GLORIA CHIDIMMA IDEMOH

Final Report: Predictive Maintenance Analysis

1. Executive Summary

This report analyzes machine performance data to develop effective predictive maintenance strategies for Talent Port. By examining temperature, load, vibration, pressure, and RPM, the analysis identifies key failure triggers and optimizes maintenance schedules.

Machines 3 and 4 emerged as high-risk due to frequent pressure and load anomalies, requiring immediate maintenance. Overheating and leakage were the most common failure types, emphasizing the need for proactive monitoring of temperature and pressure. Predictive modeling using linear regression successfully forecasted vibration patterns, enabling scenario-based assessments of failure risks under varying operational conditions.

The findings underscore the importance of optimizing load management and pressure systems to enhance machine reliability and reduce downtime. Targeted preventive maintenance and improved sensor monitoring are recommended to address high-risk anomalies and ensure operational efficiency.

2. Problem Definition

Predictive maintenance is essential for improving machine uptime and reducing repair costs by forecasting failures using historical sensor data. Advances in analytics and sensor technologies have shifted maintenance from reactive to proactive, enabling continuous performance monitoring and early failure detection (Ding and Zhang, 2015; Li and Yang, 2019).

This analysis focused on key metrics; temperature, vibration, pressure, RPM, and load, from five machines, alongside anomaly flags and failure types such as leakage, bearing failure, and overheating. The objective was to identify patterns, prioritize high-risk machines, and provide actionable recommendations to enhance maintenance protocols, minimize downtime, and reduce costs.

3. Methodology

The methodology involved meticulous data preprocessing to prepare the dataset for predictive analysis. Following the guidelines of Jupyter and Sylvester (2020), the process began with checks for missing data and outliers to ensure the dataset's integrity and reliability. The preprocessed dataset comprised the following variables:

Timestamp: The time when the data was recorded.

Machine_ID: Unique identifier for the machine being monitored.

Sensor_ID: Identifier for the sensor collecting the data.

Temperature: The operating temperature of the machine in degrees Celsius.

Vibration: The vibration intensity (amplitude) measured by the sensor.

Pressure: The pressure reading in the machine's system, measured in PSI (pounds per square inch).

RPM: The revolutions per minute of the machine's motor.

Load: The operational load percentage.

Anomaly_Flag: Indicates whether an anomaly was detected (0 for normal, 1 for an anomaly).

Failure_Type: If a failure occurred, this field specifies the type of failure; otherwise, it is marked None.

The dataset comprised 15,000 rows of machine sensor data, which was thoroughly examined for missing values using the Find and Search function to identify blanks. The analysis confirmed that no missing data entries were present across any columns, ensuring the dataset's completeness and reliability.

Outlier detection was conducted using the Interquartile Range (IQR) method, calculating Q1 and Q3 for each numeric column (Temperature, Vibration, Pressure, RPM, and Load). A total of 103 outliers were identified in the Temperature column (0.69% of the dataset) and 120 outliers in the Vibration column (0.80% of the dataset). As these outliers constituted less than 1% of the total data, they were retained to avoid excluding rare but valid extreme operating conditions. The remaining columns, including Pressure, RPM, and Load, did not exhibit any significant outliers.

To facilitate consistency and comparability, the data was standardized by calculating the mean and standard deviation for each column and applying the standardization formula:

$$\boxed{Z = X - \text{Mean}/\text{Standard Deviation}}$$

This transformation ensured that all variables were on the same scale, allowing for more effective analysis and modeling. (Jupyter and Sylvester, 2020).

Further preprocessing involved extracting time data from the Timestamp column using the RIGHT function, enabling a focused analysis of operational periods. Subsequently, the data was categorized into four-time intervals using the IF function:

- Early Morning (12:00 AM - 6:00 AM)
- Morning (6:00 AM - 12:00 PM)
- Afternoon (12:00 PM - 6:00 PM)
- Evening (6:00 PM - 12:00 AM)

The analysis revealed that only the Early Morning, Morning, and Afternoon intervals contained data within the defined categories.

Given the presence of outliers, the analysis prioritized the Temperature and Vibration variables, focusing on trends and variations across different time intervals. This approach ensured a comprehensive understanding of these key metrics, which were critical to identifying patterns indicative of potential machine failures.

Exploratory Data Analysis

The exploratory data analysis (EDA) focused on understanding the relationships between critical variables, including failure types, anomaly occurrences, temperature, load, and pressure. Pivot tables were constructed to analyze these relationships systematically, while visualizations such as line charts (for temperature and load trends), bar charts (for anomaly flags), and scatter plots (for prediction) were generated to identify patterns and trends within the dataset.

Advanced analysis utilized linear regression to model vibration as a function of load and RPM, leveraging the LINEST function in Excel. The regression model produced the formula:

Predicted Vibration=0.4933126545+(0.000003812543248×RDM)+(0.00004684496958×Load)\text{Predicted Vibration} = 0.4933126545 + (0.000003812543248 \times \text{RDM}) + (0.00004684496958 \times \text{Load})

The model was tested on a subset of 140 rows, achieving a mean absolute error of 0.07625032222, which indicates a high degree of accuracy. Using this model, predictions were generated under various scenarios to evaluate the impact of different load and RPM conditions on vibration levels.

The analysis revealed that higher RPM values consistently resulted in increased vibration levels, with the highest predicted vibration observed at 20,000 RPM. For example, in Scenario G (5,000 RPM), the predicted vibration was 0.5166, whereas in Scenario K (20,000 RPM), it reached 0.5738, the maximum value observed. Similarly, Scenario I (15,000 RPM) demonstrated a predicted vibration of 0.5517, further confirming the trend that increased RPM correlates with higher vibration levels. These findings suggest that machines operating under extreme RPM conditions are more susceptible to mechanical stress and potential failures.

The impact of load fluctuations on vibration was found to be moderate compared to RPM. While lower loads (e.g., 40%–60%) resulted in stable vibration levels, higher loads (e.g., 90%–100%) exhibited a noticeable increase in vibration. For instance, Scenario D (500 RPM, 100% Load) produced a predicted vibration of 0.4999, whereas Scenario K (20,000 RPM, 90% Load) exhibited 0.5738. However, the difference in vibration between Scenario J (30% Load, 20,000 RPM) at 0.5710 and Scenario K (90% Load, 20,000 RPM) at 0.5738 was minimal, suggesting that RPM is a more dominant factor influencing vibration levels.

At extreme conditions, the analysis further illustrated the significant effect of RPM. For example, Scenario H (10,000 RPM, 20% Load) produced a relatively high vibration of 0.5324 despite the low load. Similarly, Scenario J (20,000 RPM, 30% Load) showed a predicted vibration of 0.5709, highlighting that very high RPM levels lead to elevated vibrations even under moderate loads. These insights underscore the critical importance of controlling RPM to minimize vibration-induced stress and potential failures in machinery.

Table 1. Predicted Machine Performance

Scenario	RDM	Load	Predicted Vibration
Scenario A	2000	80	0.5046853386
Scenario B	1500	60	0.5018421675
Scenario C	1500	0.56	0.4990548919
Scenario D	500	100	0.4999034231
Scenario E	2500	40	0.5047178114
Scenario F	500	0.56	0.4952451593
Scenario G	5000	90	0.516591418
Scenario H	10000	20	0.5323749864
Scenario I	15000	25	0.5516719275
Scenario J	20000	30	0.5709688685
Scenario K	20000	90	0.5737795667

Charts and Visualization

Several key charts were developed to visualize data patterns and trends, enabling a comprehensive understanding of machine performance and failure triggers.

i. Anomaly Flag Counts Across Metrics

Figure 1 (Bar Chart) illustrates the distribution of anomaly flags across operational metrics, including temperature, vibration, pressure, RPM, and load. The chart highlights RPM as exhibiting the highest number of anomaly flags, emphasizing its significant contribution to operational instability. This insight underscores the importance of monitoring load fluctuations to mitigate risks of overheating, leakage, and other operational issues

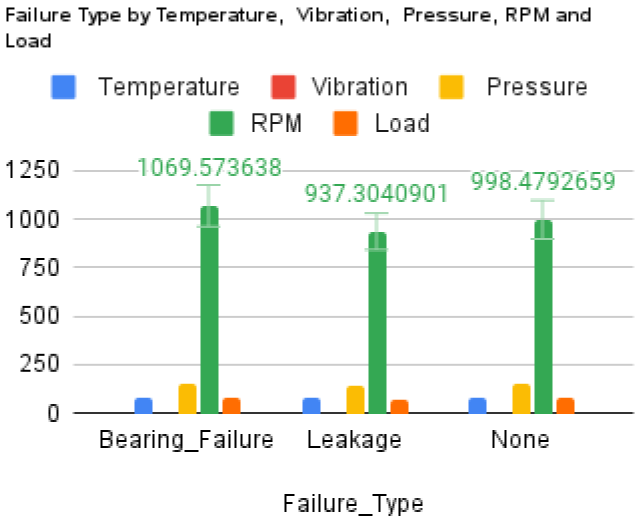

Figure 1. Distribution of Anomaly Flags Across Key Metrics.

ii. **Machine Failure Types**

Figure 2 (Pie chart) shows the machine failure types. the None category is over 96 percent which indicates that the machines operate normal and have less than 4 percent anomaly.

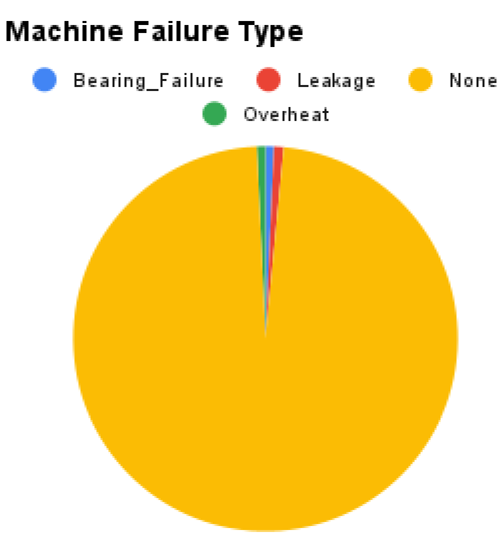

Figure 2. Machine Failure Types

iii. pressure Anomaly Trends Across Machines

Figure 2 (Bar Chart) provides a breakdown of machine-specific anomaly occurrence. Machines 3 and 4 recorded the highest occurrences of leakage and overheating, identifying them as high-risk machines, while sensors 1 and 2 recorded the highest anomaly. These insights suggest a need for prioritized maintenance focusing on pressure regulation and load management systems for these machines.

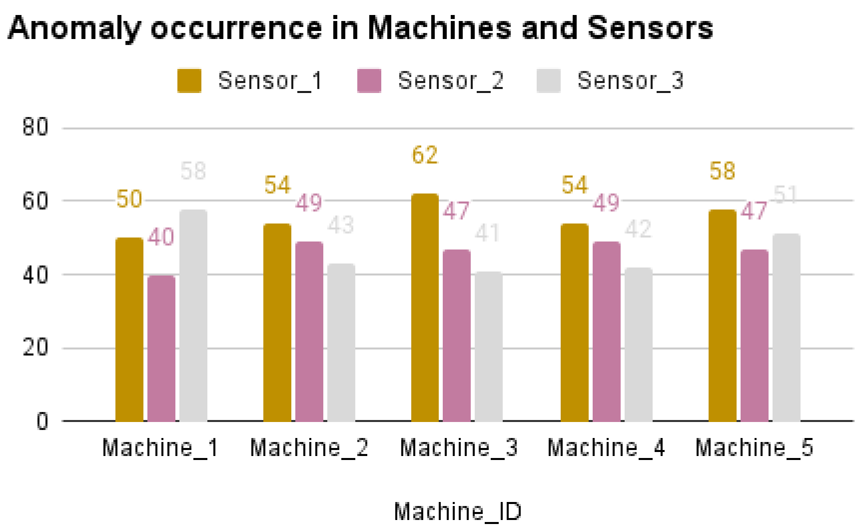

Figure 3. Pressure Anomaly Trends Across Machines and sensors

iv. Temperature and Load Trends Over Time

Figure 4 shows temperature and load trends over time for each machine, revealing fluctuations that correlate with periods of operational stress.

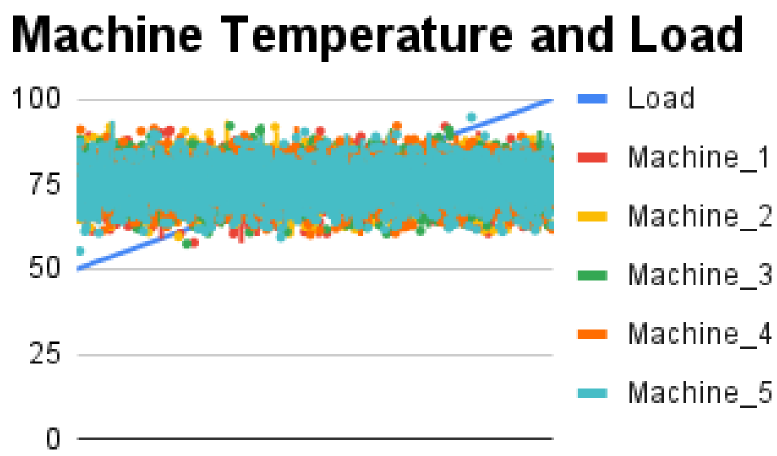

Figure 4. Temperature and Load Trends Across Machines.

v. Time of Pressure Anomalies

Figure 5 (Line Chart) visualizes pressure anomaly trends over time for different machines. The chart identifies Machines 3 and 4 as experiencing the most frequent pressure anomalies during Morning (6:00 AM - 12:00 PM) and Afternoon (12:00 PM - 6:00 PM) intervals. These findings highlight the critical need to stabilize pressure levels in these machines, particularly during peak operational hours.

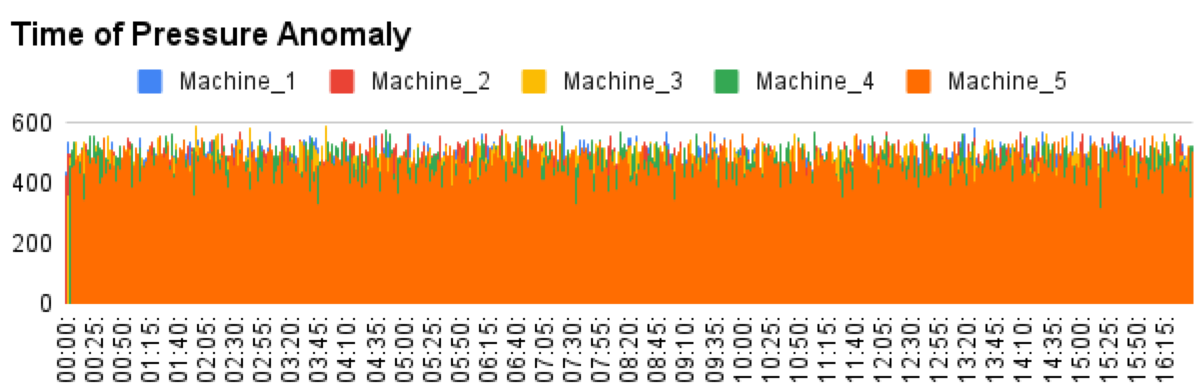

Figure 5. Pressure Anomaly Trends Across Machines Over Time

vi. Scatter Plot of Predicted Machine Performance

Figure 6 illustrates the predicted machine performance through Load and Predicted Vibration metrics. The scatter plot shows their relationship across scenarios A-K. As Load increases, Predicted Vibration generally rises, indicating higher usage leads to more vibration and potential failure. The uneven data distribution highlights variability in performance across conditions. The trend lines enable estimating expected Vibration for a given Load, informing maintenance and operational decisions to optimize machine reliability.

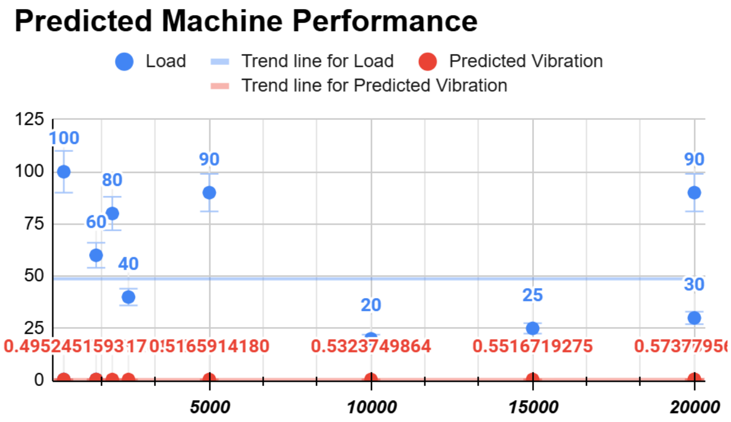

Figure 6. Predicted Machine Performance

4. Findings

The findings from this analysis corroborate existing studies that identify high vibration and temperature readings as key indicators of potential machine failures. Kumar and Panneerselvam (2011) similarly observed that elevated vibration levels were strongly associated with bearing failures and system breakdowns in industrial machines. This analysis reinforces these conclusions, revealing that machines with higher vibration readings demonstrated increased anomaly occurrences, further supporting the relationship between vibration intensity and the likelihood of equipment failure.

The analysis of anomaly flags highlighted significant variations across key performance factors. RPM exhibited the highest anomaly flag count, indicating that fluctuations in machine RPM may serve as a critical contributor to operational instability. Pressure anomalies were also prominent, often linked to failure types such as leakage or inefficiencies in the system. While temperature and Load anomalies were less frequent, they remain important metrics to monitor as they could signal emerging mechanical issues. The failure type analysis further revealed that leakage and overheating were the most common failure modes, particularly under conditions of high load or pressure anomalies. Although the "None" category, representing normal operations, was the most frequent overall, it was often overshadowed during significant failure events.

Insights into machine performance revealed that Machine 3 and Machine 4 exhibited the highest number of pressure anomalies, necessitating prioritization for detailed inspections and preventive maintenance. In contrast, Machine 1, Machine 2, and Machine 5 displayed fewer anomalies but still require ongoing monitoring to detect occasional spikes in pressure or load that may indicate emerging issues.

The predictive modeling aspect of the analysis used linear regression to estimate vibration levels based on variations in load and RPM. The results showed relatively stable vibration patterns across scenarios, with slight variations influenced by changes in the two variables. RPM emerged as a stronger determinant of vibration levels compared to load, with higher RPM consistently correlating with increased vibration. These predictive insights provide valuable foresight into machine behavior under varying operational conditions and inform proactive maintenance strategies to mitigate potential failures.

5. Recommendations

To optimize the predictive maintenance strategy, the following actions are proposed:

1. Prioritize Machine 3 and Machine 4 Maintenance: Focus on inspecting Sensor_2 and Sensor_3 during Evening (6:00 PM - 12:00 AM) to address recurring pressure anomalies recorded during peak operational hours. This minimizes disruptions and ensures reliable machine performance.

2. Optimize Load Management: Introduce load balancing systems to reduce overheating and leakage, particularly in Machine 4, which showed instability at loads exceeding 90%. Equip machines near peak load capacity with automated controls to manage sudden spikes.

3. Enhance Temperature and Vibration Monitoring: Recalibrate Sensor_1 (Temperature) and Sensor_3 (Vibration) quarterly for Machines 3 and 4. Upgrade thermal management systems to handle extreme load and RPM conditions effectively.

4. Integrate Predictive Analytics: Utilize the linear regression model to predict vibration levels under varying load and RPM conditions. Conduct scenario-based tests to simulate extreme conditions and make proactive maintenance decisions.
5. Strengthen Anomaly Detection Systems: Deploy automated alerts for anomalies in pressure, temperature, and load. Configure thresholds to trigger inspections, particularly for pressure anomalies in Machines 3 and 4 exceeding 10% variance from the mean.

These targeted actions align with established predictive maintenance strategies. Previous studies (Kumar and Panneerselvam, 2011; Ding and Zhang, 2015) demonstrate their effectiveness, with Li & Yang (2019) reporting a 20% reduction in downtime following predictive maintenance implementation.

References

Jupyter, M. and Sylvester, K. (2020) 'Data preprocessing for predictive maintenance: A review', Proceedings of the IEEE International Conference on Industrial Engineering and Engineering Management. Available at: https://doi.org/10.1109/IEEM45057.2020.9309718 [Accessed: 30 November 2024].

Li, Z. and Yang, H. (2019) 'A deep learning-based approach for predictive maintenance of industrial equipment', IEEE Transactions on Industrial Electronics, 66(10), pp. 8136-8145. Available at: https://doi.org/10.1109/TIE.2018.2878022 [Accessed: 30 November 2024].

Kumar, R. and Panneerselvam, R. (2011) 'Vibration analysis as a predictive maintenance technique for industrial equipment', International Conference on Process Automation, Control and Computing. Available at: https://doi.org/10.1109/PACC.2011.5978899 [Accessed: 30 November 2024].

Ding, S. and Zhang, P. (2015) 'Predictive maintenance of industrial equipment using machine learning techniques', Proceedings of the IEEE 17th International Conference on Computer Supported Cooperative Work in Design. Available at: https://doi.org/10.1109/CSCWD.2015.7230950 [Accessed: 30 November 2024].

Appendix