Finding balance between corrective and preventative maintenance using predictive maintenance

Authors: Anant Jain, Viral Patel

Problem Description

- The business problem for this project is to predict issues caused by component failures.
- The business question therefore is "*What is the probability that a machine goes down due to failure of a component within the next 7 days*?"
- The problem is formatted as a multi-class classification problem (multiple components per machine)
- The model is trained on historical data collected from machines.

Problem Description (contd.)

- An initial approach is to rely on **corrective maintenance**, where parts are replaced as they fail.
  - ensures parts are used completely (not wasting component life)
  - expensive in both downtime and unscheduled maintenance
- An alternative is **preventative maintenance**, where a business may track or test component failures and determine a safe lifespan in which to replace that component before failure.
  - This approach can insure no catastrophic failures.
  - The down side is components are replaced frequently, many with remaining life left.
- The goal of **predictive maintenance** is to optimize the balance between corrective and preventative maintenance. This approach only replaces those components when they are close to failure.
  - The savings come from both extending component lifespans (compared to preventive maintenance), and reducing unscheduled maintenance (over corrective maintenance).

Proposed Project

- First stage, preliminary data visualization
- data consists of five different datasets
    - telemetry data (voltage, rotation, pressure, and vibration measurements)
    - non-breaking error logs and messages which occur while the components are still operational and do not constitute as failures
    - historical maintenance records comprising of regular inspection of machines and breakdowns.
    - Machine information such as type and age
    - Component replacement records with each record having a date, time, machine ID, and failed component type

Proposed Project

- Second, feature engineering (bringing the datasets together to produce features that ideally describe a machine's health condition at any given hour).
  - The timestamps in telemetry data make it suitable for calculating lagging features.
  - Errors also come with timestamps but error IDs are categorical values so we count the number of errors of each type in a lagging window.
  - Possible features from maintenance records can be the calculation of duration since last replacement of a component as that might correlate better with component failures since the more used a component is, the more degradation it should have.
  - The machine features about the type and age can be used as they are.
- Finally, we merge all the features to obtain the final feature matrix.

Proposed Project

- Next, labelling
  - take an appropriate time window prior to the failure and label the feature records as "failure expected" if it falls in that window, while labelling the rest of the records as "normal."
  - The time window can be picked according to the business’s requirement
    - sometimes predicting failure hours is enough
    - the business might require time for the arrival of replacement parts
- Finally, modelling and evaluation
  - Naive Bayes Classifier
  - Neural Network Classifier
  - Random Forest Classifier
- Questions?