PROJECT of

DEEP LEARNING

(CSE715)

Bachelor of Technology (CSE)

By

Gauri Rana (22000382)

Heer Patel (22000384)

Dhruvi Patel (22000402)

Zainab Khokhawala (22000425)

4th Year, Semester 7

Department of Computer Science and Engineering

School Engineering and Technology

Navrachana University, Vadodara

Autumn Semester

(July- Nov 2025)

1. **What about the Project?**

   The project “Credit Card Fraud Detection using Deep Learning” focuses on developing an intelligent system capable of identifying fraudulent credit card transactions in real time.

With the growing volume of online and digital payments, detecting fraudulent activities has become essential to ensure financial security.

This project applies deep learning techniques, such as Artificial Neural Networks (ANN) or Recurrent Neural Networks (RNN), to analyze transaction patterns and distinguish between legitimate and fraudulent behavior. The goal is to train a model that can automatically detect anomalies and flag suspicious transactions, minimize financial losses and protect users' data.

2. What are Data Sources?

The dataset used for this project is typically obtained from open-source repositories such as Kaggle, which provides anonymized credit card transaction data.

A popular dataset for this purpose is the "Credit Card Fraud Detection Dataset" available on Kaggle, containing European cardholder transactions made in September 2013.

Key details of the dataset include:

- Number of transactions: 284,807
- Number of fraudulent transactions: 492 (highly imbalanced dataset)
- Features: 30 attributes (including anonymized features V1 to V28, Amount, and Time)
- Label column: Class (1 = Fraud, 0 = Genuine)

The dataset is preprocessed to handle data imbalance, scaling, and normalization before training the deep learning model.

3. Useful Libraries

To build, train, and evaluate the fraud detection model, several **Python libraries** are used:

Category	Library	Purpose
Data Handling	pandas, numpy	For data manipulation and numerical operations

Visualization	matplotlib, seaborn	For plotting graphs and visualizing fraud vs. non-fraud patterns
Preprocessing	sklearn	For scaling, splitting datasets, and handling class imbalance.
Deep Learning	TensorFlow	For building and training neural network models.
Evaluation	sklearn.metrics	For generating confusion matrix, precision, recall, and F1score.

4. Measurable Goals and Success Metrics Type of Problem

This project is a Classification Problem.
The model's task is to classify each transaction as either:

- 0 → Legitimate Transaction, or
- 1 → Fraudulent Transaction

Since the output involves categorical labels, it falls under binary classification in machine learning and deep learning.

Measurable Goals

The project aims to achieve the following measurable and trackable goals:

1. Accuracy Goal:
   Achieve an overall model accuracy of at least 95% on the test dataset.
2. Precision and Recall Goals:
   a. Precision (Fraud class): ≥ 90% — to ensure that when the model predicts "fraud," it is usually correct.

b. Recall (Fraud class): ≥ 85% — to ensure the model catches most of the actual fraud cases.

3. F1-Score Goal:
   Attain an F1-score above 90% for the fraud class, balancing both precision and recall.
4. ROC-AUC Score Goal:
   Maintain an Area Under the ROC Curve (AUC) of ≥ 0.95, indicating strong separation between fraudulent and genuine transactions.
5. Real-Time Prediction Speed:
   The system should process and classify transactions in under 1 second per record, ensuring suitability for real-time fraud detection applications.

5. What Impact Your Project Will Make

This project has significant real-world impact in the financial and cybersecurity sectors. Some key benefits include:

- Enhanced Fraud Detection Accuracy: Deep learning models can identify complex patterns in transaction data that traditional rule-based systems often miss.
- Reduced Financial Losses: Early detection of fraudulent activity minimizes financial damage for banks and customers.
- Improved Customer Trust: A secure payment system builds customer confidence and ensures smoother digital transactions.
- Scalability: The model can adapt to new fraud patterns as it continues learning from updated transaction data.

6. What is the Workflow

The project follows a systematic workflow to ensure accuracy and efficiency:

a. Data Collection: Obtain the credit card transaction dataset from Kaggle or a similar trusted source.
b. Data Preprocessing:
    a. Handle missing values and outliers
    b. Normalize or standardize features
    c. Balance the dataset using SMOTE (Synthetic Minority Over-sampling Technique)
c. Data Splitting: Divide data into training and testing sets (e.g., 80% training, 20% testing).
d. Model Design: Build a Deep Neural Network (DNN) using Keras/TensorFlow with layers, activation functions, and dropout for regularization.
e. Model Training: Train the model on the dataset using Adam optimizer and binary crossentropy loss.
f. Model Evaluation: Evaluate performance using metrics like accuracy, precision, recall, F1-score, and AUC-ROC curve.
g. Prediction & Deployment: Deploy the trained model to detect fraud in real-time transaction streams.

7. What is the desired outcome?

The desired outcome of this project is to develop a **highly accurate and reliable deep learning-based fraud detection system** that can automatically identify fraudulent credit card transactions in real time.

The system should be able to:

- **Accurately classify** transactions as *fraudulent* or *legitimate* based on learned patterns.
- **Reduce false positives**, ensuring that genuine transactions are not incorrectly flagged as fraud.
- **Adapt to new fraud patterns** through continuous learning and retraining on updated data.

- **Provide real-time alerts** or notifications for suspicious activities.
- **Assist financial institutions** in minimizing monetary losses and improving transaction security.

Ultimately, the project aims to **enhance trust and safety** in digital financial systems by leveraging deep learning to detect fraud faster and more effectively than traditional rulebased approaches.