6	18	0	0	0	0	0	0
10	M14869	M	298.5	309	1741	28	21	0	0	0	0	0	0
11	H29424	H	298.4	308.9	1782	23.9	24	0	0	0	0	0	0
12	H29425	H	298.6	309.1	1423	44.3	29	0	0	0	0	0	0
13	M14872	M	298.6	309.1	1339	51.1	34	0	0	0	0	0	0
14	M14873	M	298.6	309.2	1742	30	37	0	0	0	0	0	0
15	L47194	L	298.6	309.2	2035	19.6	40	0	0	0	0	0	0
16	L47195	L	298.6	309.2	1542	48.4	42	0	0	0	0	0	0
17	M14876	M	298.6	309.2	1311	46.6	44	0	0	0	0	0	0
18	M14877	M	298.7	309.2	1410	45.6	47	0	0	0	0	0	0
19	H29432	H	298.8	309.2	1306	54.5	50	0	0	0	0	0	0
20	M14879	M	298.9	309.3	1632	32.5	55	0	0	0	0	0	0
21	H29434	H	298.9	309.3	1375	42.7	58	0	0	0	0	0	0
22	L47201	L	298.8	309.3	1450	44.8	63	0	0	0	0	0	0
23	M14882	M	298.9	309.3	1581	30.7	65	0	0	0	0	0	0
24	L47203	L	299	309.4	1758	25.7	68	0	0	0	0	0	0
25	M14884	M	299	309.4	1561	37.3	70	0	0	0	0	0	0
26	L47205	L	299	309.5	1861	23.3	73	0	0	0	0	0	0
27	L47206	L	299.1	309.5	1512	39	75	0	0	0	0	0	0
28	H29441	H	299.1	309.4	1811	24.6	77	0	0	0	0	0	0
29	L47208	L	299.1	309.4	1439	44.2	82	0	0	0	0	0	0
30	L47209	L	299	309.4	1693	30.1	84	0	0	0	0	0	0
31	M14890	M	299.1	309.5	1339	48.2	86	0	0	0	0	0	0
32	L47211	L	299	309.4	1798	25.5	89	0	0	0	0	0	0
33	L47212	L	299	309.4	1419	48.3	91	0	0	0	0	0	0
34	L47213	L	298.9	309.3	1665	32.5	93	0	0	0	0	0	0
35	M14894	M	298.8	309.1	1559	34.7	95	0	0	0	0	0	0
36	M14895	M	298.8	309.2	1452	48.6	98	0	0	0	0	0	0
37	M14896	M	298.9	309.2	1581	36.7	101	0	0	0	0	0	0

DATA CLEANING & FEATURE ENGINEERING

- **Removed unique identifiers (UID)**
- **Applied one-hot encoding to categorical variables**
- **Normalized and scaled continuous features**
- **Addressed severe class imbalance (~3.4% failure cases) using SMOTE reduce bias toward the majority class**
- **Used a stratified train–test split**
- **Power [W]: Derived from torque and rotational speed to represent actual mechanical power.**
- **Temp Delta: Difference between process and air temperature; useful for detecting cooling system failures (HDF).**
- **Wear_per_Torque: Tool wear normalized by torque, signaling abnormal wear behavior.**
- **Speed_Torque_Ratio: High values may indicate operational stress, relevant for OSF-type failures.**

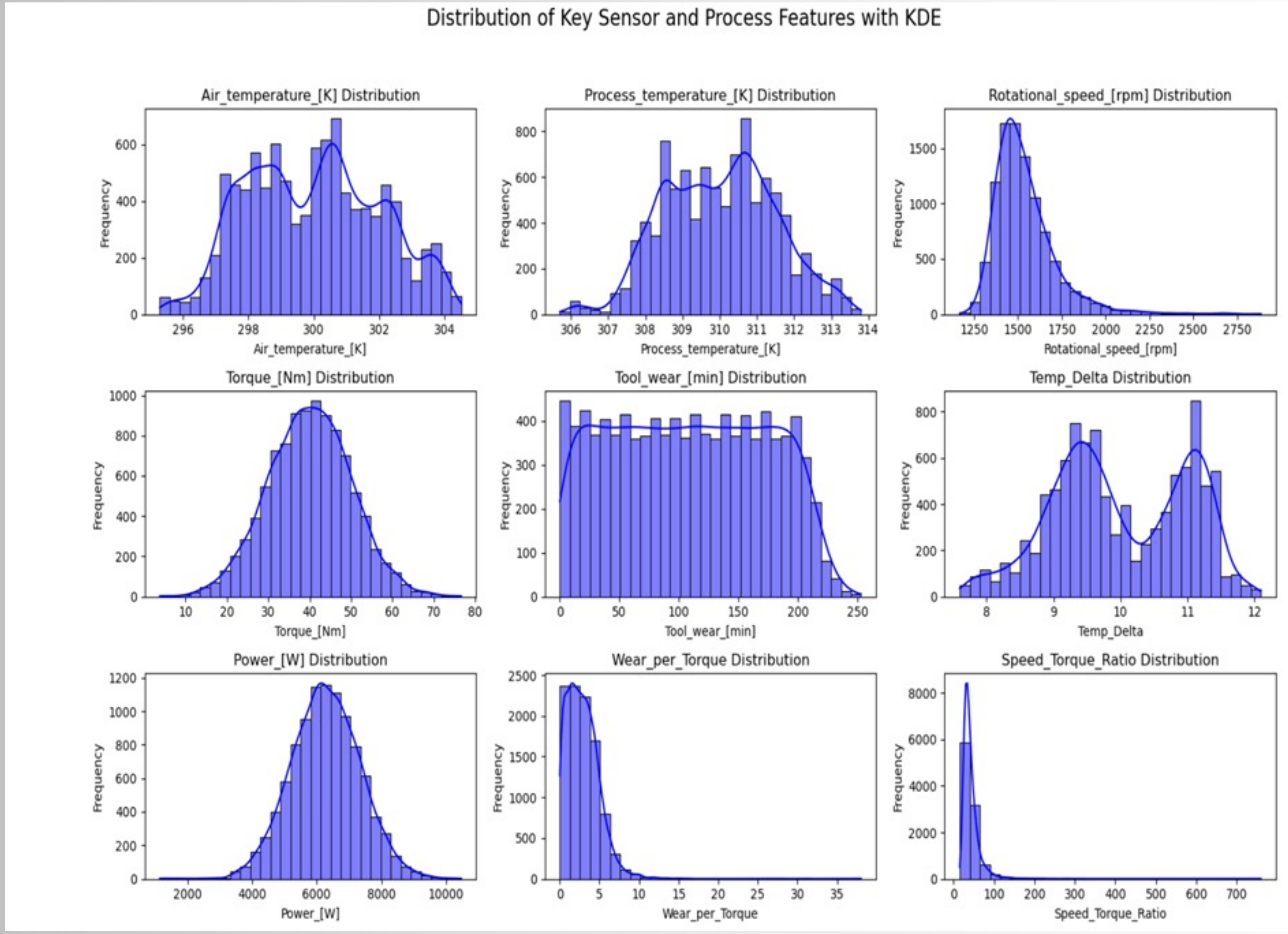
EXPLORATORY DATA ANALYSIS

Boxbplots: Highlight the spread and presence of outliers in each feature across the failure and non-failure classes.



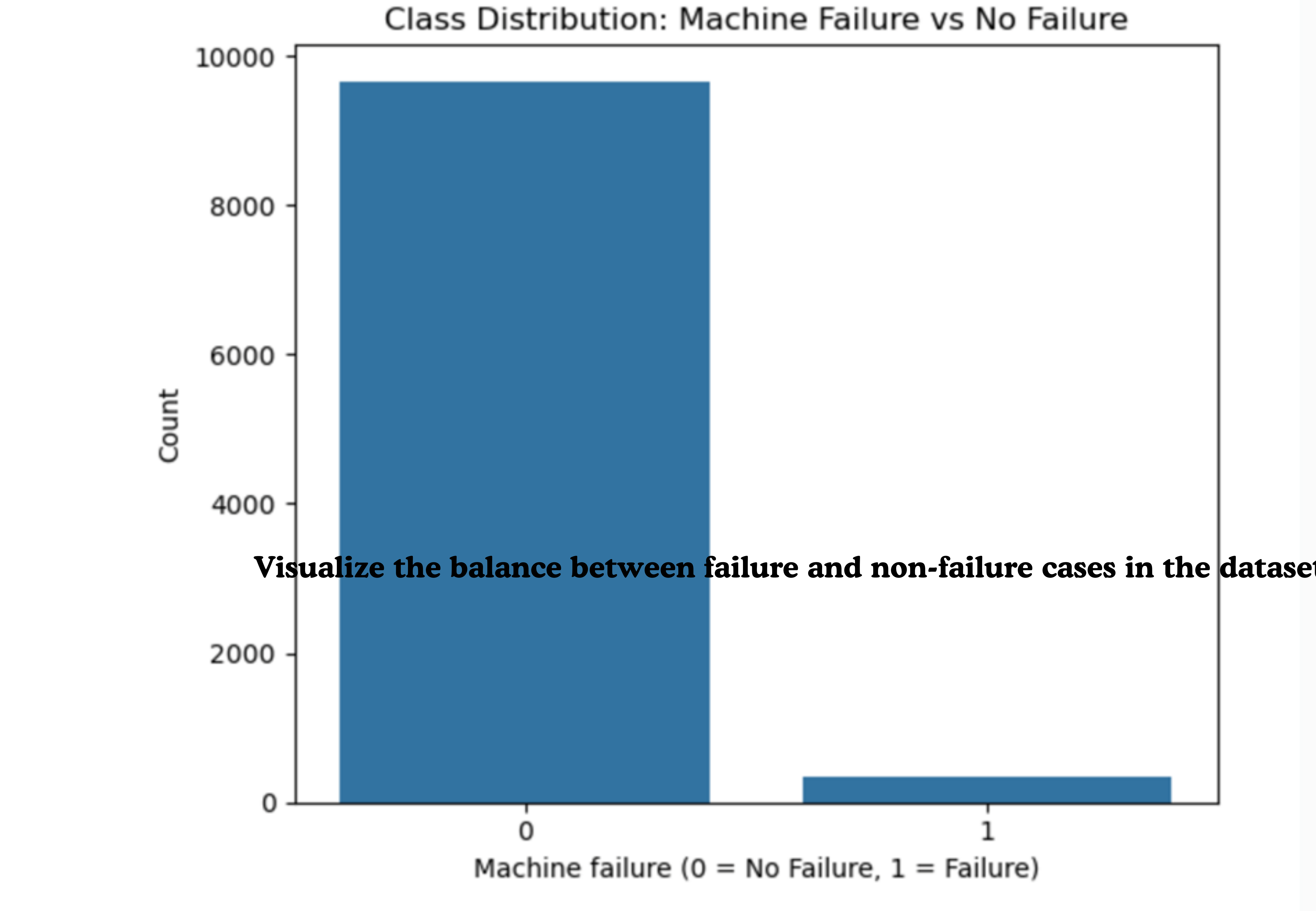
Boxplots of Sensor and Process Features Grouped by Machine Failure

HISTOGRAM OF FEATURES : **Show the distribution of key continuous features, comparing how values differ across machine failure classes.**



Distributions of Sensor and Engineered Features

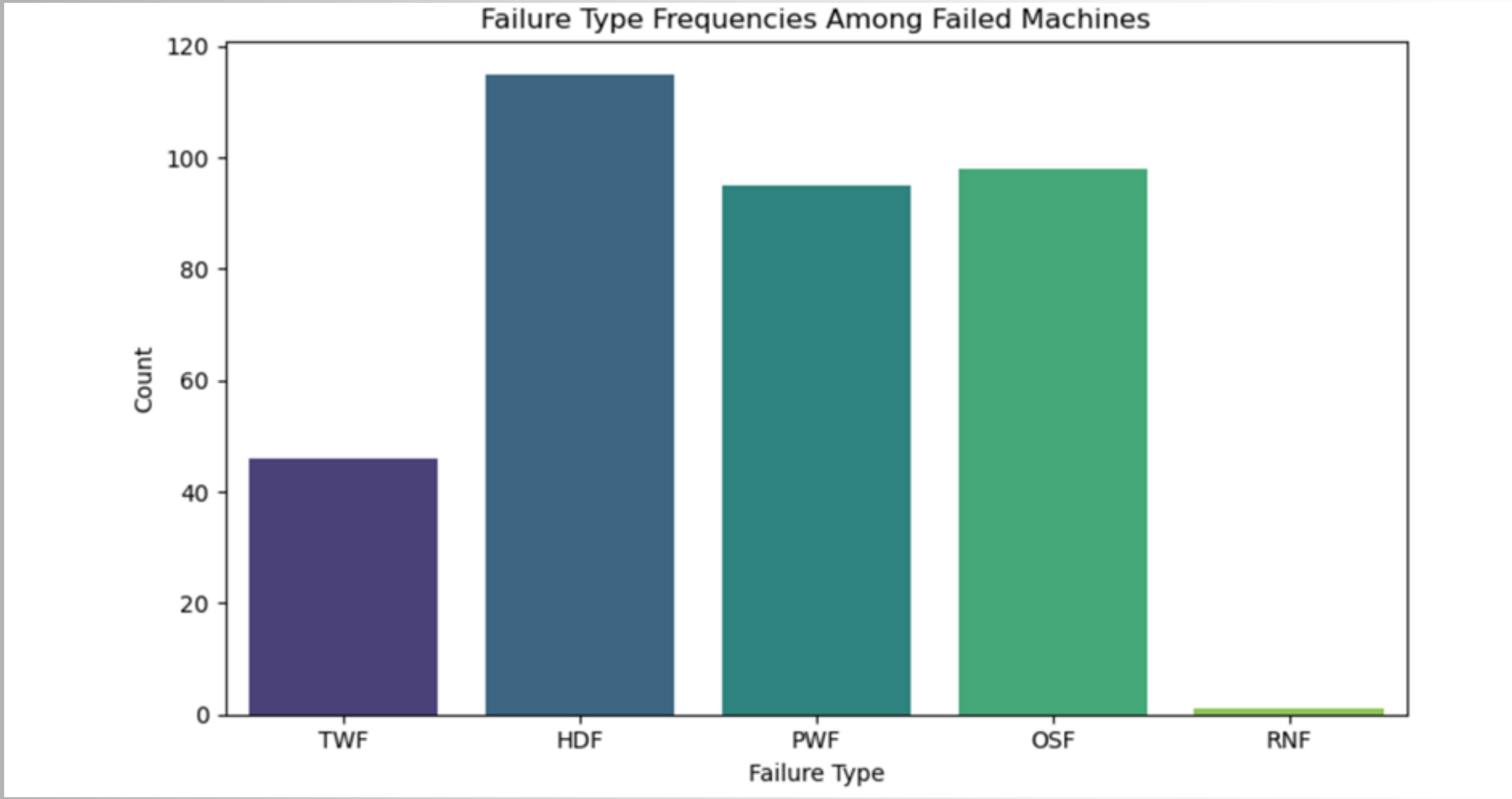
CLASS DISTRIBUTION:



Visualize the balance between failure and non-failure cases in the dataset.

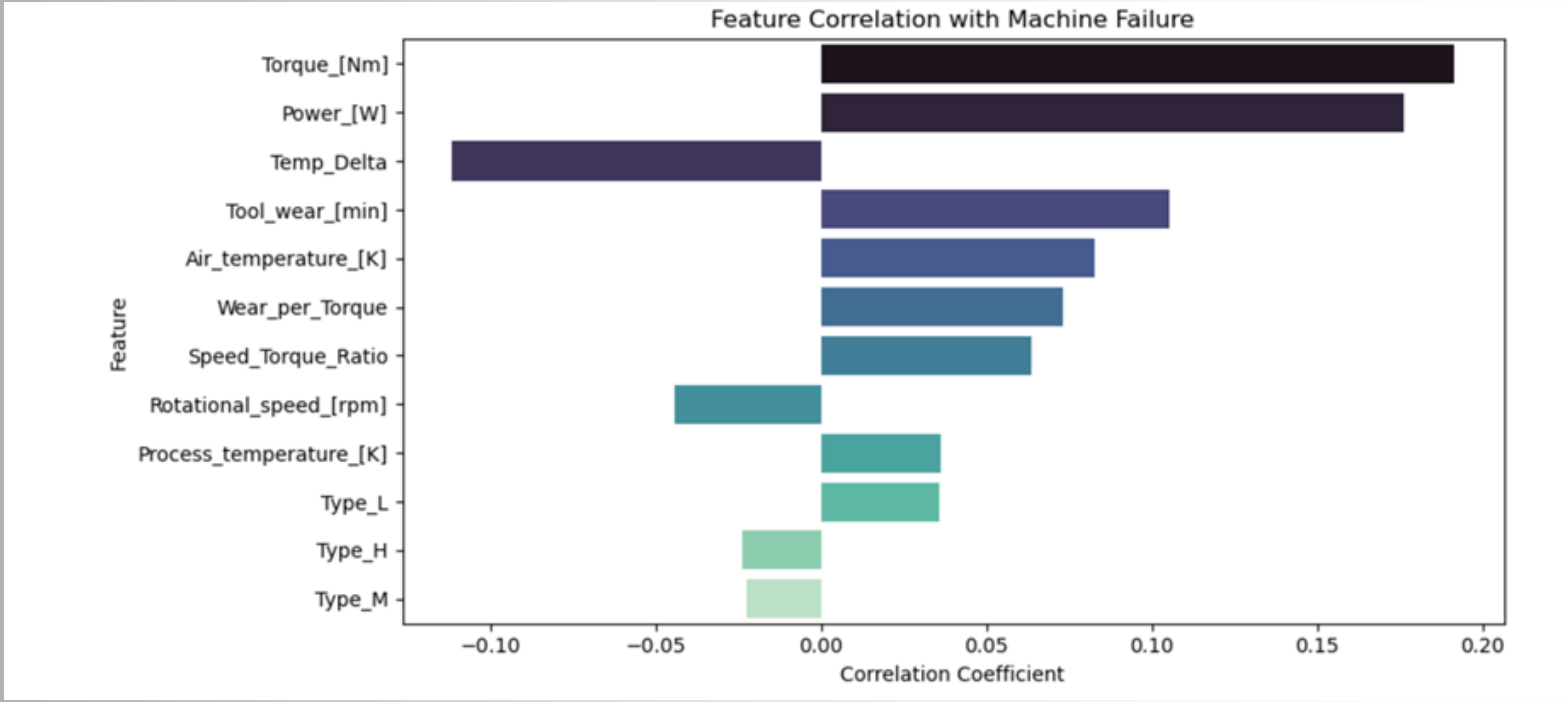
Distribution of Machine Failure vs No Failure

FAILURE TYPE DISTRIBUTION : **Break down the five underlying failure types (HDF, TWF, PWF, OSF, RNF) to understand their frequency and contribution to total failures.**



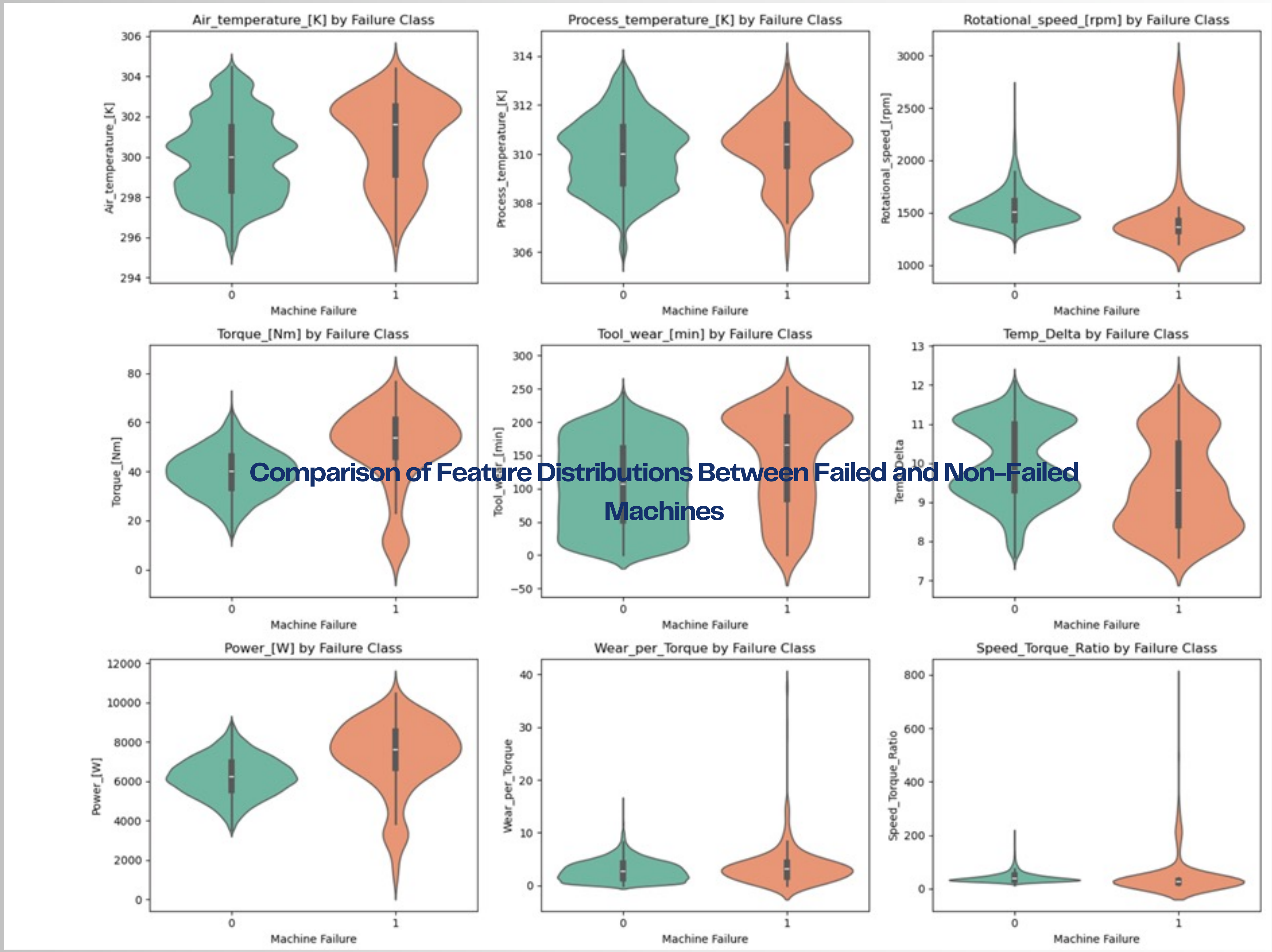
Frequency of Each Failure Type

CORRELATION WITH TARGET BAR PLOT: **Rank numeric features based on their correlation with the binary failure target. Useful for quick feature selection insights.**



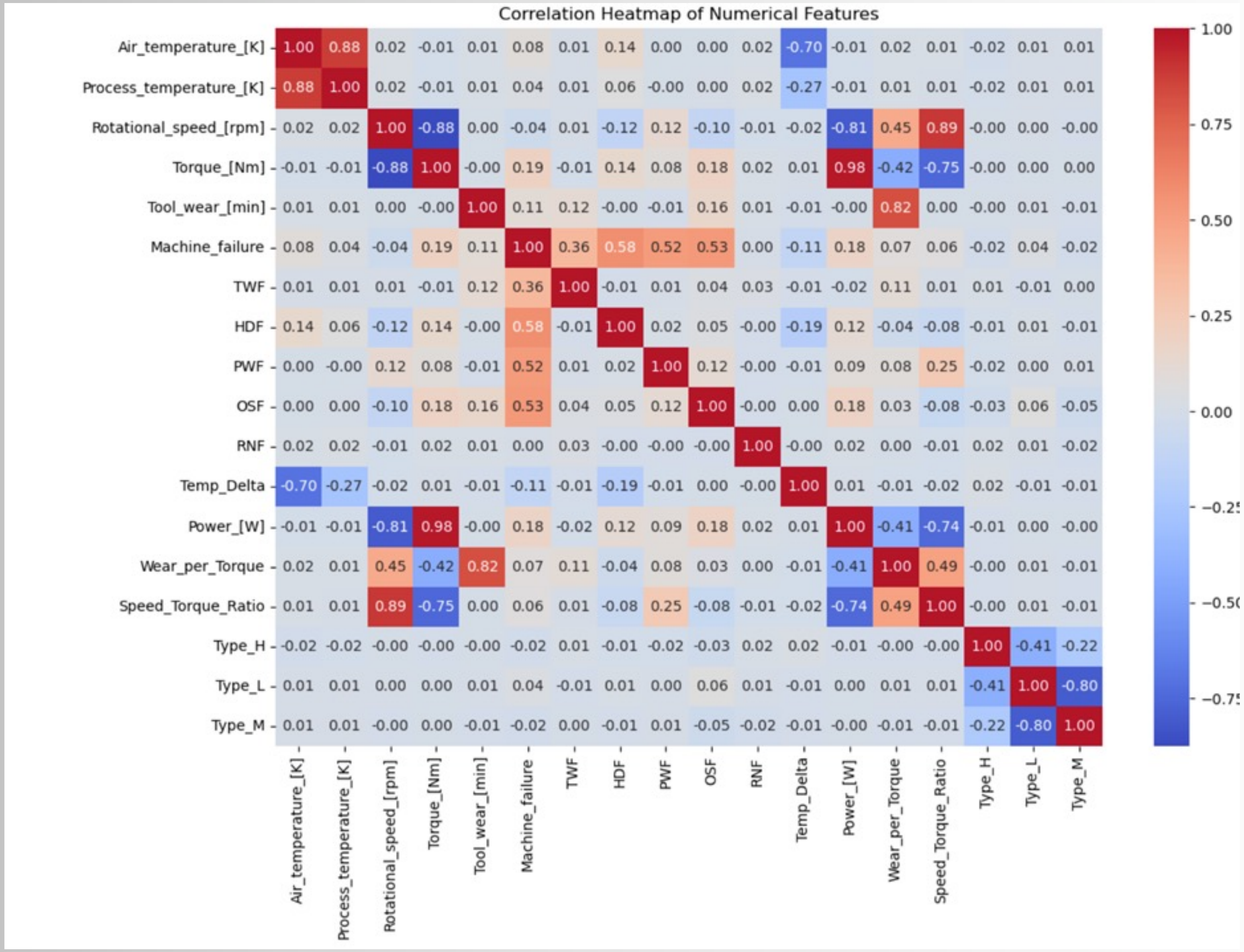
Feature Correlation with Machine Failure

VIOLIN PLOTS:Display both the distribution and density of feature values split by class, helping reveal subtle differences in variable behavior.



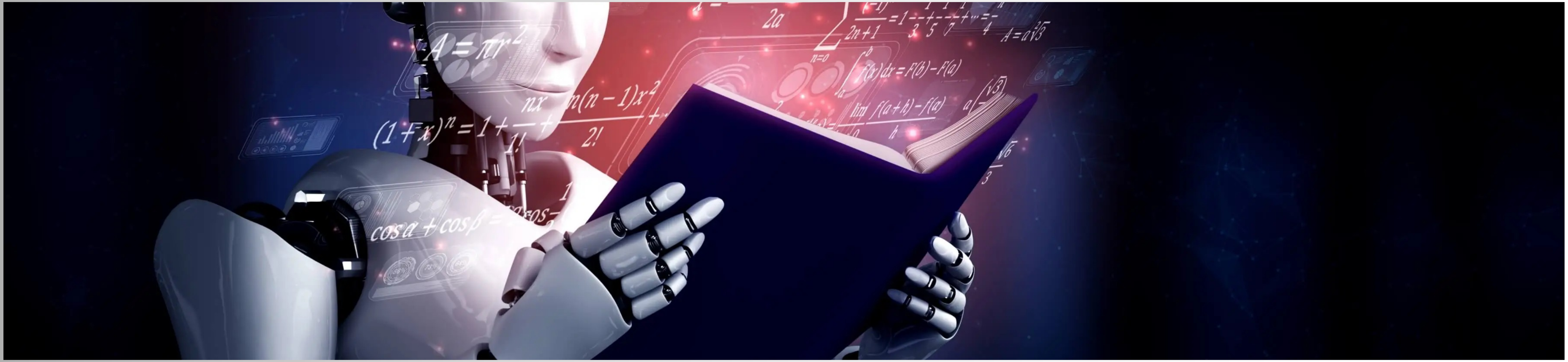
Comparison of Feature Distributions Between Failed and Non-Failed Machines

CORRELATION HEATMAP:Explore pairwise correlations among all numeric features, helping identify redundant variables or multicollinearity.



Feature Correlation Matrix with Engineered and Encoded Variables

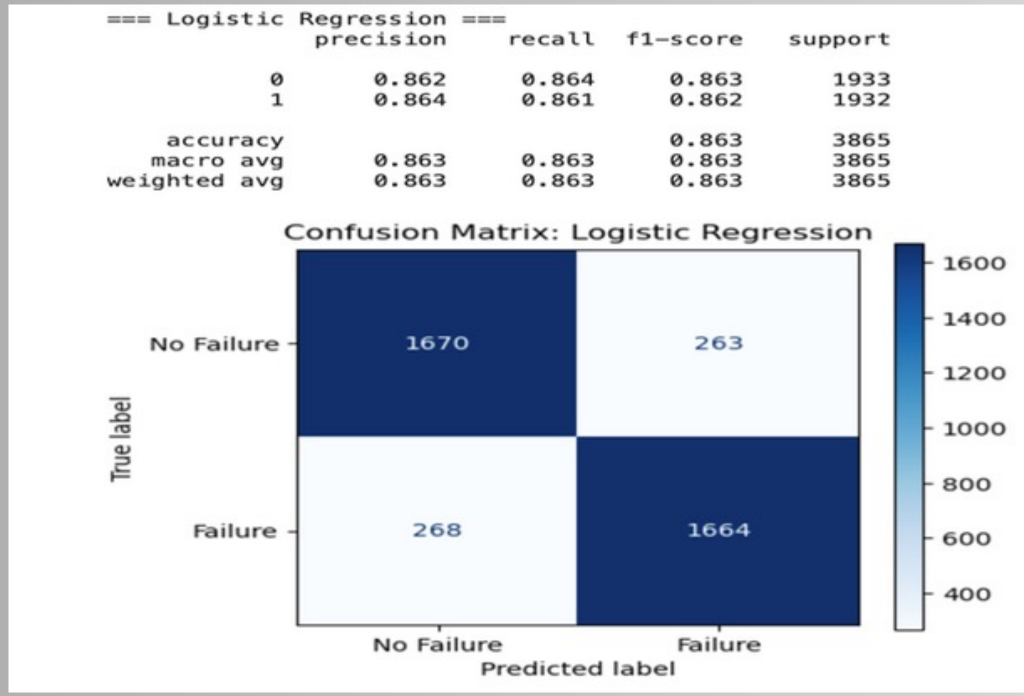
Model Training & Evaluation

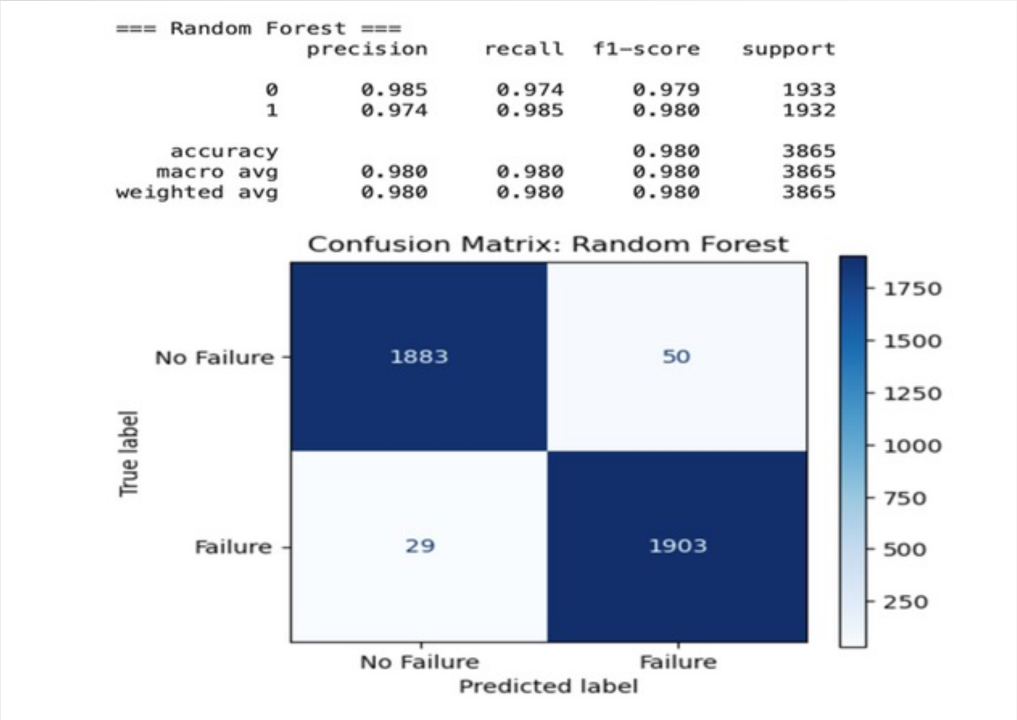


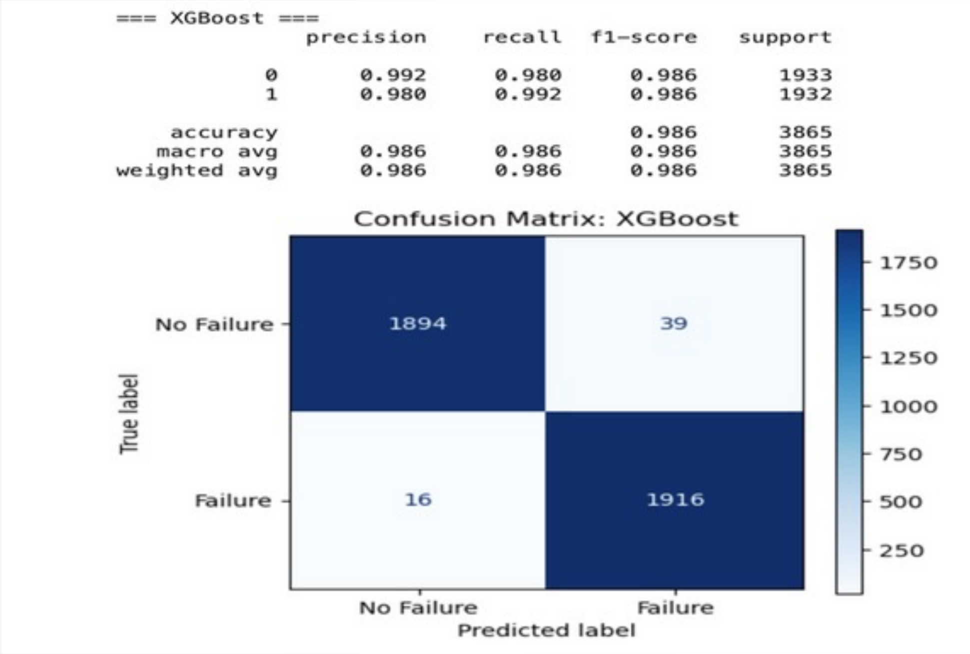
- **Logistic Regression: Basic mode easy to understand and interpret the result.**
- **Random Forest:Model based on trees and able to represent nonlinear relationships and interactions of features.**
- **XGBoost: A high performance, scalable and robust gradient boosting model on imbalanced data.**

- **Confusion Matrices to examine the true, false positives, and negative.**
- **Classification Reports that gives accuracy, recall, F1-score, and precision.**
- **Determination of discriminative ability with Receiver Operating Characteristic (ROC) curves and Area Under the Curve (AUC).**

Model Performance Comparison via Confusion Matrices and Classification Reports







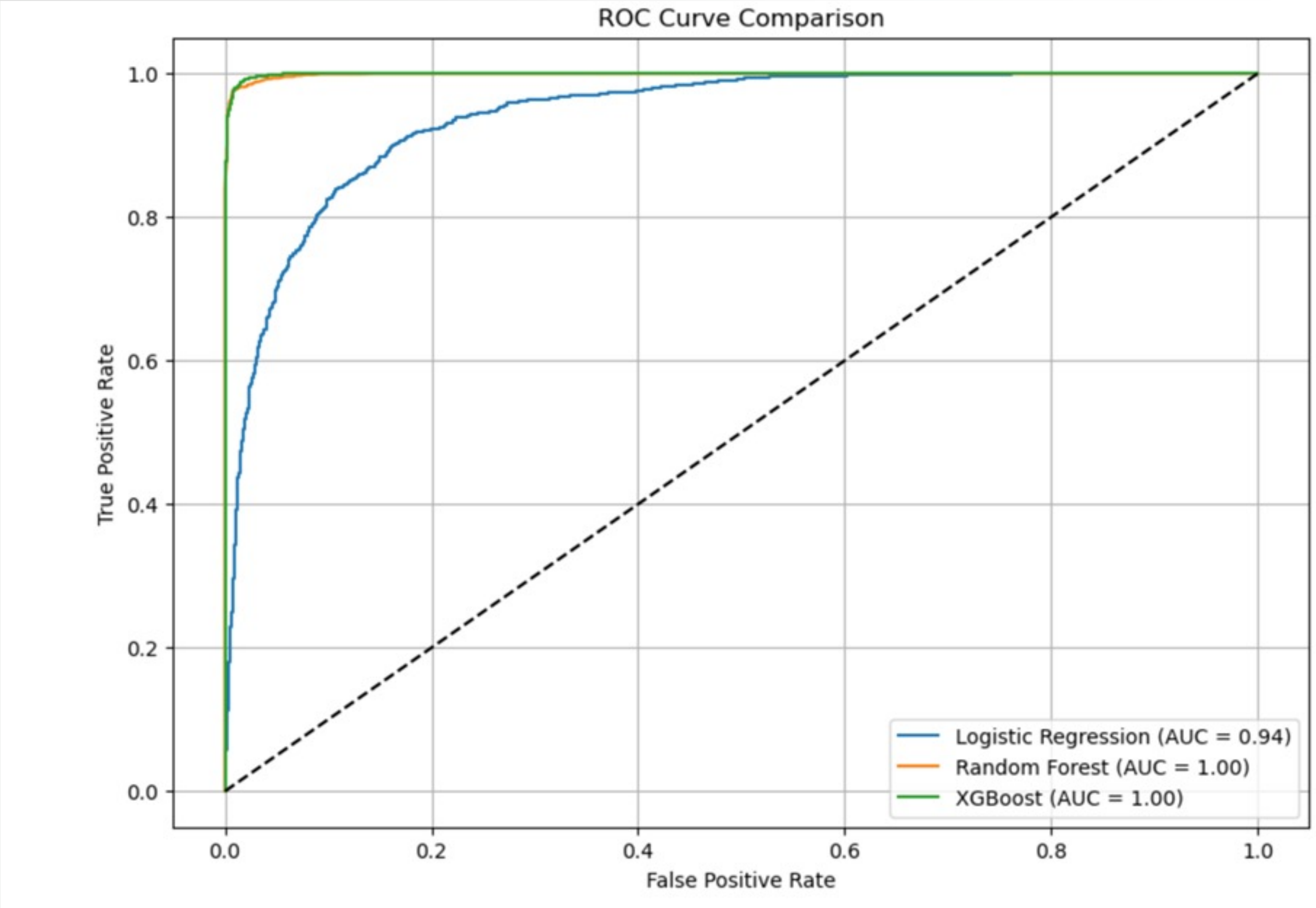
Logistic Regression: ~86% accuracy

Random Forest: ~98% accuracy

XGBoost: 98.6% accuracy

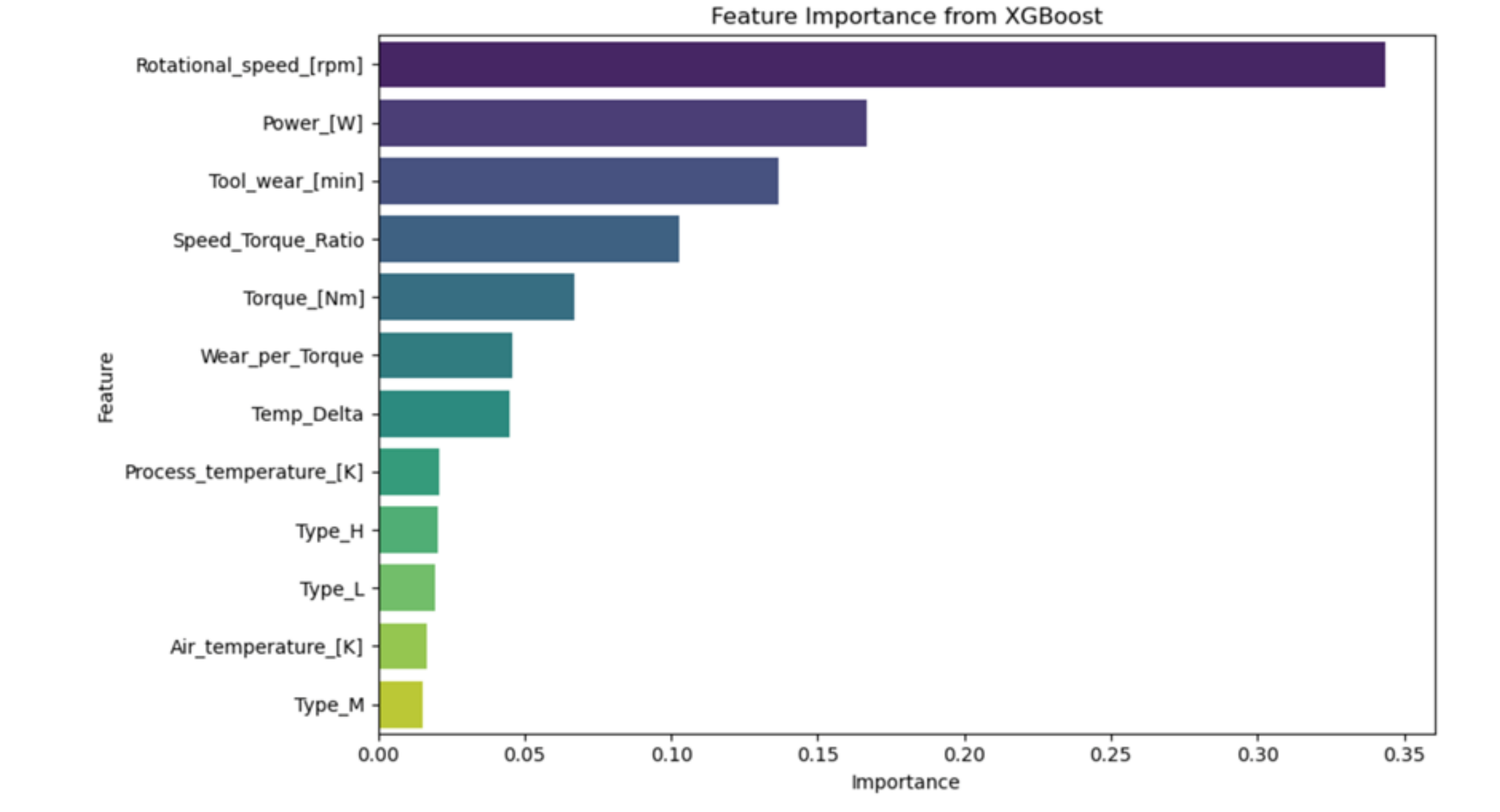
ROC CURVE ANALYSIS

Assess the discriminative power of each model. AUC values closer to 1 indicate strong separation between classes.



FEATURE IMPORTANCE OF BEST MODEL

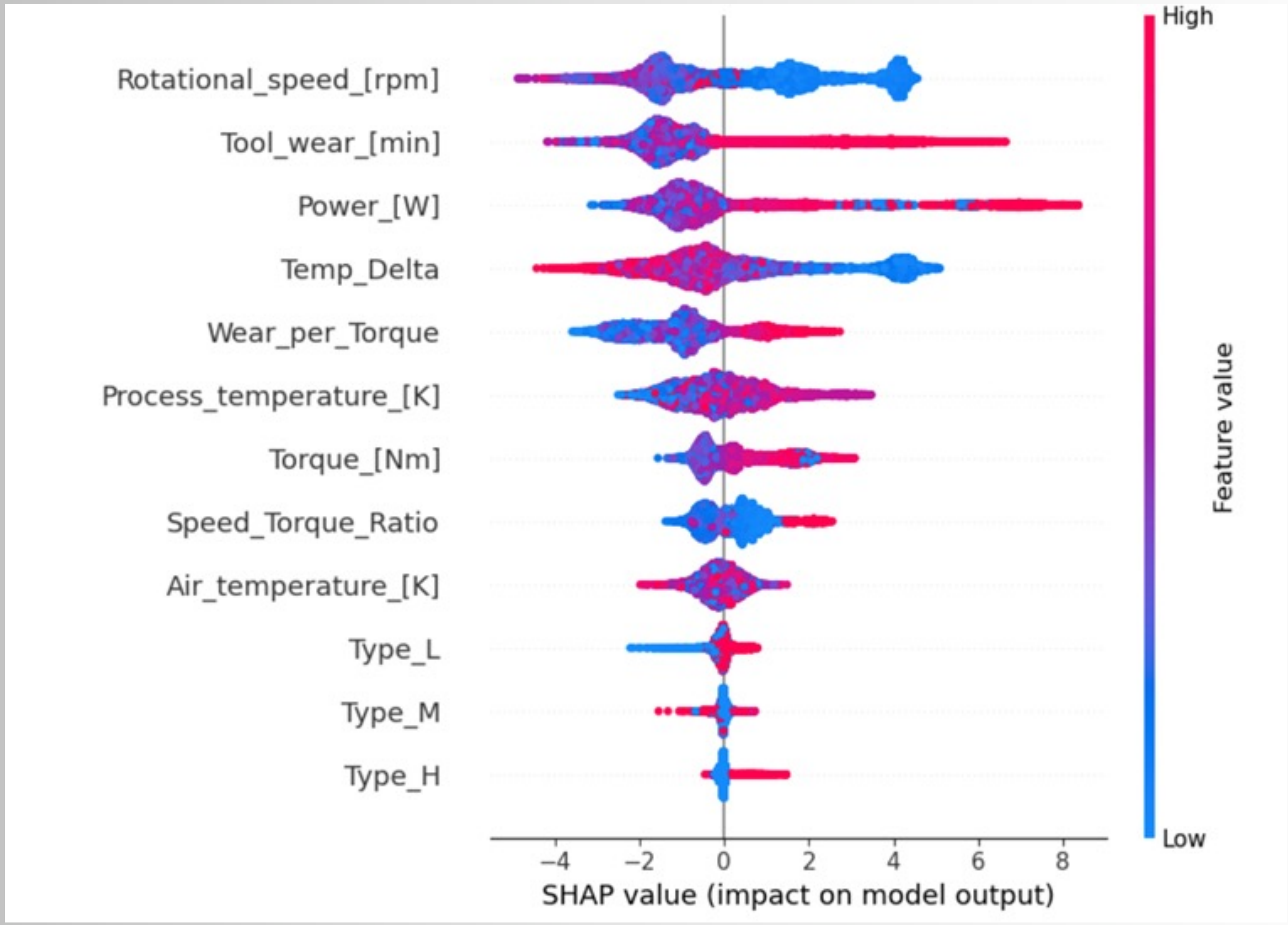
Rank features by their importance according to the XGBoost model, offering insight into what drives predictions.



Feature Importance Ranking from XGBoost

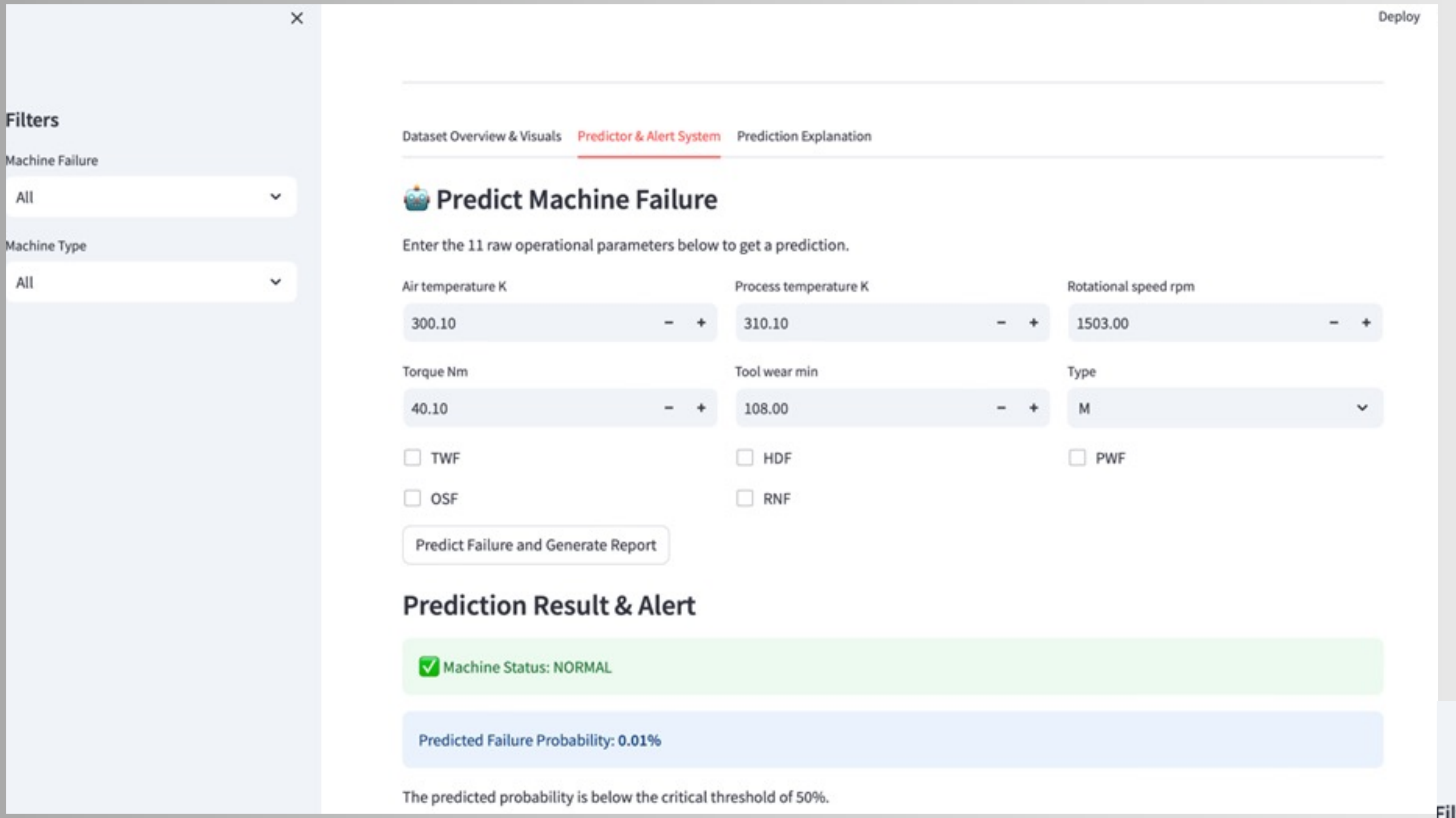
MODEL EXPLAINABILITY WITH SHAP

Summarize global feature influence using SHAP values to interpret how each variable impacts model output across the dataset.



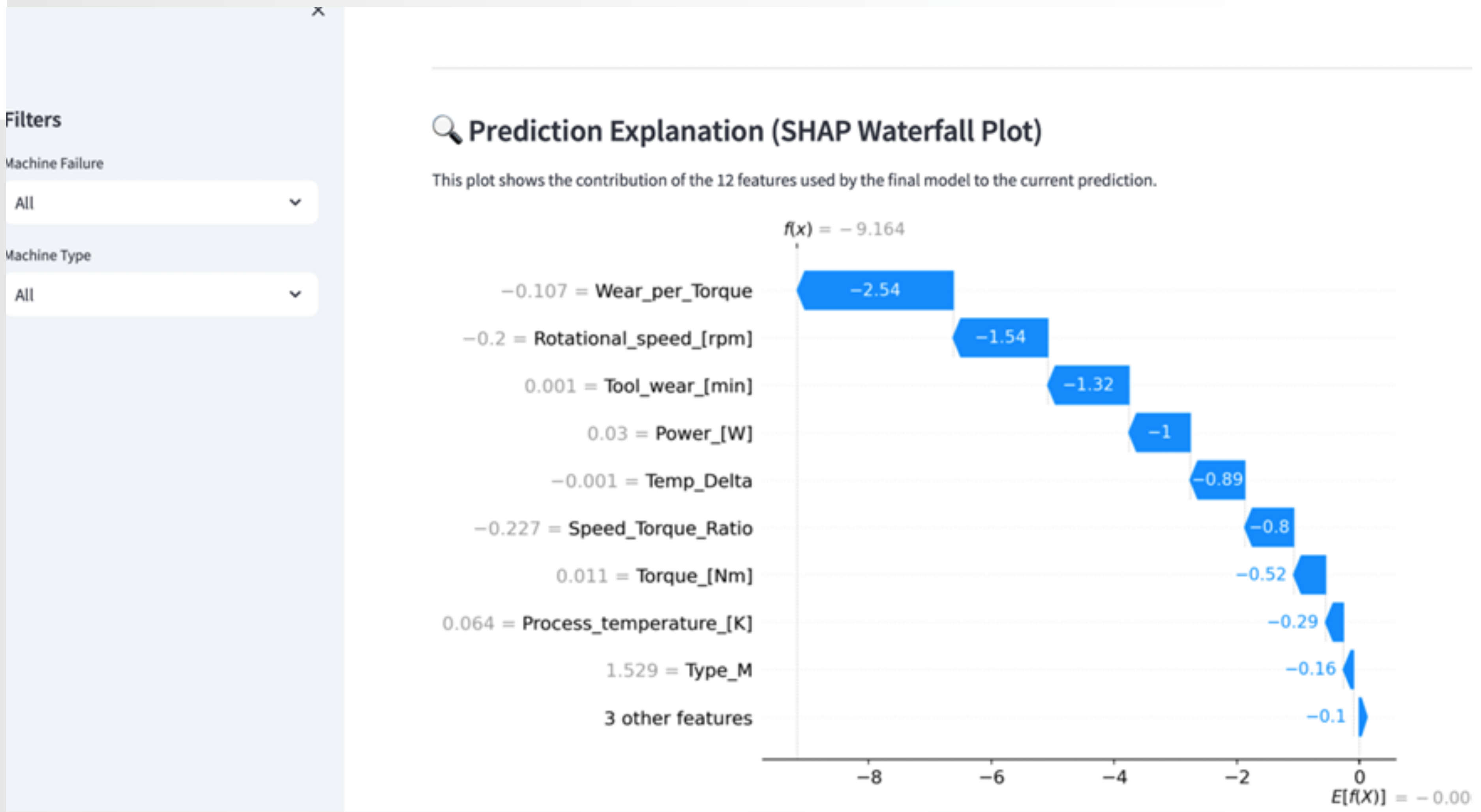
SHAP Summary Plot for Model Explainability

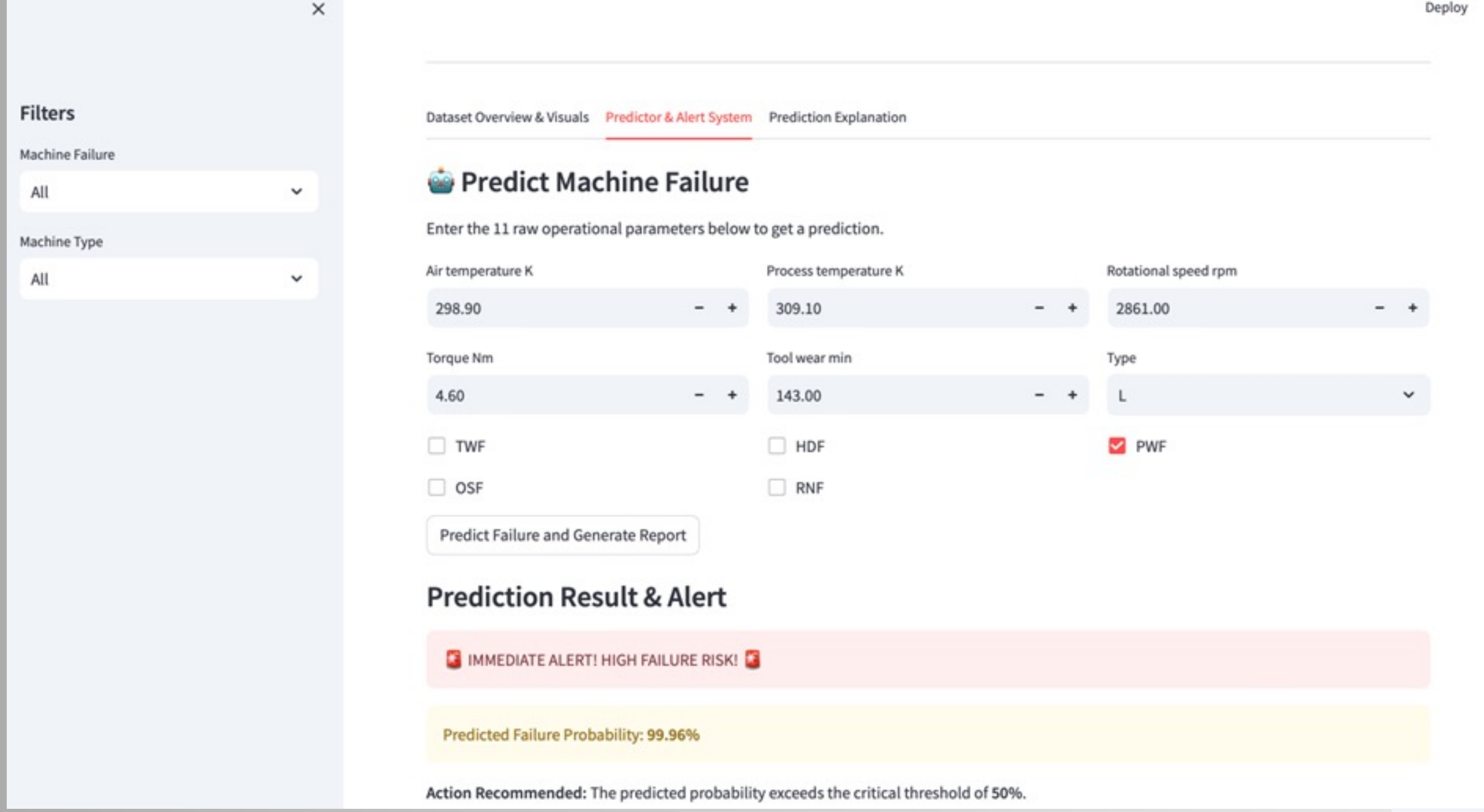
STREAMLIT DASHBOARD



- The system predicts normal machine functioning (No Failure)
- SHAP explanation indicates key features contributing negatively to failure risk

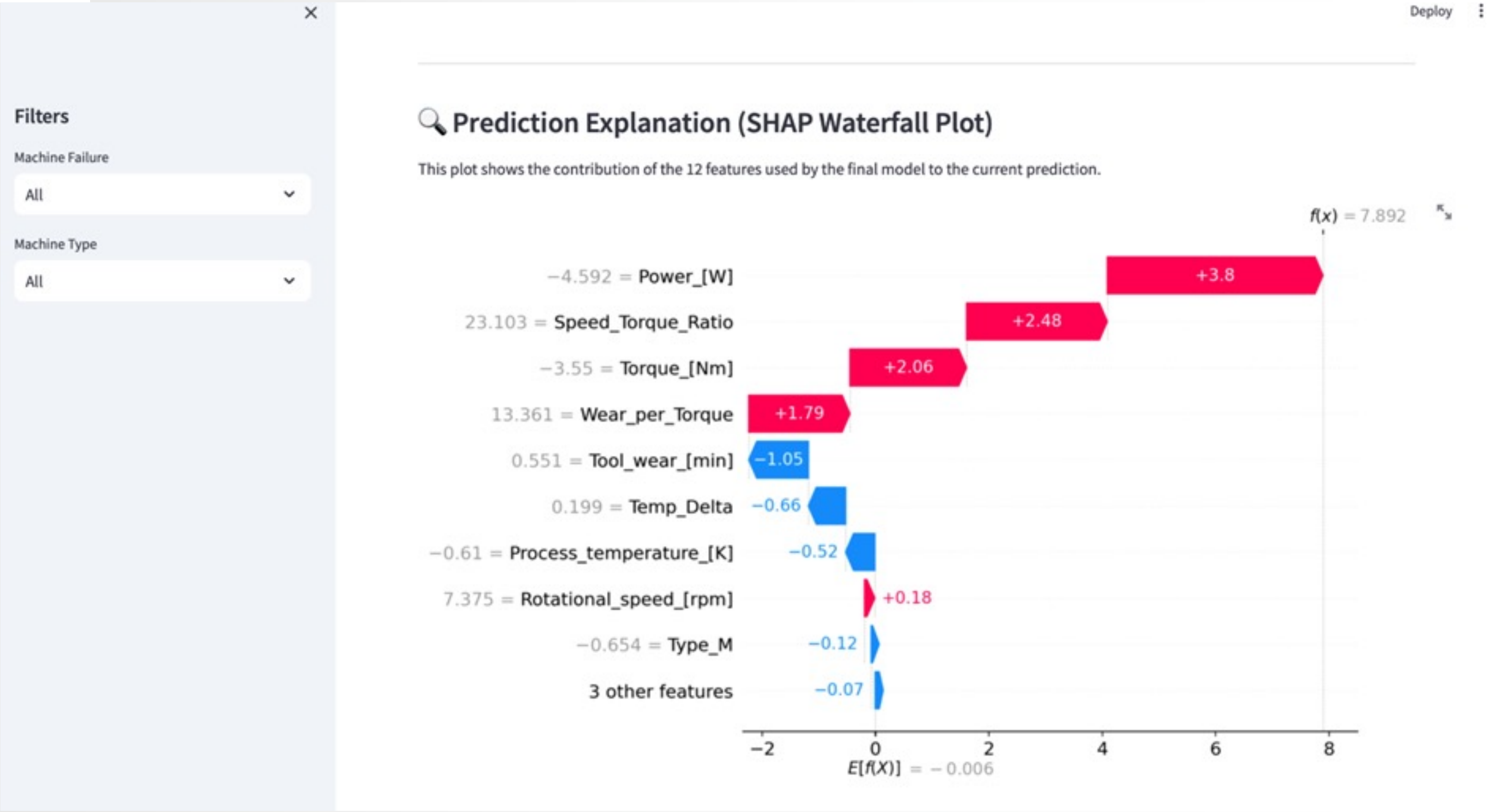
- Observed conditions include:
  - Moderate tool wear
  - Balanced torque levels
  - Stable rotational speed
  - Acceptable thermal conditions
- SHAP waterfall plot shows these features reduce the likelihood of failure





- **The system predicts machine failure based on input sensor values**
- **SHAP waterfall plot highlights features strongly contributing to failure risk.**

- **Key contributing factors include:**
  - **High tool wear**
  - **Increased mechanical stress**
  - **Unfavourable thermal conditions**
  - **Acceptable thermal conditions**
- **Wear per Torque and Speed-to-Torque Ratio show high postive SHAP values**



RESULTS AND DISCUSSION

- **The dataset contained 10,000 machine operation records, with only 3.4% failure cases, reflecting real-world class imbalance**
- **Among the evaluated models, ensemble-based approaches outperformed linear models**
- **XGBoost achieved the best performance with 98.6% accuracy, minimal misclassification, and strong generalization on imbalanced data**
- **Logistic Regression provided an interpretable baseline but was limited by its inability to capture nonlinear interactions between sensor variables**
- **Feature importance and SHAP analysis identified rotational speed, tool wear, and power consumption as the most influential predictors of failure**
- **Domain-engineered features (wear per torque, speed-to-torque ratio, temperature difference) significantly enhanced predictive capability**
- **SHAP explanations enabled both global and instance-level interpretability, improving trust and practical usability**
- **The deployment of the model as a Streamlit dashboard demonstrated real-time, explainable failure prediction suitable for industrial decision-making**

LIMITATIONS AND FUTURE SCOPE

Limitation

- Dataset is synthetic, based on predefined assumptions
- Analysis limited to a single machine type
- Failure prediction is binary, not failure-type specific
- No modeling of temporal degradation patterns
- Dataset imbalance may still affect robustness
- Streamlit dashboard is a proof of concept, not industrial-scale

Future Scope

- Validate model on real-world industrial datasets
- Extend to multi-class failure prediction
- Apply time-series models (LSTM, Transformers)
- Estimate Remaining Useful Life (RUL)
- Integrate additional sensors:
  Vibration
  Acoustic
  Electrical signals
- Enable real-time, cloud or edge deployment

CONCLUSION

- Machine learning effectively enables predictive maintenance
- XGBoost achieved 98.6% accuracy on imbalanced data
- Domain-informed feature engineering significantly improved performance
- SHAP ensured transparency at both global and local levels
- Streamlit dashboard demonstrated real-time, interpretable deployment
- The approach reduces unplanned downtime and improves operational efficiency


INVESTMENT
PROFIT
SOLUTION
BUSINESS INTERNATIONAL
SUCCESS
PRODUCTIVITY
MARKET
TEAMWORK
FUTURE
RESEARCH
DATA
TEAM
VISION
SALES

THANK YOU