MACHINE LEARNING ASSIGNMENTS

PRE-REQUISITE

1) Review the training deck available at: https://github.com/vijay-agrawal/genai-batch2/tree/main/training-materials
2) Download, execute and understand all the demo codes available at: https://github.com/vijay-agrawal/genai-batch2/tree/main/code (machinelearning.zip)

ASSIGNMENT 1

Copy and edit the following notebook:

https://www.kaggle.com/code/gpreda/credit-card-fraud-detection-predictive-models

Execute the notebook completely and review and understand al the steps.

Suggest any enhancements that can be done.

Try any model not tried in the notebook, or tune using hyperparameters and explore if you can improve the evaluation metrics

Explain how imbalanced data can be handled (as comments in your notebook)

Submission: Submit your notebook in the drive folder called ML

ASSIGNMENT 2

In clustering algorithm K-Means, do the number of clusters to be created figured out by the algorithm or it needs to be given beforehand?
If it needs to be given beforehand, how can we determine what are optimal number of clusters to give?

Explain with code any approach you can come up with to determine optimal number of clusters