**Predictive Maintenance: Schedule Maintenance by estimating the time to Failure | Trisha Winquist**

­­

**Introduction**

Predictive maintenance lets you find the optimum time to schedule maintenance by estimating the time to failure. It also pinpoints problems in your machinery and helps you identify the parts that need to be fixed. Using predictive maintenance, you can minimize downtime and maximize equipment lifetime.

Every day we rely on a wide range of machines. But the truth is that every machine eventually breaks down unless it’s being maintained. Companies follow different maintenance programs to increase operational reliability and reduce costs. One way is to do reactive maintenance, where the machine is used to its limits and repairs are performed only after machine failure.

One big challenge with preventive maintenance is determining when to do maintenance. Since you don’t know when failure will occur, you must be conservative in your planning, especially if you’re operating safety-critical equipment. But by scheduling maintenance very early, you’re wasting machine life that is still usable, and this adds to your costs.

Utilizing machine learning techniques, this research presents a study on early detection of machine failure.

**Data Acquisition:** The foundation of predictive maintenance lies in collecting relevant data. This includes information from sensors, historical records, and real-time monitoring systems.

Machine learning algorithms are the backbone of predictive modeling. These algorithms analyze historical data to create models that predict when equipment failures are likely to occur. By continuously learning from new data, these models evolve over time, enhancing their accuracy and reliability.

**Optimizing Maintenance Schedules:**

Machine learning algorithms analyze data to determine the optimal times for maintenance activities. This ensures that maintenance interventions are conducted when they are most cost-effective and least disruptive to operations. By avoiding unnecessary maintenance and focusing efforts on equipment that genuinely needs attention, organizations can achieve significant cost savings.

**Data Selection**

Since real predictive maintenance datasets are difficult to obtain, the data set was obtained from the UC Irvine Machine Learning Repository (AI4I Predictive maintenance Dataset) This 2020 Predictive maintenance Dataset is a synthetic dataset that reflects area predictive maintenance data encountered in this industry. This Dataset is a multivariate, Time-Series with 10000 instances and has 14 columns.

There are two target groups:

1. Target: Failure or Not
2. Failure Type: Type of Failure

Overall, the dataset appears to be focused on monitoring and recording data related to manufacturing or production processes, including various machines. Different types of machines are likely used in the process, each associated with specific product quality variants and tool wear characteristics.

* **UID**: unique identifier ranging from 1 to 10000
* **Product ID:** consisting of a letter L, M, or H for low (50% of all products), medium (30%), and high (20%) as product quality variants and a variant-specific serial number
* **Air Temperature [K]:** generated using a random walk process later normalized to a standard deviation of 2 K around 300 K
* **Process Temperature [K]:** generated using a random walk process normalized to a standard deviation of 1 K, added to the air temperature plus 10 K.
* **Rotational Speed [rpm]:** calculated from power of 2860 W, overlaid with a normally distributed noise
* **Torque [Nm]:** torque values are normally distributed around 40 Nm with an Ïƒ = 10 Nm and no negative values.
* **Tool wear [min]:** The quality variants H/M/L add 5/3/2 minutes of tool wear to the used tool in the process. and a 'machine failure' label that indicates whether the machine has failed in this particular data point for any of the following failure modes are true.
* **Machine Failure:** consists of five independent failure modes
* **TWF(Tool Wear Failure):** the tool will be replaced of fail at a randomly selected tool wear time between 200 â€“ 240 mins (120 times in our dataset). At this point, the tool is replaced 69 times and fails 51 times (randomly assigned).

**Variable Tables 4**

* Heat Dissipation Failure (HDF): heat dissipation causes a process failure if the difference between air- and process temperature is below 8.6 K and the toolâ€™s rotational speed is below 1380 rpm. This is the case for 115 data points.
* Power Failure (PWF): the product of torque and rotational speed (in rad/s) equals the power required for the process. If this power is below 3500 W or above 9000 W, the process fails, which is the case 95 times in our dataset.
* Overstrain Failure (OSF): if the product of tool wear and torque exceeds 11,000 minNm for the L product variant (12,000 M, 13,000 H), the process fails due to overstrain. This is true for 98 data points.
* Random Failures (RNF): each process has a chance of 0,1 % to fail regardless of its process parameters. This is the case for only 5 data points, less than could be expected for 10,000 data points in our dataset.

**Data Preparation**

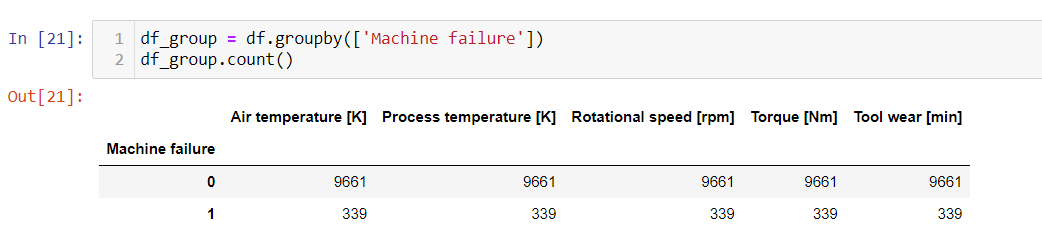
Very little data preparation was needed for the data set. The most significant decision about the data was determining which methodology for modeling the data. There was a unique ID number that had no relevance to the data so that column was dropped.

**Modeling & Methods**

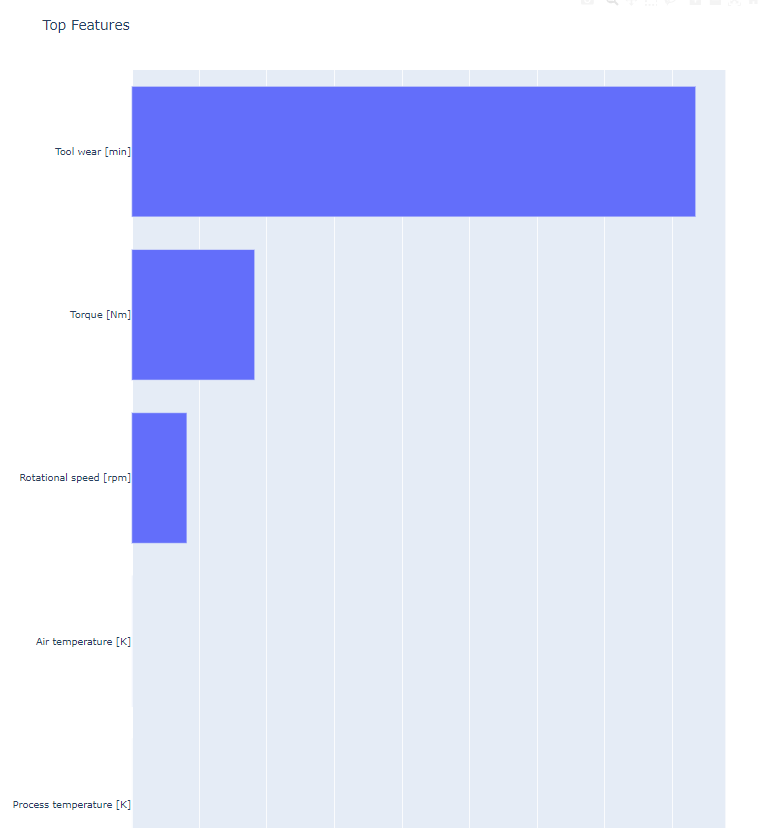
As mentioned, the objective is to predict machine failures before the machine fails. The data was summarized using logistical classification, and the results used to measure were Accuracy, Precision, Recall, and F1-Score.

As for machine learning techniques, I will be utilizing a decision tree and a Random Forest Classifier as the models.

Because we are looking for issues that cause Machine Failure, we can use pandas.get\_dummies to convert categorical variables into a dummy/indicator variable. Each variable is converted in as many 0/1 variables as there are different values. Columns in the output are each named after a value; if the output is a DataFrame, the name of the original variable is prepended to the value.

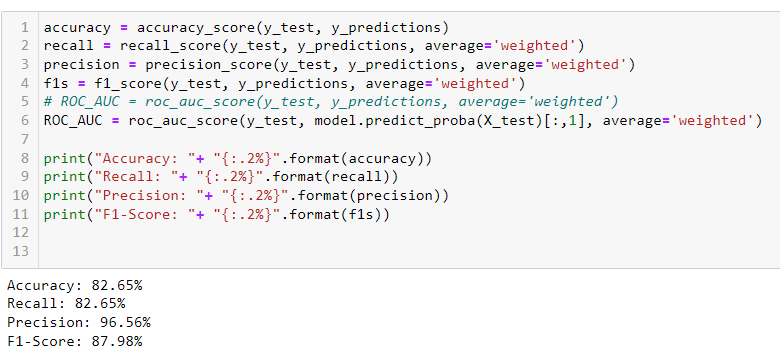


Chi-Square is to be used when the feature is categorical, the target variable is any way can be thought as categorical. It measures the degree of association between two categorical variables. Univariate feature selection works by selecting the best features based on univariate statistical tests. It can be seen as a preprocessing step to an estimator.

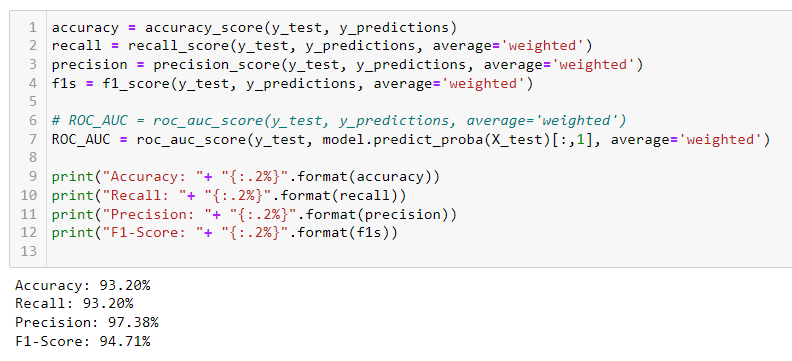


The train\_test\_split method is used to split data into features (X) and labels(y). The data frame gets divided into X\_train,X\_test, y\_train, and y\_test. X\_train and y\_terain sets are used for training and fitting the model. The X\_test and the y\_test is used for testing the model if it's predicting the right outputs/labels.

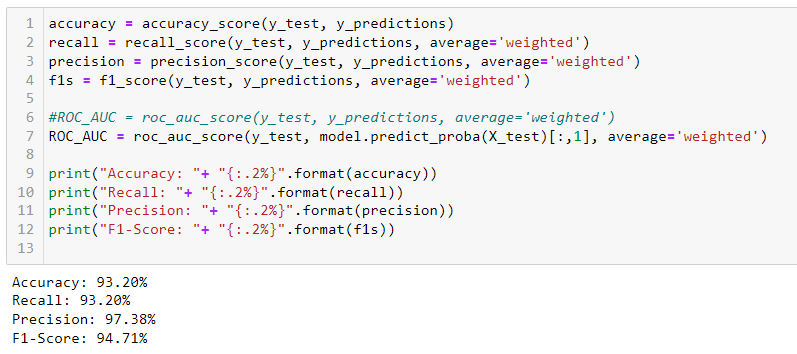
After splitting the data, I used Logistic Regression modeling by dividing the data into training and test sets.

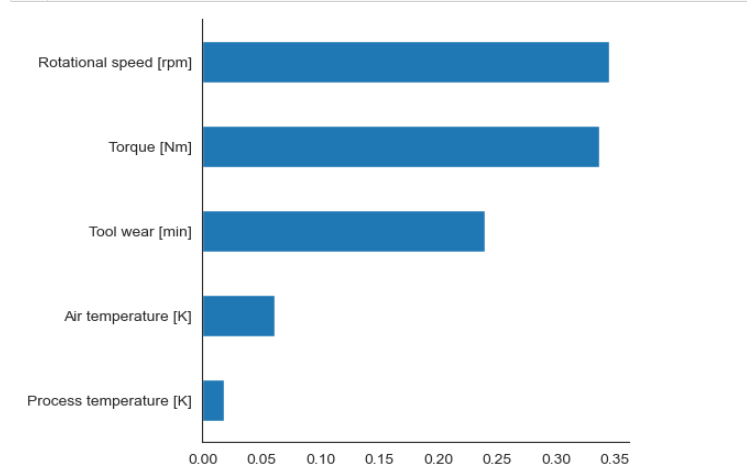


I also wanted to compare which classifier models gave the best outcome. I used the Gradient Boosting Classifier. **Gradient boosting** is a powerful technique in machine learning that aims to improve predictive models by combining the predictions of multiple weak learners.



MLP Classifier (multi-layer Perception classifier) MLPs are versatile neural networks that can handle complex data patterns and are widely used in various domains.





**Results Interpretation:**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | **Accuracy** | **Recall** | **Precision** | **F1-Score** | **ROC\_AUC** |
| **Logistic** | **82.65** | **82.65** | **96.56** | **87.9** | **96.3** |
| **Decision Tree** | **97.45** | **97.45** | **97.33** | **97.36** | **78.11** |
| **Gradient Boosting** | **91.85** | **91.85** | **97.18** | **93.82** | **97.16** |
| **MLP** | **91.60** | **91.60** | **97.09** | **93.64** | **95.79** |
| **Random Forest** | **98.10** | **98.10** | **97.99** | **98.03** | **96.07** |

**Conclusion**

Predictive maintenance, empowered by AI and machine learning, represents a paradigm shift in how industries approach asset management. The ability to predict equipment failures before they occur, optimize maintenance schedules, and reduce downtime has profound implications for efficiency, cost savings, and safety. While challenges exist, the benefits of implementing predictive maintenance with machine learning far outweigh the obstacles. As technology continues to advance, the collaboration between AI, machine learning, and predictive maintenance is set to redefine industry standards and ensure that assets operate at their optimal potential. Embracing these technological advancements is not just a choice but a strategic imperative for organizations seeking a competitive edge in the ever-evolving landscape of industrial maintenance.

In the example that is used to predict machine failures, it is possible to predict with a great degree of certainty when a machine will fail.

Based on the different models that I used, I found to have had fairly accurate data to predict future maintenance. However, based on my knowledge of Industrial maintenance failures, I would have to say that having a scheduled Preventive maintenance schedule isn’t always the most cost-effective way to save a dollar or 2. I could say with certainty, that preventive maintenance is the past and predictive maintenance is the future.

As for the directions of future work in this domain, I would like to take a certain piece of equipment and trend it over time. I think companies could save money by having someone analyze the history of every machine and not always rely on the manufacturer's recommendations.

**References:**

UCI Machine Learning Repository

<https://archive.ics.uci.edu/datasets?search=AI4I%202020%20Predictive%20Maintenance%20Dataset>

Kaggle- Predictive Maintenance

Getting to know your Data

<https://medium.com/@shanegary/getting-to-know-your-data-9e42935e7f60>

Predictive Maintenance: A Deep Dive into the Role of AI and Machine Learning

<https://medium.com/@AIreporter/predictive-maintenance-a-deep-dive-into-the-role-of-ai-and-machine-learning-97ab00038449>