Equipment Downtime Prediction

Homayoun Gerami, Spring-Summer 2021

Summary:

AltaML

- I have been assigned to a project, that my task is to build up an analytical model to predict downtime for a facility that produces ‘Potash’. We are requested to do the prediction one our ahead of actual downtown.
- This presentation briefly demonstrate the project.

What this presentation covers:

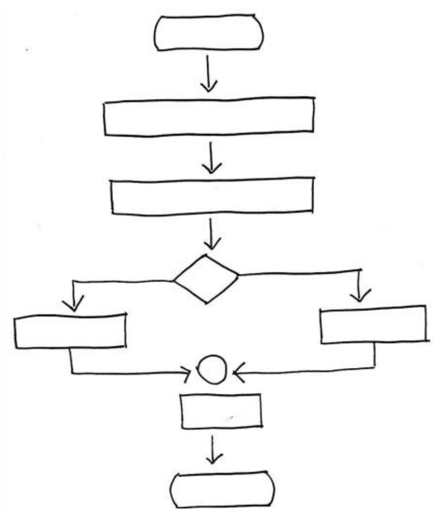
AltaML

- Why predicting a downtime is important?
- Problem definition and Goal of this project
- Exploratory Data Analysis (EDA)
- Feature Selection and Feature Engineering
- Predict potential failures ahead of time (Ongoing)
- To be done

Why predicting a Downtime is important?

AltaML

Industrial Process



Downtime can occur in any industrial process

Improve the system reliability with targeted preventive maintenance

The operators can have enough time to troubleshoot and prevent the failure

16,000 tonnes in production saving per month *estimated

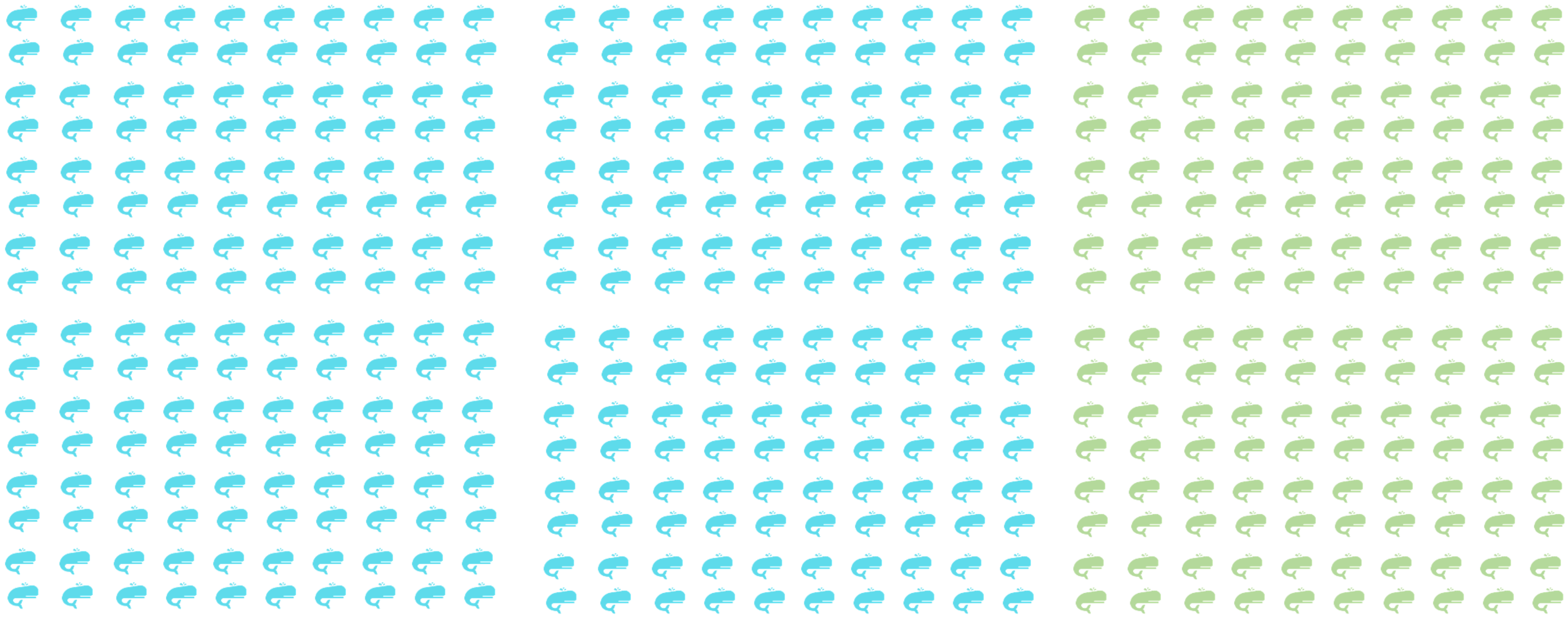
Equipment Downtime Prediction | What and Why

production loss due to downtime/month

48000 tonnes

Equipment Downtime Prediction | What and Why

AltaML

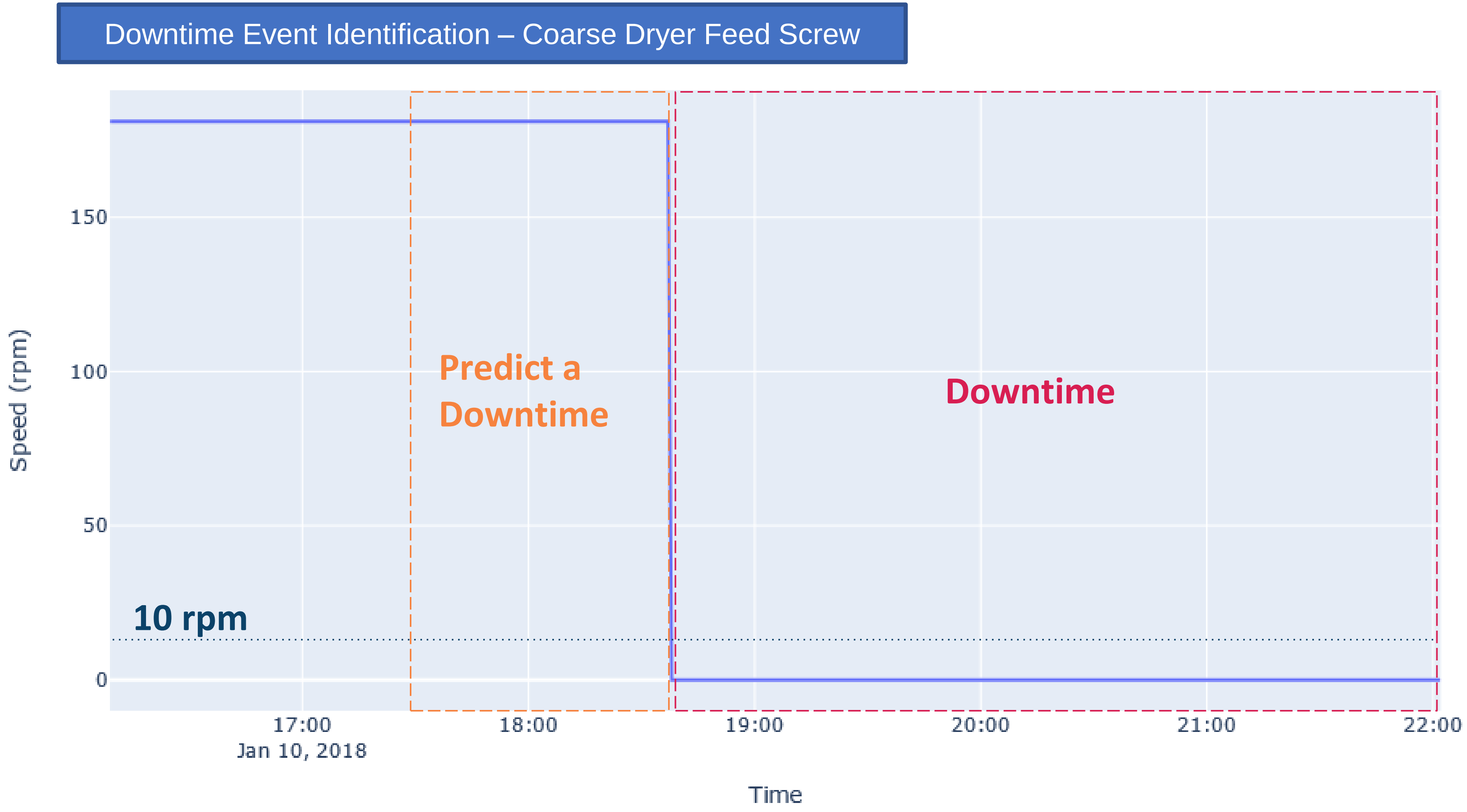


estimated to reduce 30% downtime, by predicting equipment failure using machine learning.

Leading to a recovery of

16000 tonnes of potash

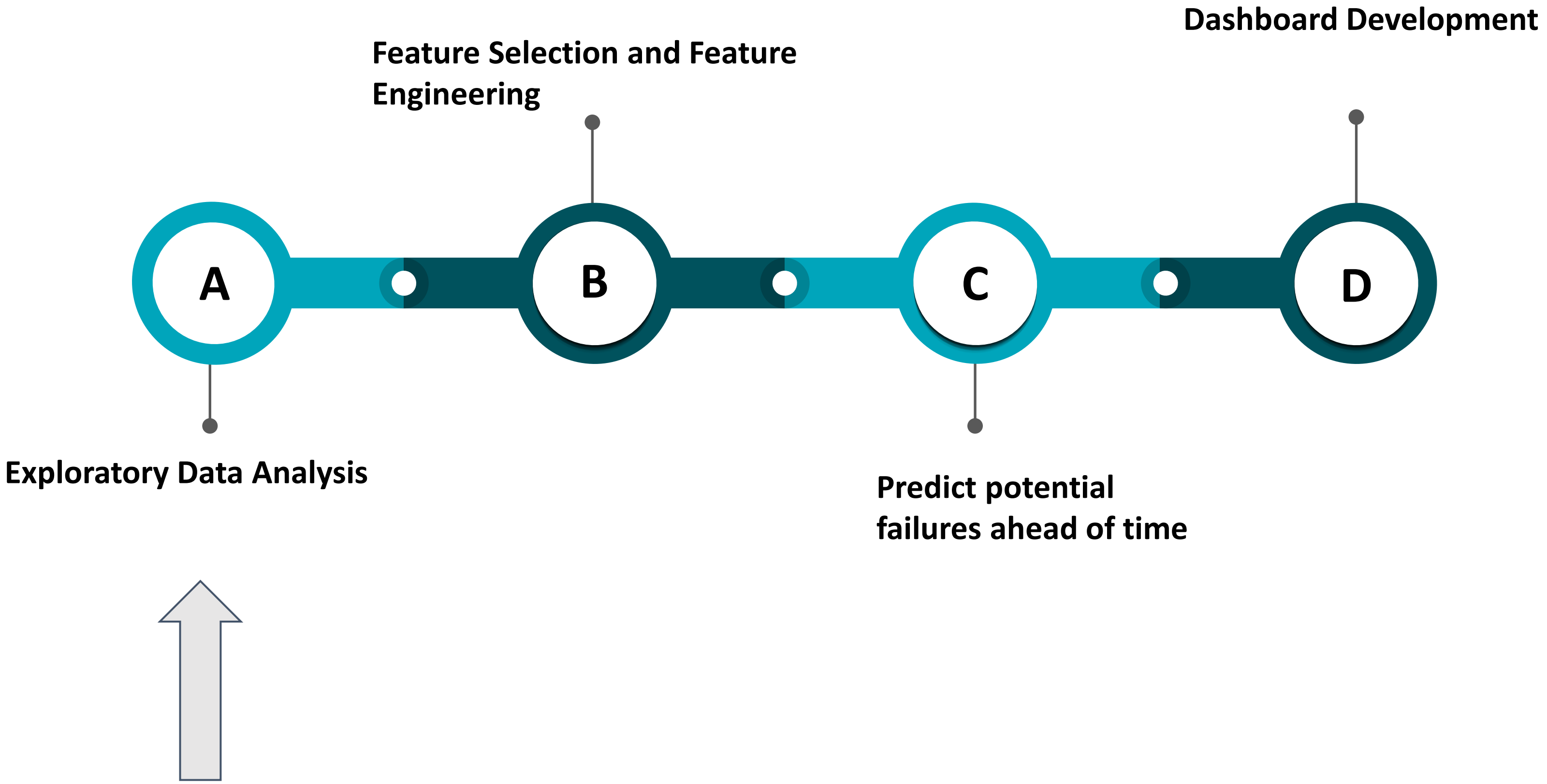
Problem Definition and Project objectiove:



Client defines downtown as any event that Coarse Dryer Feed Speed drops to 10 rpm. Our goal is to predict such event one hour ahead of happening.

Road map for the project:

Exploratory Data Analysis (EDA)



Exploratory Data Analysis (EDA)

- Available Data Review
- Data Processing & Data Cleaning
- Correlation Analysis

Tasks Performed:

- Non-time domain analysis:
  - Adjust the skew.
  - Smooth out the outliers.

- General data analysis:
  - Drop NaN values and string type values.
  - Study the distribution of failures in features.
  - Study the behavior of process before downtime.
  - Find the correlation between features.
  - Find the rate of feed screw speed changes before and after failure.

Available Data:

AltaML

95 Features

11 Tables

Including data from

2018 and **2019**

Available Data:

SPARTAN CONTROLS AltaML

11 datasets:

- Unit process instrumentation readings:
  - Drive current, Torque SP and PV, Speed etc.
- Coarse Dryer parameters:
  - Differential pressure, Speed, Temperature, Bed depth etc.
  - Set points for the unit processes.

Timestamp	CENTRFUGE_FEED-TK_L.PV	HYDROFLOAT_PROD-TK_LC.MD_x	HYDROFLOAT_PROD-TK_LC.PV_x	HYDROFLOAT_PROD-TK_LC.SP_x	Centrifuge 1 Torque PV	Centrifuge 1 Torque SP	Centrifuge 1 Drive Current	Centrifuge 1 Drive Speed	Centrifuge 2 Torque PV	...
2018-01-01 00:00:00	45.0336	Auto	29.6249	35	731.651	725	37.6984	56.136	1262.94	...
2018-01-01 00:01:00	45.3468	Auto	29.2124	35	717.425	725	37.6984	56.1638	1261.84	...
2018-01-01 00:02:00	46.2669	Auto	28.7189	35	699.428	725	35.7143	55.312	1293.91	...

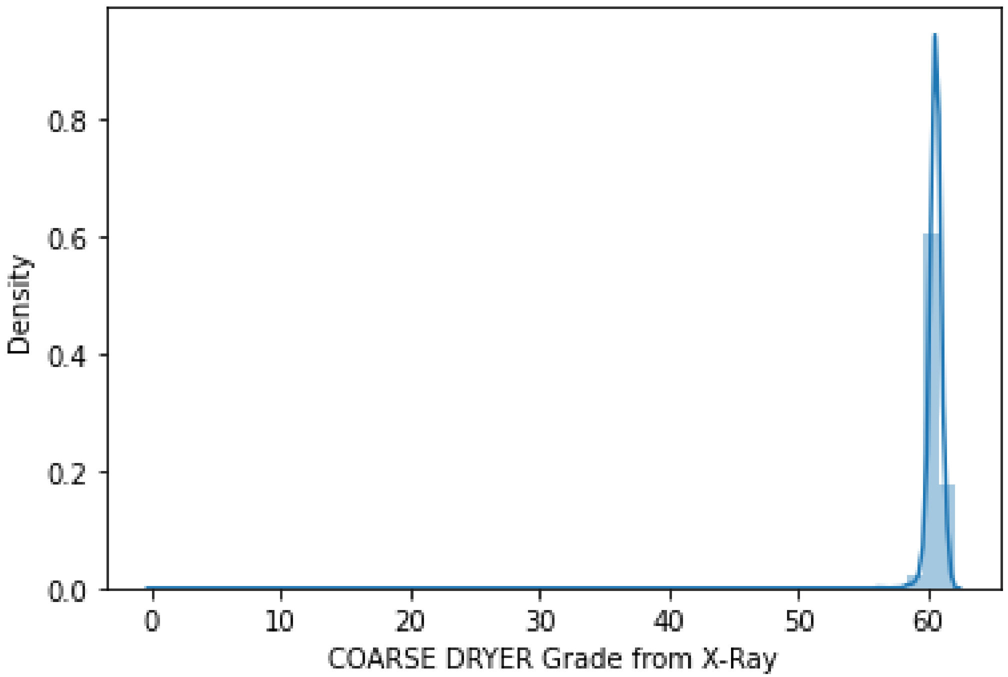
Data Processing:

The EDA for this project is separated into 3 Major Categories:

- Non-time domain analysis: Skew, Outliers, SNR, Interdependence, Feature Importance, Error features.
- General data analysis: Check NaN values, Distribution of positive shutdowns vs negative

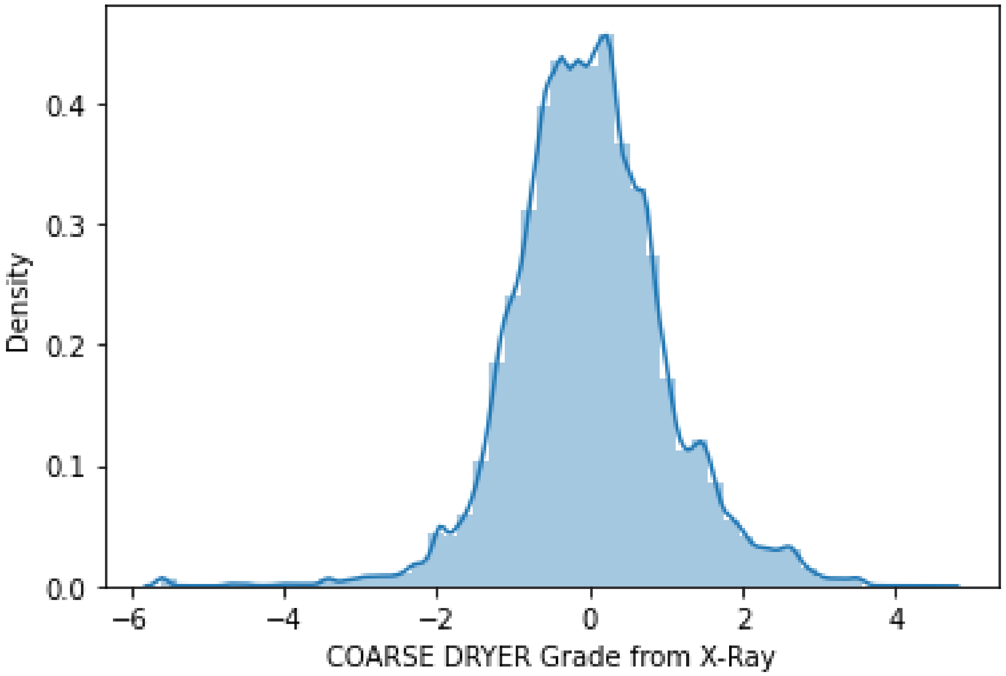
Data Processing:

Skew Adjustment



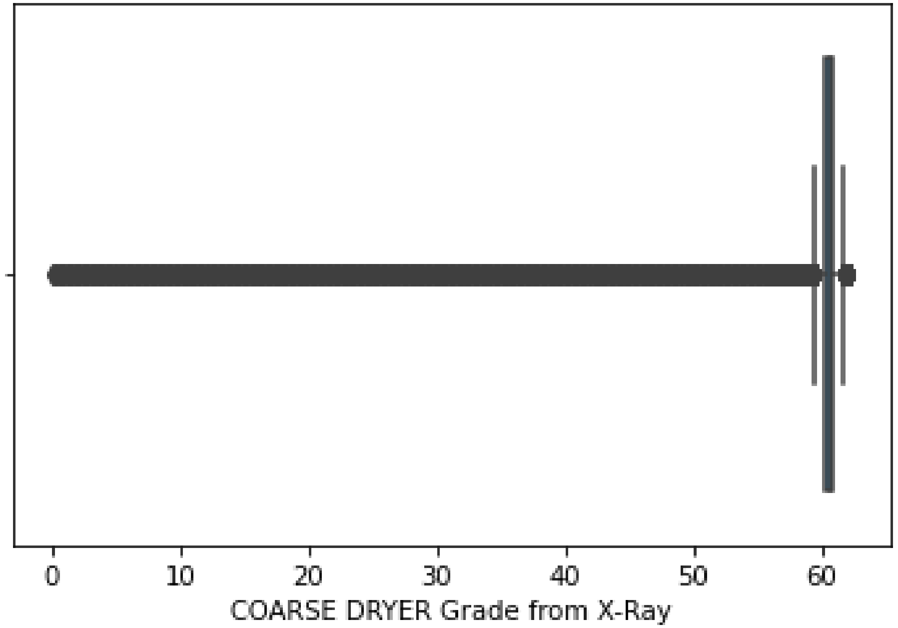
AltaML

Adjust the skew

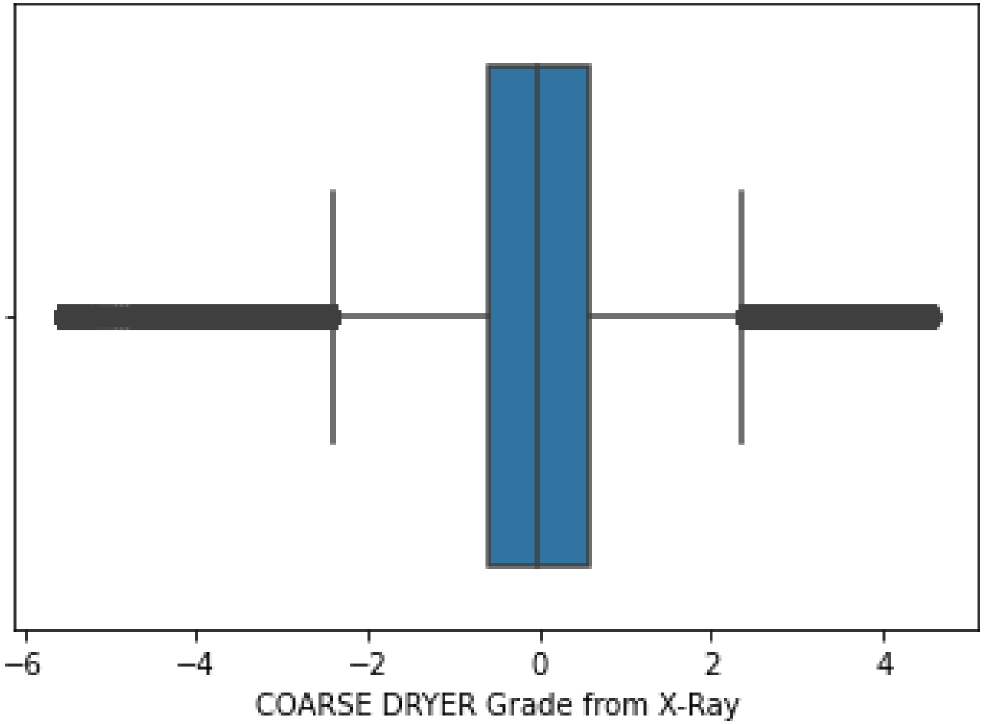


Data Processing:

Skew Adjustment



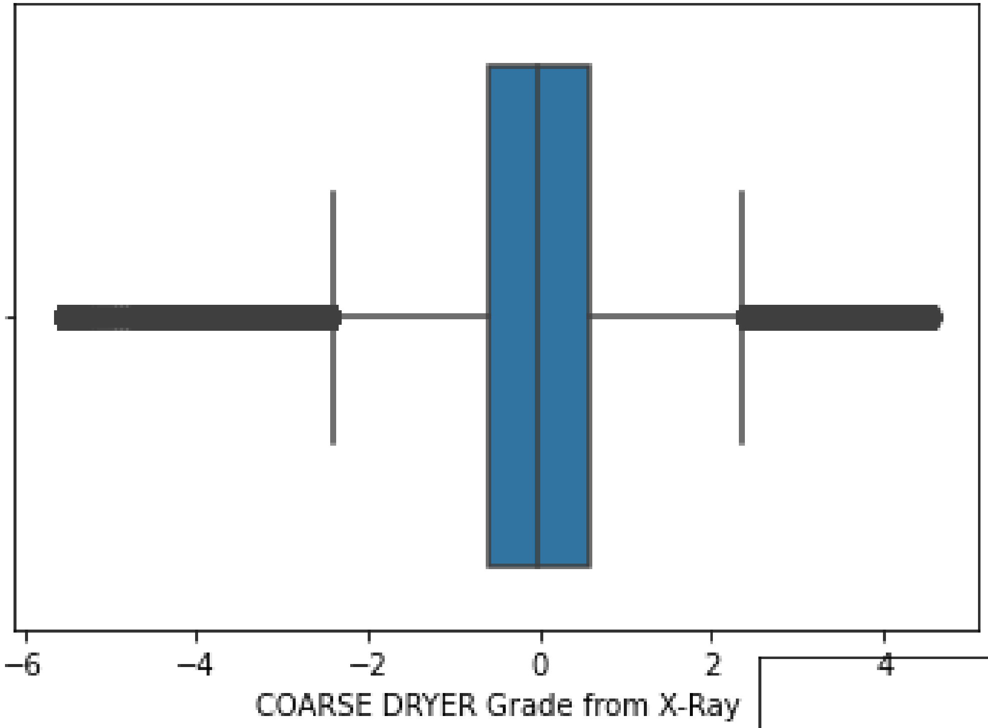
Adjust the skew

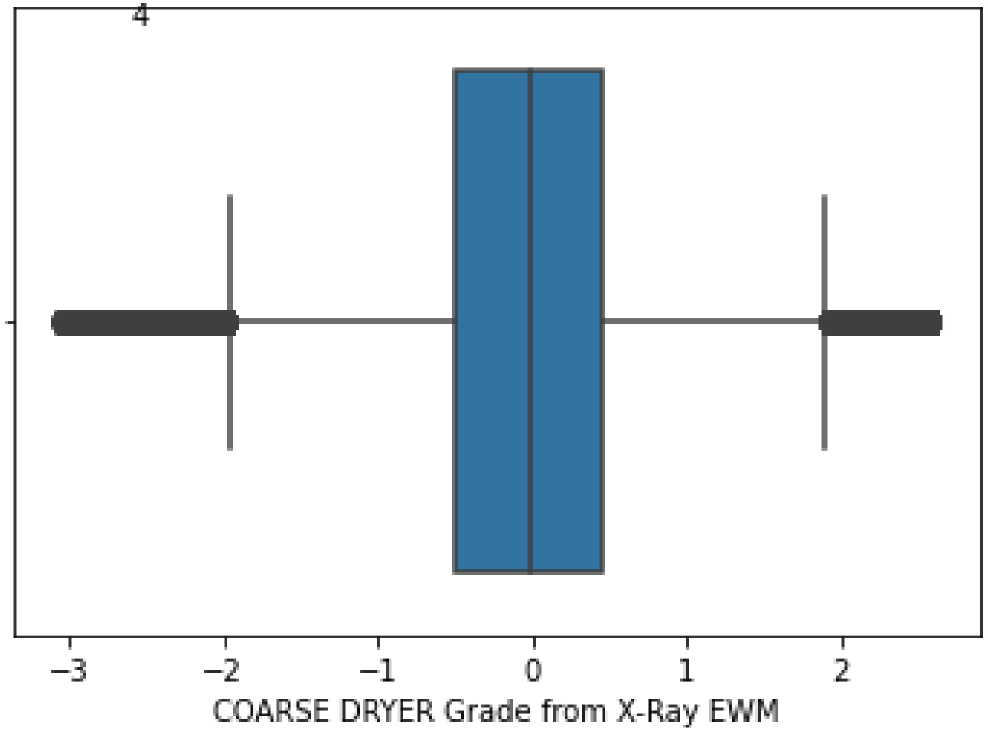


Data Processing:

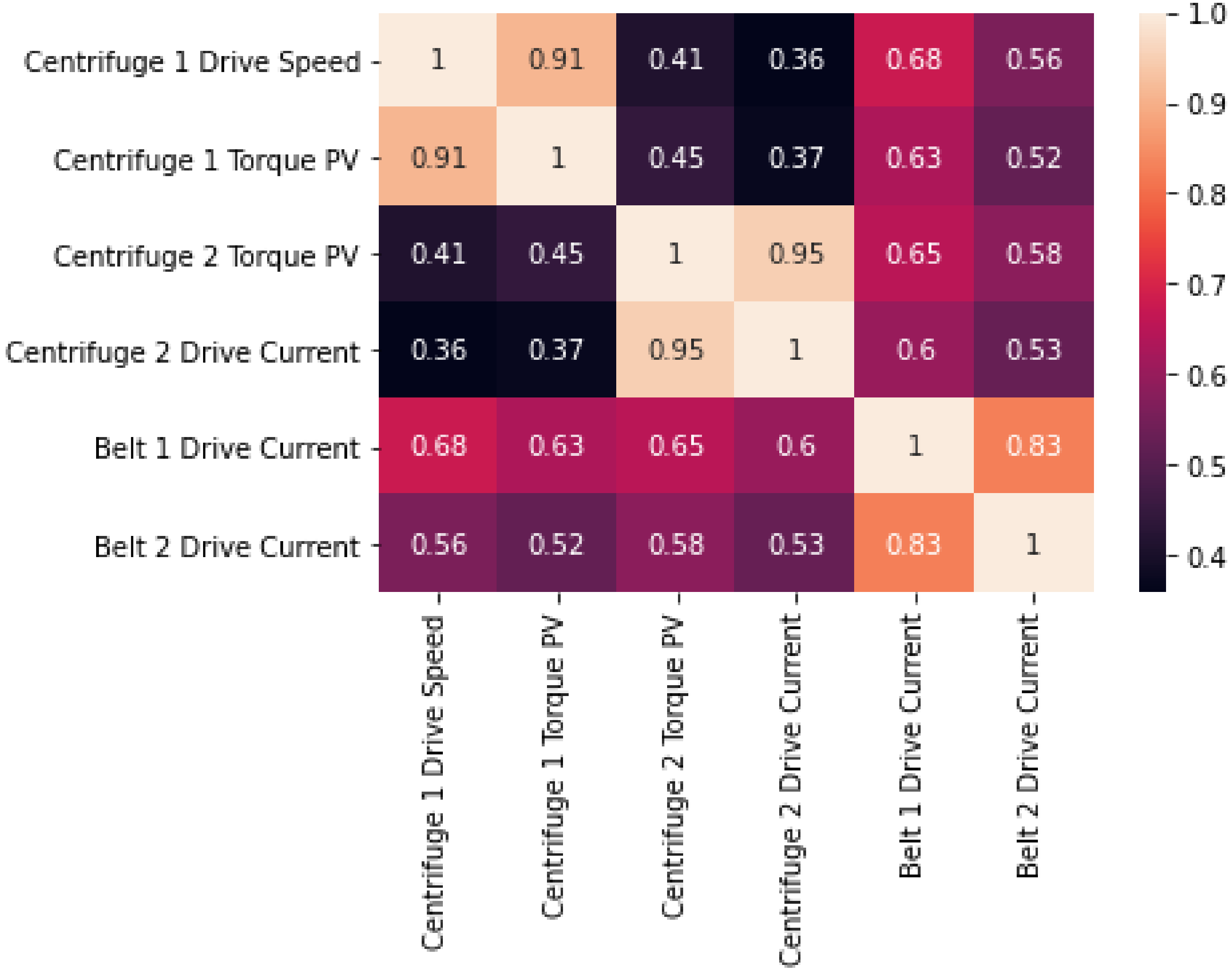
Smooth out Outliers

Smooth out the outliers





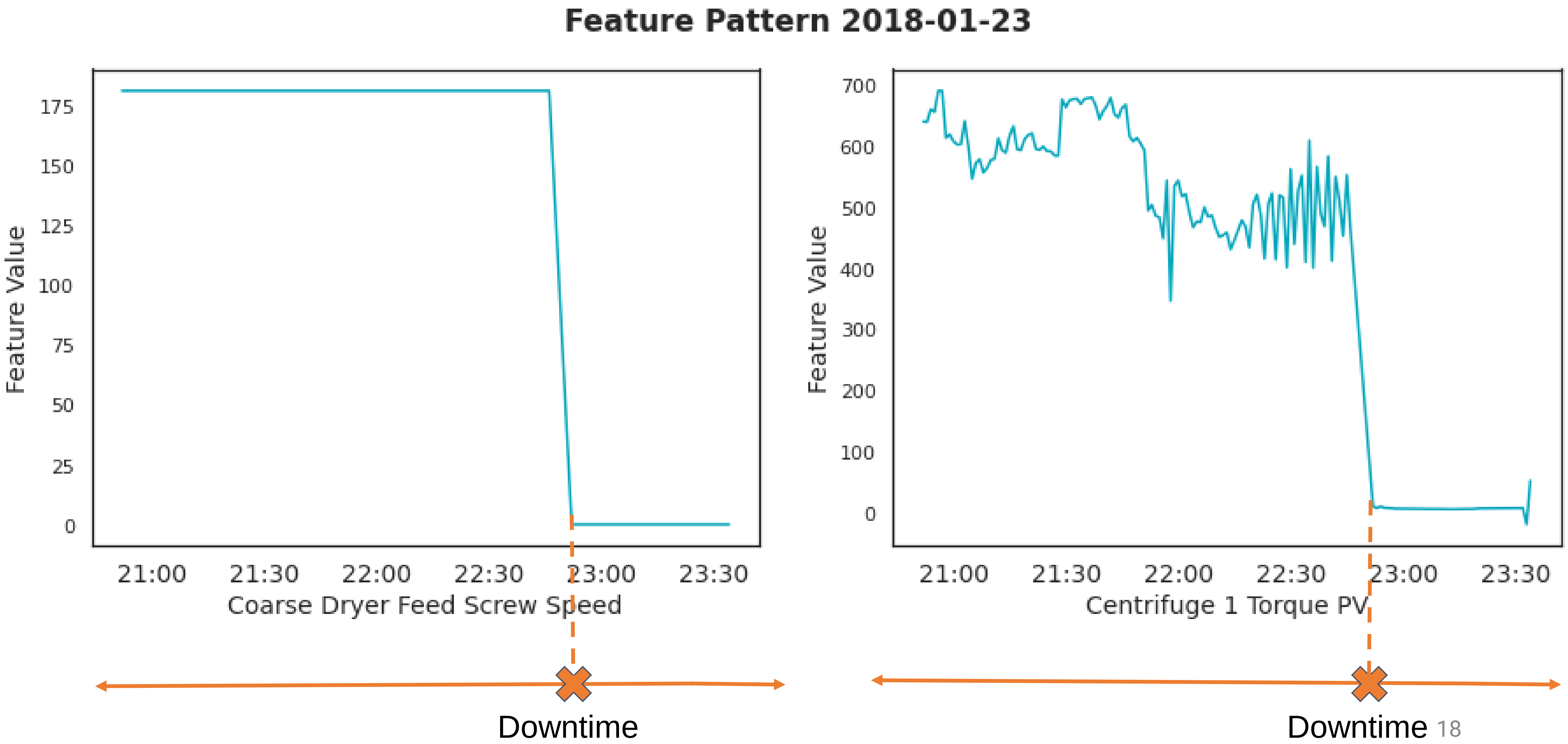
Correlation of the Features



Correlation Example:

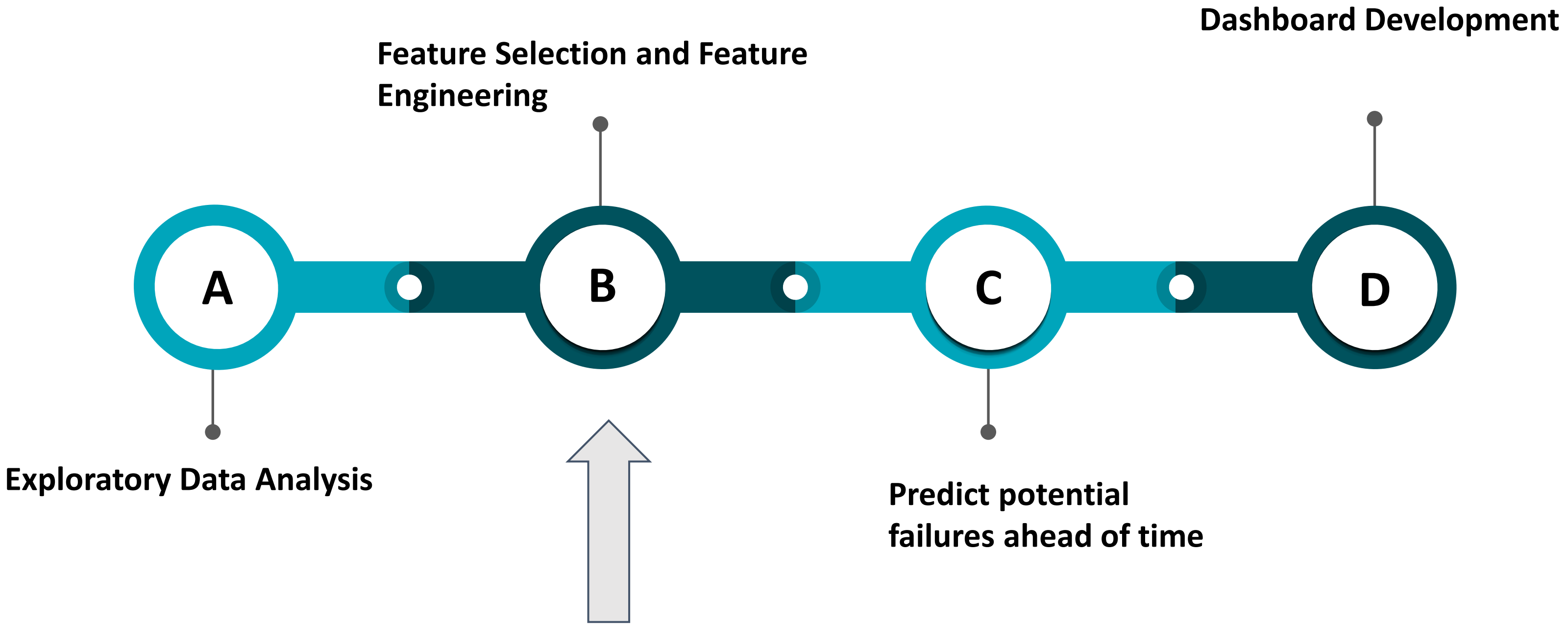
Coarse Dryer Feed & Centifuge Torque PV

Downtime= 2018-01-23 22:52:00



Road map for the project

Feature Selection and Feature Engineering



Feature Selection and Feature Engineering:

- Domain Specific Analysis:

  Identified new relationships between process values and set points to be used in the model.

- Time Domain Analysis:

  In a chemical process, upstream operations impacts downstream processes. We extracted the correlations between the features and the delayed copy of themselves.

- Downtime Analysis and Clustering:

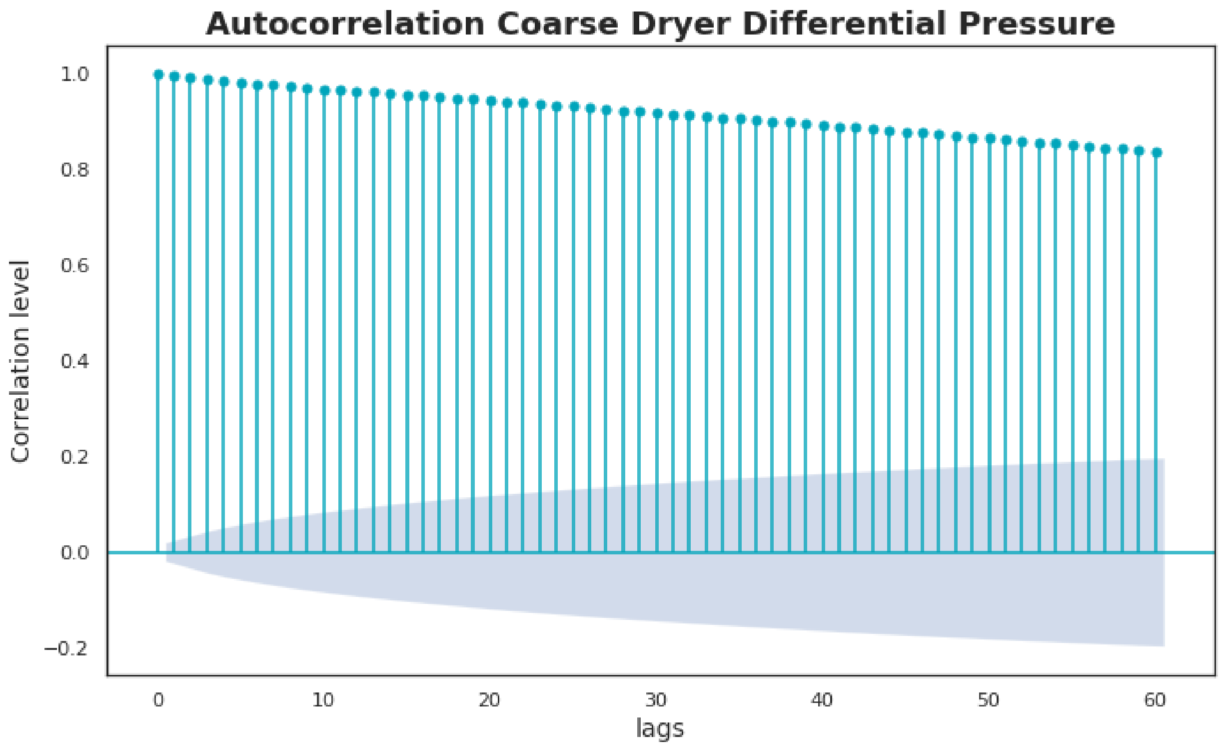
  We have discovered that the downtimes are most probably not all the same types, as the associated features represent totally different behaviors. We further demonstrated that through clustering techniques to client.

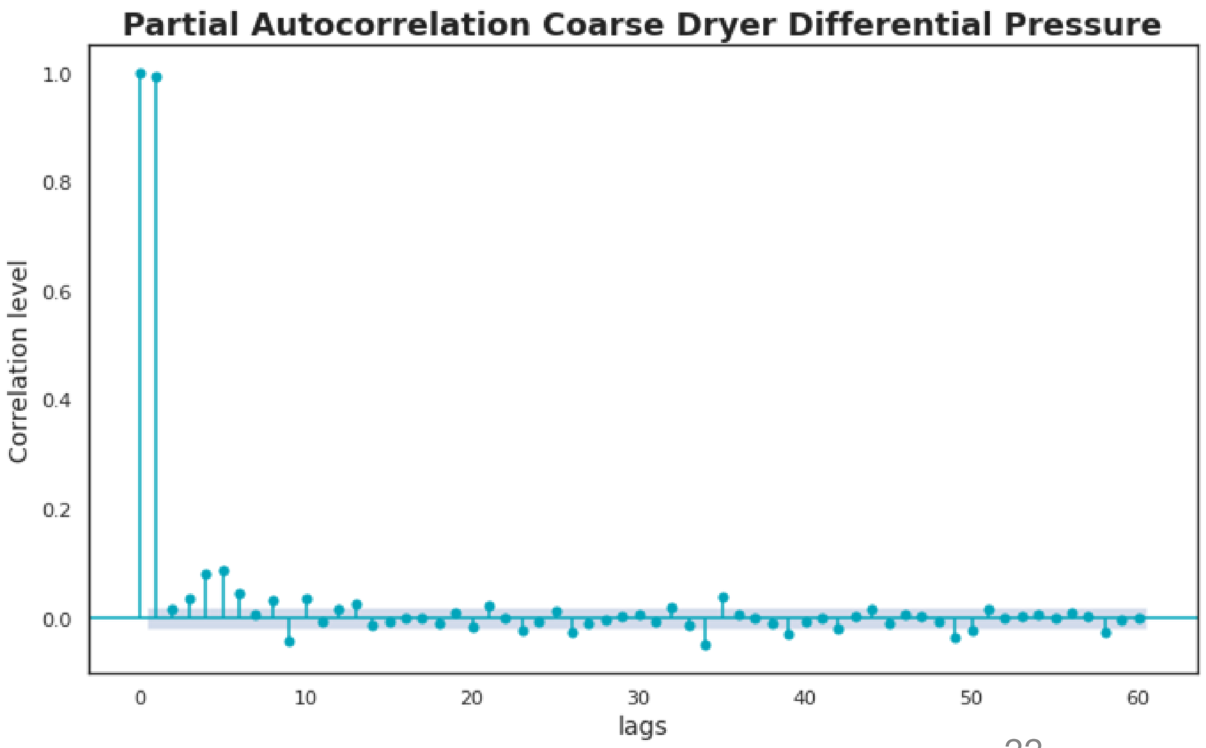
Tasks Performed:

- General data analysis:
  - Extract the correlations between the features and the delayed copy of themselves.
  - Identified new relationships between process values and set points (steady state errors) to be used in the model.
  - Identified the distribution of features change rate in the process to be used in the model.
  - Downtime analysis and clustering

Time Domain Analysis, Autocorrelation

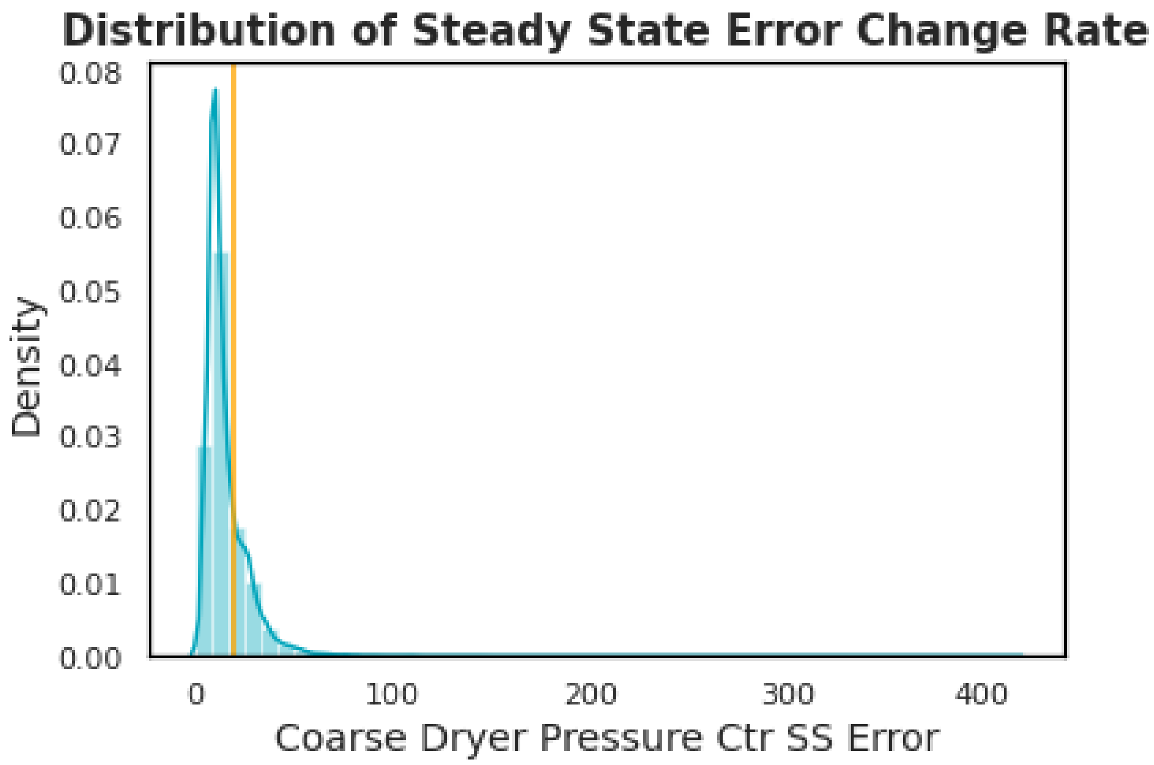
- Analyze the correlation of a signal with a delayed copy of itself.
- Select the delays or ‘lags’ which have the highest correlation to the present value as new features.

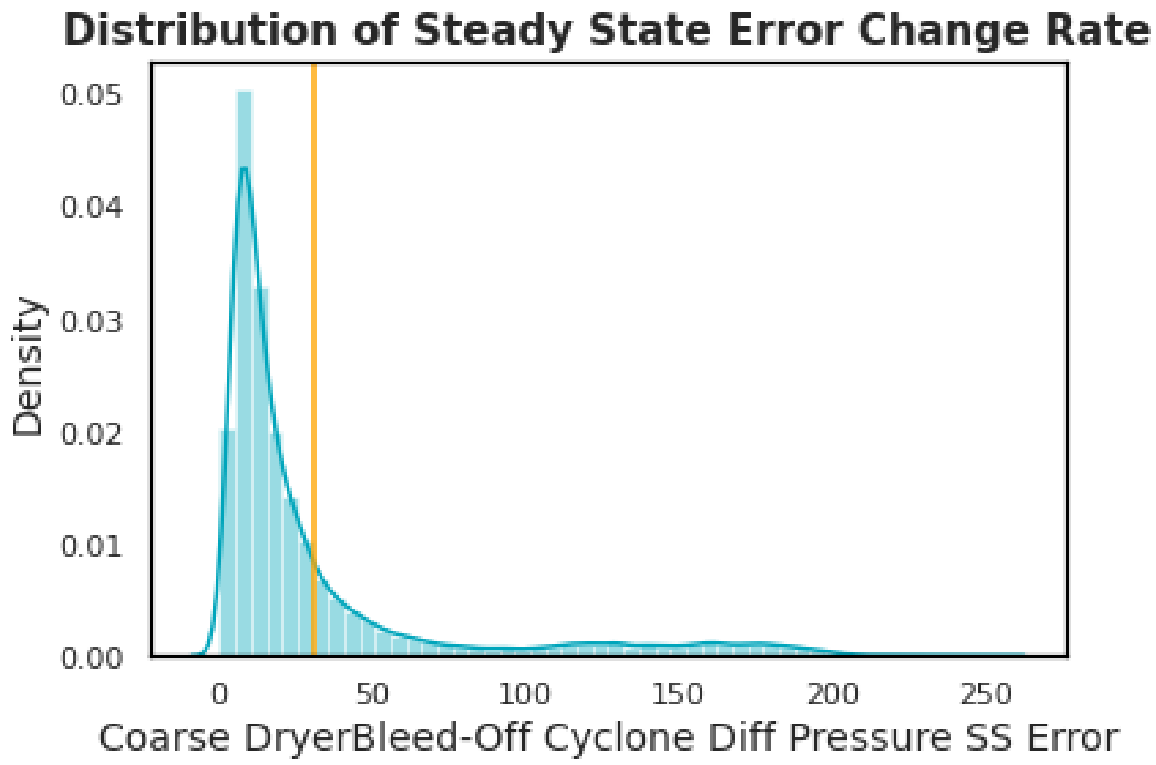




Methodology, Analysis & Results | EDA Insights

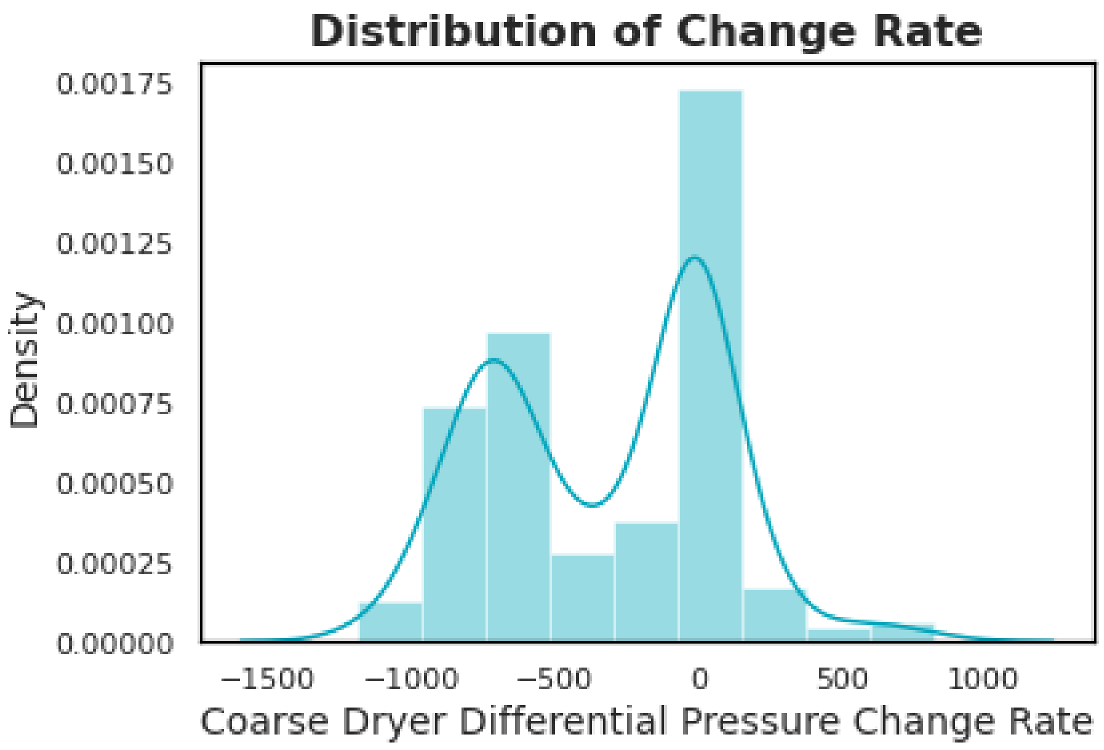
- Rate of Steady State error changes are calculated.
- It can be used as new parameters for ML model.

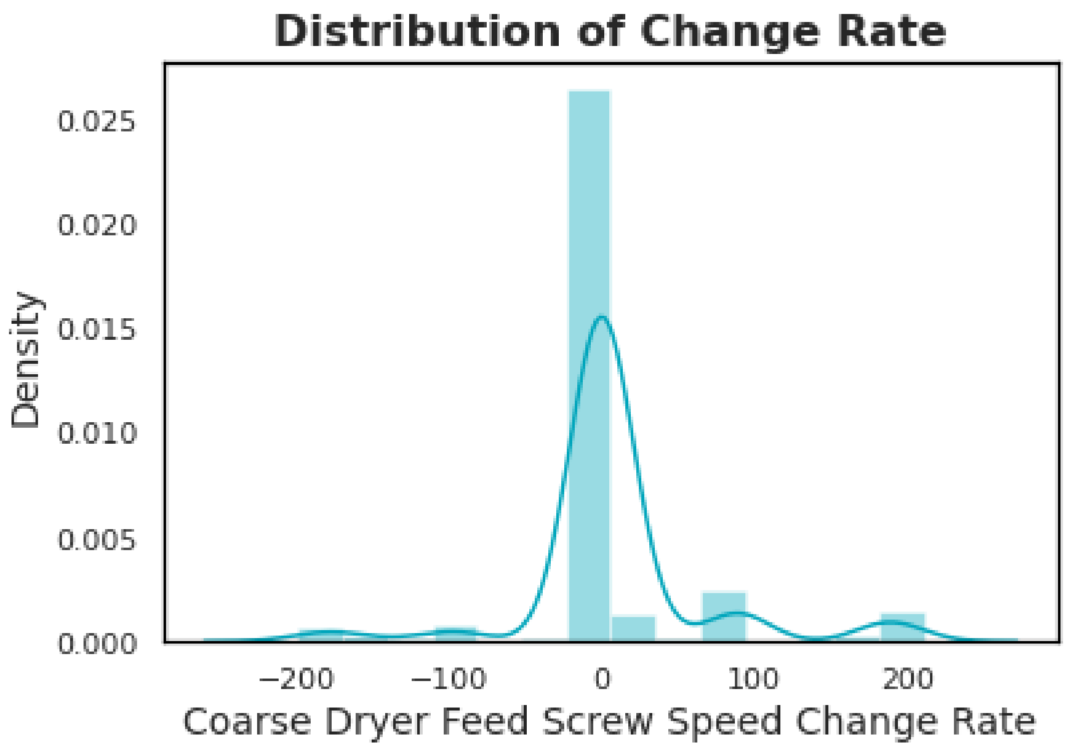




Methodology, Analysis & Results | EDA Insights

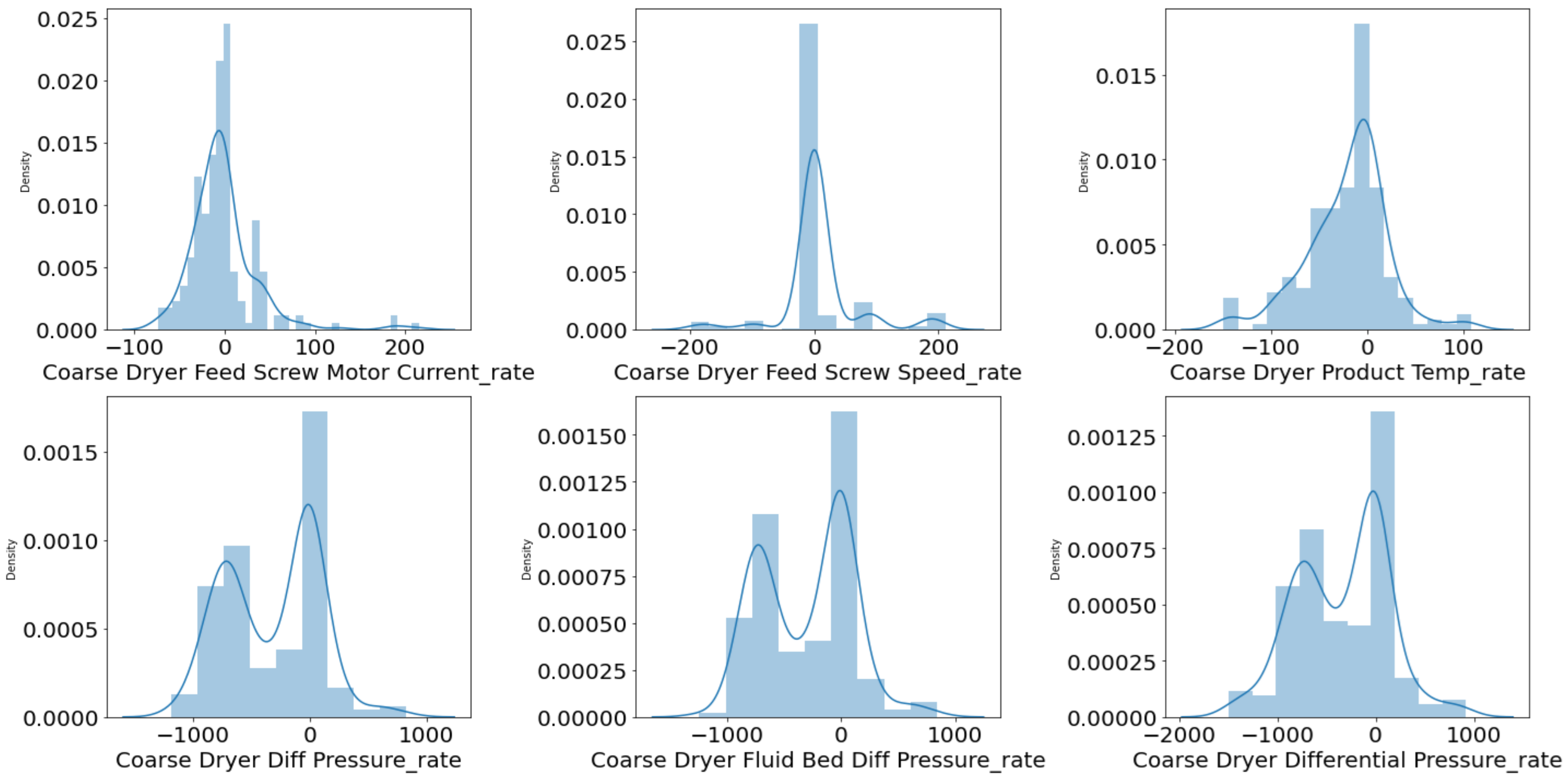
- Rate of changes for all parameters are calculated.
- It can be used as new parameters for ML model.





Distribution of Rate of Parameters Changes

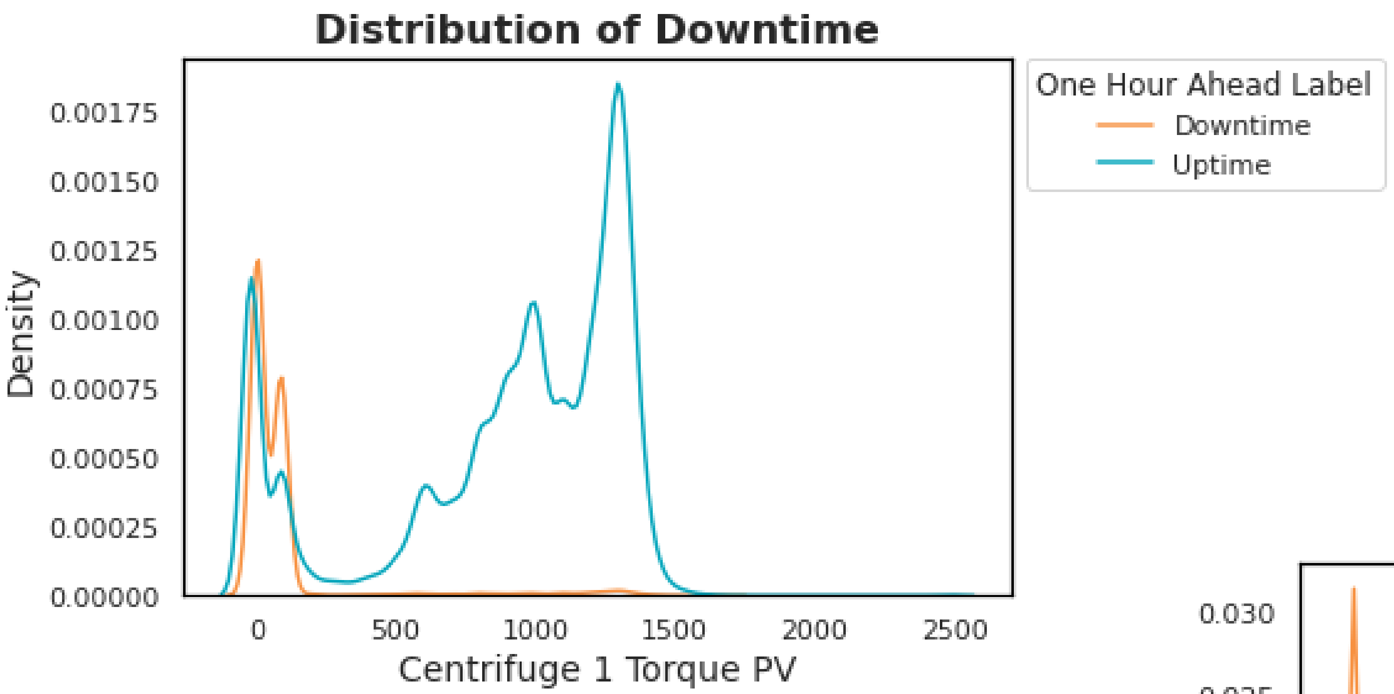
- Rate of changes for all parameters are calculated.
- It can be used as new parameters for ML model.

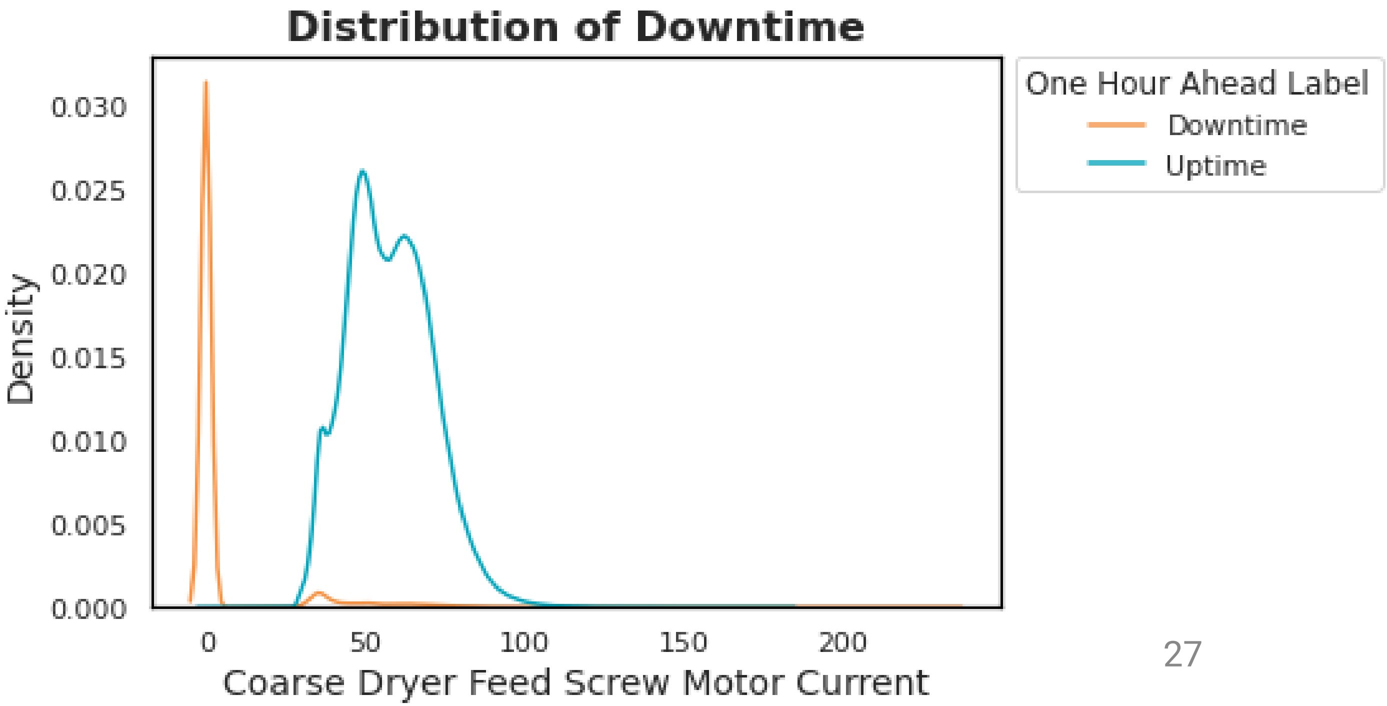


Downtime analysis and Clustering

- Downtime analysis and clustering: We have discovered that the downtimes are most probably not all the same types, as the associated features represent totally different behaviors.
- We further demonstrated that through clustering techniques to client

Downtime Analysis and Clustering:



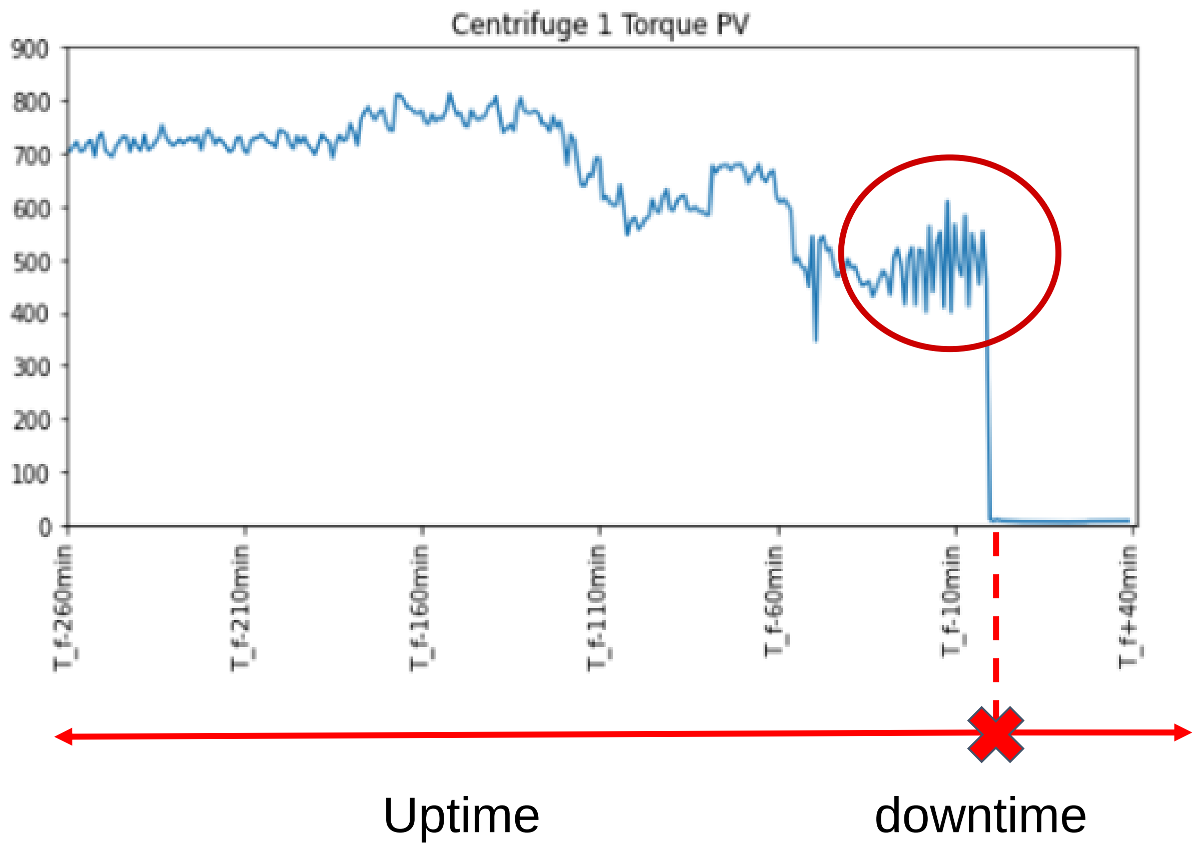


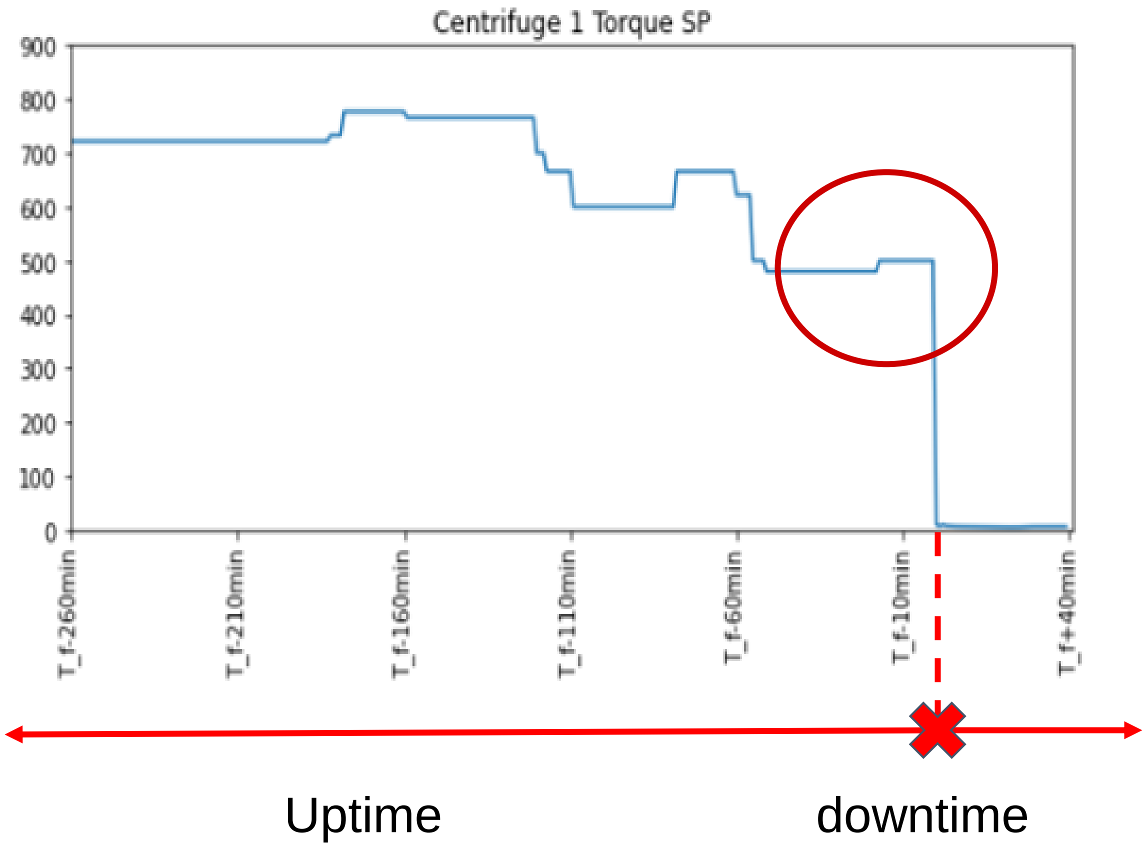
Downtime Analysis and Clustering:

Case I: Behavior of the process before downtime

Failure :9560

One hour ahead: 9500



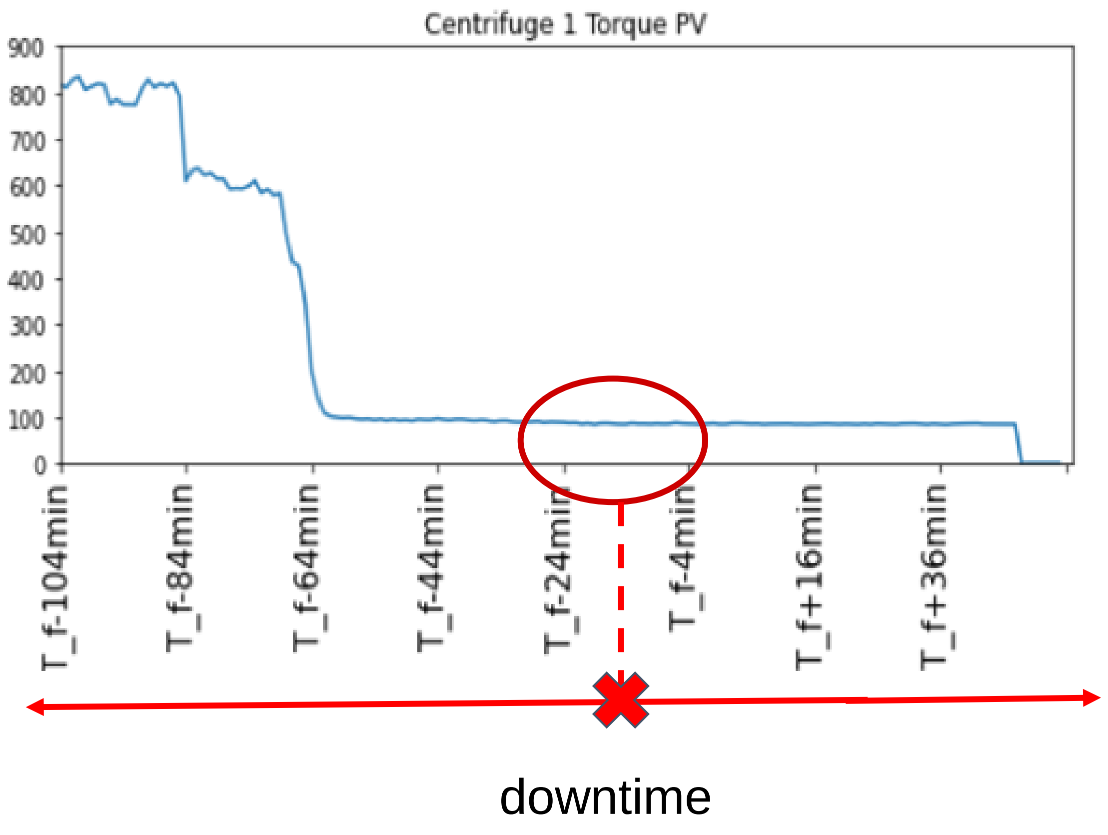


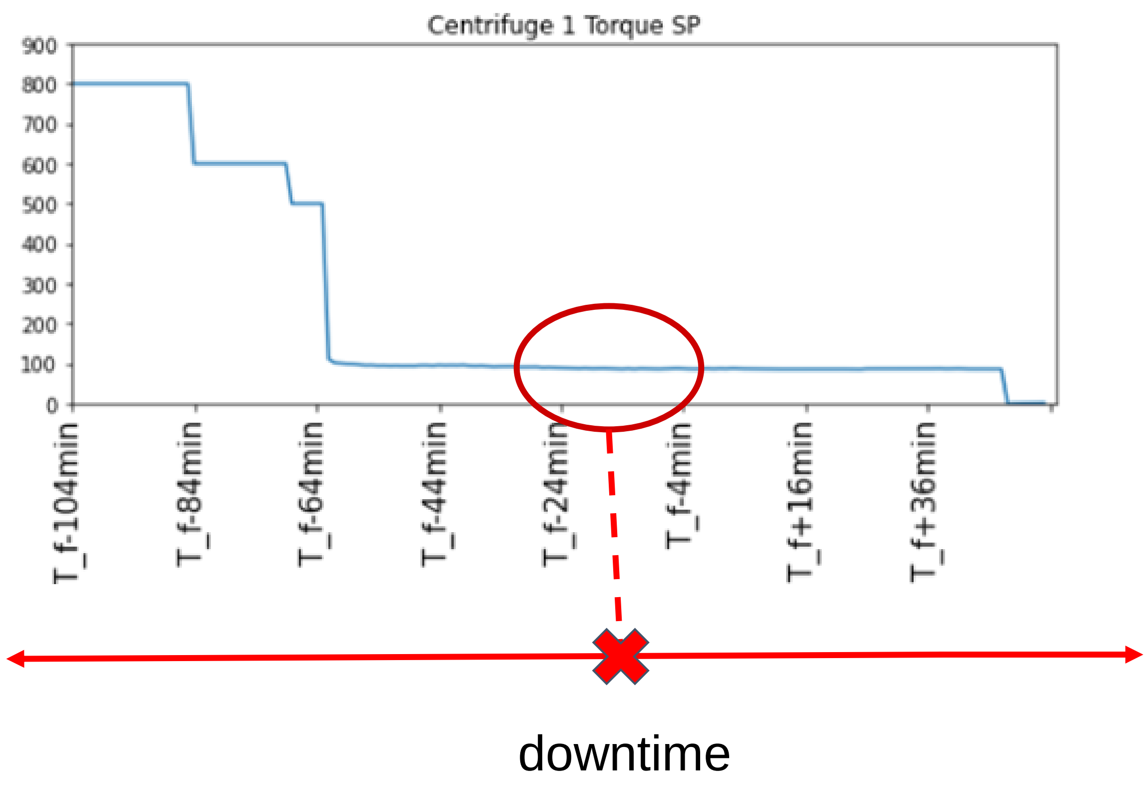
Downtime Analysis and Clustering:

Case II: Behavior of the process before downtime

Failure :12524

One hour ahead: 12464



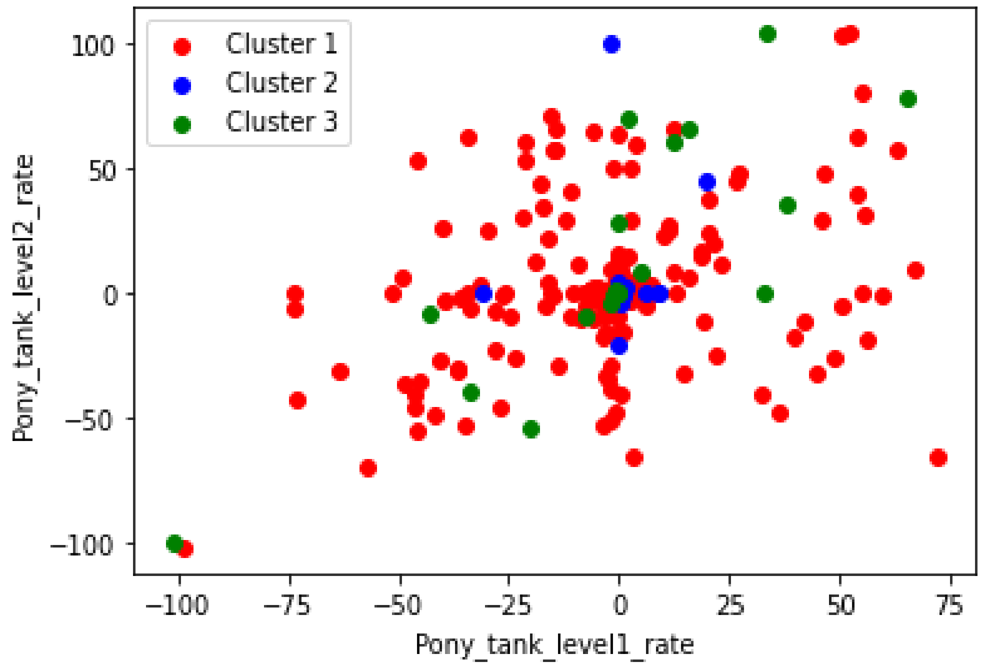


Downtime Analysis and Clustering:

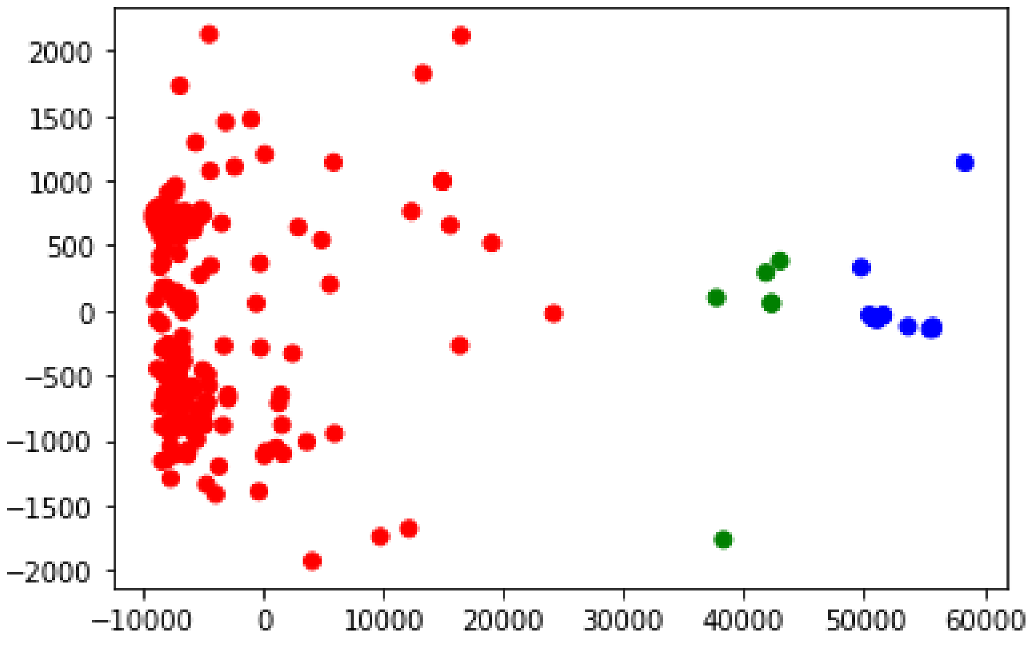
List of Parameters:

- ❑ Rate of Coarse Dryer Pressure, Temperature, Motor Current, and Tanks Levels changes.
- ❑ Error between different SPs and PVs 1 hour ahead of downtime.
- ❑ Shape of final Data set is (213,28)

K-Means



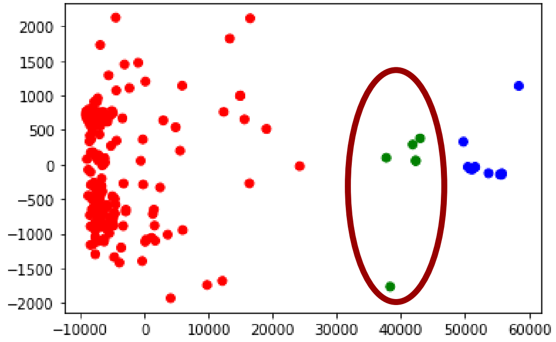
Spectral Clustering

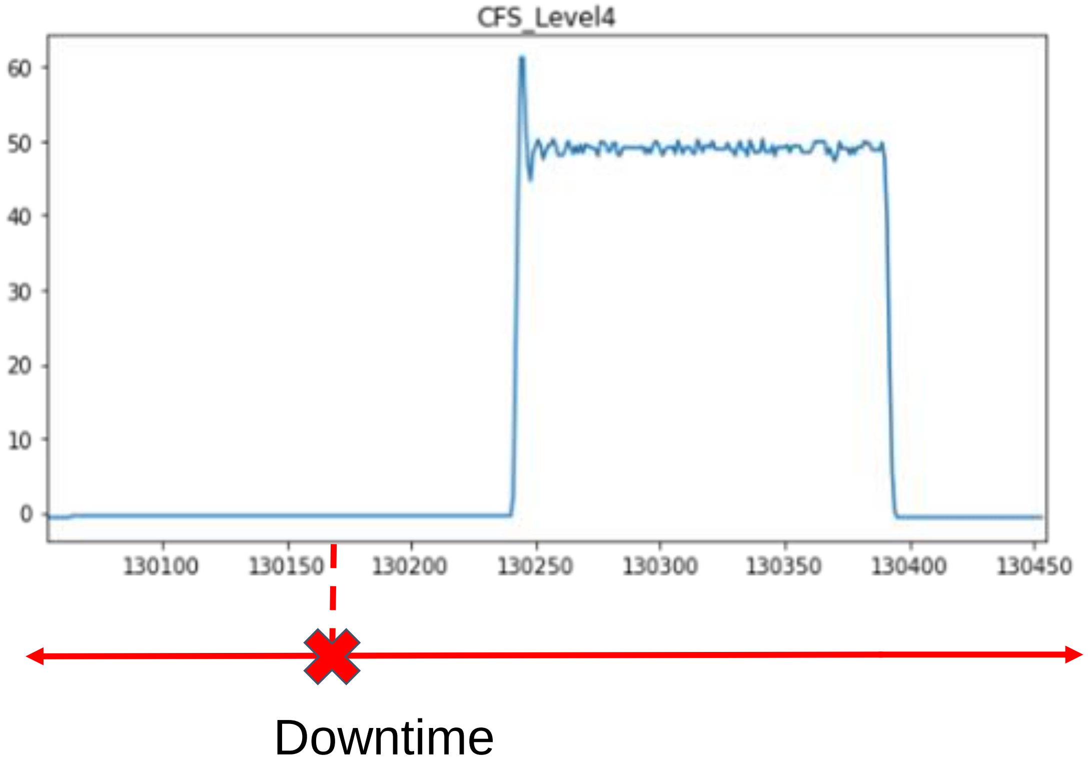


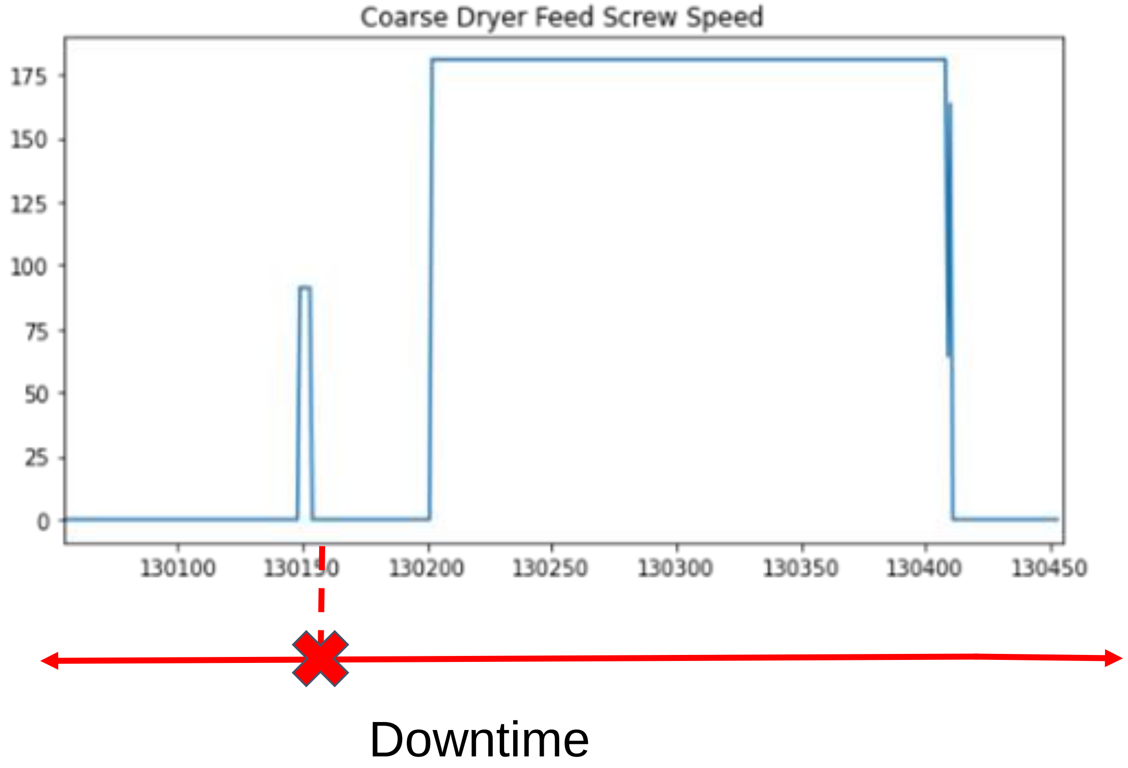
Behavior of the parameters at downtime

Cluster 2

`Downtime= 130154`



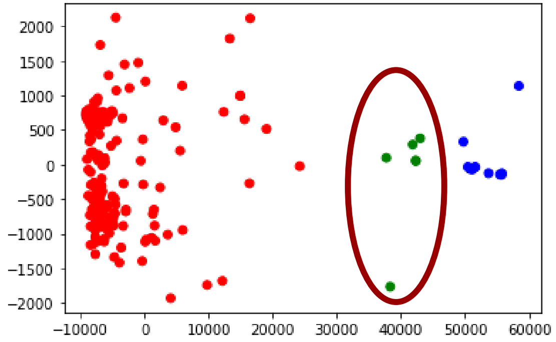


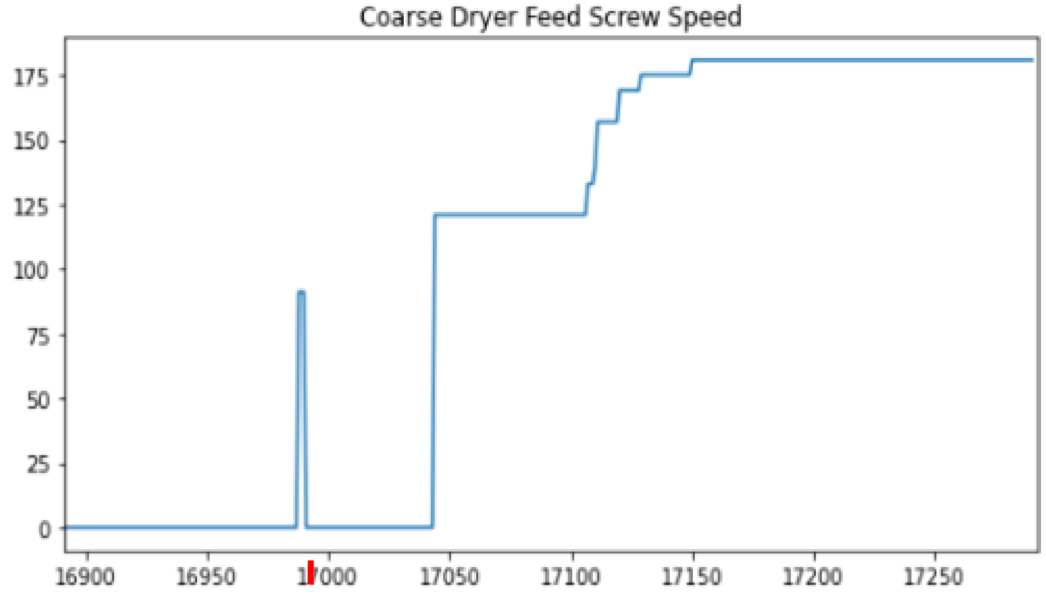


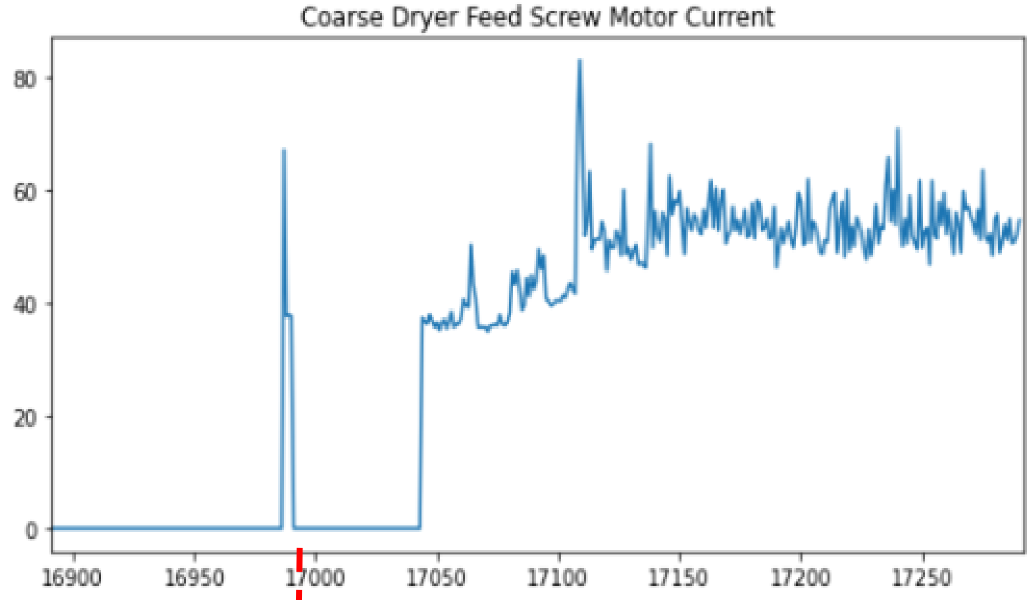
Behavior of the parameters at downtime

Cluster 2

`Downtime= 16991`







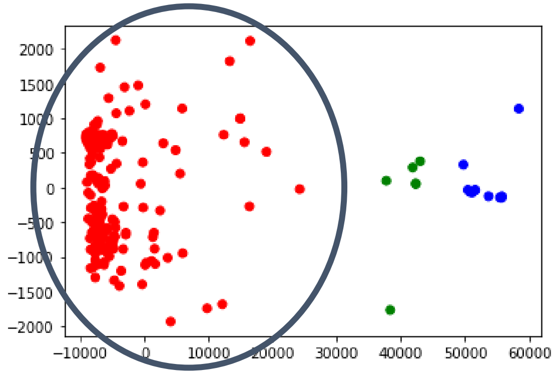
Downtime

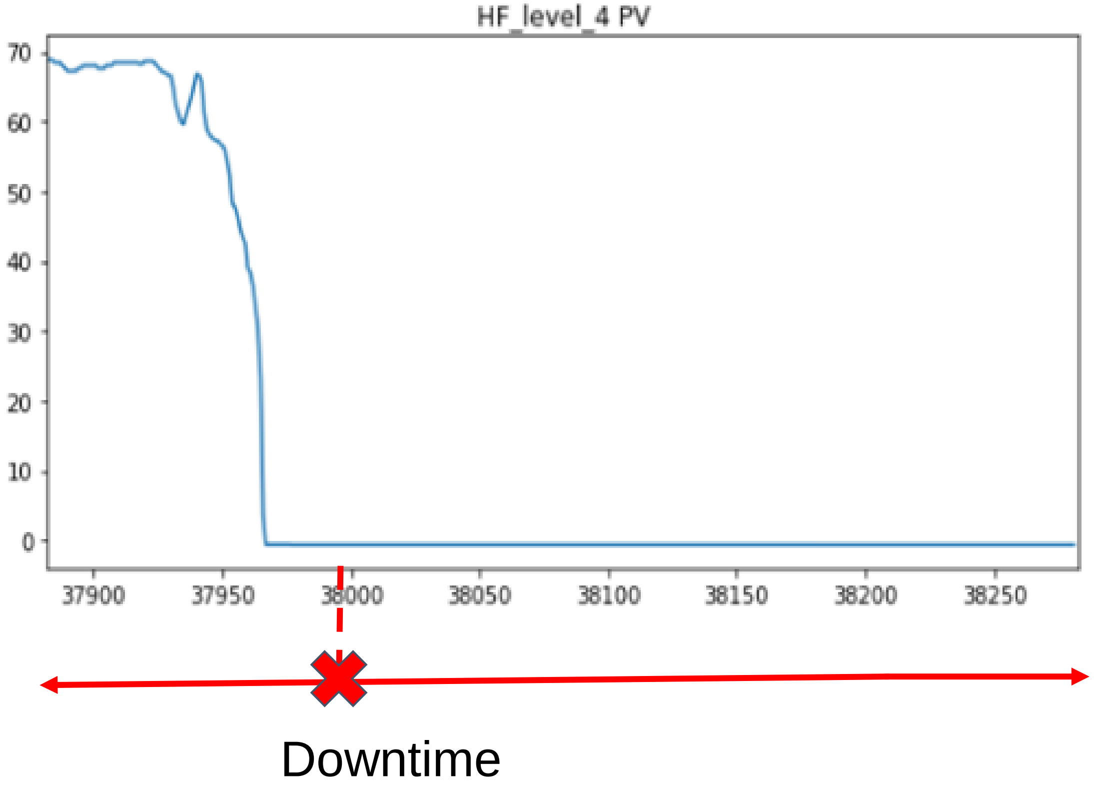
Downtime

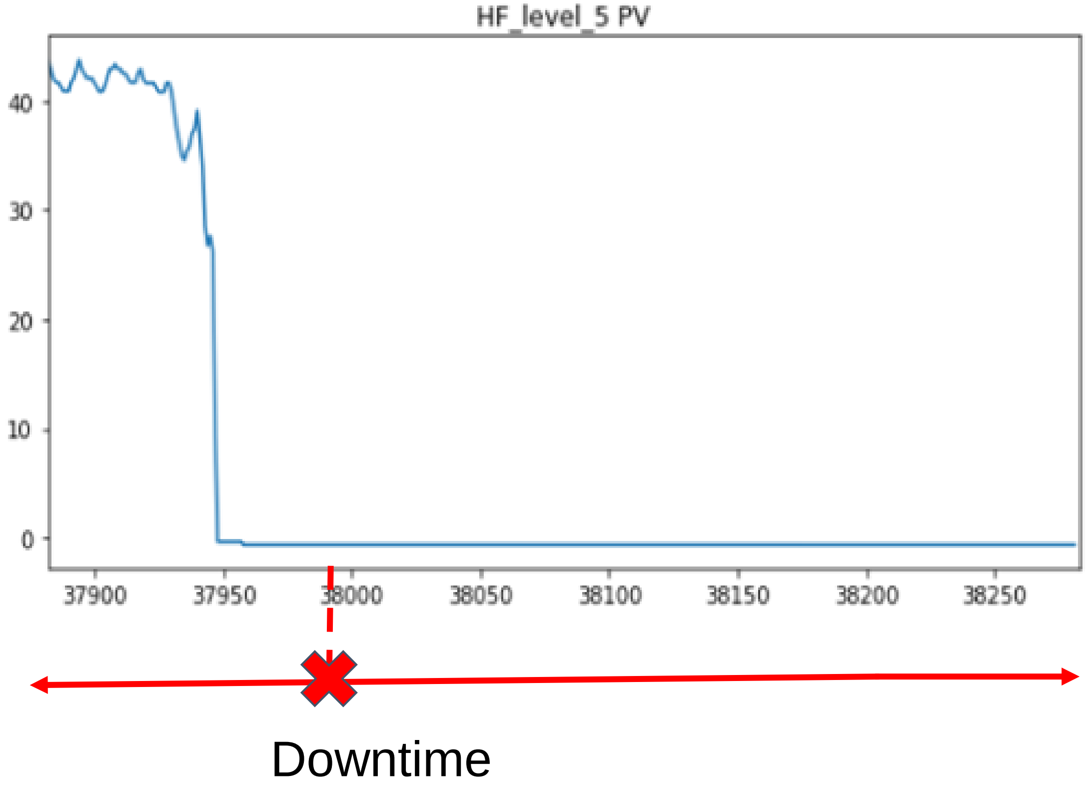
Behavior of the parameters at downtime

Cluster 1

`Downtime= 37982`



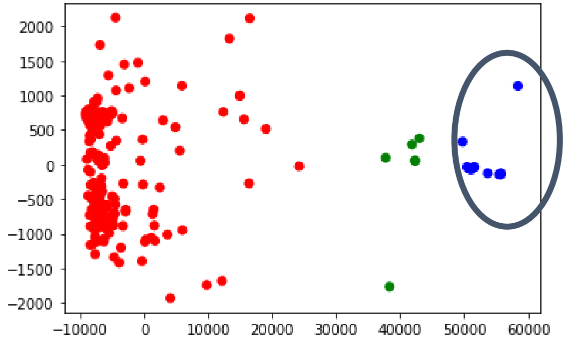


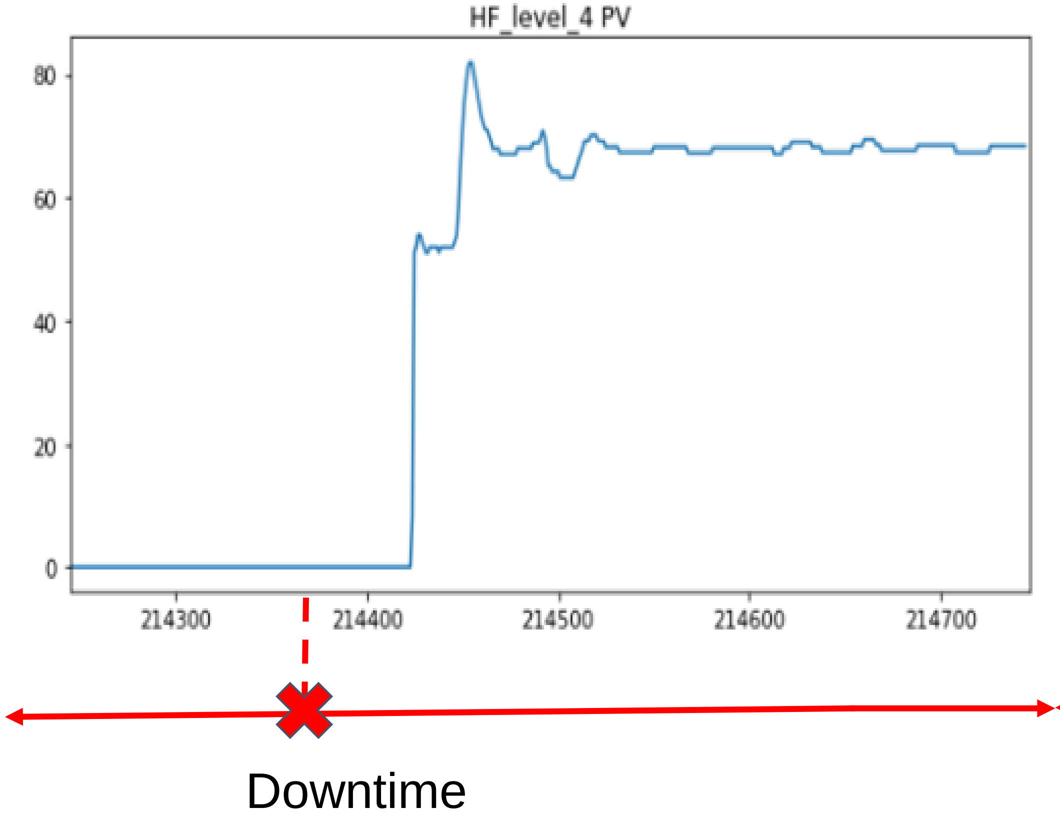


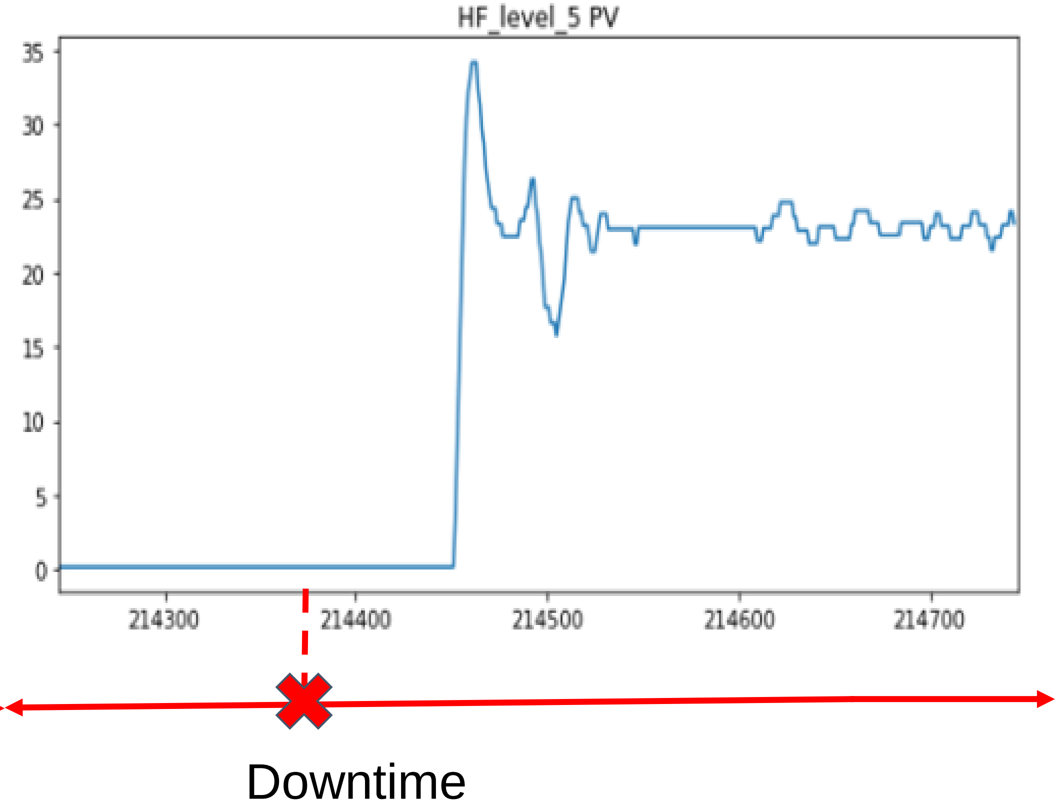
Behavior of the parameters at downtime

Cluster 3

`Downtime= 214345`







Observations on Downtime Analysis and Clustering and the way forward

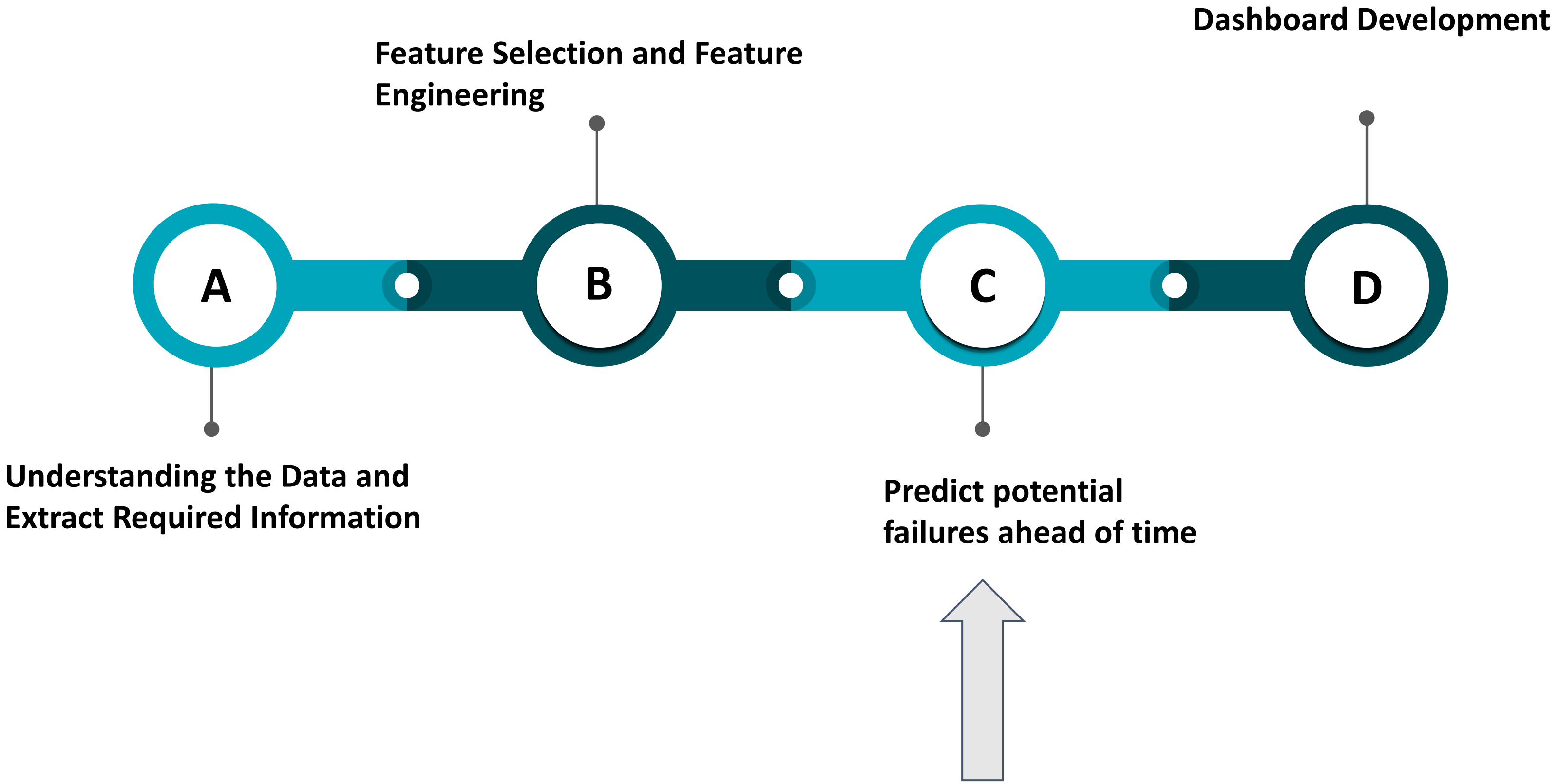
AltaML

- We have discovered that the downtimes are most probably not all the same types, as the associated features represent totally different behaviors. Clustering techniques were also used to further demonstrate that.

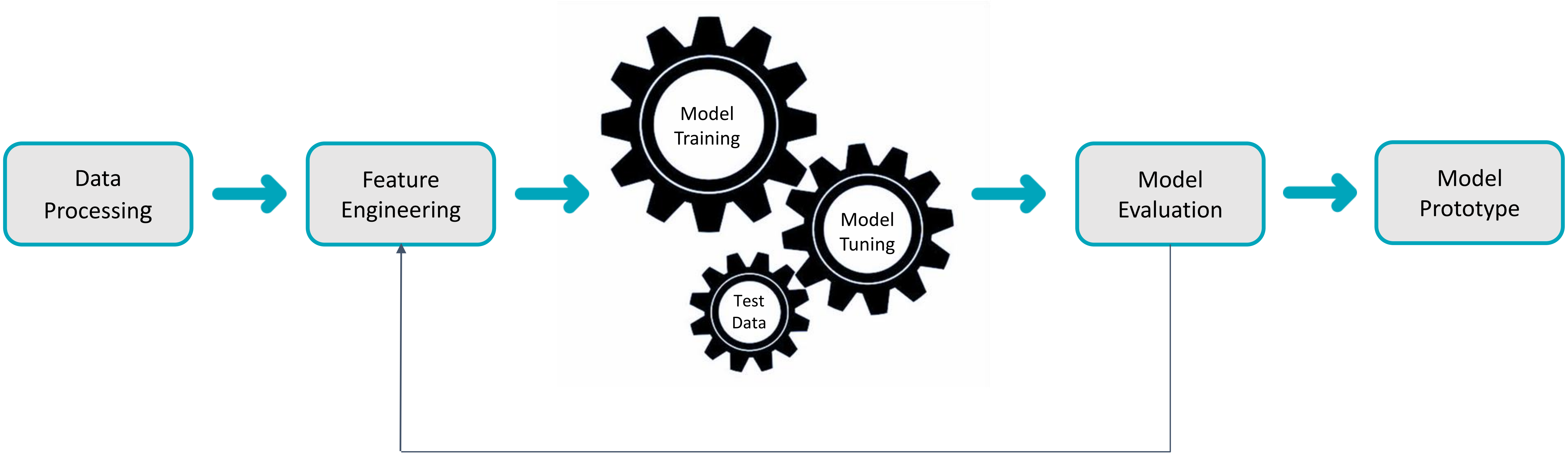
- We informed this to client and asked them if they can provide us labels for the downtimes. They appreciated the effort, however they couldn’t provide us any labels at this time and wanted to treat all the downtimes as one group in our model for this project.

Road map for the project:

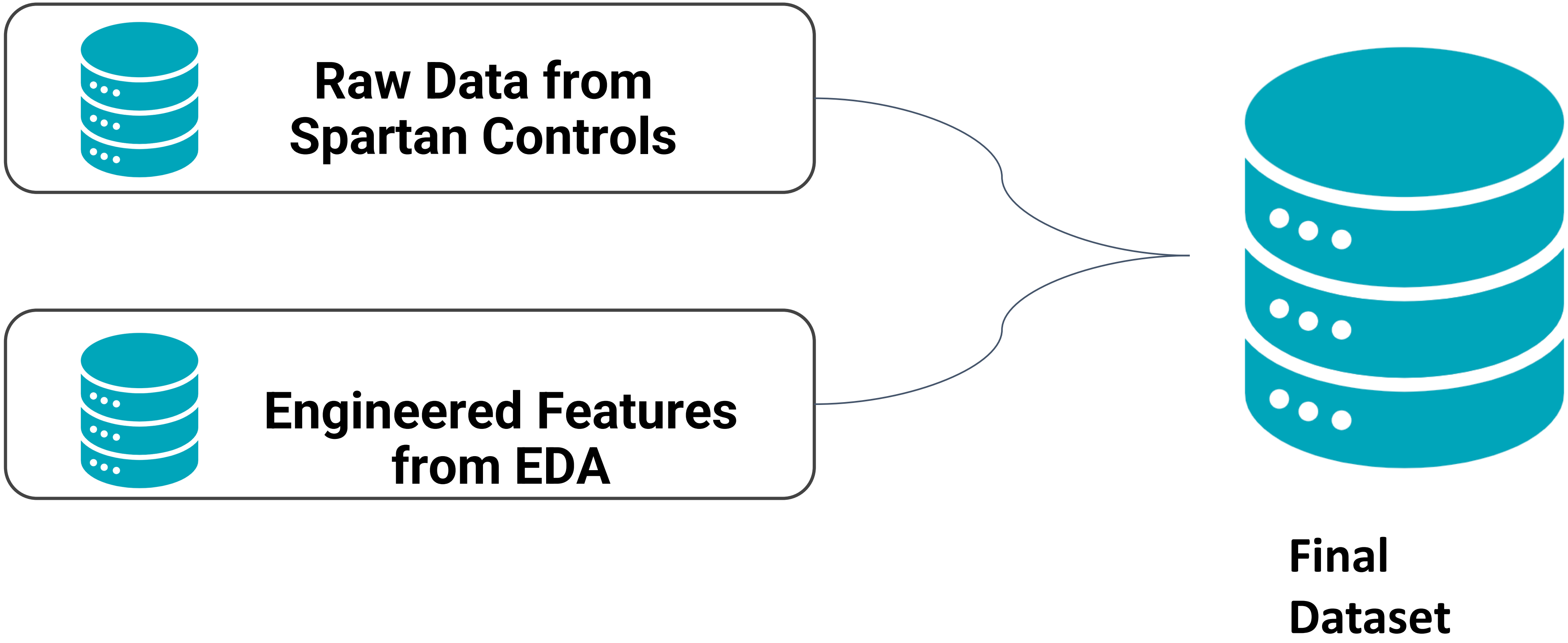
Predict potential failures ahead of time (ongoing)



Methodology, Analysis & Results | Modelling & Experiment Design



Methodology, Analysis & Results | Feature Engineering



AltaML

Model Performance

50% of the total downtime events were correctly labeled as shutdown

The Breakdown:

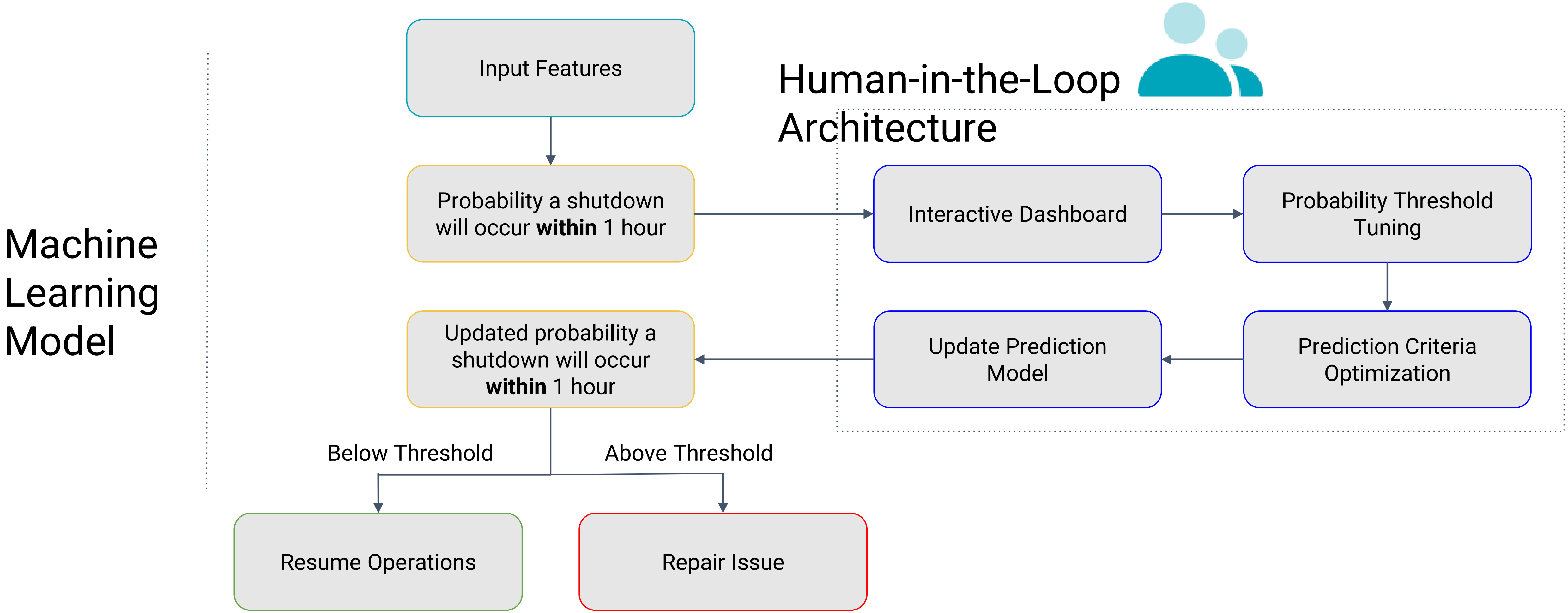
99% Correctly predicted that no downtime will occur

89% Correctly predicted that a downtime will occur.

1% A downtime was predicted, but no downtime occurred

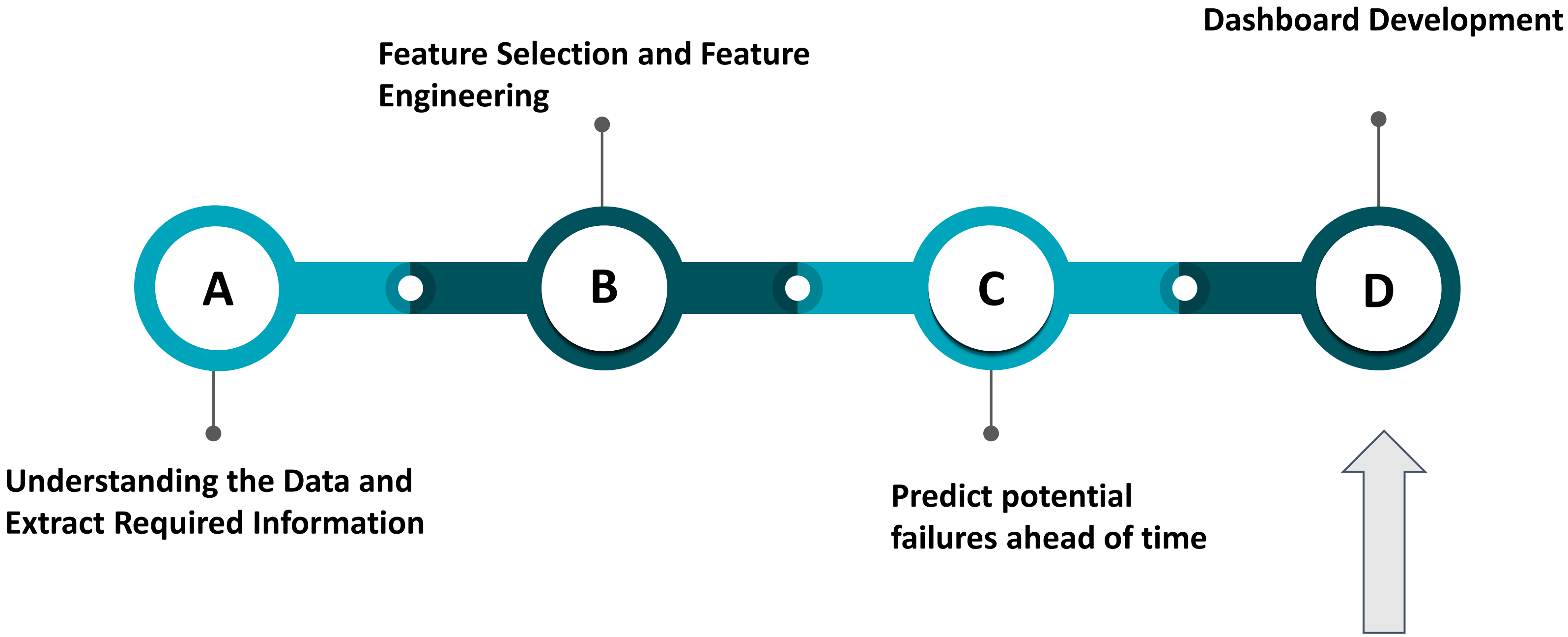
11% No downtime was predicted and a downtime occurred

Methodology, Analysis & Results | Prediction Structure



Road map for the project:

Predict potential failures ahead of time (ongoing)



Dashboard

AltaML

python + Streamlit

Recommendations & Next Steps | Equipment Downtime Prediction

AltaML

Data & Technical Next Steps	• Data labelling • Shutdown classification by type • Real-time data ingestion • **ML Model Improvement:** Additional features such as the moving average of each measurement rate of change may be included and tested to boost the XGBoost performance. • ** INVESTIGATE DOWNTIME EVENTS WHERE MODEL HAD A BAD PERFORMANCE
Business Next Steps	• **Model Integration:** The project has proven the potential of shutdown prediction. However, before investing in a software solution, it should be tested on a real industry and be adapted to real user's feedback. A pilot case is suggested. • **AI Education:** Training and educational workshops should be provided to operators and staff. • **Policy update:** As the model uses data from Spartan's client, policy updates may be required to adjust the nuances of data usage.
Use Case Recommendation	**Move to pilot phase** **We firmly believe that the shutdown prediction case has the strength to move forward to the pilot phase. It will reveal the model weaknesses, check the model adaptiveness and flexibility to other industrial processes, and improve user experience.**