

**Introduction:**

The "Credit Card Fraud Detection" project aims to develop a machine learning model that can accurately identify fraudulent credit card transactions using two popular classification algorithms: Decision Tree and Support Vector Machine (SVM). The dataset used for training contains information about credit card transactions made by European cardholders in September 2013. The goal is to classify transactions as legitimate or fraudulent based on the provided features.

**Algorithms used:**

The Decision Tree algorithm is a powerful tool for classification tasks and can handle both numerical and categorical data. It works by creating a tree-like model where each internal node represents a "decision" based on a feature, and each leaf node represents the outcome (class label). The SVM algorithm, on the other hand, is a supervised learning algorithm used for classification and regression tasks. It works by finding the hyperplane that best separates the classes in the feature space. SVM is particularly useful when dealing with complex data and is effective in high-dimensional spaces.

**Significance:**

Credit card fraud is a significant issue that can cause financial losses for both cardholders and financial institutions. Detecting fraudulent transactions accurately is crucial for preventing fraud and maintaining trust in the financial system. Machine learning models offer a promising approach to identifying fraudulent activities by analyzing patterns and anomalies in transaction data.

**New in project:**

In this project, both Scikit-Learn and Snap ML Python APIs are used to model the classification task. Snap ML is chosen for its high-performance implementations of linear and tree-based models, offering efficient CPU/GPU implementations and novel ML algorithms with best-in-class accuracy. By comparing the performance of models trained using Scikit-Learn and Snap ML, the project aims to evaluate the effectiveness and efficiency of each library in detecting credit card fraud.

**Technique used:**

To prepare the dataset for modeling, basic data preprocessing tasks such as handling missing values, scaling features, and encoding categorical variables are performed. Both Decision Tree and SVM models are trained using Scikit-Learn

and Snap ML Python APIs. The trained models are then used to make predictions on new credit card transactions, and their performance is evaluated using appropriate evaluation metrics such as accuracy, precision, recall, and F1-score.

**Summary:**

Overall, the project aims to contribute to the development of effective fraud detection systems using machine learning techniques and to assess the capabilities of different libraries in implementing such systems.