Context

A bank has hired you as their first data scientists. They have been collecting transactional data from their customers for the last 3 years.

They are not well versed in Data Science nor Machine Learning. Your team has been tasked to explore existing datasets and propose machine learning use cases that will bring value directly to the business or the end customers.

Assignment:

Each member of your team has to define a business use case from the shared datasets, train relevant Machine Learning algorithms and showcase results achieved. You have to provide at least one use case for each type of the following data mining problem:

- Regression analysis
- Classification analysis
- Unsupervised learning such as clustering, anomaly detection, etc

So each student has to define a business use case, translate it into a data mining problem and then build Machine Learning solutions accordingly. Group members need to collaborate in order to assign different business use cases, share code and insights.

Here are some examples of potential use cases:

- Helping customers to better budget their finances by predicting their total spending amount for the next month (regression or classification)
- Helping Compliance Team to identify fraudulent behavior by predicting if a transaction is a fraud or not (classification)
- Helping Sales Team to target customers most likely to have a significant increase in spending in the next 3 months (regression or classification)
- Helping Marketing Team to send customized marketing emails to groups of customers presenting similar spending behaviors (clustering)
- Helping Customer Support team to reach out to customers with abnormal behaviors from their usual spending patterns (anomaly detection)

You will write a final report presenting the business use cases for each students, assessing the benefits and drawbacks of each model predictions, the dataset used for training, the approaches you took during your experimentation phase, best results achieved, analysis on correct and incorrect predictions, business impacts and ethical consideration of your model and final recommendation.

Dataset:

Dataset can be downloaded here: https://drive.google.com/file/d/1ipwbg0VClq9OGYP-B9wnFegKxPneMFug/view?usp=drive_linkLinks to an external site.