Identification of Credit Card Fraud Based on Machine Learning and Deep Learning Approaches

Mehenaz Afrin and Priyanka Singla
Department of Computer Science
Boise State University
mehenazafrin@u.boisestate.edu priyankasingla@u.boisestate.edu

Abstract

Nowadays, digital transactions have become prevalent across the globe which as well raises alarming concerns in the security sector. Unauthorized transactions become frequent in the crowd of digital transactions. Identifying fraudulent credit card transactions is crucial to minimize monetary losses and protect consumers' trust in the financial sector. Therefore, this study focuses on developing an efficient and accurate model to identify fraudulent transactions using machine learning techniques and ensemble techniques (combination of Artificial Neural Network (ANN with base classifiers). By analyzing a dataset containing labeled transactions, this approach applies data preprocessing (standardization technique), and algorithmic optimization to differentiate between legitimate and fraudulent activities. We have done a comparative analysis among all the classifiers for fraud detection. To validate the study result, a cross-validation technique has been applied. Among all the classifiers, Random Forest classifiers show the highest performance level by achieving 95% precision, 90% recall 92% f1 score, and 99% accuracy. In the comparison of base classifiers and their ensemble techniques, we found ANN with Gaussian Naïve Bayes (GNB) performs better than usual GNB.

1 Introduction

In today's digital age, where financial transactions have become a cornerstone of daily life, ensuring the security of these transactions has never been more critical (Chaudhary et al., 2012). Credit card fraud detection is a pressing challenge faced by financial institutions worldwide, as fraudulent activities can lead to significant monetary losses and erode trust in digital payment systems (Raj and Portia, 2011). The need for efficient and accurate fraud detection systems is paramount to safeguard both consumers and organizations. This project focuses on building an intelligent system for credit card fraud detection using machine learning and deep learning techniques. The dataset employed comprises anonymized credit card transactions, which include features such as transaction amount, time, and a binary label distinguishing fraudulent from legitimate transactions. A notable challenge of this dataset is its highly imbalanced nature, where legitimate ones significantly outnumber fraudulent transactions. To address this, we implemented a comprehensive approach that combines traditional machine learning models, such as Random Forest, Logistic Regression, K-Nearest Neighbors, and Decision Trees, with the power of artificial neural networks (ANN). By integrating the probabilistic outputs of these classifiers as input features for the ANN, we aim to leverage the strengths of both methodologies. We sincerely appreciate Jerome H. Friedman, Trevor Hastie, and Robert Tibshirani, authors (Friedman, 2009) of The Elements of Statistical Learning: Data Mining, Inference, and Prediction. Their seminal work has provided a solid theoretical foundation for understanding and implementing advanced statistical and machine-learning models. In particular, their discussions on ensemble methods and model optimization significantly informed our use of the Random Forest algorithm in this project for detecting credit card fraud. Preprocessing steps such as data normalization, feature engineering, and techniques to mitigate class imbalance are carefully incorporated into the pipeline of our project. This project not only demonstrates the application of hybrid modeling techniques to real-world problems but also evaluates their effectiveness using metrics such as accuracy, precision, recall, and F1-score. Through this work, we aim to contribute to the development of robust, scalable, and accurate solutions for credit card fraud detection in today's fast-evolving financial landscape. Specifically, in this project, we aim to explore, manipulate, and evaluate the credit card dataset for investigating authorized or unauthorized. we did an extensive analysis of every classifier in

the detection system meticulously. Overall, Our contributions can be summarized as follows:

- Analyze the characteristics of the dataset through Countplot, Kernal Density Estimation(KDE).
- Employ v normalization techniques to preprocess the data
- Several machine learning approaches along with ensemble techniques(ANN with base classifiers) have been applied for detecting credit card fraud or non-fraud transactions precisely.

2 Methodology

We employed Random Forest, Gaussian Naive Bayes (GNB), Artificial Neural Networks (ANN), Support Vector Machine(SVM), Decision Tree(DT),K nearest neighbor (KNN) and Logistic Regression(LR) in this analysis due to their ability to deliver accurate results and reliable performance. Below, we provide some explanation of why each algorithm was selected, along with relevant references and justifications for their use.

1) Random Forest Algorithm: Random Forest is used for its robustness, ability to handle large datasets, and its effectiveness in improving accuracy by combining multiple decision trees to reduce overfitting (Segal, 2004).

2) Logistic Regression: Logistic Regression used for its simplicity, efficiency, and ability to model the probability of binary outcomes, making it ideal for classification tasks. This reference helped in applying logistic regression to classify transactions as fraud or no fraud with simplicity and interpretability (Hosmer Jr et al., 2013).

3) K-Nearest Neighbour: KNN is used for its simplicity and effectiveness in classification tasks, as it classifies data points based on the majority vote of their nearest neighbors, making it suitable for both small and large datasets.This reference guided our application of KNN for transaction classification, helping to determine the most likely outcome based on surrounding data points (Cover and Hart, 1967).

4)Decision Tree: DT is used for its interpretability and ability to handle both numerical and categorical data, making it effective for classification tasks and providing clear decision-making pathwaysIt also highlighted the importance of pruning to prevent overfitting, which we applied when using decision trees to classify transactions effectively (Breiman, 2017).

5) Aritificial Neural Networks: ANN are used for their ability to model complex, non-linear relationships in data, making them highly effective for tasks requiring advanced pattern recognition and predictive accuracy. The concept of Artificial Neural Networks (ANN) stems from the work of Rumelhart et al., who introduced a method for training these networks to learn complex patterns effectively (Rumelhart et al., 1986).

6) Gaussian Naive Bayes (GNB):GNB was used for its simplicity and efficiency in probabilistic classification, especially when dealing with high-dimensional data.his method assumes that the features follow a Gaussian (normal) distribution, which often results in strong performance even with limited data. This reference provided valuable insights into optimizing the speed and scalability of GNB, which we applied to our analysis to ensure efficient handling of large volumes of data in classification tasks (Ontivero-Ortega et al., 2017).

7) Support Vector Machine (SVM): Support Vector Machine(SVM) was used for its ability to efficiently handle high-dimensional spaces and perform well with both linear and non-linear classification tasks. SVM works by finding the optimal hyperplane that maximizes the margin between different classes, making it particularly effective in separating complex data (Cortes, 1995).

In this section, we have done data exploration, data preprocessing and finally data evaluation with several machine learning and deep learning algorithms. The detailed analysis is discussed below.

2.1 Data Exploration

We have used the existing dataset from Kaggle regarding fraudulent and non-fraudulent transactions. The dataset is enriched with credit card-related information containing 31 columns and 284807 rows. As the dataset is huge, it consumes 67.4 MB of memory. With the help of the describe function, we measure the summary of statistical calculations for numerical columns in this dataset. Later we checked the null values and found none. By using the count plot() (Waskom et al., 2015) function in Figure (1) we understand the distribution of fraud (positive class) versus non-fraud (negative class) transactions in this dataset.

We have not performed any outlier treatment since the outliers are already taken care of by trans-

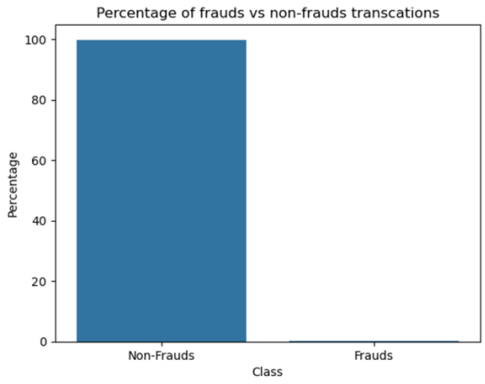

Figure 1: Percentage of fraud vs non-fraud transactions

forming the data (Kurita, 2019). They were already PCA transformed.

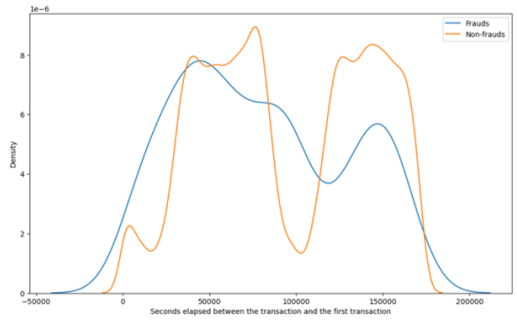

Figure 2: Elapsed between the transaction and the first transaction in seconds

The above Figure 2 demonstrates the distribution of transaction times for fraudulent and non-fraudulent transactions, measured as the number of seconds elapsed between a transaction and the first transaction. It uses kernel density estimation (KDE) to represent the density of these occurrences. blue curve: non-frudulent transactions, orange curve: fraudulent transactions.

2.2 Data Manipulation

In this section, we try to manipulate the dataset according to its nature. After properly understanding the characteristics of the dataset, we have utilized data standardization techniques for scaling down the features in a particular range thus it will help the model to converge quickly. It transforms all the features into one unit form so the machine can learn so easily (Afrin et al., 2022). we have used cross-validation technique for removing the imbalance of the dataset and feed each part of the data in the model.

2.3 Data Evaluation

For accomplishing our project, we have initially used machine learning classifiers such as Random Forest (RF), Decision tree (DT), K nearest neighbor (KNN), Gaussian naïve bayes (GNB), Logistic regression (LR) classifier to predict the accuracy of both transactions from the credit card dataset. Later we have employed Artificial Neural Network with each of the above base classifiers including Support Vector Machine (SVM). We have applied four evaluation matrix such as accuracy, precision, recall, and f1 score. We have also used a cross-validation (Berrar et al., 2019) technique as we learned that our dataset is imbalanced from the exploration part. This technique will validate our result as it will take each portion of the dataset in the training and testing part.

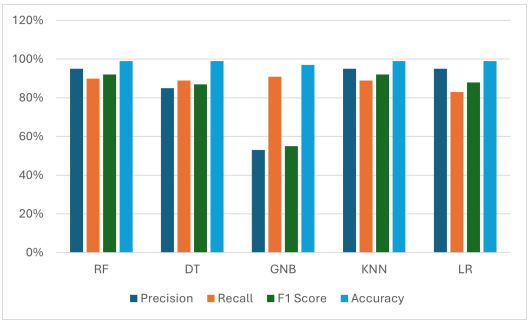

Figure 3: Performance level of machine learning classifiers in terms of precision, recall, f1 score and accuracy

In the fraud detection system, among all the base classifiers, random forest classifier outperformed. It reduces overfitting by averaging the predictions of many decision trees, each trained on a random subset of data and features (Zakariah et al., 2014). RF gives the highest performance levels 95%,90%, and 92% respectively in precision, recall, and f1 scores. We found KNN is our second efficient classifier giving us 95% precision, 89% recall, and 92% f1 score. KNN thrives in scenarios where the proximity of data points in this feature space reflects their similarity (Reddy and Kanimozhi, 2022) (e.g., fraudulent vs. legitimate transactions). KNN classifies a point based on the majority class of its neighbors, making it well-suited for identifying localized patterns of fraud. Other than GNB, all the base classifiers give a 99% accuracy level. GNB offers us the lowest precision (53%) and f1 score (55%).

For the implementation of hybrid approaches in our project we follow the criteria:

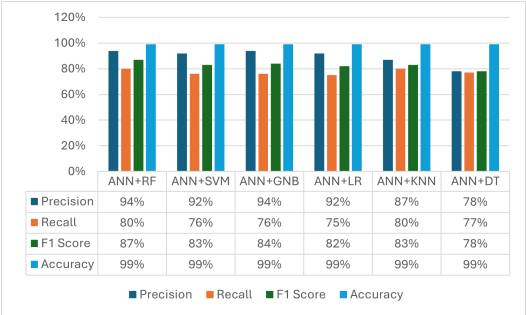

Figure 4: Performance level of hybrid classifiers in terms of precision, recall, f1 score and accuracy

- Initially we train the base classifiers.
- Then combine the original features with base classifier probabilities for the ANN.
- Train the ANN with the input layer.
- Get predictions from ANN and combine them with the base classifiers for the ensemble.
- Convert probabilities to binary predictions.
- Calculate metrics for this fold.
- Calculate average scores across all folds

With the combination of Artificial Neural Network with the classifiers, after setting epochs: 10 and batch size 32 in Figure 4, We got 94%, 80%, and 87% performance levels. Strangely the combination somewhat decreases the performance level. We are assuming that it happens because of the dataset characteristics. Accuracy is identical to the usual classifiers, but we found discrepancies in other evaluation matrices.

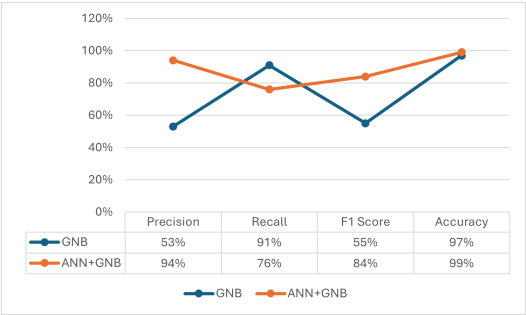

Figure 5: Comparison between GNB and ANN+GNB

Ensemble method here improves the performance level compared to the usual GNB. From figure 5 we can see, it increases the accuracy level by 2% as GNB makes strong assumptions (e.g., independence and Gaussian distribution of features). ANN's feature learning can transform data into a space where these assumptions hold better, improving GNB's performance (Chandra et al., 2024). ANN's output simplifies complex patterns into a more structured format that aligns with GNB's probabilistic nature, leading to better synergy (Kedia and Bhushan, 2022). Also, we get better precision and f1 score.

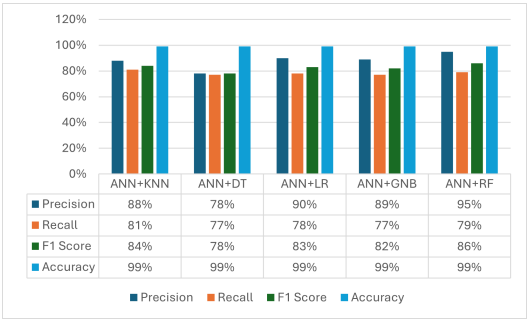

Figure 6: Performance level of hybrid classifiers in terms of precision, recall, f1 score, and accuracy with different settings.

After setting different epochs: 20 and batch size: 64, we again analyze the performance level of ensemble methods to see any differences. With this setting, ANN+RF gives better results by increasing 1% precision score and 2% f1 score than the previous result. This setting also helps to increase the precision and f1 score of the KNN and ANN combination by 1%. With all the combinations we get a 99% accuracy level. While the setup and ensemble techniques help to higher some evaluation matrices, it decrease the Recall score than the previous setup.

3 Conclusion

For securing digital banking, we focus on detecting unauthorized or legitimate transactions. Detecting unauthorized movement earlier will prevent the loss of earned income. In this study, We introduce 6 machine learning base classifiers and 7 combinations as in Artificial intelligence with different base classifiers for identifying the exact category of transactions. Among all the combinations RF most accurately detects the fraud transaction with 99% accuracy, 95% precision, 90% recall, and 92% f1 score. The combination of ANN with other base classifiers was supposed to perform better than the usual classifiers. ANN with GNB gives us better performance than GNB. We will consider applying SMOTE for data preprocessing and weight initialization techniques for assigning weight properly in every layer in our further research to enhance the performance level, especially the recall score.

A APPENDIX

References

Mehenaz Afrin, Salma Akter Asma, Nazneen Akhter, Jaheed Hasan Ridoy, Sazida Sharmila Sauda, and Kazi Abu Taher. 2022. A hybrid approach to investigate anti-pattern from source code. In *2022 25th International Conference on Computer and Information Technology (ICCIT)*, pages 888–892. IEEE.

Daniel Berrar et al. 2019. Cross-validation.

Leo Breiman. 2017. *Classification and regression trees*. Routledge.

T Bharath Chandra, A Sujan Reddy, A Adarsh, MA Jabbar, and BN Jyothi. 2024. Diabetes prediction using gaussian naive bayes and artificial neural network. In *2024 International Conference on Distributed Computing and Optimization Techniques (ICDCOT)*, pages 1–5. IEEE.

Khyati Chaudhary, Jyoti Yadav, and Bhawna Mallick. 2012. A review of fraud detection techniques: Credit card. *International Journal of Computer Applications*, 45(1):39–44.

Corinna Cortes. 1995. Support-vector networks. *Machine Learning*.

Thomas Cover and Peter Hart. 1967. Nearest neighbor pattern classification. *IEEE transactions on information theory*, 13(1):21–27.

Jerome Friedman. 2009. The elements of statistical learning: Data mining, inference, and prediction. *(No Title)*.

David W Hosmer Jr, Stanley Lemeshow, and Rodney X Sturdivant. 2013. *Applied logistic regression*. John Wiley & Sons.

Shyam Kedia and Megha Bhushan. 2022. Prediction of mortality from heart failure using machine learning. In *2022 2nd International Conference on Emerging Frontiers in Electrical and Electronic Technologies (ICEFEET)*, pages 1–6. IEEE.

Takio Kurita. 2019. Principal component analysis (pca). *Computer vision: a reference guide*, pages 1–4.

Marlis Ontivero-Ortega, Agustin Lage-Castellanos, Giancarlo Valente, Rainer Goebel, and Mitchell Valdes-Sosa. 2017. Fast gaussian naïve bayes for searchlight classification analysis. *Neuroimage*, 163:471–479.

S Benson Edwin Raj and A Annie Portia. 2011. Analysis on credit card fraud detection methods. In *2011 International Conference on Computer, Communication and Electrical Technology (ICCCET)*, pages 152–156. IEEE.

Kummita Sravan Kumar Reddy and KV Kanimozhi. 2022. Novel intelligent model for heart disease prediction using dynamic knn (dknn) with improved accuracy over svm. In *2022 international conference on business analytics for technology and security (ICBATS)*, pages 1–5. IEEE.

David E Rumelhart, Geoffrey E Hinton, and Ronald J Williams. 1986. Learning representations by back-propagating errors. *nature*, 323(6088):533–536.

Mark R Segal. 2004. Machine learning benchmarks and random forest regression.

Michael Waskom, Olga Botvinnik, Paul Hobson, Jordi Warmenhoven, John B Cole, Yaroslav Halchenko, Jake Vanderplas, Stephan Hoyer, Santi Villalba, Eric Quintero, et al. 2015. seaborn: v0. 6.0 (june 2015). *Zenodo*.

Mohammed Zakariah et al. 2014. Classification of large datasets using random forest algorithm in various applications: Survey. *International Journal of Engineering and Innovative Technology (IJJEIT)*, 4(3).