How To Run This Project?

Since the project is composed of multiple Jupyter Notebooks, as long as the environment is installed, each notebook can be executed individually.

Summary of Each Notebook

exploratory-data-analysis.ipynb:

This notebook visualizes and explores the data. It implements multiple plots to view distributions, trends etc.

preprocessing.ipynb:

This notebook implements preprocessing steps that mentioned in the preliminary report. This implementation is old and was used in the preliminary report.

preprocessing-2.ipynb:

This notebook implements the latest preprocessing steps in the final report.

reduced_features.ipynb:

This notebook implements XGBoost, TabNet, and hyperparameter optimization with the dataset that has reduced feture space.

tab2img.ipynb:

This notebook implements CNN and uses tab2img library to convert tabular data into images.

xgboost.ipynb:

This notebook implements the initial XGBoost mentioned in preliminary report.

xgboost-categorical.ipynb:

This notebook implements the XGBoost model that trained on high dimensional data mentioned on the final report.

classicalmethods.ipynb:

This notebook implements the logistic regression model that is mentioned in preliminary report.

Since project is composed of Jupyter Notebooks, most of the codes and steps are self-explanatory.