University of Niagara Falls Canada

Predictive Maintenance Analytics for Industrial Equipment

Student's Name:

Burno Mascarenhas da Silva

Chandan Rana

Faizanfarid Malek

Roshan Tuladhar

Submission Date: 25 January 2026

Professor: Hany Osman

Project Idea

Our capstone project uses real machine sensor data from a worldwide context to use data analytics in the manufacturing and industrial equipment sector. This project's main goal is to show how predictive analytics may be used to spot possible machine breakdowns before they happen. Unexpected breakdowns are a major problem for contemporary companies since they can cause production delays, financial losses and safety hazards in many manufacturing settings.

This research mimics the operation of a real-world predictive maintenance system by examining operational and sensor-based data. It demonstrates how proactive and well-informed maintenance decisions may be supported by data-driven insights. With this strategy, businesses can increase overall operational performance, decrease downtime and make better use of their resources. The subject is particularly pertinent in the current industrial environment, as Industry 4.0 and smart manufacturing are transforming traditional operations by making them more data-driven, intelligent, and connected.

Problem Identification

In many industrial environments, maintenance decisions are still based on reactive responses or fixed service schedules rather than the actual operating condition of machines. Equipment is often serviced too early, too late or only after a failure has already occurred. This approach leads to unnecessary maintenance costs, unexpected downtime, safety risks and disruptions to production processes.

A major challenge faced by organizations is the lack of data-driven tools that can detect early warning signs of machine failure. Without predictive insights, maintenance teams struggle to prioritize critical machines, anticipate high-risk operating conditions, and allocate resources efficiently. As a result, operations managers and plant supervisors must make maintenance decisions under uncertainty, balancing cost, performance and safety with limited actionable information.

Aim of Investigation

The aim of this investigation is to develop and evaluate predictive analytics models that estimate the likelihood of machine failure using operational and sensor-based data. Specifically, the study seeks to identify which sensor variables significantly influence failure risk and to compare the predictive performance of logistic regression and tree-based machine learning models.

By quantifying failure probability and highlighting key risk factors, this research intends to support data-driven maintenance decision-making. The findings are expected to demonstrate how predictive maintenance analytics can help organizations move from reactive maintenance strategies toward proactive and condition-based maintenance planning.

RESEARCH QUESTION:

Which operational and sensor-based variables significantly influence machine failure, and do tree-based machine learning models outperform logistic regression in predicting machine failure using the AI4I 2020 dataset?

Hypothesis 1: Predictive Capability

- **$H0_1$ (Null Hypothesis):**

  Operational and sensor-based variables do not significantly predict machine failure.

- **$H1_1$ (Alternative Hypothesis):**

  Operational and sensor-based variables significantly predict machine failure.

Hypothesis 2: Sensor Impact

- **$H0_2$ (Null Hypothesis):**

  Temperature, torque, rotational speed and tool wear have no significant association with machine failure.

- **$H1_2$ (Alternative Hypothesis):**

  Temperature, torque, rotational speed and tool wear have a significant association with machine failure.

Hypothesis 3: Model Performance Comparison

- **$H0_3$ (Null Hypothesis):**

  Tree-based machine learning models do not outperform logistic regression in predicting machine failure.

- **$H1_3$ (Alternative Hypothesis):**

  Tree-based machine learning models outperform logistic regression in predicting machine failure.

Proposed Data Sources

The AI4I 2020 Predictive Maintenance Dataset, which was acquired from Kaggle, will be used in this research. In addition to a binary indicator of machine failure, the dataset contains about 10,000 records of machine operating conditions, including temperature, rotational speed, torque, tool wear and product type. The data is a perfect starting point for this research because it is publicly accessible, well-structured and suitable for supervised machine learning and predictive analytics.

Intended Outcome

The creation of a trustworthy prediction model that can recognize high-risk machine failure scenarios prior to malfunctions is the project's desired result. The research will identify the most significant causes of failures and offer precise, fact-based suggestions for setting maintenance priorities.

These insights can help industrial planners, operations teams, and maintenance managers implement proactive maintenance strategies, minimize unscheduled downtime and make better use of the resources at their disposal. All things considered, the study shows how data analytics may be a useful tool for making decisions in contemporary industrial settings.

References

Kaggle. (2020). Predictive Maintenance Dataset (AI4I 2020). https://www.kaggle.com

Mobley, R. K. (2002). An introduction to predictive maintenance (2nd ed.). Butterworth-Heinemann.

Zonta, T., da Costa, C. A., da Rosa Righi, R., de Lima, M. J., da Trindade, E. S., & Li, G. P. (2020). Predictive maintenance in the industry 4.0: A systematic literature review. Computers & industrial engineering, 150, 106889. https://doi.org/10.1016/j.cie.2020.106889

Carvalho, T. P., Soares, F. A. A. M. N., Vita, R., Francisco, R. P., Basto, J. P., & Alcalá, S. G. S. (2019). A systematic literature review of machine learning methods applied to predictive maintenance. Computers & industrial engineering, 137, 106024. https://doi.org/10.1016/j.cie.2019.106024