BrainStation Final (Capstone) Data Science Project

The goal of my project was to detect fraudulent credit card transactions using machine learning models.

The models trained were:

- Logistic Regression
- K-Nearest Neighbours (KNN)
- Decision Tree
- eXtreme Gradient Boosting (XGBoost)
- Random Forest

Contents

In this folder you will find the following:

- **Business Report**

- **Presentation Slides**

- **Jupyter Notebooks:**
    - Notebook #1 - Cleaning and Exploratory Data Analysis
        - In this notebook, I loaded the raw dataset from the source, conducted data cleaning where applicable, and dived into exploratory data analysis to obtain insights from the data prior to preprocessing and modelling.

    - Notebook #2 - Preprocessing
        - In this notebook, I loaded the cleaned up and feature engineered dataset from Notebook #1, and transformed it to be model ready. This includes transforming categorical columns to numeric columns.

    - Notebook #3 - Modelling
        - In this notebook, I loaded the preprocessed, model ready dataset and fitted it on the models mentioned above. I then optimized the best performing base models to get even better model performance.

- **Tableau File** - contains the visualizations made for the project.