Assignment- Problem Statement:

https://www.kaggle.com/datasets/mlg-ulb/creditcardfraud

In this assignment, delve into credit card fraud detection using a high-dimensional dataset. Address data imbalance by employing oversampling or undersampling techniques, followed by data normalization. To tackle high dimensionality, apply Principal Component Analysis (PCA) for dimension reduction, enhancing model efficiency. Experiment with advanced anomaly detection algorithms, including Isolation Forest, Local Outlier Factor (LOF), and Autoencoders. Employ cross-validation for robust model evaluation, utilizing metrics such as Precision, Recall, F1-Score, and ROC-AUC to measure performance effectively.

Advance your skills by delving into hyperparameter tuning using grid search or Bayesian optimization, fine-tuning models to optimize detection accuracy. Visualization plays a critical role: plot data distribution in reduced dimensions and visualize anomaly patterns for insightful interpretation. Craft a comprehensive report encapsulating your methodology, challenges faced, and insights gained. The report should detail the experimentation process, including code snippets, and explain the rationale behind your decisions. To complement the report, create a dynamic presentation that concisely communicates key findings, model performance, and visualization outcomes. Submission entails a detailed PDF report accompanied by a Jupyter Notebook (or equivalent) containing your code.