**Group 7**

|  |  |
| --- | --- |
| Name | PRN |
| Manish Parate (TL) | 240344525016 |
| Mihir Dalvi | 240344525017 |
| Shoaib Sayed | 240344525035 |
| Sudarshan Sangale | 240344525039 |
| Vaishnavi Kapde | 240344525042 |
| Vedant Mahale | 240344525043 |
| Ishan Kulkarni (Mentor) |  |
| Akash Chandrashekharan (Mentor) |  |

**Project Topic** :- ML /AI (Machine Learning / Artificial Intelligence)

**Dataset** :- [IBM Transactions for Anti Money Laundering (AML)](https://www.kaggle.com/datasets/ealtman2019/ibm-transactions-for-anti-money-laundering-aml?resource=download)

**Purpose:** The project aims to develop and evaluate machine learning models to detect money laundering activities in financial transactions using synthetic data generated by IBM. This project seeks to address the high false positive and false negative rates in existing automated AML systems and create robust models that accurately identify laundering transactions.

**Insights from the Data: The project will extract insights such as:**

* Identification of money laundering patterns and behaviors.
* Evaluation of various machine learning algorithms' effectiveness.
* Analysis of the impact of different transaction volumes and illicit ratios on model performance.
* Insights into the entire money laundering cycle: placement, layering, and integration stages.
* Understanding how laundering transactions propagate through the financial system.

**Architecture:**

**Data Ingestion:**

* Load, preprocess, and clean synthetic transaction data.

**Feature Engineering:**

* Extract and transform relevant features, encode categorical variables, and create additional derived features.

**Model Development:**

* Implement and test machine learning models, including supervised (Decision Trees, Random Forests, Gradient Boosting, Neural Networks) and unsupervised (Clustering, Anomaly Detection) methods.

**Pattern Analysis:**

* Identify and analyze laundering patterns, comparing detected patterns with provided ones.

**Evaluation:**

* Assess model performance using precision, recall, F1 score, and AUC-ROC, and perform scalability testing.