

Data Science Mini Project:

Background:

Ever since Covid-19 pandemic, OLG has moved all the casino games online to the PlayOlg platform. With increased volume of transactions online, the executive team wants to know if the company can identify whether a transaction is suspicious or not.

Here are some transactional data and your goal is to find all the possible suspicious transactions. Please follow the instructions below to complete this mini project.

What you need to do:

You are given 2 .csv files: “historical_transaction.csv” and “current_transaction.csv”. Both files contain ‘x0’ – ‘x99’, which are the features that describe the transactions. Only the “historical_transaction.csv” file contains the ‘y’ column, which indicates whether a transaction is suspicious or not, with 0 being “no” and 1 being “yes”. Your goal is to build a model with Python by using the data in “historical_transaction.csv”, then make a prediction for each row in “current_transaction.csv” by classifying whether the transaction is suspicious or not in the ‘y’ column.

Note: The data in the “historical_transaction.csv” file needs to be cleaned. All the Python scripts need to be commented.

Please send all the Python scripts and files to alu@olg.ca as a zip file.