Journal of Computer Science and Technology Studies
ISSN: 2709-104X
DOI: 10.32996/jcsts
Journal Homepage: www.al-kindipublisher.com/index.php/jcsts

JCSTS
AL-KINDI CENTER FOR RESEARCH AND DEVELOPMENT

| RESEARCH ARTICLE

Explainable AI in Credit Card Fraud Detection: Interpretable Models and Transparent Decision-making for Enhanced Trust and Compliance in the USA

Md Rokibul Hasan[1] ✉ Md Sumon Gazi[2] and Nisha Gurung[3]
[123]MBA Business Analytics, Gannon University, USA
Corresponding Author: MD Rokibul Hasan, **E-mail**: prorokibulhasanbi@gmail.com

| ABSTRACT
Credit Card Fraud presents significant challenges across various domains, comprising, healthcare, insurance, finance, and e-commerce. The principal objective of this research was to examine the efficacy of Machine Learning techniques in detecting credit card fraud. Four key Machine Learning techniques were employed, notably, Support Vector Machine, Logistic Regression, Random Forest, and Artificial Neural Network. Subsequently, model performance was evaluated using Precision, Recall, Accuracy, and F-measure metrics. While all models demonstrated high accuracy rates (99%), this was largely due to the dataset's size, with 284,807 attributes and only 492 fraudulent transactions. Nevertheless, accuracy solely did not provide a comprehensive comparison metric. Support Vector Machine showed the highest recall (89.5), correctly identifying the most positive instances, highlighting its efficacy in detecting true positives. On the other hand, the Artificial Neural Network model exhibited the highest precision (79.4, indicating its capability to make accurate identifications, making it proficient in optimistic predictions.

| KEYWORDS

Credit Card Fraud; Machine Learning; Python; Support Vector Machine (SVM); Artificial Neural Network (ANN); Random Forest (RF); Logistic Regression (LR).

| ARTICLE INFORMATION

ACCEPTED: 15 March 2024 **PUBLISHED:** 06 April 2024 **DOI:** 10.32996/jcsts.2024.6.2.1

1. Introduction

According to Bora (2022), fraud detection has become an instrumental aspect in various sectors, ranging from e-commerce, and finance to healthcare, where the employment of artificial intelligence (AI) systems has demonstrated promising outcomes. Nonetheless, the employment of Artificial Intelligence in fraud detection frequently experiences challenges associated with compliance and trust because of the black-box nature of many Artificial Intelligence models. By conducting a comprehensive review of existing literature, this paper intends to provide insights regarding the benefits of AI and machine learning in fraud detection and intends to provide suggestions for future research and implementation. This study examines the significance of explainable AI (XAI) in Credit fraud detection, concentrating on interpretable models and transparent decision-making to reinforce compliance and trust. Particularly, it delves into the essence of interpretability in Artificial Intelligence systems and presents various methods for attaining explainability in fraud detection.

Fraud presents substantial challenges across various industries, including insurance, healthcare, finance, and e-commerce. As fraudsters progressively modify their strategies, companies should adopt advanced technologies to pinpoint and combat fraudulent activities efficiently. Artificial intelligence (AI) has surfaced as an imperative tool in fraud detection, consolidating machine learning algorithms to evaluate large volumes of data and detect anomalous trends indicative of fraudulent behavior (Bora, 2022). Nevertheless, the innate opacity and complexity of Artificial Intelligence models raise issues concerning transparency, trust, and compliance.

Copyright: © 2023 the Author(s). This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC-BY) 4.0 license (https://creativecommons.org/licenses/by/4.0/). Published by Al-Kindi Centre for Research and Development, London, United Kingdom.

2. Literature Review

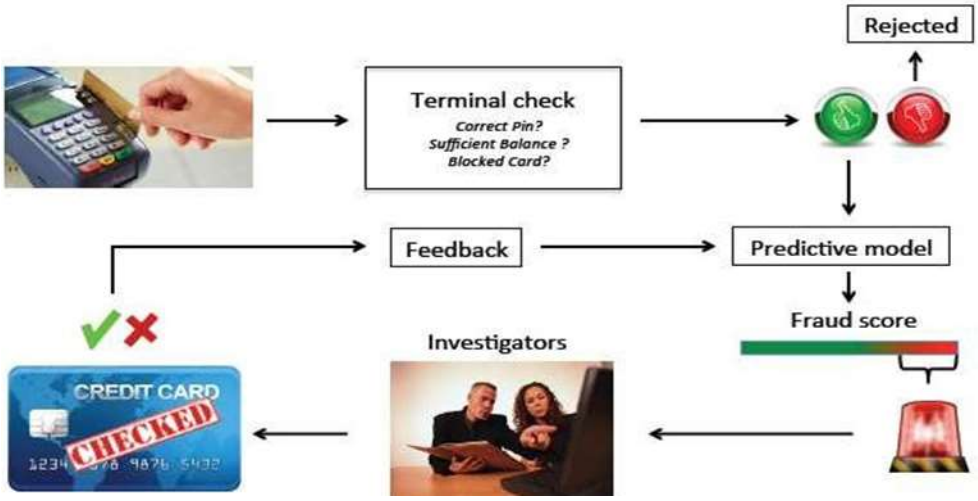
2.1 Credit Fraud

As per Jayaram (2021), Credit card fraud revolves around unauthorized usage of someone else's card information or credit card, to withdraw funds or make purchases without their consent. This fraudulent breaching may comprise compromised card details, stolen physical cards, data wrongly obtained through hacking or phishing, or other types of identity theft. Gn (2018), indicated that credit card fraud presumes different types, in cash withdrawal from another individual's credit card without their permission via an Automated Teller Machine (ATM). It also includes making purchases with a credit card that does not belong to the individual making the transaction. perpetrators may perform these unauthorized activities by physically stealing someone's credit card or obtaining their credit card information, entailing the number and private details needed to make illegal transactions (Ijcsis, 2019).

While associated, credit card fraud varies from identity theft. Credit card theft, where the physical card is stolen, is a common form of identity theft. Nevertheless, identity theft more generally denotes stealing someone's private credentials. Credit card fraud particularly encompasses using another person's credit card details like their ID number to illegally get funds or make purchases without permission (Firdous, 2023) Retrospectively, credit card fraud can destroy financial institutions and cardholders, making detection and prevention of such criminal acts an important objective.

2.2 Credit Fraud Detection

Krilavičius (2021), contended that credit fraud detection denotes the process of pinpointing and mitigating fraudulent activities associated with credit cards or financial transactions. It entails the employment of various techniques, algorithms, and technologies to detect unauthorized or suspicious transactions, account takeover attempts, identity theft, and other forms of fraudulent behavior. Credit fraud detection mechanisms analyze user behavior trends, transactional data, and other pertinent information to detect unusual activities or anomalies that may signify fraudulent activities. These mechanisms aim to safeguard users, financial organizations, and merchants from financial losses and reputational damage related to fraudulent transactions.



2.3 Importance of Interpretability in AI:

Massoud (2023), contended that the interpretability of Artificial Intelligence models revolves around the capability to understand and articulate how the model arrives at its decisions or predictions. In the setting of fraud detection, interpretability is pivotal for several reasons. First, it assists stakeholders, comprising auditors, investigators, and regulators, to comprehend the protocol behind the model's outputs, therefore reinforcing trust in the system. Second, interpretable models enable compliance with standards and regulations, as companies must defend their decisions and affirm accountability. Furthermore, interpretability elevates cooperation between AI systems and humans, facilitating domain professionals to offer feedback and insight to enhance model performance.

As per Saputra (2019), a myriad of machine learning algorithms have been proposed particularly to prioritize model interpretability as compared to pure predictive accuracy for applications sensitive decisions like fraud detection. Decision trees are a prevalently used interpretable model that functions by recursively categorizing data premised on decision rules that can be simply adhered to and understood by humans. Rule-based protocols take a similar dimension by learning decision rules in an IF-THIS-THEN-THAT format. Another option is generalized additive frameworks which articulate forecasting as a sum of interpretable base learners like regression trees. These inherently interpretable frameworks enable fraud analysts to analyze how attributes contribute to predictions to validate decisions.

2.4 Transparency in the Decision Process

Beyond mere interpretable frameworks, consolidating transparency into the entire decision process is also imperative from a compliance perspective Tools and measures should be tailored to document how scenarios are escalated, reviewed, and adjudicated by human analysts. Dashboards showcasing key performance metrics, framework monitoring of data and concept drift, and auditable operator activity logs help affirm accountability and combat issues associated with bias, unfair treatment, and errors (Younas, 2021). Regulated sectors demand disclosure of framework specifics and inspection of how sensitive decisions are made. Explainable Artificial Intelligence can incorporate these compliance requirements via additional interface elements alongside the explanatory models and core predictive.

2.5 Machine Learning Techniques

According to Younas (2021), Machine learning (ML) is a complex scientific discipline that facilitates computer-powered frameworks to learn and develop independently through the evaluation of historical data. It plays a pivotal role in facilitating machines to effectively handle and process large volumes of data. Nonetheless, as situations change, interpreting data exclusively based on present information becomes difficult. This causes technical obstacles in retrieving meaningful insights from the data. As a consequence, the application of Machine Learning methods becomes paramount. Subsequently, companies that seek to optimize advanced technology pinpoint the significance of implementing machine learning to retrieve relevant and valuable data. The growing consciousness of the possibility of Machine Learning methods and the benefits they provide have motivated critical economic sectors to embrace and integrate them into their operations.

Jayaram (2021), in their study, illustrated the applicability of a myriad of classification models for credit card fraud detection. Particularly, the researcher assessed three models: neural networks, decision trees, and logistic regression. These three frameworks were employed for the issue of credit card fraud detection. The research ascertained that among the three frameworks evaluated, logistic regression and neural network outperformed decision trees as regards classification accuracy. logistic regression and Neural networks generated superior outcomes when employed in the credit card fraud detection issue, based on the comparisons and analysis described in their research.

On the other hand, Firdous (2023), suggested a probability theory framework for decision-making during uncertainty. Their study started by examining Bayesian theory as the theoretical premise. Subsequently, as part of assessing this framework, the researchers applied and deployed two classifier algorithms to a credit card transaction dataset: Particularly, the k-nearest neighbor (k-NN) and the naïve Bayes (NB)classifier. Both of these classifiers— k-NN and naïve Bayes —were chosen based on their capability to approximate probabilities under the Bayesian model suggested. By adopting these two probabilistic classifiers to credit card data, the research aimed to illustrate how a probability theory perspective could inform decisions for fraud detection in a scenario with inherent uncertainty.

3. Methodology

The dataset was retrieved from *kaggle.com*. It contained credit card transactions made by American cardholders in January 2024. It includes transactions documented over a timeline of two days, comprising 284,807 transactions, out of which 491 were pinpointed as fraudulent. Because of the highly imbalanced aspect of the dataset, with fraudulent transactions attributing to approximately 0.172% of the overall transactions, and to ensure client confidentiality, particular attributes were transformed into Principal Component Analysis (PCA). Particularly, features labeled as V1, V2, V3, through V21 portray the transformed variables using PCA, while other characteristics such as class, time, and amount remain unaltered as showcased in Table 1 below:

S/No.	Characteristics	Descriptions
1	Class	Transaction amount
2	Time	Time in seconds to indicate the timeline used between the present transaction and the previous one.
3	Amount	1-fraud 0-not fraud
4	LIMIT_BAL	Refers to the limit of the credit card

3.1 Data Preprocessing

The data set was subjected to several phases to refine the data selection. Because of the presence of a huge number of characteristics (887) and a relatively small sample size (995 declarations), specific procedures were employed to filter out less informative characteristics. Firstly, data sets with over 50% missing values were eliminated, since they lacked adequate data to present meaningful insights. Furthermore, characteristics with similar values were eliminated, since they did not contribute to the

variability of the data. Moreover, categorical variables and text attributes with more than 30 categories were also excluded from the analysis.

3.2 Feature Engineering Selection

After performing data explorations, it became apparent that the transaction data displayed a hierarchical structure, encompassing various degrees of granularity such as money depositing, money withdrawals, and money transfer. This hierarchical aspect of the data indicates that adopting a hierarchical time series framework could provide an efficient approach for modeling and forecasting fraud. In a hierarchical time-series framework, the data set is modeled and analyzed at multiple levels of aggregation. This comprises using separate frameworks for every level, facilitating the data analyst to detect different trends and patterns of variation present at every level. Successively, these frameworks account for the association and interdependencies among the different levels.

3.3 Models and Metrics

3.3.1 Logistic Regression (LR):

As per Firdous (2023), logistic models are mostly applied in predictive and classification analytics tasks. Logistic regression evaluates the probability of an incidence happening based on the independent variables, such as whether an individual votes or not. Subsequently, the dependent variable is restricted between 0 and 1, as it portrays a probability. Logistic regression converts the odds, stipulated as the probability of success sub-divided by the probability of failure, employing a logit transformation. This logistic function is represented by the following formulas:

$$Logit(pi) = 1/(1 + exp(-pi))$$

$$ln(pi/(1-pi)) = Beta\_0 + Beta\_1 * X\_1 + \ldots + B\_k * K\_k$$

3.3.2 Support Vector Machine (SVM):

As regards SVM, an N-dimensional hyperplane (where N is the number of attributes) is employed to distinguish data points. Multiple hyperplanes can be adopted to partition the two categories of data points. By elevating the margin distance, analysts can detect a hyperplane with the most substantial space or margin between data points from both classes (Ijcsis, 2019). This dimension strengthens the model by enabling accurate classification of future data points.

3.3.3 Artificial Neural Network (ANN):

Artificial Neural Network (ANN) is a branch of artificial intelligence motivated by biological systems, specifically the human brain. Emulating the brain's structure, an ANN is a computational network composed of linked neurons, akin to the neurons in the human brain. These neurons are arranged into distinct layers, sharing a substantial resemblance to the interrelation model observed in the human brain (Deep, 2022).

3.3.4 Random Forest:

According to Bora (2022), the Random Forest (RF) algorithm integrates the outputs of distinct decision trees to produce a unified outcome. Prominent for its adaptability and simplicity, it has witnessed widespread integration because of its capability to tackle both classification and regression tasks effectively.

3.4 Experimental Results

Python was adopted as the principal tool for both modeling and preprocessing tasks. Its popularity and adaptability in data science, specifically in web development, are greatly acknowledged. The robust use of Python in data science is unquestionable via the presence of specialized libraries designed for modeling and data processing. In particular, the research adopted popular data science libraries such as pandas for effective manipulation and data analysis, such as matplotlib for generating various data visualizations such as charts and graphs as well as NumPy for scientific calculations. Furthermore, the sci-kit-learn library played a pivotal role in developing computational models.

3.5 Importing Libraries

Output:

In [1]:
```
import numpy as np
import seaborn as sns
import pandas as pd
import warnings
warnings.filterwarnings("ignore")
```

In [3]:
```
df = pd.read_csv('creditcard.csv')
df
```

Out[3]:

	Time	V1	V2	V3	V4	V5	V6	V7	V8	V9	...	V21	
0	0.0	-1.359807	-0.072781	2.536347	1.378155	-0.338321	0.462388	0.239599	0.098698	0.363787	...	-0.018307	0
1	0.0	1.191857	0.266151	0.166480	0.448154	0.060018	-0.082361	-0.078803	0.085102	-0.255425	...	-0.225775	-0
2	1.0	-1.358354	-1.340163	1.773209	0.379780	-0.503198	1.800499	0.791461	0.247676	-1.514654	...	0.247998	0
3	1.0	-0.966272	-0.185226	1.792993	-0.863291	-0.010309	1.247203	0.237609	0.377436	-1.387024	...	-0.108300	0
4	2.0	-1.158233	0.877737	1.548718	0.403034	-0.407193	0.095921	0.592941	-0.270533	0.817739	...	-0.009431	0
...	...	...	...	...	...	...	...	...	...	...	...	...	
284802	172786.0	-11.881118	10.071785	-9.834783	-2.066656	-5.364473	-2.606837	-4.918215	7.305334	1.914428	...	0.213454	0
284803	172787.0	-0.732789	-0.055080	2.035030	-0.738589	0.868229	1.058415	0.024330	0.294869	0.584800	...	0.214205	0
284804	172788.0	1.919565	-0.301254	-3.249640	-0.557828	2.630515	3.031260	-0.296827	0.708417	0.432454	...	0.232045	0
284805	172788.0	-0.240440	0.530483	0.702510	0.689799	-0.377961	0.623708	-0.686180	0.679145	0.392087	...	0.265245	0
284806	172792.0	-0.533413	-0.189733	0.703337	-0.506271	-0.012546	-0.649617	1.577006	-0.414650	0.486180	...	0.261057	0

284807 rows × 31 columns

3.6 Data Loading and Exploration

Before progressing with further transformations, the data was loaded into the Python system. During the loading process, structural transformations were employed to match the data with the input demands of every model. During the Initial stage, the dataset consisted of rows representing transaction I.D. and, time and amount transacted Nonetheless, since the chief goal was to forecast fraudulent activities, the investigator aggregated the data by combining transaction figures from the financial organization and days to obtain the overall transaction. This aggregation generated a consolidated monthly transaction figure, facilitating streamlined analysis.

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 284807 entries, 0 to 284806
Data columns (total 31 columns):
 #   Column  Non-Null Count   Dtype  
---  ------  --------------   -----  
 0   Time    284807 non-null  float64
 1   V1      284807 non-null  float64
 2   V2      284807 non-null  float64
 3   V3      284807 non-null  float64
 4   V4      284807 non-null  float64
 5   V5      284807 non-null  float64
 6   V6      284807 non-null  float64
 7   V7      284807 non-null  float64
 8   V8      284807 non-null  float64
 9   V9      284807 non-null  float64
 10  V10     284807 non-null  float64
 11  V11     284807 non-null  float64
 12  V12     284807 non-null  float64
 13  V13     284807 non-null  float64
 14  V14     284807 non-null  float64
 15  V15     284807 non-null  float64
 16  V16     284807 non-null  float64
 17  V17     284807 non-null  float64
 18  V18     284807 non-null  float64
 19  V19     284807 non-null  float64
 20  V20     284807 non-null  float64
 21  V21     284807 non-null  float64
 22  V22     284807 non-null  float64
 23  V23     284807 non-null  float64
 24  V24     284807 non-null  float64
 25  V25     284807 non-null  float64
 26  V26     284807 non-null  float64
 27  V27     284807 non-null  float64
 28  V28     284807 non-null  float64
 29  Amount  284807 non-null  float64
 30  Class   284807 non-null  int64  
dtypes: float64(30), int64(1)
memory usage: 67.4 MB
```

In [5]:

Out[5]:

	Time	V1	V2	V3	V4	V5	V6	V7	
count	284807.000000	2.848070e+05	2.848070e+05	2.848070e+05	2.848070e+05	2.848070e+05	2.848070e+05	2.848070e+05	2.8480
mean	94813.859575	1.168375e-15	3.416908e-16	-1.379537e-15	2.074095e-15	9.604066e-16	1.487313e-15	-5.556467e-16	1.213
std	47488.145955	1.958696e+00	1.651309e+00	1.516255e+00	1.415869e+00	1.380247e+00	1.332271e+00	1.237094e+00	1.1943
min	0.000000	-5.640751e+01	-7.271573e+01	-4.832559e+01	-5.683171e+00	-1.137433e+02	-2.616051e+01	-4.355724e+01	-7.3216
25%	54201.500000	-9.203734e-01	-5.985499e-01	-8.903648e-01	-8.486401e-01	-6.915971e-01	-7.682956e-01	-5.540759e-01	-2.086
50%	84692.000000	1.810880e-02	6.548556e-02	1.798463e-01	-1.984653e-02	-5.433583e-02	-2.741871e-01	4.010308e-02	2.235
75%	139320.500000	1.315642e+00	8.037239e-01	1.027196e+00	7.433413e-01	6.119264e-01	3.985649e-01	5.704361e-01	3.273
max	172792.000000	2.454930e+00	2.205773e+01	9.382558e+00	1.687534e+01	3.480167e+01	7.330163e+01	1.205895e+02	2.0007

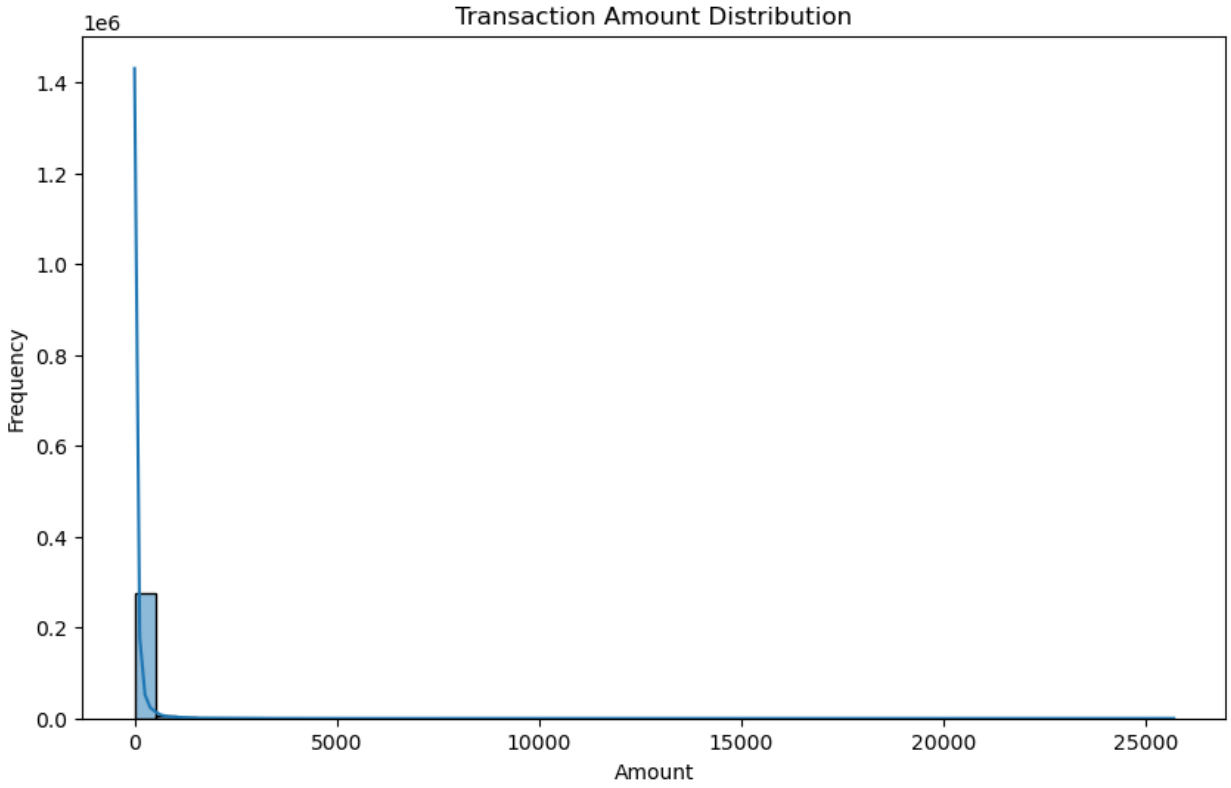
8 rows × 31 columns

Output:

Concerning the data frame, each row was adjusted to represent transaction amount distribution for a specific month. This implies that for every month, the row displayed the maximum transaction value attained among all ATMs. By arranging the data in this way, the investigator could decipher insights into the peak transaction activities for every month and evaluate the trends and patterns related to ATM transactions across the branches.

In [8]:
```
# Transaction amount distribution
plt.figure(figsize=(10, 6))
sns.histplot(df['Amount'], bins=50, kde=True)
plt.title('Transaction Amount Distribution')
plt.xlabel('Amount')
plt.ylabel('Frequency')
plt.show()
```

Output:



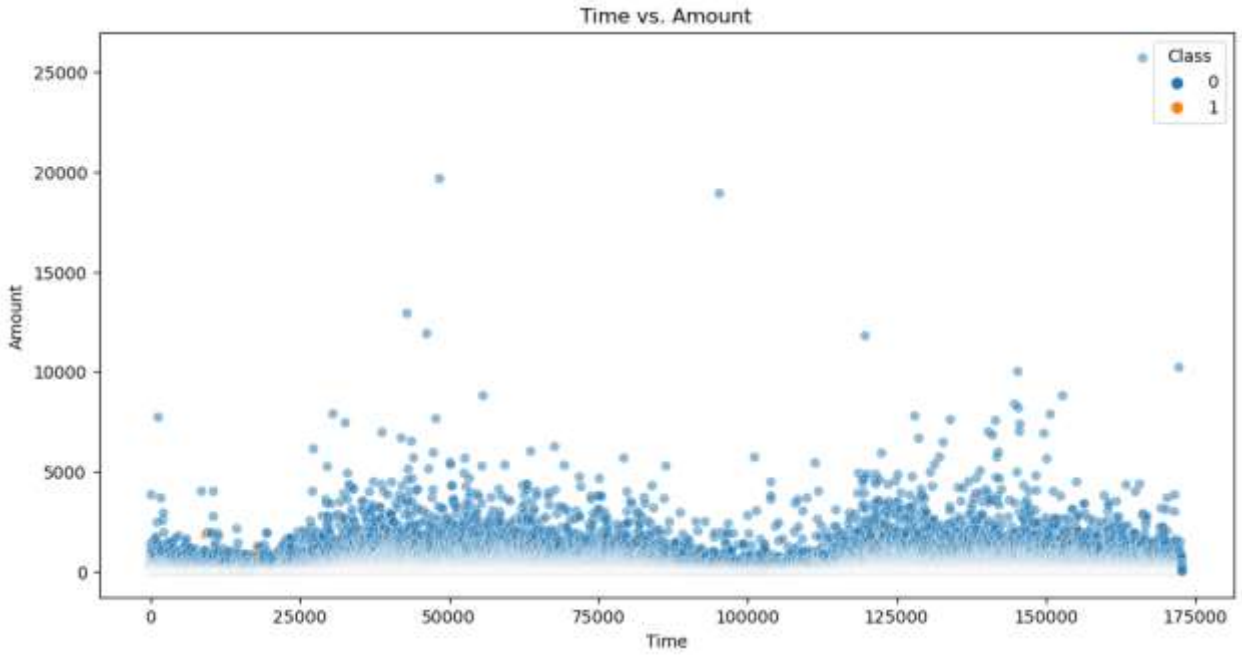
By referring to the above chart, it was evident that the majority of transactions were below 5,000 million dollars. Implying that the majority of the transactions fall below this amount. It is important to mention that, the distribution seemed to be right-skewed. Implying that the majority of transactions fell on the left side of the distribution (lower amounts) as compared to the right side (higher amounts).

The analyst further computed the variance between the transactions of each month and modified it as a new column in the data frame. Particularly, this conversion was conducted to make the data stationary, which assisted in modeling and assessing time series data effectively. Particularly, the time Vs. amount () function was performed to present information regarding the time and amount. The function represented transaction date in terms of months, and days, therefore providing an extensive understanding of the period covered by the transaction data.

In [9]:

```
# Time vs. Amount
plt.figure(figsize=(12, 6))
sns.scatterplot(x='Time', y='Amount', data=df, hue='Class', alpha=0.5)
plt.title('Time vs. Amount')
plt.xlabel('Time')
plt.ylabel('Amount')
plt.legend(title='Class')
plt.show()
```

Output:

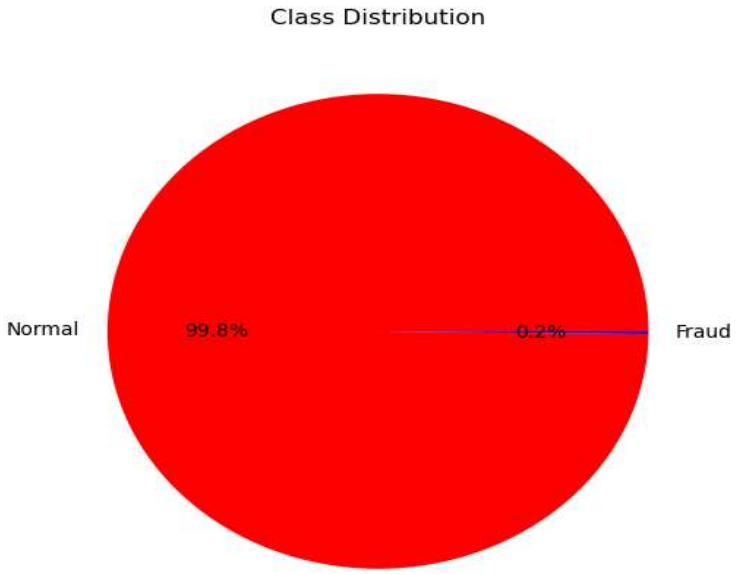


The graph above demonstrates the association between time and some variable amount. The time was measured in seconds, ranging from 0 seconds to 175, 000 seconds. Overall, there seemed to be a strong positive correlation between amount and time. As time increased along the x-axis, the general amounts generally increased along the y-axis. This indicated that greater time intervals outcome in higher measured amounts.

Subsequently, a code snippet was employed to create a pie chart to visualize the distribution of classes in a panda's data frame. The code snippet aimed at generating a chart with two slices displaying the normal and fraudulent classes colored blue and red respectively. The proportion of the slices was intended to correspond to the proportion of every class in the data.

```
In [11]:
# Pie chart for class distribution
class_counts = df['Class'].value_counts()
plt.figure(figsize=(8, 6))
plt.pie(class_counts, labels=['Normal', 'Fraud'], autopct='%1.1f%%', colors=['red', 'blue'])
plt.title('Class Distribution')
plt.show()
```

Output:



The chart above displays the distribution of two classes, notably, normal and fraud. From the chart above it was evident that the majority of incidences, 99.8%, belonged to the normal class, while on the other hand, 0.2% of the incidences belonged to the fraud class, colored red in the chart.

Confusion Matrix

	Predicted False	**Predicted True**
Actual False	False Positive (FP)	True Negative (TN)
Actual True	True Positive (TP)	False Negative (FN)

Performance Metrics

- The **precision** refers to the value of the TP component over TP and FP, utilizing the confusion matrix can be computed as follows:

$$Precision = \frac{TP}{TP + FP}$$

- The **recall** denotes the value of the TP component over TP and FN, computed as follows:

$$Recall = \frac{TP}{TP + FN}$$

- The **f-measure**, as per the Confusion Matrix, balances the recall and precision values as follows:

$$F - measure = 2 \times \frac{Recall \times Precision}{Recall + Precision}$$

- The **accuracy** as per the Confusion Matrix, is the ratio of correct predictions to the sample size which can be calculated as follows:

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

Experimental Results

Algorithm	Precision	Recall	Accuracy	F-Measure
SVM	74.7	89.5	99	81.5
LR	59	82	99	69
ANN	79	65	99	71.4
RF	78	65	99	71.4

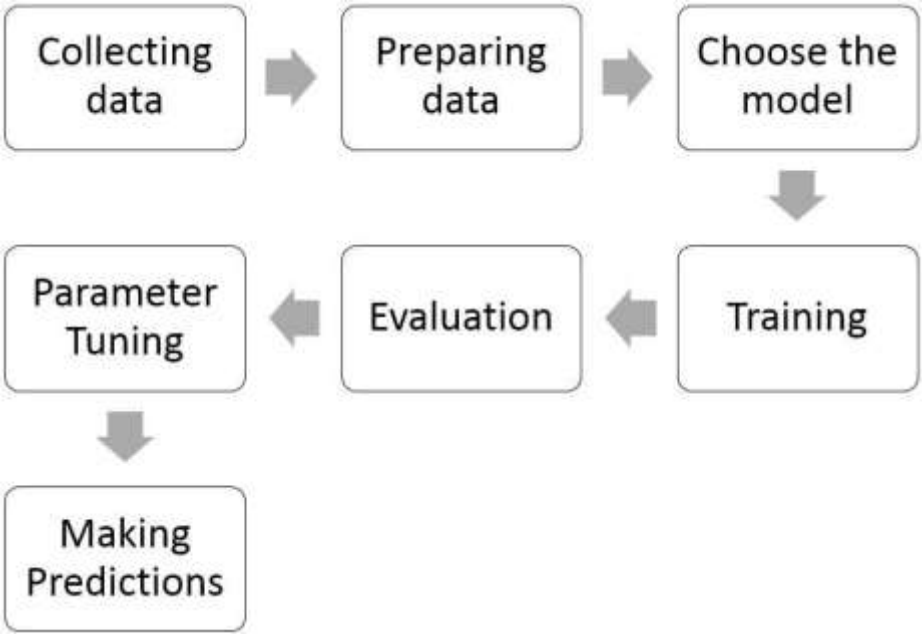
By referring to the table above, all models displayed relatively high accuracy rates (99%). As such, this high accuracy percentage was attributed to the huge number of components in the dataset. With 284,807 attributes, and only 492 transactions detected as fraudulent (0.14%), However, accuracy solely may not be the most valuable metric for comparison. SVM illustrated the greatest recall value (89.5) among the models. This demonstrated that SVM accurately detected the greatest proportion of positive class samples out of the overall samples for that class. Therefore, SVM displayed a solid performance in terms of correctly detecting actual positive incidences. On the other hand, The ANN model demonstrated the greatest precision. This indicated that ANN generated the greatest level of optimistic predictions (79), making it ideal in terms of generating precise positive class identifications.

3.8 Business Impact

Normally, a significant number of financial organizations in the USA have been employing rule-based systems which involve employing thresholds and predefined rules to flag possible fraudulent transactions. For instance, transactions occurring in high-risk locations, transactions beyond a specific amount, or deviating from the cardholder's normal spending behavior may elicit alerts for further inspection. While rule-based systems are straightforward to execute, they lack adaptability to emerging fraud patterns and may generate a high number of false positives. In that light, the proposed SVM and ANN models are suitable for combating and mitigating these shortcomings and weaknesses of conventional models. By adopting the proposed SVM and ANN models financial organizations can benefit in the following ways:

- ✓ **Real-time Fraud Detection:** By adopting SVM and ANN models, financial organizations in the USA can assess incoming credit card transactions in real time. The proposed models monitor and evaluate various features of the transactions, such as location, transaction amount, and consumer behavior, to detect suspicious patterns. In a scenario where a transaction is flagged as possibly fraudulent, instant rapid actions can be adopted to validate the transaction or block it, thereby preventing financial losses.
- ✓ **Risk Scoring**: The SVM and ANN models can assign risk scores to independent transactions or clients based on the predicted probability of fraud. Financial organizations in the USA can set levels for risk scores and utilize them to prioritize investigation efforts. High-risk transactions that have attained high scores can be subjected to rigorous review or further verification, while low-risk transactions can be processed more effectively. This technique optimizes resources and minimizes the impact on legitimate customers.
- ✓ **Adaptive Learning:** SVM and ANN models can progressively learn and adapt from new data to remain informed regarding evolving fraud patterns. As cyber criminals develop new approaches, the models can be retrained using the current information. This adaptive learning capacity affirms that the models remain efficient in terms of pinpointing emerging fraud trends, therefore improving overall fraud detection accuracy.

How to Use the Proposed Model



- **Training Data Preparation:** Companies in the USA need to collect a huge dataset of historical credit card transactions, entailing both legitimate and fraudulent transactions. This dataset should include different transaction attributes, such as time, location, transaction amount, and client behavior. Subsequently, the dataset should then be subdivided into a training set and a test set.
- **Model Training:** Successively, the SVM and ANN models should be trained to utilize the labeled data from the training set. As a result, the models learn to detect relationships and patterns between the transaction attributes and the fraud labels. Throughout the training process, the models adapt their internal parameters to reduce errors and maximize accuracy.
- **Feature Engineering and Selection**: Businesses can perform feature engineering and selection to elevate the models' performance. This entails pinpointing the most relevant attributes that contribute to fraud detection and transforming new features that capture valuable information. Feature engineering and selection assist in enhancing the models' capability to distinguish between fraudulent and legitimate transactions.
- **Model Validation and Optimization**: After training the ANN and SVM models, businesses need to validate and optimize them. This entails tweaking the models' hyperparameters, such as the kernel regularization and function boundaries for SVM, and the quantity of layers and neurons for ANN. Optimization affirms that the models are operating at their best. Validation is performed utilizing the test set to evaluate the models' performance and generalization capability.
- **Real-time Fraud Detection**: After the models are properly trained, validated, and optimized, companies can execute them for real-time fraud detection. Incoming credit card transactions are inputted into the models, which evaluate the transaction characteristics and offer predictions on whether the transactions are legitimate or fraudulent. Transactions flagged as possibly fraudulent can then be imposed for further investigation or declined if necessary.
- **Model Monitoring and Maintenance**: Organizations should progressively monitor and evaluate the performance of the ANN and SVM models in production. This entails tracking strategic metrics, such as false positive rate detection accuracy, and false negative rate. If the models exhibit a decline in performance or fail to adjust to new fraud trends, companies need to update or retrain the models with fresh data to ensure their effectiveness.

3.9 Benefits to the USA Economy

- ✓ **Reduced Financial Losses**: Effeccient credit card fraud detection utilizing the SVM and ANN models assists in combating financial losses for both financial organizations and clients. By combating fraudulent transactions, the economy preserves finances that would otherwise be lost to fraud. Consequently, this contributes to overall economic growth and stability.
- ✓ **Enhanced Operational Efficiency**: SVM and ANN models will empower financial organizations to streamline their fraud detection protocols. By automating the detection of fraudulent transactions, the American government can minimize manual effort, improve operational efficiency, and optimize national resources. This can result in overall cost savings and enhanced profitability.

4. Conclusion

The chief objective of this research was to assess the efficiency of Machine Learning techniques in terms of detecting credit card fraud transactions. Four Machine Learning approaches were adopted, most notably: Support Vector Machine, Logistic Regression, Random Forest, and Artificial Neural Network. Model performance was evaluated based on Recall, Precision, F-measure, and Accuracy metrics. The proposed approach was tested using a real dataset obtained from Kaggle.com. All models exhibited notably high accuracy rates. However, this high accuracy was largely influenced by the extensive dataset, comprising 284,807 attributes, with only 492 identified as fraudulent transactions. Nonetheless, accuracy alone may not provide the most comprehensive comparison metric. Among the models, SVM demonstrated the highest recall value, accurately identifying the largest proportion of positive class samples. This underscores SVM's effectiveness in correctly detecting actual positive instances. Conversely, the ANN model exhibited the highest precision, indicating its proficiency in generating precise positive class identifications, making it particularly adept at optimistic predictions.

Funding: This research received no external funding.

Conflicts of Interest: The authors declare no conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers.

References

[1] Bora, S. (2022). Credit Card Fraud Detection using Machine Learning Framework. *www.academia.edu*. https://www.academia.edu/71139567/Credit_Card_Fraud_Detection_using_Machine_Learning_Framework

[2] Deep, U. (2022). Insurance fraud detection using machine learning. *www.academia.edu*. https://www.academia.edu/68589812/Insurance_Fraud_Detection_Using_Machine_Learning

[3] Gn, B. (2018). Machine learning approaches for credit card fraud detection. *Vit-in*. https://www.academia.edu/36810759/Machine_Learning_Approaches_for_Credit_Card_Fraud_Detection

[4] Ijcsis, J. O. C. S. (2019). A comparative study of credit card fraud detection using machine learning for United Kingdom Dataset. *www.academia.edu*. https://www.academia.edu/40932542/A_Comparative_Study_of_Credit_Card_Fraud_Detection_Using_Machine_Learning_for_United_Kingdom_Dataset

[5] Jayaram, R. (2021). Fraud Identification of Credit Card using Machine Learning. *www.academia.edu*. https://www.academia.edu/93452208/Fraud_Identification_of_Credit_card_using_Machine_Learning

[6] Journal, I., & Firdous, Z. (2023). Credit card fraud detection using machine learning. *Irjet*. https://www.academia.edu/104422949/Credit_Card_Fraud_Detection_Using_Machine_Learning

[7] Krilavičius, T. (2021). Machine learning approaches for customs fraud detection. *www.academia.edu*. https://www.academia.edu/87912540/Machine_Learning_Approaches_for_Customs_Fraud_Detection

[8] Massoud, M. (2023). Credit card fraud detector based on machine learning techniques. *Jinan*. https://www.academia.edu/105781326/Credit_Card_Fraud_Detector_Based_on_Machine_Learning_Techniques

[9] proAIrokibul. (2024). Fraud-Detection-And-Prevention/Model/Fraud Detection. ipynb at main · proAIrokibul/Fraud-Detection-And-Prevention. GitHub. https://github.com/proAIrokibul/Fraud-Detection-And-Prevention/blob/main/Model/Fraud%20Detection.ipynb

[10] Saputra, A. (2019). Fraud Detection using Machine Learning in e-Commerce. *www.academia.edu*. https://www.academia.edu/107003335/Fraud_Detection_using_Machine_Learning_in_e_Commerce

[11] Younas, M., (2021). Credit Card Fraud Detection using Machine Learning Algorithms. *www.academia.edu*. https://www.academia.edu/45004971/Credit_Card_Fraud_Detection_using_Machine_Learning_Algorithms