Production Quality Alerting System
Jan Majewski 2023/01

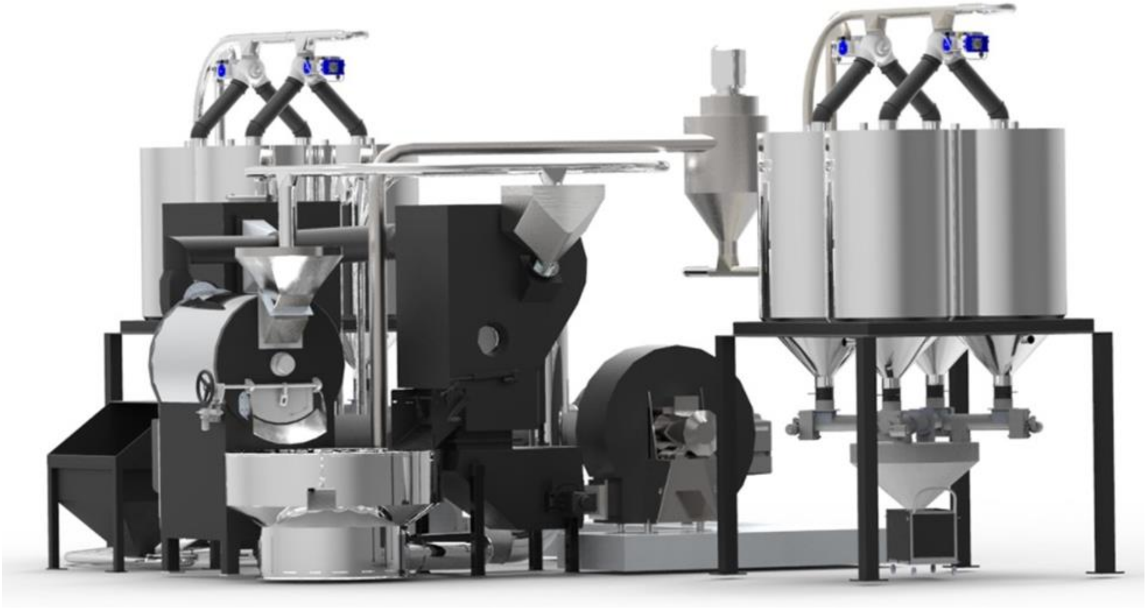
Photo by Viviane Okubo on Unsplash

Agenda

- Problem statement
- EDA
- Data Transformation
- Baseline model and Feature Importance
- RNN regressor and classifier
- Next steps

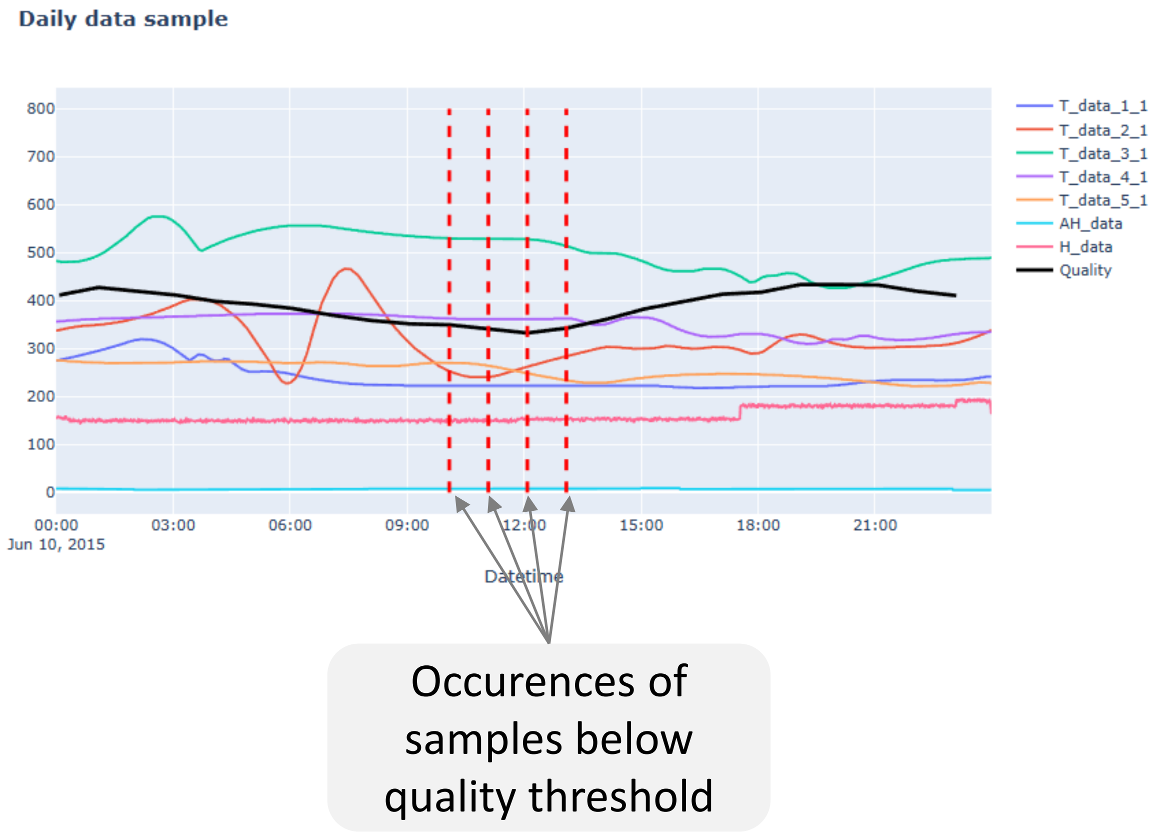
Problem statement

- The aim is to create a mechanism, which will alert machine operators of an expected product quality drop within 5 mins.
- The forecasting model is based on a series of sensors measuring temperatures within the roasting machine and raw materials' heigh and moisture.
- I will use recall us our primary metric, as our key goal is to detect quality drops below 350. To avoid too many false alerts we will also use precision as our secondary metric.
- Product quality is measured for a product sample every hour. This requires testing in the laboratory. At the prototype stage I assumed that product quality from previous samples (e.g. last hour) cannot be used as a model feature, as it would require test results to be ready within 55 minutes after every sample.



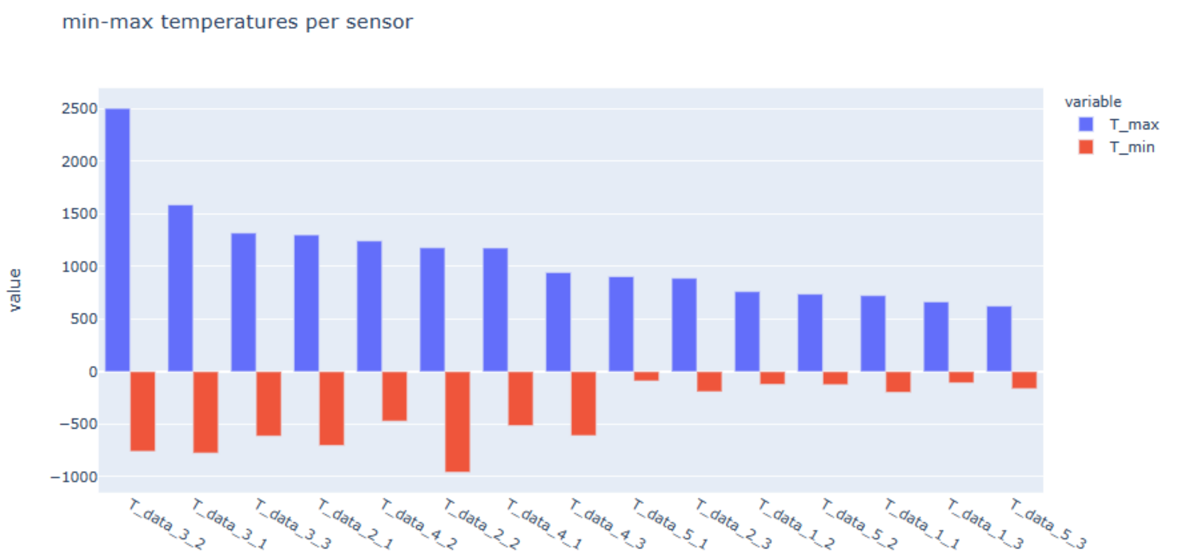
Exploratory Data Analysis

- There are 1461 continuous days of data for temperatures and raw materials parameters, which are measured every minute
- Quality is measured every hour for 1215 days
- Despite hourly sampling of quality, the production process itself is continuous – it also seems to have strong inertia with temperature swings usually taking several hours
- Merging features and labels will require further transformation due to their varying granularity



Temperature sensor failures seem to be the main data quality issue

- All feature records have 100% data completeness
- There is no significant quality drift through the 3 year period
- Time of day and day of the week do not seem to have significant effect on quality
- All raw material height and moisture records fall within reasonable values
- There are some highly improbable temperature reading swings, which are probably caused by sensor malfunctions

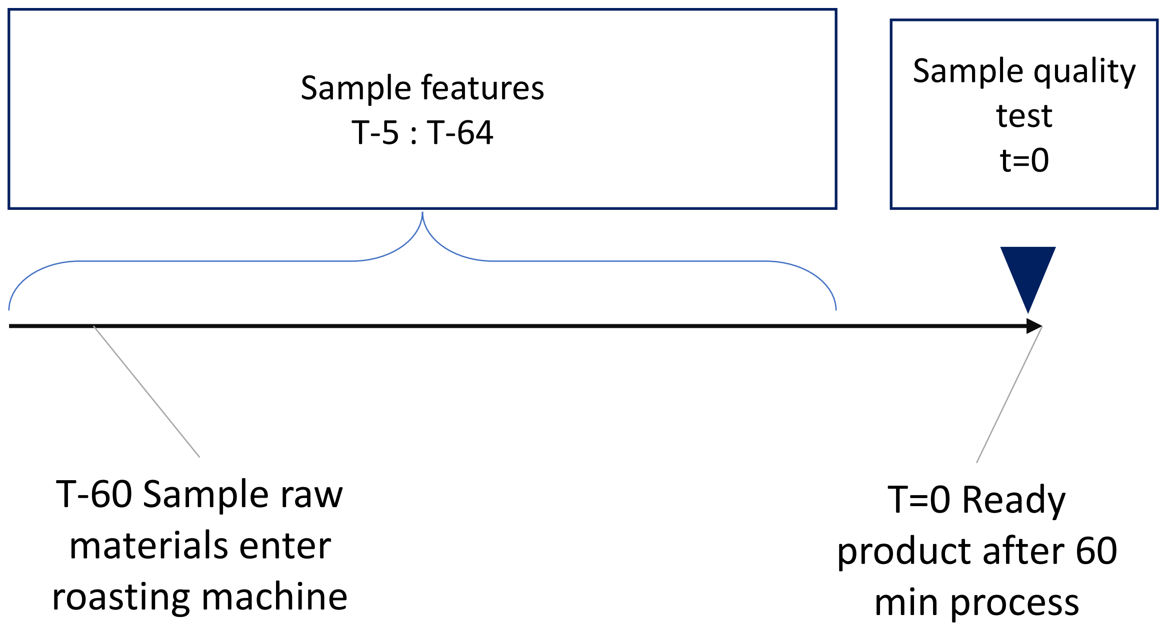


- For most sensors minimal temperature readings fall below -273ºC, which is impossible
- Sensors in chamber 3 have readings above 1500ºC, which are more likely to be seen in a gas turbine than a roasting machine
- For the prototype only the median of 3 sensors will be used to mitigate this data quality risk, this is not optimal as some information is lost
- In the next steps it would be worth investigating the production process and feasible temperature ranges. This would allow us to detect actual spikes and sensor failures

Merging features and labels is the key transformation challenge due to different granularity

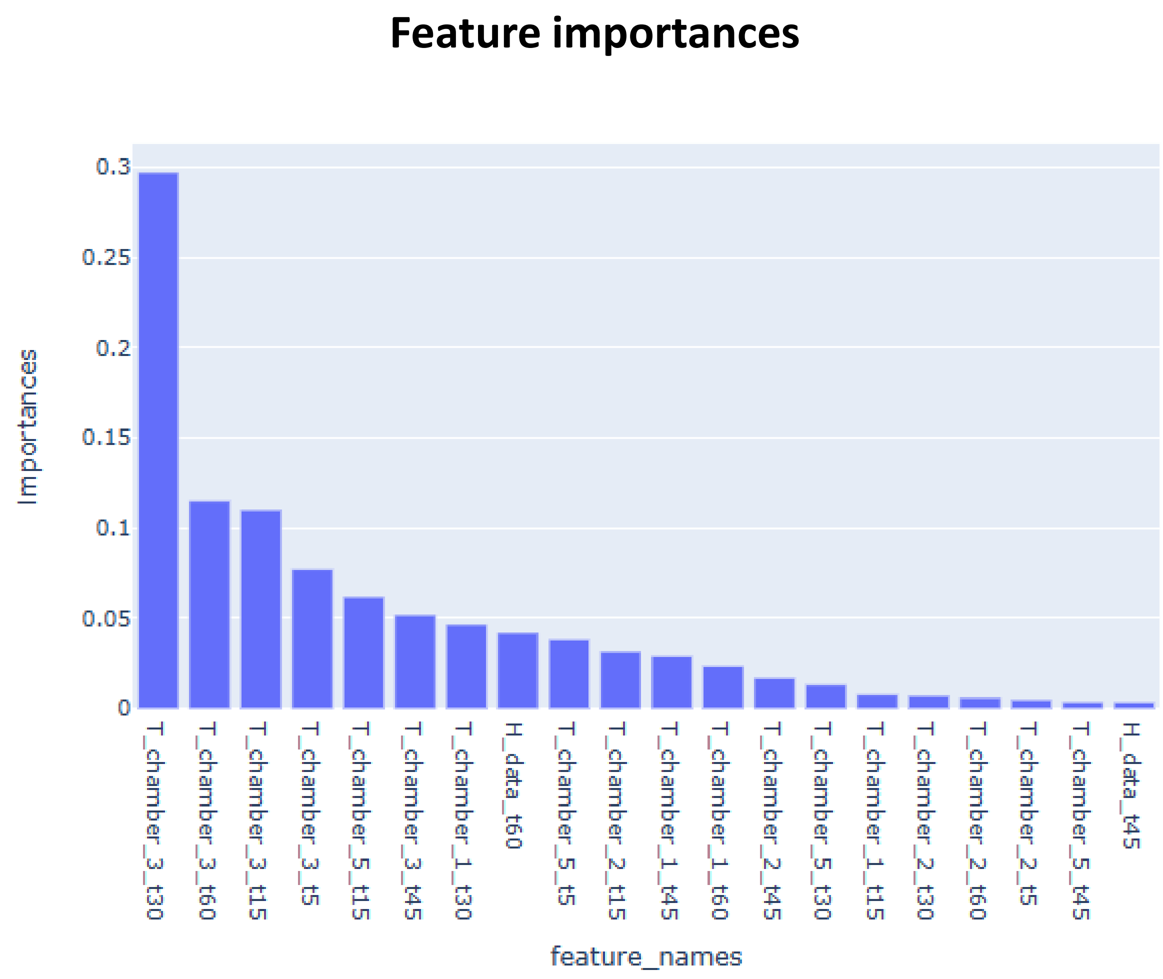
- In the prototyping process data is split into hourly batches corresponding to each measured sample (labels)
- This approach allows capturing the state of the machine through (nearly) the whole production process as well as a few minutes before raw materials enter the machine
- As our task is to predict a drop 5 mins before its occurrence we do not use the last 5 mins before measurement in our training process
- As there is no overlap between samples, the risk of data leakage, which is common in time-series tasks is mitigated and train and test sets can be taken from the whole timespan

Data sampling diagram



Baseline model shows dominating importance of temperature within chamber 3

- To understand feature importance I created a baseline XGBoost model, which relied only on data in intervals of [5,15,30,45,60] minutes before the measurement
- This model achieved a recall score of 85% on test set
- There is a high chance that malfunctions and temperature spikes within chamber 3 are the key quality issue
- A dummy model with one rule – predict failure within 5 mins if chamber 3 Temperature exceeds 550^{o}C is able to correctly predict ¾ of failures (recall = 75%, accuracy 65%)



Performance

Recall = 85.1%
Precision = 87.5%

Experiments with RNN regression and classifier models

	Regression model	Classifier model
Advantages	• Can be used for other tasks like quality prediction • Better suited for continuous processes with high inertia	• Prediction can be adjusted as failure probability, which is more understandable • Recall/accuracy balance can be adjusted with the decision threshold
Disadvantages	• Prone to outliers in labels • Might struggle with predicting labels with low frequencies	• Requires retraining if quality acceptance threshold changes • Looses some quality information due to binary labels
Performance	**Recall = 83.7%** **Precision = 90.5%**	**Recall = 86.7%** **Precision = 84.9%**

Next steps

Data quality and transformation

- Investigating feasible sensor temperatures and improving temperature outliers handling
- Using longer date spans of data – due to high temperature inertia, temperature trends from previous hours might also have valuable information
- Validating quality lab test results delay and if they can be used as input for following time period – due to process inertia quality drops often exceed 1 hour

Models

- Improving XGBoost model results with Hyper Parameter tuning + more granular data
- Experimenting with larger DNN models
- Experimenting with Transformer based models, which might deal better with data from longer time spans due to the attention mechanism
- Experiment with a decision threshold based on business cost of false alerts (FP) and missed failures (FN)