## Project Title: Hidden Markov Model for Credit Fraud Detection

Project’s Github: [https://github.com/kunsergio117/CreditFraudDetectionHMM.git]

Team Profile

Team Members:

1. Sergio Gabriel Jiawei Kun

2. Joseph Lee Judkins - [Qualifications and Strengths]

Team Leader: Sergio Gabriel Jiawei Kun

### Project Description

#### 1. Introduction

Credit fraud is a significant and growing issue that impacts individuals and organizations globally. Current systems often struggle to keep up with sophisticated fraudulent activities. This project proposes to develop a system using Hidden Markov Models (HMM) to effectively detect credit fraud in transactions.

#### 2. Problem Diagnosis

The problem domain involves detecting fraudulent transactions in credit systems. The issues with current practices include:

- High False Positive Rates: Legitimate transactions are often flagged as fraudulent, leading to user dissatisfaction.

- Inability to Adapt: Current systems may not adapt quickly to evolving fraudulent techniques.

- Data Volume: Financial transactions generate massive datasets that are challenging to analyze in real-time.

\*Example scenarios of these issues might include interviews with financial institutions that discuss their current challenges with fraud detection.\*

#### 3. Proposed Treatment

To address the diagnosed problems, we propose the following interventions:

- Implementation of HMM that models the sequence of transactions to identify unusual patterns indicative of fraud.

- Transaction data will be ingested by program in the form of csv files (Real-time monitoring functionality can be added in the form of API calls or access to real-time databases).

- User-Friendly Dashboard for analysts to view flagged transactions and insights.

- [https://www.kaggle.com/datasets/mlg-ulb/creditcardfraud] : This Kaggle dataset will be used as our training and testing grounds for our HMM model.

**Metrics for success** will include:

Reduced false positives,

increased detection rates,

and improved analyst response times.

Example scenarios could illustrate how the HMM will flag a suspicious transaction for further investigation.

### Plan for the following weeks

#### 1. Initial Steps:

- Research existing HMM algorithms and their application to fraud detection. Python already has HMM libraries for this purpose. Knowledge of training and testing such a model is necessary.

- [https://www.kaggle.com/datasets/mlg-ulb/creditcardfraud] Testing of the model on this dataset will allow us to score its ability, additionally since it is stated that 492 frauds out of 284,807 transactions are contained allowing us to score it accurately.

- The dashboard, which will provide smoother user-interaction with the application, needs to be done, some front-end frameworks need to be explored such as Flask or Django, or even just a simple table file to be output.

#### 2. Functional Features:

- Transaction Analysis: Detect patterns in transaction sequences by the statistically trained model

- Reporting Tools: Generate reports showing detection rates, false positives, and other relevant metrics. This can be a another text file for debugging and optimization purposes.

Each team member will be responsible for specific functionalities, such as:

- Sergio Kun: Development of the transaction analysis module (for now this is assumed to be done in python as I am most familiar with statistical analysis with it.

- Lee Judkins: Implementation of the alerts generation mechanism (the front-end or user facing segments.

#### 3. Product Ownership:

Each team member’s contributions will be clearly defined to ensure accountability this can be clearly tracked by our github collaboration on the repository.

### Conclusion

This proposal aims to leverage the HMM to provide a robust solution to credit fraud detection. By thoroughly diagnosing the problem, prescribing a targeted treatment, and outlining a clear plan of work, we aim to create a beneficial system for financial institutions.