FACULDADE DE CIÊNCIAS E TECNOLOGIA
UNIVERSIDADE NOVA DE LISBOA

SRCIM 19/20 - Predictive Maintenance

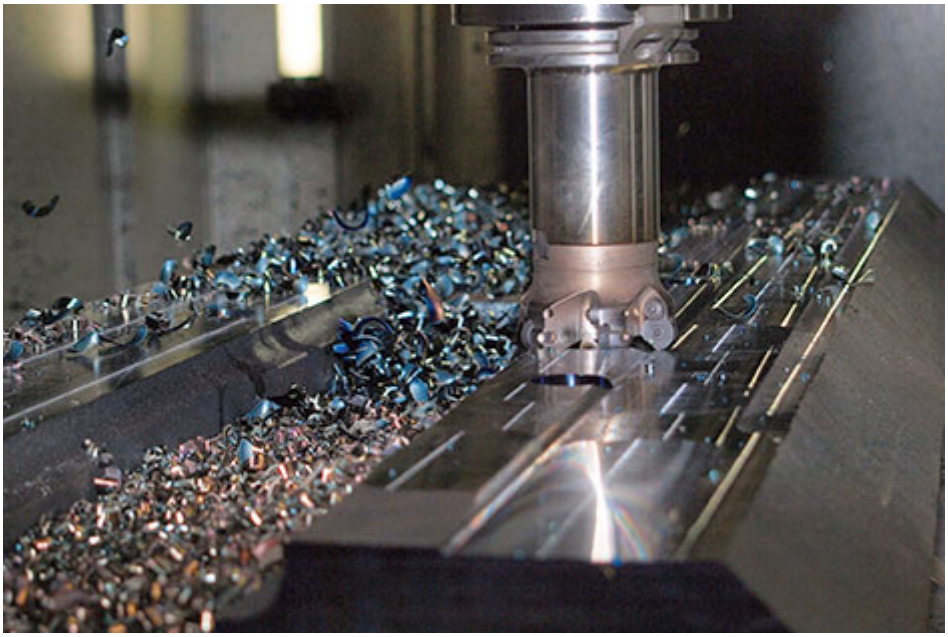
Problem Description

Unexpected downtime is one of the major cost drivers in several industries, making predictive maintenance an extremely appealing topic to explore. A recent survey of manufacturers in the automotive industry indicates that downtime costs on average USD 22,000/minute, going as high as USD 50,000/minute in some cases.

Therefore, observing the health and condition of equipment through sensors and telemetry data becomes crucial to enable predictive maintenance to be carried out for in-service equipment before critial failures occur. In this context, machine learning can be used to learn the complex relationships and patterns between sensor values and historical failure data in order to predict failures in the future. Examples of this could be the application of regression algorithms to predict the equipment's remaining useful life (RUL, the number of cycles remaining before failure), or classification algorithms to predict if a resource will faile within a specific time window (binary classification). This assignment will be focused on the latter.

Data:

The .csv files contain data related to equipment run-to-failure experiments, encompassing operational settings measurements from 20 different sensors. For the context of this problem it is assumed that an equipment's degradation pattern is reflected in its sensor measurements.



Requirements:

If you don't have a Python environment already configured in your machine it is recommended that you download and install Python 3.7 via the Anaconda distribution.

Packages: *Pandas* and *Numpy* (data structures and manipulation), *Scikit-Learn* (Machine Learning), *Matplotlib* and *Seaborn* (visualization).

Submission Guidelines and Deadline:

- Completed projects should be submitted by e-mail to **ricardo.peres@uninova.pt** before the end of the deadline.
- Projects should be executed in groups of 2 or 3 (maximum, no exceptions) students. Individual projects are allowed, but not encouraged.
- The project should be submitted as a **single .rar** file, named following the template *"studentNumber1_studentNumber2_studentNumber3.rar"** (e.g. *31444_31445_31446.rar*) containing:
  - The complete jupyter notebook contemplating the data analysis / machine learning part of the assignment. You can use the template provided in the CLIP platform which simultaneous serves as a guideline, project template and report.
  - Python project related to the REST API server.
- Deadline is **14 of December, 23:59 GMT**.

Evaluation Criteria

All of the goal/value pairs listed below are based on the assumption that a correct implementation is submitted.

Feel free to fill in the *Completed* column in accordance to your submission for the discussion (replace "-" with "*X*" when suitable).

Goal	Value	Completed
💾 Loading train/test data	1	-
🤖 Training **at least 3** different classifiers	5	-
💯 Evaluating each classifier using adequate metrics	4	-
📈 Plotting the results for comparison	3	-
💰 Estimating the expected value	2	-
💬 Discussing the results	1	-
📓 Building a REST API to deploy and serve the best model	2	-
❔ Additional bonus features	2	-

Please refer to the lab staff for additional info regarding possible bonus features. Examples include for instance performing feature extraction on the original dataset to check if newly created features improve performance, tuning the models' hyper-parameters (please refer to the documentation which includes several examples) to optimize performance, or complementing the model evaluation with a precision-recall analysis based on the classification thresholds.

Lab Planning

- Lab 1 (08/11/19) - Intro and Data Ingestion
- Lab 2 (15/11/19) - Helper Functions
- Lab 3 (22/11/19) - Binary Classifiers
- Lab 4 (29/11/19) - Expected Value Calculation and Discussion of Results
- Lab 5 (06/12/19) - REST API to serve the best model
- Lab 6 (13/12/19) - Bonus Features