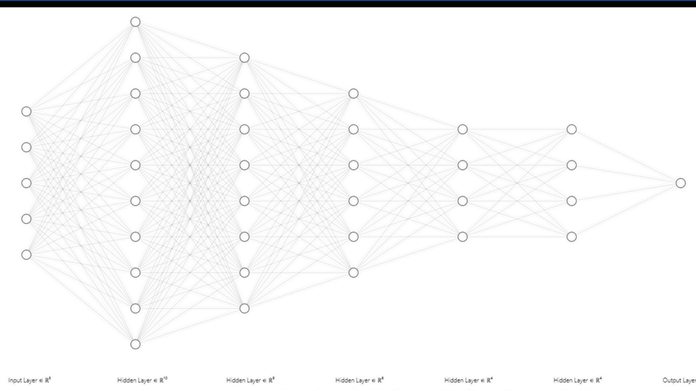
FRAUD DETECTION USING ML

Presented by
Abdellatif Itani
Abdullah Itani
Mohammed Omar Mkhallati

Project Statement & Background

Financial fraud and cybercrime are currently some of the most prevalent crimes in our digital day and age. There are always new and innovative ways used by criminals to manipulate their way into a person's bank account, whether it be through hacking, or through social engineering. This is a growing issue with cybercrime expecting to cost up to $10.5 trillion annually by 2025. Our goal through this project is to develop a system that will use AI and machine learning in order to properly label a transaction as fraud or not. We will use the classic available methods and improve upon them. We will also utilize the customers history of transactions in order to determine whether a transaction is fit with what is usually taking place or if it is an intruder.

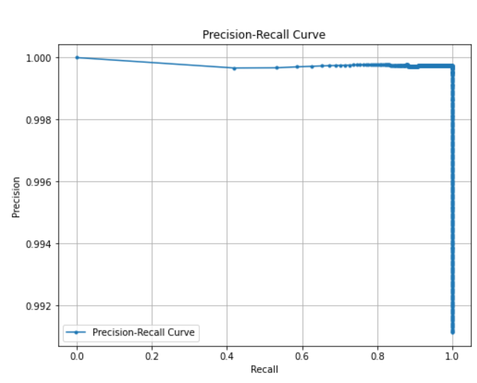
Key Methods

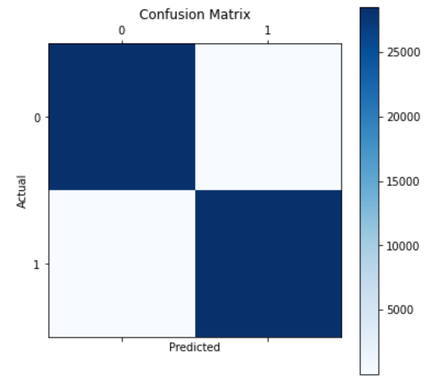


1. Dataset Preparation :
   - Source: A dataset with labels (0=Legitimate, 1=Fraud)
   - Processing: Remove missing values and shuffle the dataset
2. Neural Network Model
   - An input of 20 PCA-transformed features for anonymity. The hidden layers go from 512-256-128-64-32 neurons. We used Leaky ReLU and the output is 1 neuron with a sigmoid activation for binary classification.
3. Training & Evaluation
   - We used binary cross entropy loss with the Adam optimizer and a learning rate of 0.001.
4. Visualization
   - We displayed a confusion matrix and precision recall curve to show our results and our accuracy.

Results & Explanation

The credit card fraud detection model demonstrates excellent performance, with validation and test accuracies of 99.91% and 99.96%, respectively, and low losses. It handles the balanced dataset effectively, achieving a precision of 99.93%, recall of 99.98%, and F1-score of 99.96%. These metrics highlight the model's reliability in identifying fraudulent transactions accurately while minimizing false positives and false negatives, making it highly suitable for real-world application.






```
Validation Loss: 0.0057, Validation Accuracy: 0.9991
Test Loss: 0.0041, Test Accuracy: 0.9996
Number of 0s: 284307.0
Number of 1s: 284313.0
Precision: 0.9993
Recall: 0.9998
F1-Score: 0.9996
...
    accuracy                           1.00     56862
   macro avg       1.00      1.00      1.00     56862
weighted avg       1.00      1.00      1.00     56862
```

Future Direction

We aim to develop a user-friendly Web or Mobile Interface to enable real-time fraud detection, allowing users to upload transaction data, view risk scores, and generate detailed reports. Future enhancements include adapting the model to diverse datasets for better generalization and enabling continuous learning to stay updated with evolving fraud patterns, making the solution scalable and practical for real-world applications.

Conclusion

Our project highlights a highly effective system capable of identifying fraudulent financial transactions with impressive accuracy. By leveraging a neural network model with PCA-transformed features for anonymity, robust training methods, and precise evaluation metrics, the model achieves outstanding performance, including a near-perfect F1-score of 99.96%. This reliability ensures its suitability for practical applications.