Energy Data Activity State Classification

Introduction:

Activity state classification is the process of categorizing different states of a system based on its observed behavior or performance. In this task, you are provided with energy consumption data from a manifacturing factory, and your goal is to classify the activity state of the factory into four different phases: "Power-up", "Production", "Power-down", and "Non-production".

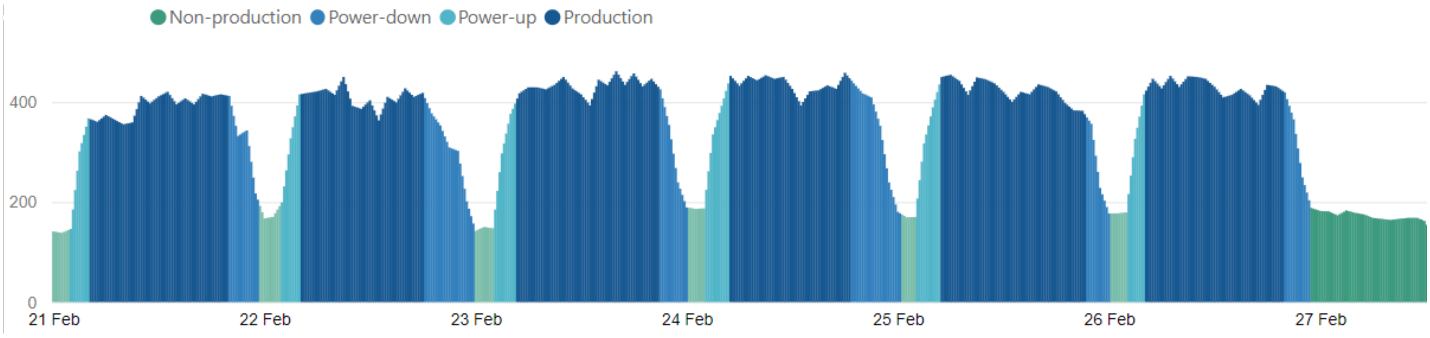

Figure: Current state of the art classification

Problem Statement:

The dataset provided is a .csv file containing time-series data with four columns: "time", "kWh", "Measure" and "label"

- The "time" column represents the time at which the measurement was taken. The data follows the standard format "*yyyy-mm-dd hh:mm:ss*", and it is sampled every 15 minutes. In the previous figure, this would be the x-axis.
- The "kWh" column represents the amount of energy consumed over the span of the sampling time, clearly in kWh. In the previous figure, this would be the y-axis.
- The "Measure" column contains a name that uniquely represents a specific part of the factory, where the consumption was measured.
- The "label" column contains the activity state of the factory at the time of the measurement. The four possible values are "Power-up", "Production", "Power-down", and "Non-production". A big part of data is not labeled. Unlabeled parts of data have a label "unclassified".

Your task is to develop an algorithmic approach that can accurately classify the activity state of the factory based on the energy consumption data provided. Due to the potential variability of the measures in the training data, we recommend considering a non-supervised (or semi-supervised) approach for this task. However, all submitted approaches will be evaluated on their performance, regardless of whether they are supervised or unsupervised. It is important to keep in mind that the algorithm may need to classify shapes that were not seen during training, so the approach should be designed to handle such scenarios. Additionally, the time of the day is not usable as a feature; however, data extrapolated from it (i.e., day of the week, month...) can be used. You may assume that the three main production stages (i.e. Power-up, Production, Power-down) always follow this order.

Bonus task

In addition to the main task of classifying the activity state of the factory, we also offer an optional bonus task of detecting anomalies in the consumption data. An anomaly can be:

- a point that significantly deviates from the normal behavior
- an activity state within a measure with an unusual (e.g. unusually high) consumption
- an activity state within a measure starting or ending at an unusual time

Please prioritize the main task, as it is the primary focus of the challenge. Only consider attempting the bonus task if you are confident in your main task performance. Ultimately, the quality of the main task submission will have the most significant impact on the evaluation.

Deliverables

To evaluate your results, we will execute your code on a test dataset. Please develop a function that accepts a file path for the test data, which will adhere to the previously discussed format. Your function should replace the unclassified labels with your predicted labels and save the updated results in a separate output file.

Evaluation Metrics:

In this task, you will be evaluated based on the accuracy of your model in classifying the activity state of the factory on a human-labeled test dataset. Specifically, we will use the F1 score as the evaluation metric.

The F1 score is a measure of the balance between precision and recall, and is defined as the harmonic mean of the precision and recall scores. It ranges from 0 to 1, with a score of 1 indicating perfect precision and recall.

It is important to note that different F1 scores may be calculated for each of the four activity states (i.e. "Power-up", "Production", "Power-down", and "Non-production"). Therefore, your solution should be able to accurately classify all four activity states with high precision and recall.

Tips

- Note that while the shape of different measures might be very similar, you should pay attention to the amplitude of the signal.
- We have attached a file "holdout.csv" containing a measure that shouldn't be used for training, and that will act as your personal benchmark, as it differs from the training data enough to understand if your model is able to generalize. Note that using this measure for training is forbidden, and will be checked (even though, clearly, it differs from the actual testing data that will evaluate your performance).

Good Luck!