Automated machine learning-based credit card fraud detection

Report submitted to the SASTRA Deemed to be University as the requirement for the course

INT424: Algorithmic Trading

Submitted by

NITISH K S
(Reg.no 125150036)

May 2024

SCHOOL OF ARTS SCIENCE AND HUMANITIES

THANJAVUR, TAMIL NADU, INDIA – 613 401

SCHOOL OF ARTS, SCIENCES, HUMANITIES AND EDUCATION

THANJAVUR – 613 401

Bonafide Certificate

This is to certify that the report titled “**Automated machine learning-based credit card fraud detection**” submitted as a requirement for the course, **INT424: ALGORITHMIC TRADING** for M.Sc. Data Science programme, is a bona fide record of the work done by **Mr.NITISH K S(Reg. No.125150036**) during the academic year 2023-24, in the School of Arts, Sciences, Humanities and Education, under my supervision.

Signature of Project Supervisor :

Name with Affiliation : Ashok Palaniappan

Date : 01/05/2024

Project *Viva voc*e held on _01-May-2024

Examiner 1 **Examiner 2**

SCHOOL OF ARTS, SCIENCES, HUMANITIES AND EDUCATION

THANJAVUR – 613 401

Declaration

I declare that the report titled "**Automated machine learning-based credit card fraud detection**" submitted by me/us is an original work done by me/us under the guidance of **Dr Ashok Palaniappan, Associate Professor, School of Chemical and Biotechnology, SASTRA Deemed to be University** during the third semester of the academic year 2023-24, in the **School of Arts Science, Humanities and Education**. The work is original and wherever I have used materials from other sources, I have given due credit and cited them in the text of the report. This report has not formed the basis for the award of any degree, diploma, associate-ship, fellowship or other similar title to any candidate of any University.

Signature of the candidate(s) : K.S. Nitish

Name of the candidate(s) : **NITISH K S**

Date : **11.07.2022**

Acknowledgements

I would like to express my sincere gratitude to everyone who contributed to the completion of this individual project. Firstly, I am deeply thankful to Dr Ashok Palaniappan, Associate Professor, School of Chemical and Biotechnology, SASTRA Deemed to be University for their guidance, support, and invaluable feedback throughout the project. Their expertise and encouragement were instrumental in shaping the project's direction and ensuring its success.

I would also like to thank SASTRA Deemed to be University for accepting my project in my desired field of expertise.

I would also like to thank all of the other supporting personnel who assisted me by supplying the equipment that was essential and vital, without which I would not have been able to perform efficiently on this project.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABSTRACT

Credit card fraud has experienced an alarming surge in recent times as a larger portion of the population opts for credit card transactions, propelled by advancements in technology and the widespread adoption of online commerce. This uptick has led to substantial financial losses attributed to fraudulent activities. To counteract such losses, it becomes imperative to devise and implement an efficient system for fraud detection. While machine learning techniques are frequently utilized for automating fraud detection, they often neglect to account for deceptive strategies or behavioral nuances that could trigger alerts. This research endeavors to confront this challenge by introducing a novel approach: a Long Short-Term Memory Network (LSTM) tailored specifically for identifying credit card fraud. Moreover, an attention mechanism has been seamlessly integrated to bolster performance further. In arenas like fraud detection, where sequential data exhibit intricate interconnections, models structured in this manner have exhibited significant efficacy. The efficacy of the LSTM model is pitted against alternative classifiers including Naive Bayes, Gated Recurrent Unit (GRU), and Artificial Neural Network (ANN). Empirical findings demonstrate that our proposed model delivers robust results with a commendable level of accuracy.

CHAPTER 1

INTRODUCTION

1.1 Algorithmic Trading

Algorithmic trading is the process of executing orders using automated pre-programmed trading instructions that account for variables such as time, price, and volume. This style of trading seeks to take advantage of computers' greater speed and processing power than human traders. In the twenty-first century, algorithmic trading has gained popularity among both individual and institutional traders. According to a 2019 study, trading algorithms, rather than humans, performed approximately 92% of all Forex trading.

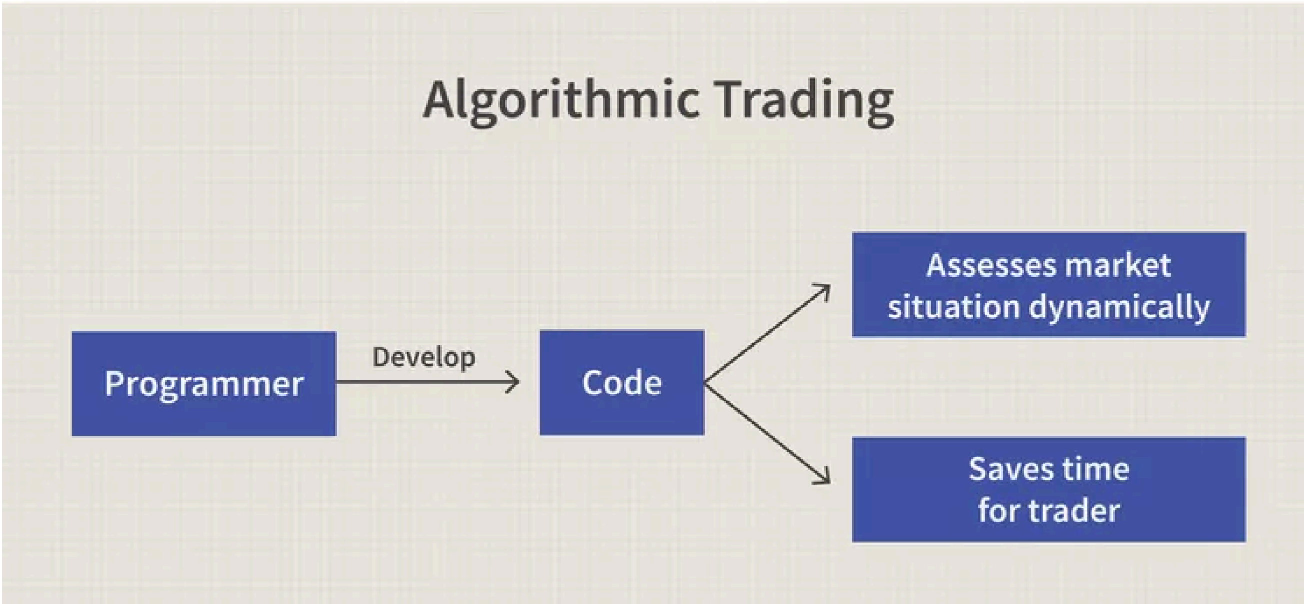

Fig 1.1: Workflow of Algorithmic Trading

1.2 About the Credit Card Fraud Detection

The challenge is to recognize fraudulent credit card transactions so that the customers of credit card companies are not charged for items that they did not purchase.

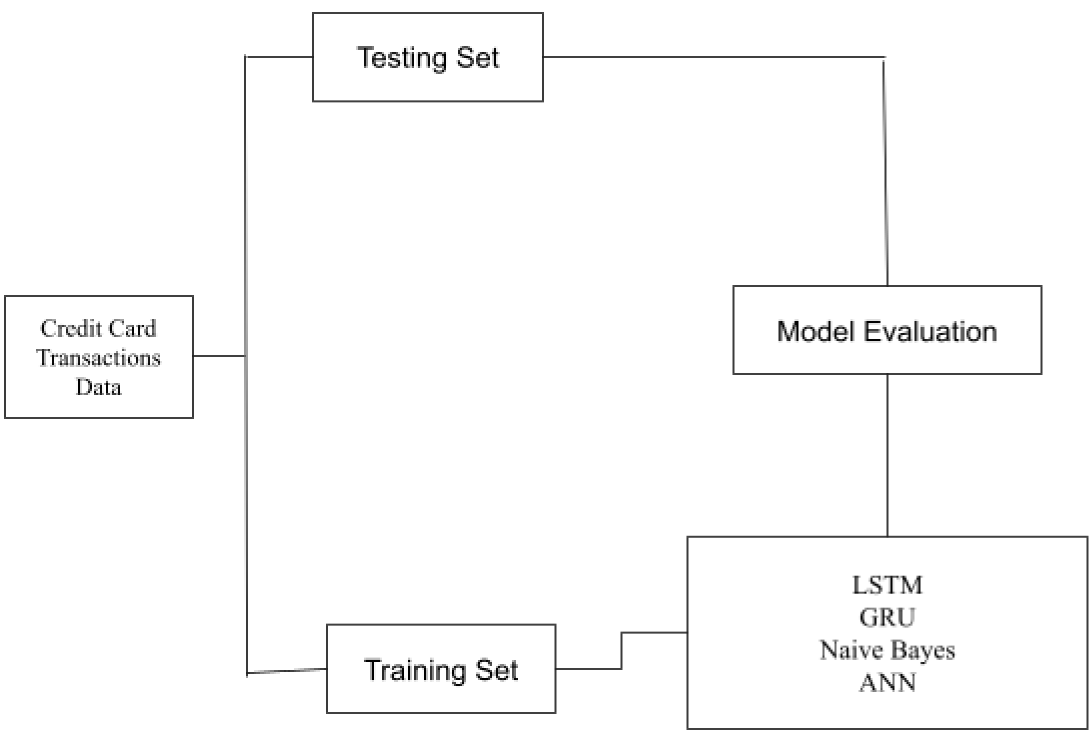

Fig 1.2: Credit Card Fraud Detection

1.3 Difference between Credit Card Fraud Detection and prevention

Aspect	Fraud Detection	Fraud Prevention
Objective	Identifying suspicious or fraudulent activities in real-time or post-transaction.	Implementing measures to proactively stop fraudulent activities.
Methods	Relies on algorithms, machine learning, anomaly detection, pattern recognition.	Uses security protocols, authentication methods, and fraud prevention tools.
Focus	Identifies potentially fraudulent transactions for further investigation.	Aims to stop fraudsters from carrying out their schemes or mitigate their impact.
Timing	Occurs during or after the transaction process.	Implemented before or during the transaction process.
Examples	Anomaly detection, Pattern recognition, Predictive modeling.	Two-factor authentication, Address verification systems (AVS), - Card verification value (CVV) checks, Tokenization.

Table 1.1 Credit Card Fraud Detection and Prevention

1.4 Motivation

Automated machine learning-based credit card fraud detection is driven by the urgent need to counter the rising threat of credit card fraud in today's digital landscape. With the surge in online transactions and the vulnerabilities accompanying technological progress, combating credit card fraud has become imperative, given its widespread impact on individuals and organizations. The paper's primary aim is to develop a proactive and efficient solution for detecting and preventing fraudulent activities in credit card transactions. Its motivations encompass several key objectives. Firstly, it seeks to bolster financial security by implementing an effective fraud detection system to safeguard consumers from potential financial losses. Secondly, the paper advocates for leveraging machine learning techniques to enable real-time detection of fraudulent transactions, offering a more sophisticated and automated approach compared to traditional methods. Thirdly, it aims to minimize the financial impact of credit card fraud on consumers and financial institutions alike, by accurately identifying and preventing fraudulent transactions, thus reducing associated costs. Lastly, the research endeavors to contribute to both academic knowledge and practical applications by proposing a novel approach that combines machine learning algorithms to address the challenges posed by fraudulent activities in the digital payment ecosystem. In essence, the motivation behind this paper lies in enhancing financial security, leveraging advanced technologies for fraud detection, minimizing financial losses, and advancing the field of credit card fraud detection through innovative research and practical solutions.

CHAPTER 2

LITERATURE SURVEY

The literature analysis of the research paper "Automated machine learning-based credit card fraud detection" provides a thorough examination of existing studies and approaches for credit card fraud detection. It makes numerous important revelations: Previously conducted inquiries have examined data mining techniques with the goal of detecting trends and anomalies suggestive of credit card fraud. Second, the use of various algorithms for fraud detection, including Naive Bayes, Gated Recurrent Unit (GRU), Artificial Neural Network (ANN), and Long Short-Term Memory-Recurrent Neural Network (LSTM-RNN), highlights the rise of machine learning as a significant tool in this field. Third, difficulties like imbalanced datasets, possibly misleading statistics, and the dynamic nature of fraud tendencies are highlighted, as well as solutions to them. Additionally, a variety of approaches, including supervised, unsupervised, and hybrid algorithms, have been utilized in earlier research projects. Examples of these are Gaussian Mixture Models, Risk-Based Ensemble models, and Logistic Regression (LR). Finally, the significance of assessment metrics such as the Matthews Correlation Coefficient (MCC) is underscored in determining how successful fraud detection strategies are. Overall, the literature review offers a thorough understanding of the state of research on credit card fraud detection, highlighting the importance of data mining and machine learning as well as the complex difficulties involved in stopping fraudulent activity in financial transactions.

CHAPTER 3

PROPOSED METHODS

The proposed method in the research paper "Automated machine learning-based credit card fraud detection" focuses on utilizing a Long Short-Term Memory-Recurrent Neural Network (LSTM-RNN) with an attention mechanism for detecting credit card fraud.

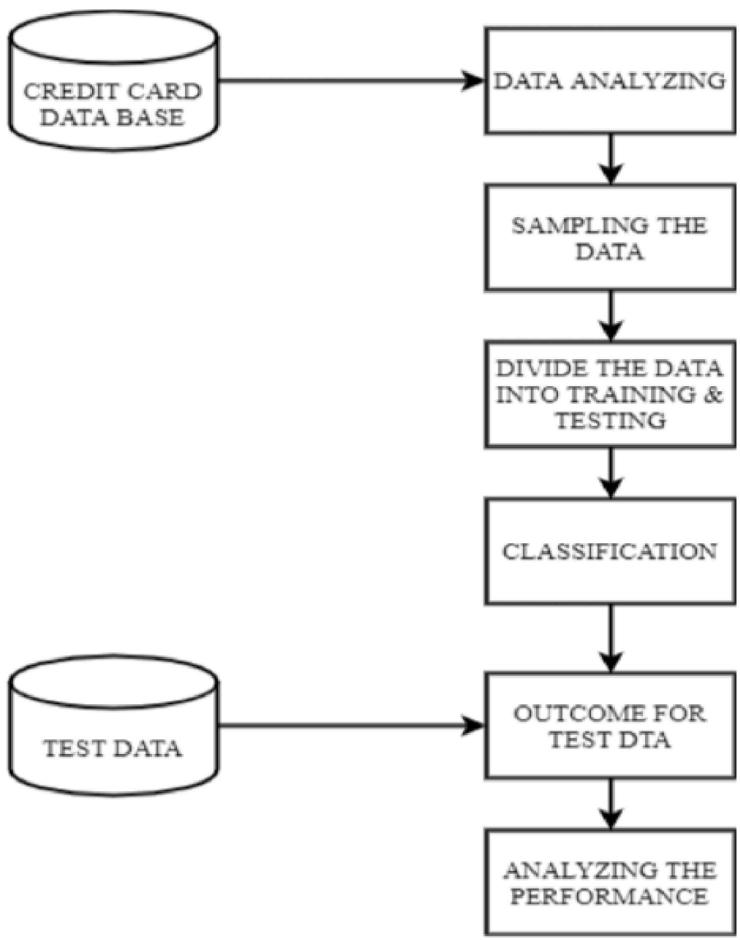

Fig 3.1: Proposed Methods

3.1 About the Data

The credit card transactions done by European cardholders in 2023 are included in this dataset. It has more than 550,000 records in total, and to safeguard the identity of the cardholders, the data has been anonymized. This dataset's main goal is to make it easier to create models and algorithms for fraud detection that may be used to spot possibly fraudulent transactions. The data download from kaggle.

	id	v1	v2	v3	v4	v5	v6	v7	v8	v9	...	v21	v22	v23	v24	v25	v26	v27	v28	amount	class
444612	444612	-0.135668	0.143539	-0.125757	0.149152	-0.112737	-0.293202	-0.134130	0.052249	-0.368581	...	0.114087	0.065991	-0.077962	-0.004476	0.128249	-0.236378	0.297167	0.501526	19303.20	1
434081	434081	-0.940288	0.576981	-0.784847	0.726264	-0.818733	-0.792234	-0.728505	0.466670	-0.812064	...	0.327106	0.053817	-0.344381	0.302694	0.454806	-0.418135	-0.005106	1.473079	857.54	1
128732	128732	-0.095787	-0.084154	0.868350	-0.156671	0.349952	0.756563	0.553955	-0.018983	0.247897	...	-0.112324	0.060142	-0.158643	-1.257074	0.017015	-0.389956	0.119398	0.200436	9136.47	0
366633	366633	-0.395590	-0.024874	0.211438	-0.171183	-0.145649	-0.004703	-0.109203	-0.016461	-0.009131	...	0.164014	0.638978	-0.171162	-0.004704	0.368778	-0.897788	-0.580749	-0.369836	16027.90	1
445479	445479	0.687725	-0.276342	0.732660	0.068296	0.352411	0.626679	0.445431	-0.111812	0.197903	...	-0.182238	-0.417094	0.053227	0.387801	0.075094	0.011084	-0.188125	-0.074410	20506.68	1
492411	492411	-1.850524	1.495123	-1.914219	2.242035	-1.358952	-0.881976	-1.942857	-0.283099	-2.135017	...	-1.169405	0.716528	-0.154954	-1.619172	1.272098	1.610659	3.246000	-3.061320	15410.75	1
160152	160152	1.831652	-0.519512	-0.038740	-0.549715	0.276608	-0.056017	0.429587	-0.183764	1.303539	...	-0.014542	0.601193	-0.001515	1.132825	0.395725	-1.062779	-0.252800	-0.231117	12744.70	0
509713	509713	-1.061050	0.721177	-1.152321	1.311062	-0.746202	-1.068439	-0.960112	0.784263	-1.297715	...	0.332041	-0.224077	0.128016	-0.746153	0.120265	1.569655	0.520760	-1.404379	2228.37	1
183112	183112	-0.262897	-0.905259	0.229759	-0.453147	0.829374	-0.684549	0.829007	-0.274951	0.241411	...	-0.038528	0.691234	0.815897	0.206791	-0.280736	2.328956	-0.149595	0.408187	21101.74	0
168022	168022	1.667807	-0.846371	-0.368812	-0.999391	0.563910	0.220738	0.663676	-0.239833	-0.140738	...	-0.113227	-0.011794	-0.376290	-0.495235	0.812466	2.010501	-0.429119	-0.281758	4316.39	0

Fig 3.2: Dataset sample

3.2 About the Model Used

3.2.1 Long short-term memory (LSTM)

Long Short-Term Memory (LSTM), a recurrent neural network (RNN) version, explicitly addresses the vanishing gradient problem that is common in standard RNN architectures. Its unique benefit over hidden Markov models, conventional RNNs, and other sequence learning techniques is its ability to retain sensitivity throughout a range of gap lengths. With its ability to provide a short-term memory mechanism in RNNs that lasts for thousands of timesteps, LSTM is well named as "long short-term memory."

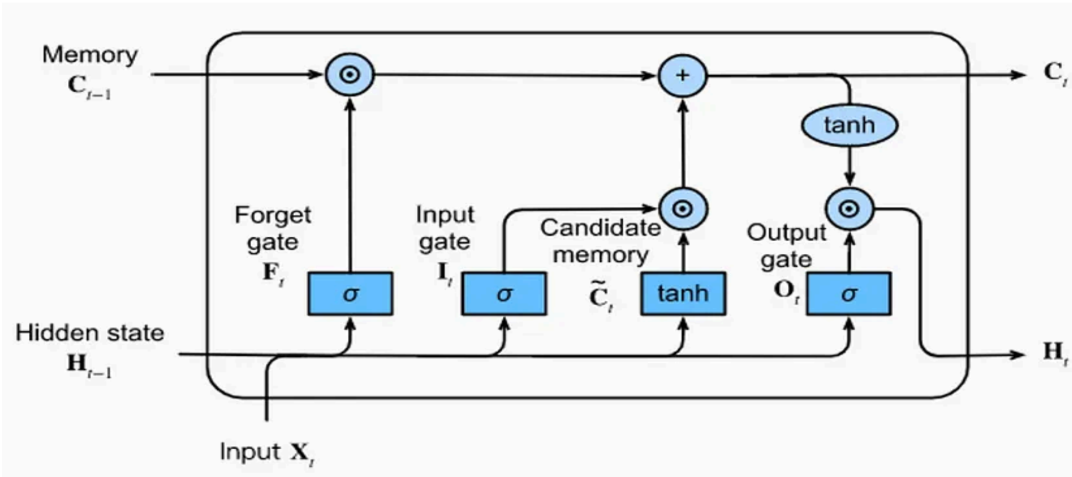

Fig 3.3: Long Short-Term Memory (LSTM)

3.2.2 Gated Recurrent Unit (GRU)

Gated Recurrent Unit (GRU) is a recurrent neural network (RNN) architecture akin to LSTM (Long Short-Term Memory), devised for modeling sequential data while enabling selective retention or discard of information over time. Although GRU shares LSTM's objective, it boasts a simpler structure with fewer parameters, rendering it potentially easier to train and computationally more efficient. The primary distinction between GRU and LSTM lies in their treatment of the memory cell state. While LSTM maintains the memory cell state separately from the hidden state, updated through three gates (input, output, and forget), GRU substitutes the memory cell state with a "candidate activation vector," updated via two gates (reset and update).

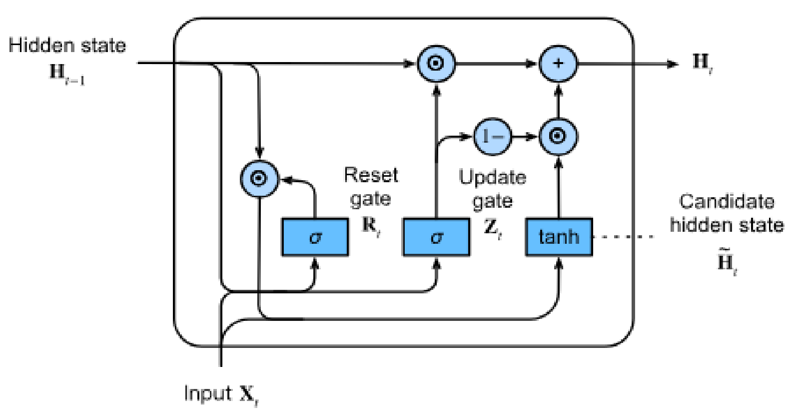

Fig 3.4: Gated Recurrent Unit (GRU)

3.2.3 Artificial Neural Networks (ANN)

An Artificial Neural Network (ANN) represents a machine learning paradigm inspired by the intricate neural architecture of the human brain. It consists of interconnected nodes, or neurons, arranged into layers, through which data traverses, dynamically adjusting connection weights to discern patterns and formulate predictions.

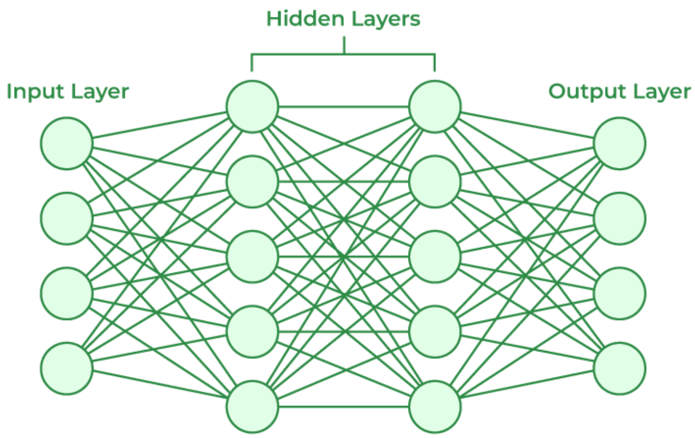

Fig 3.5: Artificial Neural Network (ANN)

3.2.4 Naive Bayes (NB)

Naive Bayes classification operates on Bayes' Theorem, relying on an independence presumption between predictors. Essentially, it assumes that within a class, the existence of one feature is independent of the existence of any other feature.

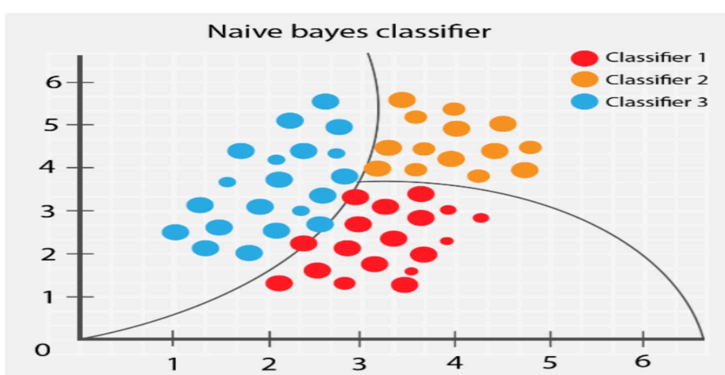

Fig 3.6: Naive Bayes (NB)

3.3 Data Understanding

In the Data Understanding phase of the project, particular emphasis is placed on exploratory data analysis (EDA), scrutinizing the distribution patterns of critical features, as well as the segmentation and scaling of data. Moreover, the division of data into subsets for training and testing purposes is a pivotal aspect of this phase. Additionally, the utilization of SMOTE (Synthetic Minority Over-sampling Technique) to address class imbalance, while typically conducted during preprocessing, remains closely linked to comprehending the intrinsic characteristics of the data.

3.3.1 Exploratory Data Analysis

Exploratory Data Analysis (EDA) encompasses the comprehensive examination and condensation of fundamental attributes within the dataset. This encompasses deriving summary statistics, scrutinizing data distributions, conducting correlation analyses, and visually representing pivotal features through various visualization techniques. Additionally, EDA involves delving into the intricacies of the data to unveil underlying patterns, trends, and potential outliers, facilitating a deeper understanding of its nuances and informing subsequent analytical decisions. Furthermore, through EDA, one can elucidate relationships between variables and discern any potential biases or anomalies, ensuring a robust foundation for subsequent data processing and modeling endeavors.

3.3.2 Distribution of Transaction id

Analysis of the distribution of the 'id' feature within the dataset entails a thorough investigation into the temporal patterns of transaction occurrences. This analytical process often involves the creation of histograms or density plots, offering visual insights into the frequency and dispersion of transaction times. By scrutinizing these plots, one can discern any discernible trends or anomalies in transaction timing, aiding in the identification of potential patterns associated with fraudulent activities. Additionally, this analysis provides valuable context regarding the temporal dynamics of the dataset, facilitating informed decision-making during subsequent stages of data processing and modeling.

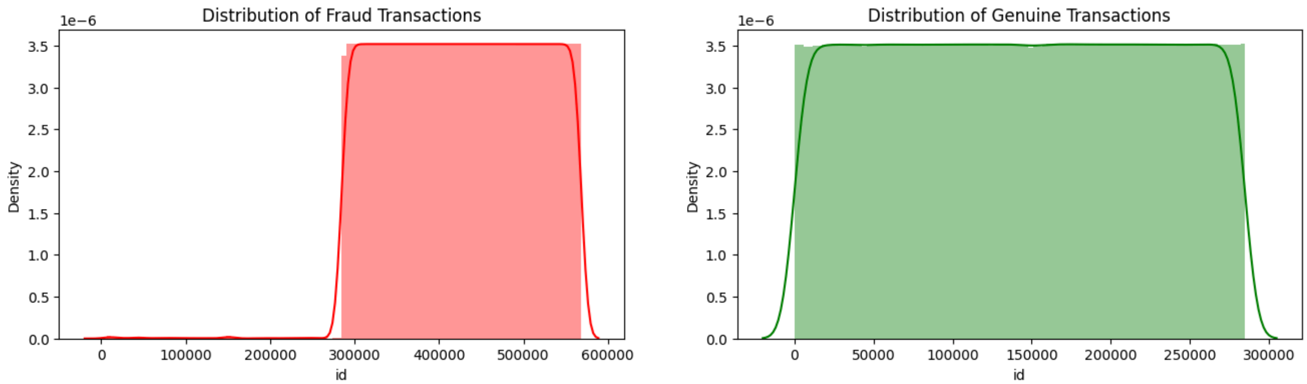

Fig 3.7 : Distribution of Fraud and Genuine Transactions using feature 'id'

3.3.3 Distribution of amount

Exploring the distribution of the 'amount' attribute within the dataset involves a detailed examination of the monetary values associated with transactions. Similar to the analysis of the 'Time' feature, this step may entail employing various visualization methods such as histograms or density plots to depict the distribution of transaction amounts effectively.

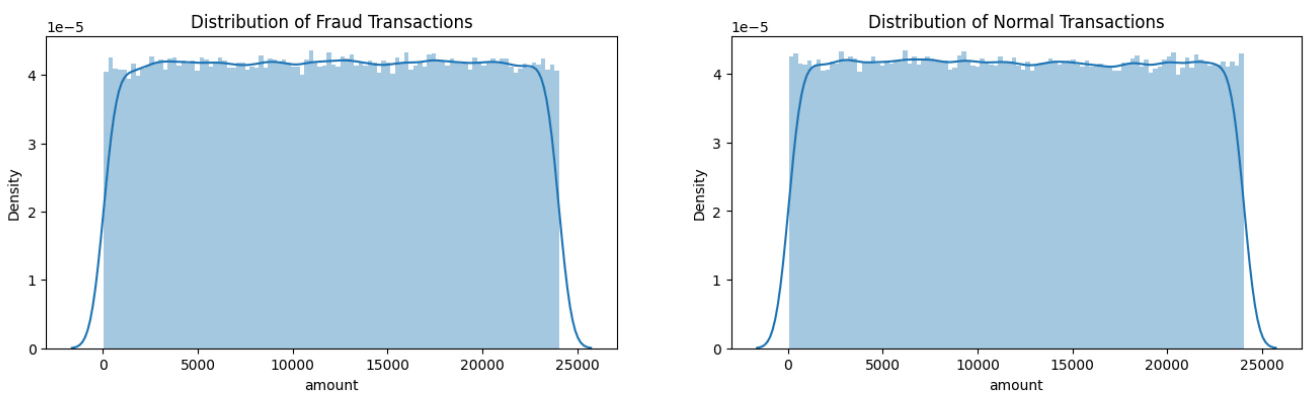

Fig 3.8 : Distribution of Fraud and Normal Transactions using feature 'amount'

3.3.4 Splitting the Data

Splitting the dataset into training and testing sets. This is typically done to evaluate the performance of machine learning models on unseen data. Common splits include 80/20 for training and testing, respectively.

3.3.5 Scaling the data

Scaling the data to ensure that all features have the same scale. This is important for certain machine learning algorithms, such as support vector machines and k-nearest neighbors, which are sensitive to feature scales.

3.3.6 SMOTE

Using the SMOTE technique to address class imbalance in the dataset. SMOTE generates synthetic samples for the minority class to balance the class distribution, thereby improving the performance of machine learning models.

3.4 Proposed classification model

In this project, the proposed classification model has been tested against a referred LSTM model, as well as ANN, GRU, and Naive Bayes. After evaluating these models on the updated credit card dataset from 2023, containing over 550,000 records, it was found that the ANN model achieved the highest accuracy of 0.9994987953502278. Therefore, it can be concluded that the ANN model outperforms others for the credit card dataset of 2023.

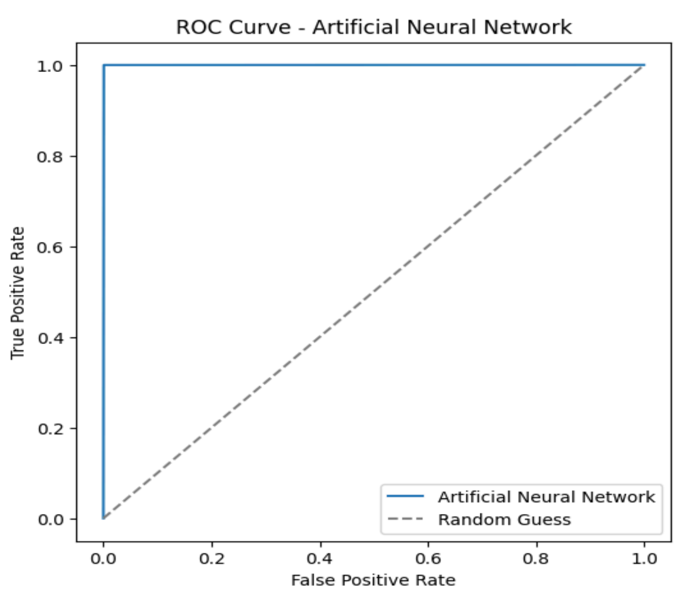

Fig 3.9 : ROC Curve - ANN

3.6 Performance metrics

The experiments are conducted using four basic metrics: TPR, TNR, FPR, and FNR rates measures.

TP (True Positive): This metric represents the share of suspicious transactions that were correctly classified as such.

$$TPR = \frac{TP}{Actual\ Positive} = \frac{TP}{TP + FN}$$

TN (True Negative): This metric represents the share of regular transactions that were accurately identified as non-suspicious.

$$TNR = \frac{TN}{Actual\ Negative} = \frac{TN}{TN + FP}$$

FP (False Positive): This metric indicates the proportion of non-fraudulent transactions that were incorrectly labeled as suspicious.

$$FPR = \frac{FP}{Actual\ Negative} = \frac{FP}{TN + FP}$$

FN (False Negative): This metric is the percentage of fraudulent transactions that were mistakenly categorized as regular transactions.

$$FNR = \frac{FN}{Actual\ Positive} = \frac{FN}{TP + FN}$$

CHAPTER 4

RESULTS AND DISCUSSION

4.1 Effects of Credit Card Fraud Detection

The impact of credit card fraud detection, especially with sophisticated methods like the Long Short-Term Memory (LSTM) proposed in the research paper, resonates across various stakeholders with multifaceted effects. Firstly, bolstering financial security through robust fraud detection systems engenders trust among consumers and financial entities, fostering a more secure payment ecosystem and curbing financial losses. Secondly, the adept detection and prevention of fraudulent activities translate into tangible cost savings for both consumers and financial institutions, mitigating the adverse financial implications of unauthorized transactions. Thirdly, the implementation of stringent fraud detection mechanisms cultivates customer confidence and loyalty, nurturing sustained usage of credit cards and online payment platforms. Fourthly, adherence to regulatory standards concerning fraud prevention becomes more attainable, ensuring compliance and averting penalties for non-conformity. Additionally, the adoption of automated fraud detection systems enhances operational efficiency by swiftly pinpointing suspicious transactions, diminishing the necessity for manual intervention and streamlining processes. Furthermore, a robust credit card fraud detection framework elevates the reputation of financial institutions and payment service providers, positioning them as leaders in security and trustworthiness within the industry landscape. Ultimately, leveraging innovative technologies for fraud detection not only fortifies financial security but also contributes to cost reduction, customer trust, regulatory compliance, operational efficacy, and industry standing, fostering a resilient and secure payment environment for all stakeholders involved.

4.2 ROC Curve Comparison

The ROC curve is a graphical representation of the true positive rate (TPR) against the false positive rate (FPR) at various threshold settings. It helps in understanding the trade-off between sensitivity and specificity of a classifier across different thresholds.

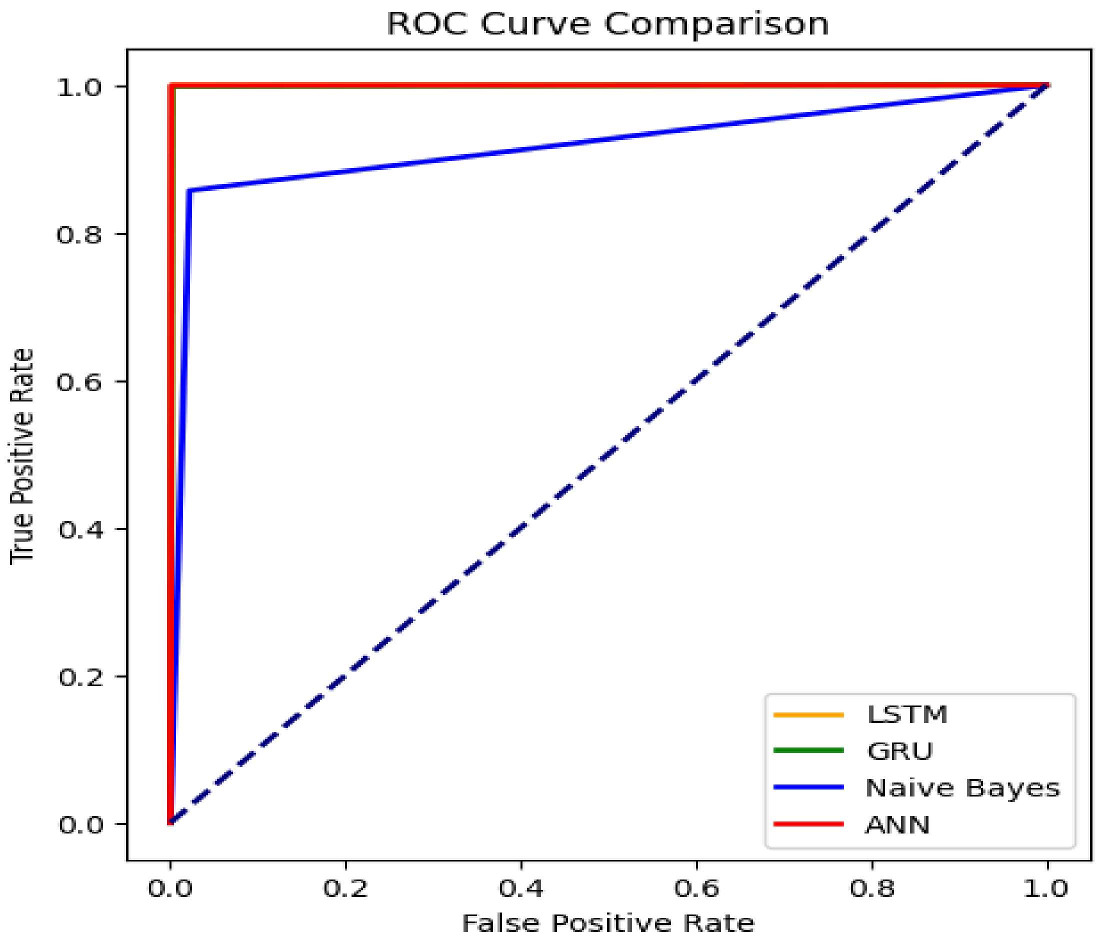

Fig 4.1 : ROC Curve Comparison with different methods

4.3 Evaluation Metric Comparison among Different Models

During the model training phase, a test dataset is employed to assess the performance of the trained model on novel data instances. This evaluation entails the analysis of the model's efficacy using various crucial metrics, thereby gauging its effectiveness in handling unseen data.

Accuracy: Accuracy measures the proportion of correctly classified instances out of the total number of instances. It's calculated as:

$$\text{Accuracy} = \frac{(TP + TN)}{(TP + FP + TN + FN)}$$

Precision: Precision measures the proportion of true positive predictions out of all positive predictions. It's calculated as:

$$\text{Precision} = \frac{True\ Positive}{True\ Positive + False\ Positive}$$

$$= \frac{True\ Positive}{Total\ Predicted\ Positive}$$

Recall (Sensitivity): Recall measures the proportion of true positive predictions out of all actual positive instances. It's calculated as:

$$\text{Recall} = \frac{True\ Positive}{True\ Positive + False\ Negative}$$

$$= \frac{True\ Positive}{Total\ Actual\ Positive}$$

F1-score: The F1-score is the harmonic mean of precision and recall. It balances both precision and recall and is calculated as:

$$\text{F1} = 2 \times \frac{Precision * Recall}{Precision + Recall}$$

ROC Curve (Receiver Operating Characteristic Curve): The ROC curve plots the true positive rate (TPR) against the false positive rate (FPR) at various threshold settings. It helps in understanding the trade-off between sensitivity and specificity.

AUC-ROC (Area Under the ROC Curve): AUC-ROC quantifies the overall performance of a classification model by calculating the area under the ROC curve. A higher AUC indicates better model performance.

Confusion Matrix: A confusion matrix is a table that summarizes the performance of a classification model. It shows the counts of true positives, false positives, true negatives, and false negatives.

Specificity: Specificity measures the proportion of true negative predictions out of all actual negative instances.

$$\text{Specificity} = \frac{\text{True Negatives}}{\text{True Negatives} + \text{False Positives}}$$

Algorithm	Precision	Recall	ROC AUC	F1 Score	Accuracy
LSTM	0.9978	0.9999	0.9988	0.9988	0.9988
GRU	0.9981	0.9983	0.9982	0.9982	0.9982
Naive Bayes	0.9752	0.8568	0.9175	0.9122	0.9173
ANN	0.9992	0.9997	0.9994	0.9995	**0.9994**

Table 4.1 Performance based on confusion matrix

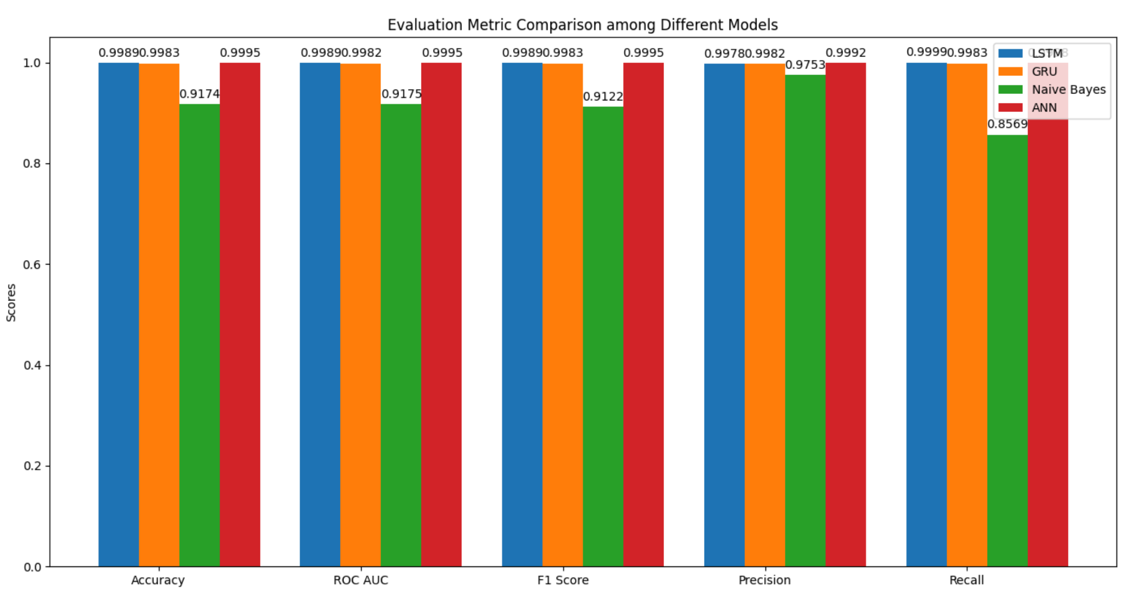

Fig 4.2 : Evaluation Metric Comparison among different methods

CHAPTER 5

CONCLUSIONS AND FUTURE WORK

Conclusion:

The experiments conducted in this project compared the performance of various classification models, including LSTM, ANN, GRU, and Naive Bayes, on an updated credit card dataset from 2023, comprising over 550,000 records. Among these models, the ANN model demonstrated the highest accuracy, achieving an impressive accuracy score of 0.9994987953502278. This result indicates that the ANN model is particularly effective for classifying transactions in the credit card dataset of 2023.

Future Work:

While the ANN model showed promising results in this study, there are several avenues for future research and improvement. One potential direction is to explore ensemble methods that combine the strengths of multiple models to further enhance classification accuracy and robustness. Additionally, it may be beneficial to investigate feature engineering techniques to identify and extract more informative features from the dataset, potentially improving the performance of the classification models. Furthermore, ongoing monitoring and updating of the model will be essential to adapt to evolving trends and patterns in credit card transactions, ensuring continued effectiveness in fraud detection.

REFERENCES

1. Algorithmic trading. (2024, March 29). Wikipedia. https://en.wikipedia.org/wiki/Algorithmic_trading
2. Chen, J. (2024, March 11). Algorithmic Trading: Definition, How It Works, Pros & Cons. Investopedia. https://www.investopedia.com/terms/a/algorithmictrading.asp
3. Roseline, J. F., Naidu, G., Pandi, V. S., Rajasree, S. A. A., & Mageswari, D. (2022, September 1). Autonomous credit card fraud detection using machine learning approach☆. Computers & Electrical Engineering. https://doi.org/10.1016/j.compeleceng.2022.108132
4. Calzone, O. (2022, April 10). An Intuitive Explanation of LSTM - Ottavio Calzone - Medium. Medium. https://medium.com/@ottaviocalzone/an-intuitive-explanation-of-lstm-a035eb6ab42c
5. 10.2. Gated Recurrent Units (GRU) — Dive into Deep Learning 1.0.3 documentation. (n.d.). https://d2l.ai/chapter_recurrent-modern/gru.html
6. G. (2023, June 2). Artificial Neural Networks and its Applications. GeeksforGeeks. https://www.geeksforgeeks.org/artificial-neural-networks-and-its-applications/
7. Chaudhuri, K. D. (2023, August 25). Building Naive Bayes Classifier from Scratch to Perform Sentiment Analysis. Analytics Vidhya. https://www.analyticsvidhya.com/blog/2022/03/building-naive-bayes-classifier-from-scratch-to-perform-sentiment-analysis/
8. Credit Card Fraud Detection Dataset 2023. (2023, September 18). Kaggle. https://www.kaggle.com/datasets/nelgiriyewithana/credit-card-fraud-detection-dataset-2023

Plagiarism Checker X - Report

Originality Assessment

0%

Overall Similarity

Date: May 1, 2024
Matches: 0 / 2677 words
Sources: 0

Remarks: No similarity found, your document looks healthy.

Verify Report:
ew er f a e Onl ne