Unsupervised multivariate anomaly detection

Who are we?

Students
Computer Science

AI enthusiasts

Foodora-lovers 🥙

Agenda

The problem

Proposed solution

Models

Future work

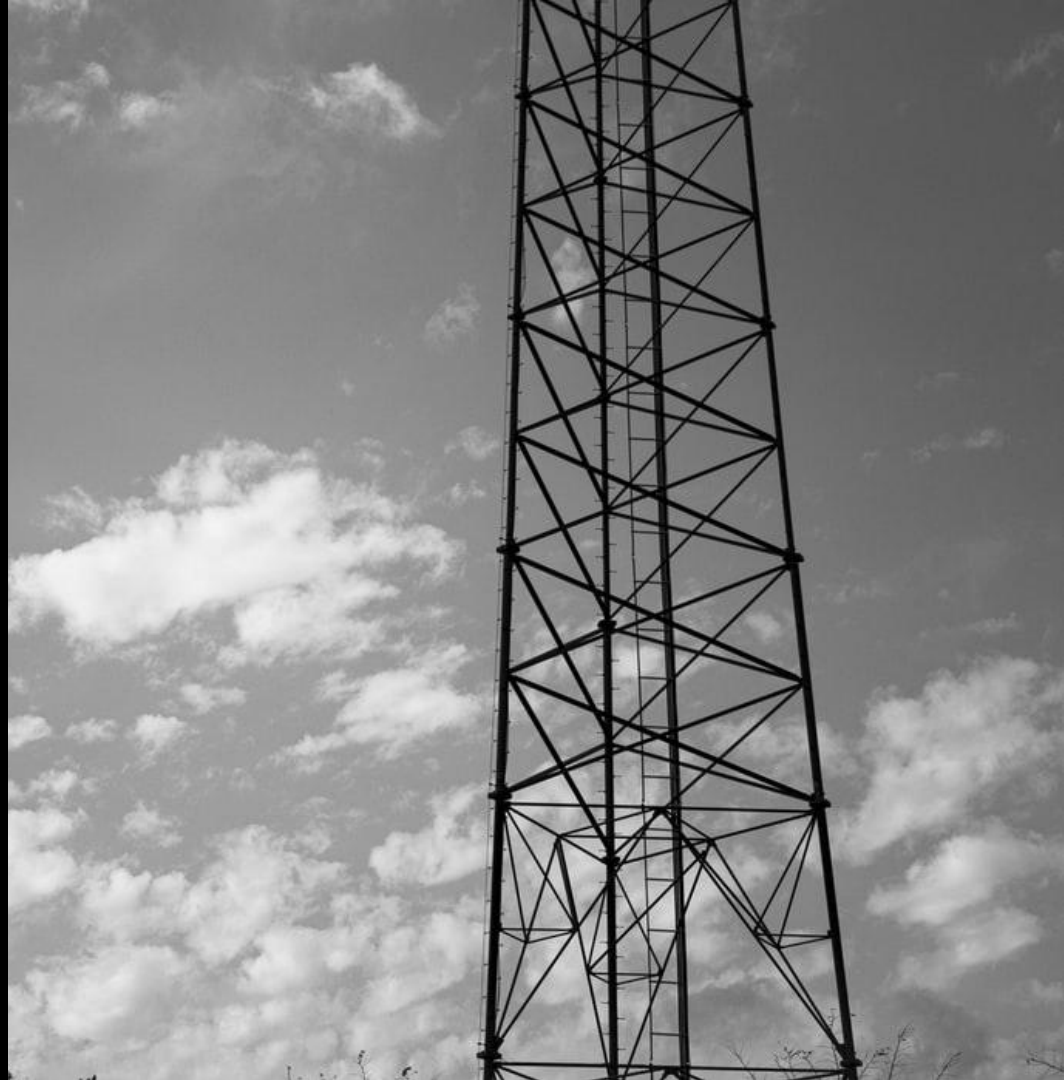
The problem

Unsupervised Anomaly Detection for Telenor Network Data

Radio Access Networks are prone to anomalies which affect the telecommunication grid. Discovering these anomalies is difficult, as domain expert knowledge is required to classify them.

Panel data

A number of cells on different cell towers. For each timestep, each cell is measuring numerous performance indicators.



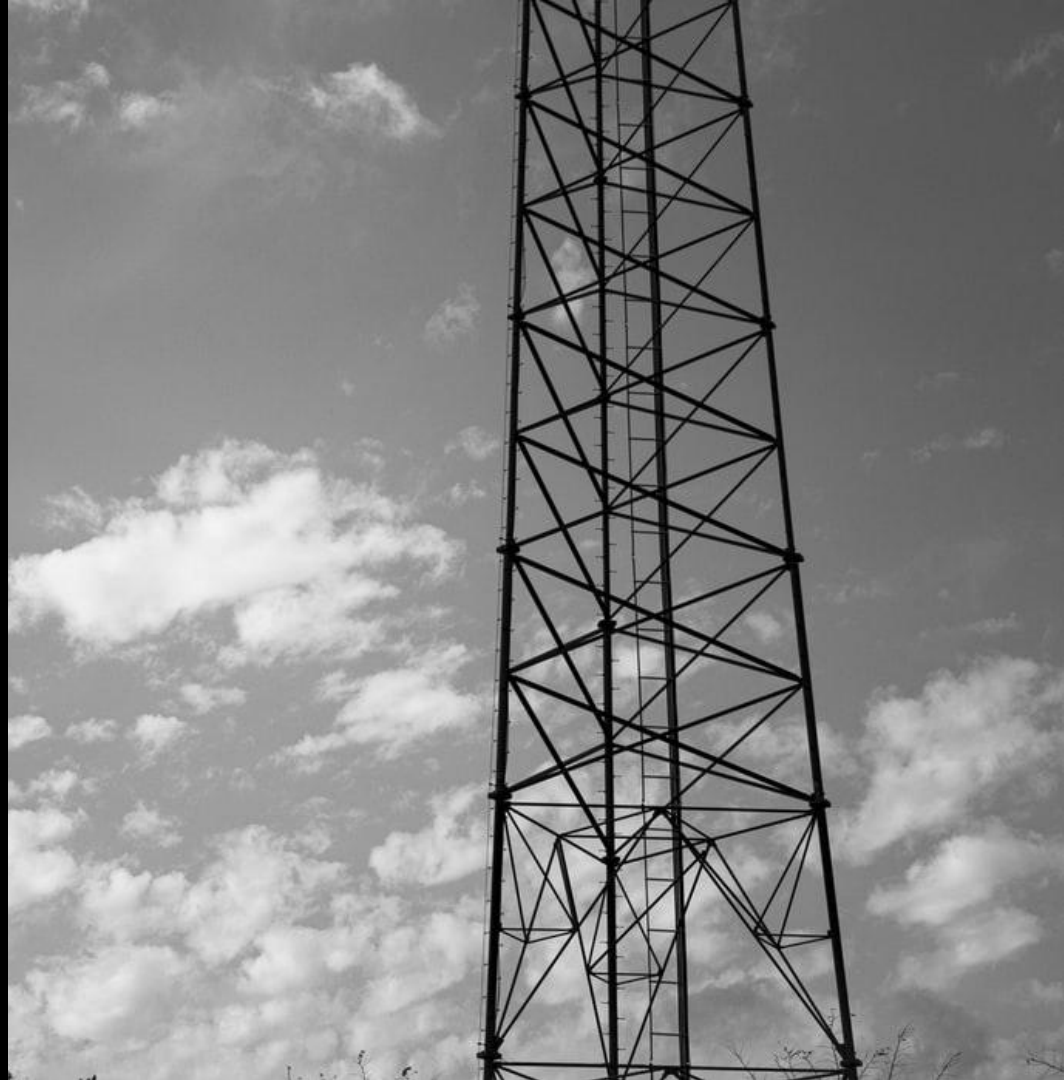
Proposed solution

Feature engineering

Extracting relevant features to ensure that relevant information is provided to an autonomous model

Ensemble of AI models

Three, computationally inexpensive, different models are compared to each other with respect to Telenor's RAN-dataset



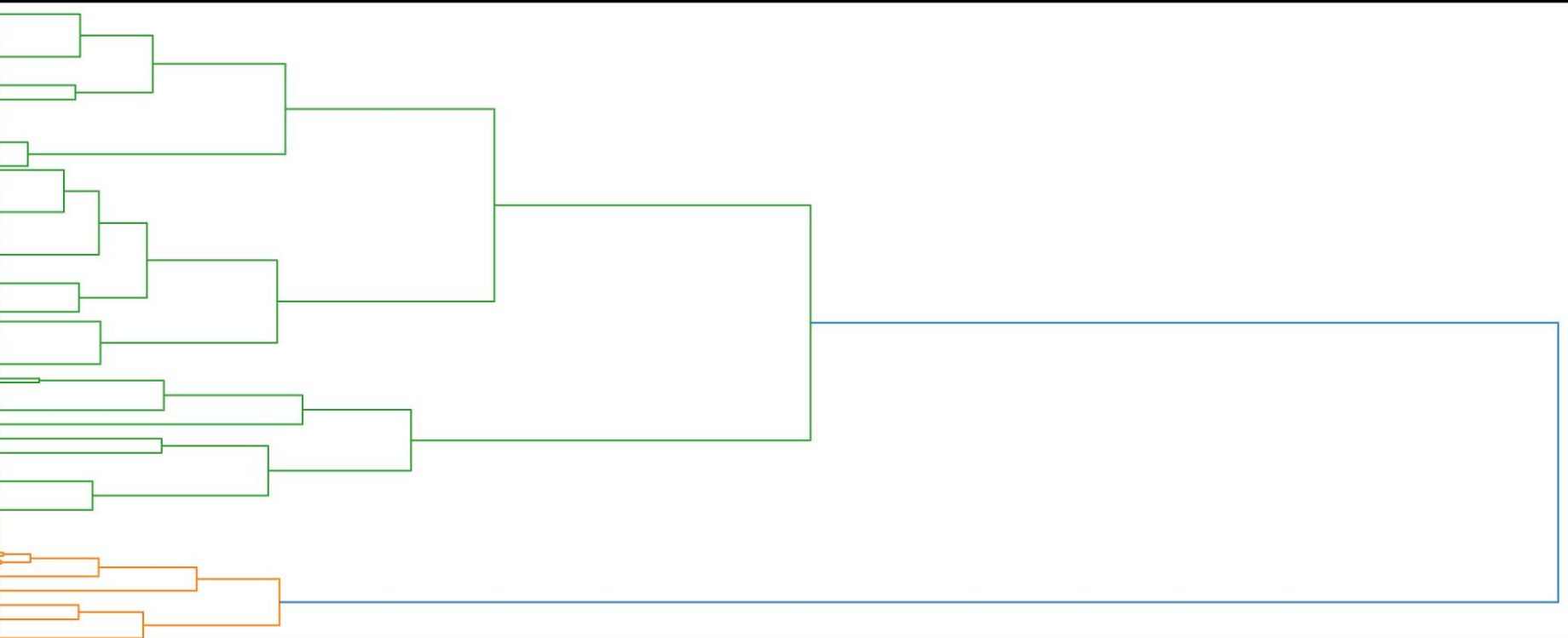
Proposed solution - Feature engineering

Cell-wise replacement of missing values

Adding general features

- Describe if certain columns originally had missing values
- Extract Technology and Frequency codes from cell-name
- Spatial features describing location of cell - hierarchical clustering

Spatial features using hierarchical clustering



Proposed solution - Ensemble of AI models

STD

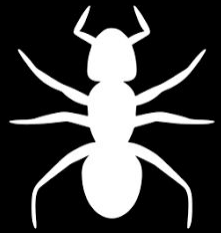
Isolation Forest

DeepAnT

Models

Models - Strengths and Weaknesses

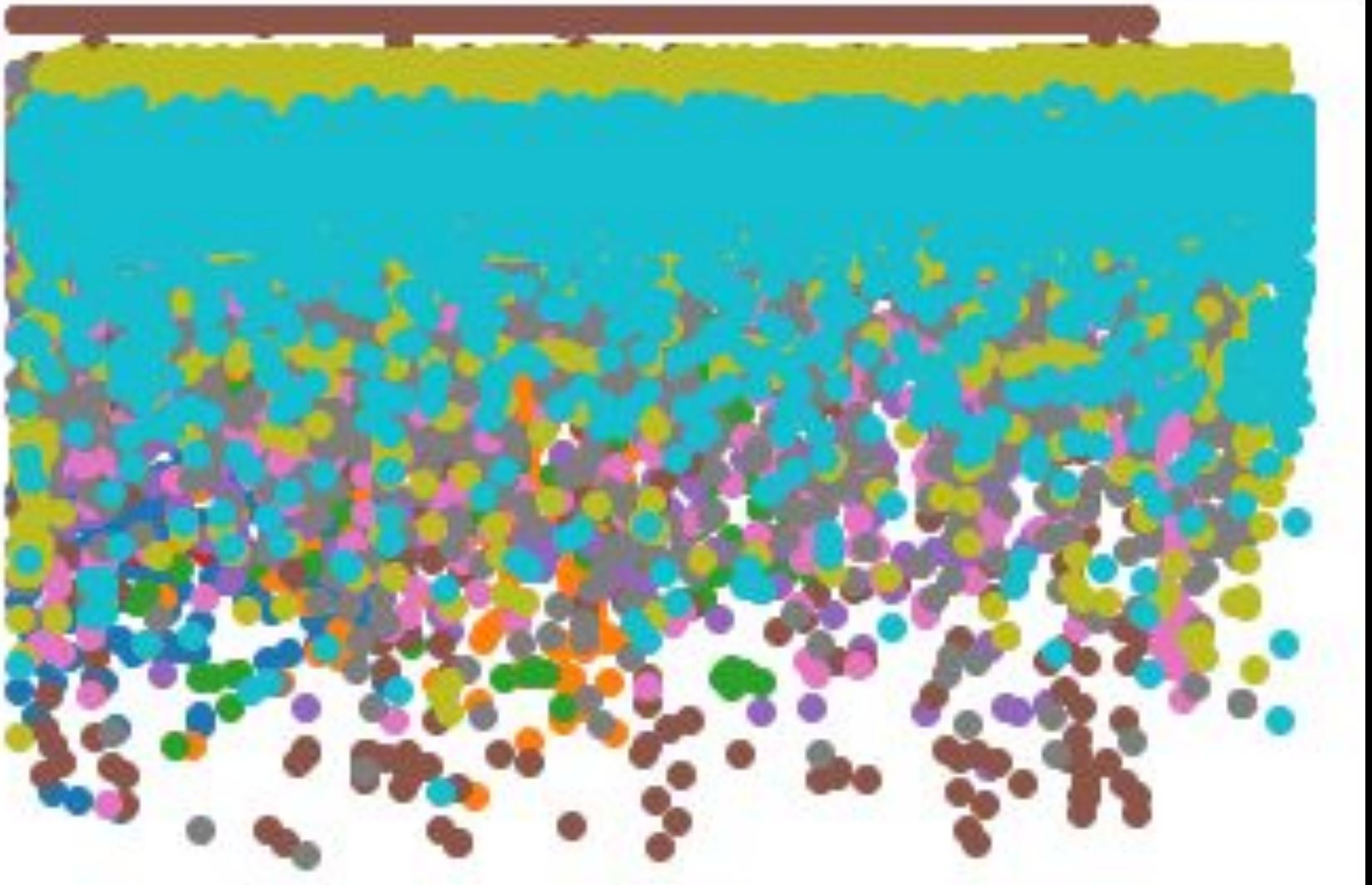
σ



+ Detects simple anomalies
- Does not detect feature-correlation or time-perspective

+ Detects feature correlation
- Not made for time-series
- Less explainable

+ Works on time-series
- Complex



Models - Ratio of anomalies found (49 possible)

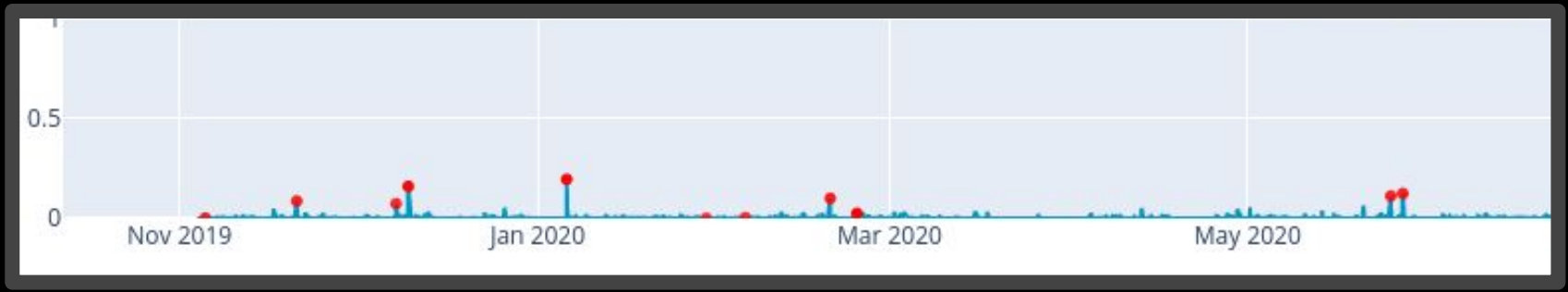
	STD	iForest	DeepAnT	Combined
Recall	**43%**	**61%**	**98%**	**98%**

Models - Visualization Previews

Feature-wise predictions

Models - Visualization Previews

Timestep-wise predictions



Future work

Irregular measurements

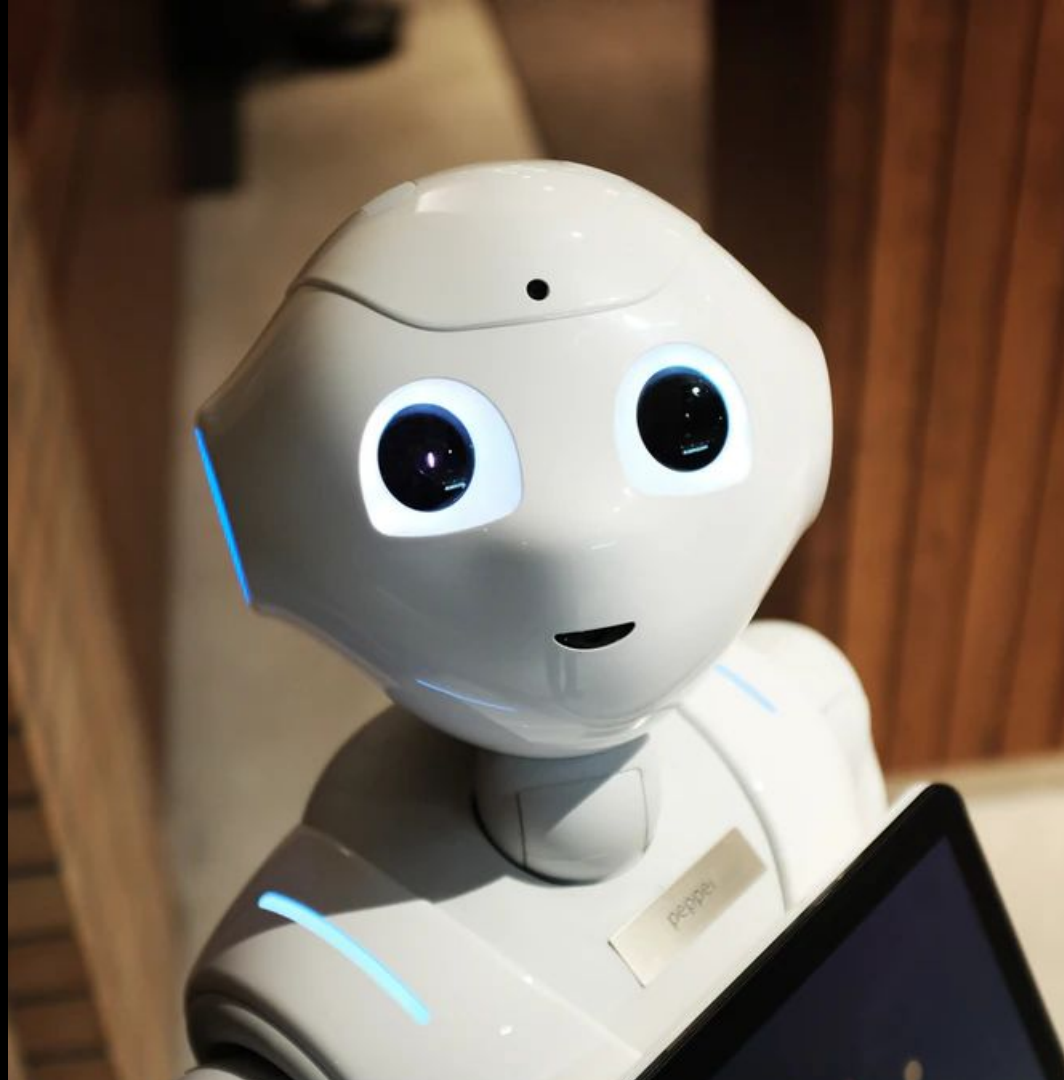
Handling inconsequent measurements-times, by incorporating this information into features

Spatial-temporal dynamics

Improve the encoding of the relationship between location and time by utilizing larger zones of cells. DeepAnT can input data with more dimensions, but this requires handling the irregular measurements-problem

We recommend

Active learning to further utilize domain-knowledge
Graph Neural Networks for spatio-temporal modelling



Telenor - Unsupervised Anomaly Detection for Telenor Network Data

Hackathon 2021 - Group 5

Installation

In this project we used requirements.txt file. To install the packages please write:

```
cd brain-cogito-hackathon-2021/
pip install -r requirements.txt
```

The requirements.txt file includes the following Python libraries installed::

- NumPy
- Pandas
- matplotlib
- scikit-learn
- pandas-profiling
- plotly
- torch
- tqdm
- statsmodels

You will also need to have software installed to run and execute a Jupyter Notebook.

If you do not have Python installed yet, it is highly recommended that you install the Anaconda distribution of Python,

Thank you for having us!

Questions?