

National University of Computer and Emerging Sciences



Big data Project Proposal Fraud Detection System

Group Members:

Name	Roll Number	Section
Abdul Moiz	22L-7468	C
Nouveen Leghari	22L-7495	C

Department of Data Science
FAST-NU, Lahore, Pakistan

1. Objective

The main objective of this project is to develop a machine learning model that can accurately detect fraudulent transactions.

2. Project description

2.1. Data collection

Gather a dataset that includes both fraudulent and non-fraudulent transactions. This data will be used to train and test our machine learning model.

2.2. Data Exploration:

Conduct an initial exploration of the dataset to understand its structure, features, and distribution. This will help us identify any patterns or anomalies in the data.

2.3. Data Preprocessing:

Clean the data by handling missing values, scaling features, and encoding categorical variables. This step is crucial to ensure that our machine learning model can process the data effectively.

2.4. Model Development

Train various machine learning algorithms on the dataset. The goal is to identify the most suitable model for detecting fraudulent transactions.

2.5. Model evaluation

Assess the performance of each model using relevant evaluation metrics, such as accuracy, precision, recall, and F1-score. This will help us determine which model performs best at detecting fraud.

2.6. Model Optimization:

Explore techniques for model optimization, such as feature selection. This can help enhance the performance of our model.

2.7. Results Interpretation

Provide insights and interpretations derived from the model outputs. This will aid in understanding the effectiveness and limitations of the developed models.