**DETAILED PROJECT REPORT (DPR)**

**Fraud Transaction Detection**

Index

1 . Introduction

2. Architecture

3. Workflow

3.1 Data Ingestion and Storage

3.2 Data Preprocessing

3.3 Data Transformation

3.4 Model Building

3.5 Prediction

**1 . Introduction**

**1.1 Objective:**

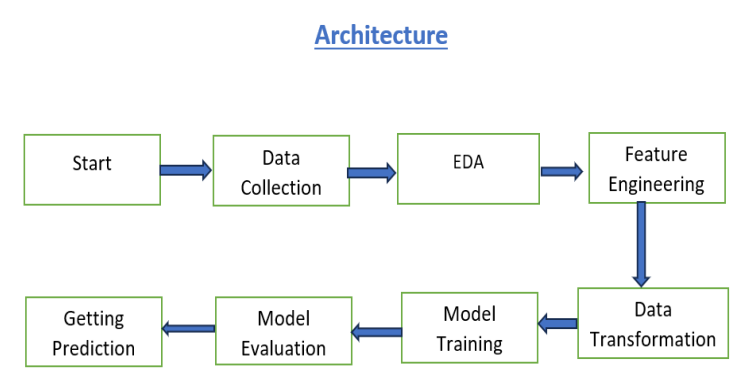
Development of a predictive model for predicting fraud transaction. The model will determine whether a customer’s transaction is placing a fraudulent or not.

**1.2 Abstract:**

In the contemporary digital era, online transactions have become a predominant mode of payment, with credit cards being one of the most widely used methods. However, this convenience is accompanied by a significant risk of fraudulent activities. Recently, there has been an observable increase in credit card fraud cases, underscoring the need for effective detection mechanisms.

This project seeks to address this issue by developing a web application for detecting fraudulent transactions using machine learning algorithms. Our dataset, containing 284,807 transactions from European countries over two days, includes 492 instances of fraud. The primary goal is to enable credit card companies to identify and prevent fraudulent activities, ensuring customers are not incorrectly charged.

**2. Architecture**



**3. Workflow**

**3.1 Data Ingestion and Storage**

1.1 Data Sources**:**

Data may come from various sources like databases, e-commerce platforms, or external sources. Ingest data from these sources. This is project of iNeuron so they provided the data for this project.

1.2 Data Storage:

We can store the data in a database or data warehouse. We can consider using technologies like PostgreSQL, MySQL, or cloud-based solutions like Amazon Redshift or Google BigQuery for scalability.

**3.2 Data Preprocessing:**

Implementing ETL (Extract, Transform, Load) processes to clean and preprocess the data. This

includes handling missing values, outliers, and transforming the data into a usable format.

* Exploratory Data Analysis (EDA):

Exploratory data analysis is one the most important stage of project. It can help identify obvious errors, as well as better understand patterns within the data, detect outliers or anomalous events, find interesting relations among the variables. Tools (python libraries) like Pandas, Matplotlib, or Seaborn are used to perform EDA.

* Feature Engineering:

Feature engineering is a crucial aspect of building an effective fraud transaction detection system. In the context of fraud detection, feature engineering involves selecting, transforming, and creating relevant features (variables) from raw data to improve the performance of machine learning models. These features help the model distinguish between legitimate and fraudulent transactions.

**3.3 Data Transformation :**

Explaining the data transformation phase involves detailing how raw data is processed and modified to make it suitable for consumption by downstream systems or applications.

**3.4 Model Building :**

* Model Training : Explaining the model building phase involves detailing how machine learning models are developed and trained to achieve specific objectives. In this project for fraud detection, logistic regression and decision tree algorithm is used for classification model.
* Model Validation : The model validation stage is a critical step in the machine learning lifecycle, and ML model is validated in this stage. This stage is designed to assess the performance and generalizability of a machine learning model.

**3.5 Prediction :**

The preprocessed and transformed features are input into the trained fraud detection model. The model applies its learned patterns and algorithms to generate a prediction or a score indicating the likelihood that the given transaction is fraudulent where 0 is a valid transaction and 1 is fraud transaction