JOURNAL OF BASIC SCIENCE AND ENGINEERING

EVOLUTIONARY INSIGHTS INTO ENSEMBLE LEARNING MODELS FOR ADVANCED ONLINE FRAUD DETECTION

Pralad Upreti
University: Om Sterling Global University
Hisar,India

Abstract

One of the largest existing problems in the financial industry is “Credit Card Fraud” (CCF) as it incurs significant financial losses and erodes consumer confidence. Accordingly, this paper aims at describing how the future online fraud detection system particularly in online “Credit Card Fraud Detection” (CCFD) which incorporates “Deep Learning” (DL), “Machine Learning” (ML), and “Ensemble Learning” (EL) models. The algorithms assessed include: “Logistic Regression”, “Random Forest” and “Decision Trees”, due to their capability of handling big data and model interpretability. Consequently, the DL approaches, including CNNs, LSTM networks, and GANs, are suggested to exhibit better results in the recognition of complex fraud features. Hence, the performance of various EL approaches such as soft voting, hybrid, and weighted methods are investigated to understand their suitability for model fusion. Also, this review offers an understanding of feature engineering, data preprocessing, and anomaly detection, all of which are essential in enhancing the performance of fraud detection systems. Therefore, based on the recent techniques of ML, DL, and EL, the study aims at providing actionable insights to the researchers and practitioners regarding the existing fraud detection strategies. These results further vindicate the need to recalibrate models and integrating human factor in combating the intricate strategies of the defrauders in enhanced measures of preventing monetary loss.

Key Terms: ML, CCFD, Anomaly Detection, DL, EL

Introduction

A well-known and rising phenomenon that has become a serious concern for the banking sector and its clients is Credit Card Fraud (CCF), which results in significant monetary losses and a loss of reputation. As more consumers turn to e-commerce, the measures required to prevent online scams remain important. This article aims to propose the utilization of EL, DL, and ML algorithms in combating CCF that occurs online while also highlighting the strengths and limitations of each. It also aims to discuss how the use of the internet in offering financial products and services, as well as utilizing the internet to shop, has revolutionized business. However, this has also led to easy access to exploitation and other chances of carrying out fraudulent activities among other unlawful acts. FTC records show that there were about 1,579 data breaches in the current fiscal year affecting around 179 million people, and the most reported form of financial

fraud was CCF (Bagga et al., 2020). This fact calls for appropriate strategies to be put in place to prevent fraud in a bid to safeguard users from monetary frauds.

There are various kinds of CCF, which include account takeover, new account, cloned card, and cards-not-present. Fraudsters employ techniques such as phishing, skimming, and data theft to access credit card information. The cost of such fraud is high, with global losses expected to cost firms $43 billion in the next five years. This figure alone requires management to implement efficient systems that can help prevent acts of fraudulence (Du et al., 2024).

First-generation expert systems, statistical methods, and basic forms of ML are no longer sufficient to address various modern fraud patterns. One main disadvantage of using rule-based systems is that they are not very good at identifying new trends, often generating a large number of false positives. Statistical methods are not very useful when the data is not normally distributed and also when highly dimensional. Simple classifiers such as logistic regression and support vector machines have problems with imbalanced data and require much time to be trained (Dornadula & Geetha, 2019).

Consequently, the advancement of ML, DL, and other EL techniques offers new hope in fraud detection. Techniques like DT, RF, and NB classifiers have been applied for fraud detection because of their ability to handle big data and create explainable models. However, these methods are not very effective for thorough guidance in detecting newer forms of fraud easily and are contaminated by the quality and type of data. Since the CCF datasets are skewed that there are far many fewer cases of fraud as compared to actual genuine transactions, this results in low detection and high false positives (Afriyie et al. , 2023).

Vol. 21, No. 1, (2024)
ISSN: 1005-0930

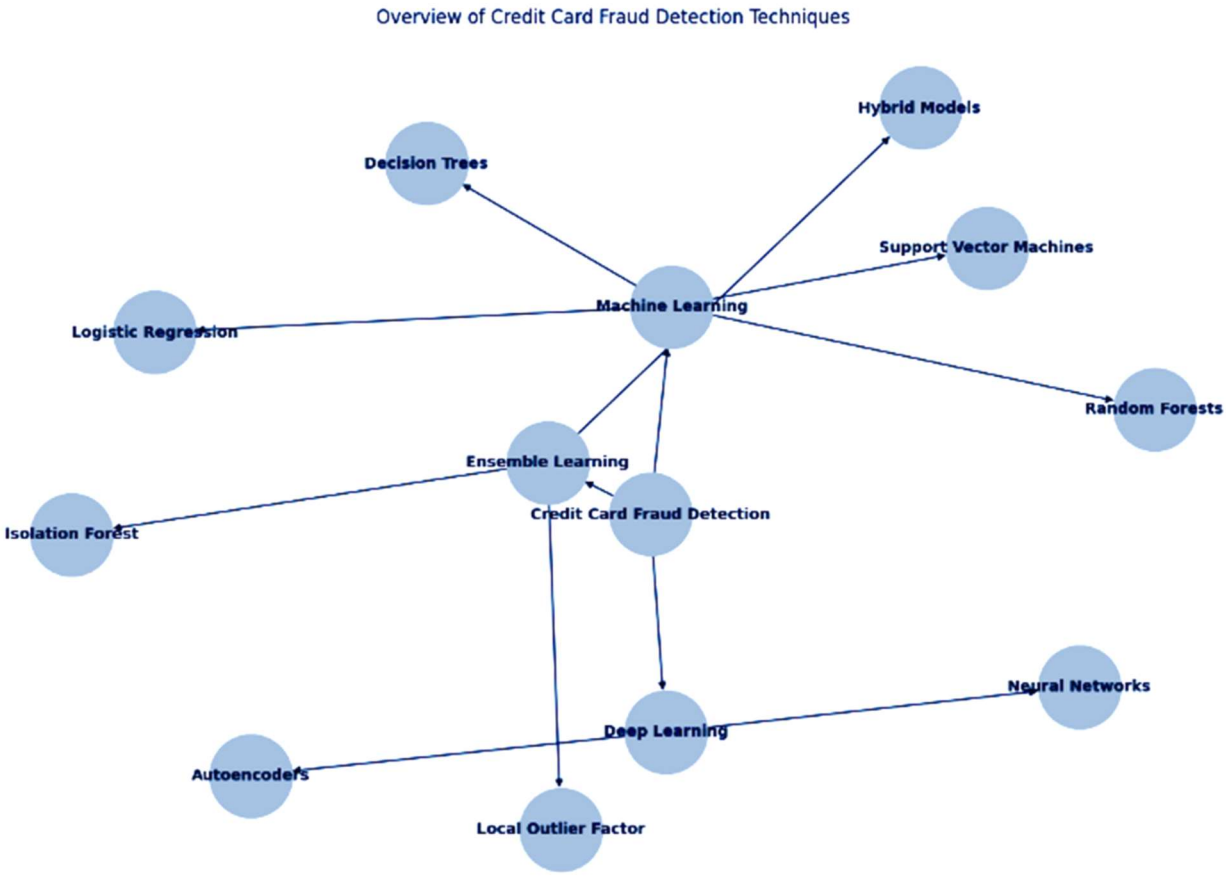

Fig.1 Overview of CCFD Techniques

Hence, other deep learning methods like neural networks and autoencoders which are under the FOR classification are more effective in identifying complex fraud patterns than other approaches. They are most suitable for use in datasets that are large and concern mass fraud and they can easily be updated with any new emerging fraud pattern as and when they are developed (Karkaba et al., 2023). However, the parameters in DL models are numerous, and the development and maintenance of these models are computationally expensive. However, the collection of large labeled training data is itself a challenge because of privacy and security concerns, and because financial transaction data cannot be easily shared publicly (Sharma et al., 2022).

Other high level classification decisions are random forest gradient boosting and hybrid models. These models can construct several pattern using several algorithms so as to arrive at a better conclusion with regards to fraudulent assessment. They also help reduce model complexity and increase its ability to generalize, which is useful when it comes to the identification of fraudulent activities (de Souza & Bordin Jr., 2021). But they also bring additional computational overhead, and optimizing them for getting good results is not a walk in the park. It can also make the system complex and not easy to train and summarize because there are different models (Afriyie et al., 2023).

Literature Review

Mimusa Azim Mim et al. (2024) present a soft voting EL that is used in identifying CCF from imbalanced data. The paper measures and compares the strengths and weaknesses of oversampling, undersampling, and hybrid sampling in addressing the imbalance in the class. The following experimental result demonstrates that the soft voting ensemble approach yields a higher level of F1-score, recall, and precision than individual classifiers. The authors point out that integration of multiple models enables better detection of targets and minimizing false negatives. But the study also emphasizes that the ensemble approach is computationally intensive and requires more resources than the other methods.

Zorion et al. (2023) have discussed and implemented a CCFD system using DL artificial neural network; the authors compared the result of convolution and recurrent neural networks with other ML algorithms. From the sources, it is clear that both CNNs and RNNs are more accurate and provide less false positive as compared to normal ML algorithms. The study identifies that feature selection, normalization and preprocessing have the most significant influence on improving these models. However, the authors also pay attention to the computational needs and resource consumption in DL models as well.

Ileberi et al. (2022) emphasized that the GA to optimize feature selection for enhancing the aforementioned classifiers. Their research enables an understanding of how the use of the proposed detection engine through ML that works with dataset acquired from the European cardholders is more efficient and effective than current systems. The advantage of using DT and RF is the ability of both algorithms to work on big data while maintaining some level of interpretability. However, the same study also points out that there is a problem of data distribution skewness in the CCF datasets hence resulting in poor detection and high false positive rates.

Thirunavukkarasu et al. (2021) evaluates the performance of ANN and NB model, used for CCF prediction. Their work identifies the target indicators that are the basis of the ML model developed in the study to screen the fake and real transactions. NB because it is easy to implement for large data whereas ANN has the unique ability to learn various pattern form the data set. This work posits that the combination of NB and ANN enhances the classification performance and lowers the probabilities of false positive. The authors also agree with the fact that such models' performance always depends on the quality and variety of data to learn at any given time.

Tiwari et al. (2021) has explained that even though LR is very basic and easy to understand and implement but it is very effective in binary classification issues and on the other hand it is seen that SVM is very effective in high dimensional space. Nevertheless, both examined techniques have certain drawbacks, particularly concerning the handling of imbalanced data and requiring significant computational resources to train. According to the authors of the study, the long rule

Vol. 21, No. 1, (2024)
ISSN: 1005-0930

algorithm, as well as the second algorithm representing support vector machine are qualitatively powerful to distinguish previously identified fraud patterns.

Dornadula and Geetha, (2019), emphasize that the DL approaches assist in the identification of these anomalous features by using the input data to reconstruct it and hence normalizing it to a normal or even an ideal state. While autoencoder is used in generating models for the unlabeled data we have, RBMs are useful in modeling dependencies of the data. As presented in the study, these models provide better accuracy compared to the usual ML algorithms, but they are also more time-consuming, require greater computing capacity, and involve big annotated data or training sets.

Sohony et al. (2018) intend the algorithmic model for CCFD that can be described as the combined ML, which utilizes the RF and NN. However, the work does show that if the two models are used in an ensemble fashion, the combined model is more accurate and gives less false positive than the above introduced models. According to this work, it can be appreciated that the implementation of the RF can address the challenge of presence of noisy data in the dataset while the use of the NN can help in detecting complicated fraud scenarios. However, the authors emphasize that the ensemble model is more computationally demanding, and the hyperparameters have to be adjusted properly.

Research Gap

One of the major issues is the problem of dealing with extreme class imbalance, when volumes of actual fraudulent transactions are considerably overshadowed by the overall ratios of legitimate ones. Such a discrepancy inevitably results in the creation of partial models that are not always capable of identifying potentially fraudulent cases. Also, fraud strategies are dynamic, and this calls for constant model upgrades and adjustments that most existing frameworks fail to offer. More research is also required in order to enhance the interpretability and decision-making capabilities of model-based fraud detection, especially when human experts are incorporated into the process. Filling these gaps is important as it helps in creating sound and efficient systems for identifying fraud.

Problem Statement

The biggest issue with CCF is that the fraudsters are also evolving and there is no sense that the new approaches to the CCFD could be truly ready to address the emerging threats. While statistically-based ML, DL and EL technique based methodologies offer robust methods for identification and prevention of fraud each methodology has its demerits. Among these approaches, ML methods can be less efficient when faced with new or different fraud schemes, while DL methods are compute-intensive, and using EL approaches complicates the system. Comparing the techniques in this article aims at identifying aspects such as effectiveness and

Vol. 21, No. 1, (2024)
ISSN: 1005-0930

limitation of these techniques in addressing FOC-CC fraud. For this reason, this article seeks to review previous studies on online CCFD and synthesize the results to facilitate positive action as well as contribute to expanding knowledge in this subfield.

Objectives:

- To ex a synthesis of EL, DL, and ML models' performance in detecting and preventing CCF-related online fraud to assist researchers and practitioners in making informed decisions.
- To analyze the advancement of knowledge in online CCFD by synthesizing existing research findings and proposing actionable insights for researchers and industry professionals.

Methodology

Data Collection and Preparation

The collection of research data began with transaction history data, which included; users, their profiles, and account data. This was necessary to exclude the loss of data or differences between them, which could complicate the further work during the next stages. This was done to reduce all day to day differences that might affect the data such as missing value and duplication. Moreover, some of the information due to privacy restrictions was partially masked to ensure compliance with privacy laws applicable to use of such data and, in general, to protect the identities of users of the corresponding platforms. The dataset which has been used in this research covers the data of European cardholders which were useful to establish a large number of transactions involving many types of categories and regions. These two types of samples were beneficial in making the model diverse because the model was able to identify different patterns for the activities of the users participating in legitimate transaction as well as the fraudsters.

Feature Engineering

Feature selection involved identifying and implementing major variables derived from the collected data. More specifically, the following items were extracted and transformed for a better predictive model; Transaction amount, time of transaction, usage pattern or usage frequency of the merchant and the category to which the merchant belongs. For example, the 'Hour of Day' and ''Day of Week' fields suggest that some suspicious activity may occur at night, which is often indicative of fraud. Similarly, 'User Transaction Frequency' and 'User Average Transaction Amounts' are valuable parameters to define users and distinguish ordinary and malicious actions. The methodologies applied in feature engineering enabled the raw data set to be prepared and ingested into ML models in a more accurate and robust manner.

Vol. 21, No. 1, (2024)
ISSN: 1005-0930

Model Selection and Training

As a result of the issue, the suitable ML, DL, and El models along with the available variables and data sets were identified. All the classification models were built with labelled training dataset and performance metrics were tested with standard tools such as precision, accuracy, F1 score and recall. The research applied "Random Forest" (RF), "Decision Trees" (DT), "Logistic Regression" (LR), "Artificial Neural Networks" (ANN), and "Naïve Bayes" (NB). Using such an approach, each model was selected based upon its best features with respect to the flowchart. For instance, Random Forests was chosen due to its large data applicability and because of the fully interpretable models that it can generate, while on the other hand ANN were chosen because of its ability to learn complex implicit patterns in the data set. The training process involved data preprocessing with the aim of splitting the training dataset into the sets that were employed in evaluation of the models' performance. The models were then set to learn further, make them more sensitive in order to identify the fraud cases but not at the risk of high true and false positive charges.

Anomaly Detection

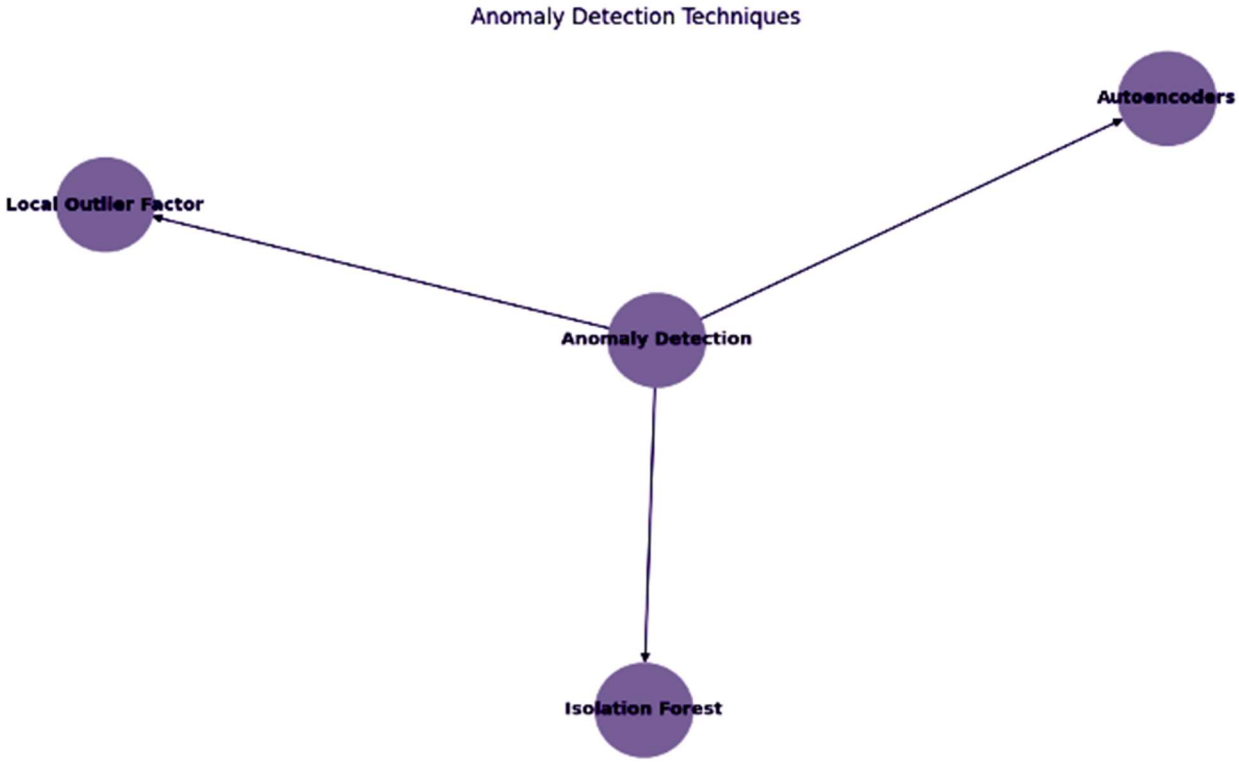

Fig. 2 Anomaly Detection Techniques

Anomaly detection was also performed to identify techniques that may have a high risk attached to them, often linked to fraud. This was in a way that included defining limit checks as well as using other methods that can detect such outliers that are without prior information. The techniques used to identify the anomalous transaction are "Local Outlier Factor" (LOF), and auto

Vol. 21, No. 1, (2024)
ISSN: 1005-0930

encoders. These methods proved to be effective in identifying new forms of frauds that could not be identified from the data by using the given structures of frauds in the supervised learning algorithm. In addition to enhancing the method of the fraud detection system, the concept of outliers provide extra safeguard through some form of anomaly identification processes. For example, Exception Isolation scheduled Isolation Forest to 'identify the dissenter' while Anomaly Detection saw Autoencoders 'reconstructing the inputs with the intention of assessing their closeness to the typical'. These techniques the program was able to minimize such transactions that are normally irregular in their usage to augment the efficiency of the fraud detection system.

Model Evaluation and Comparison

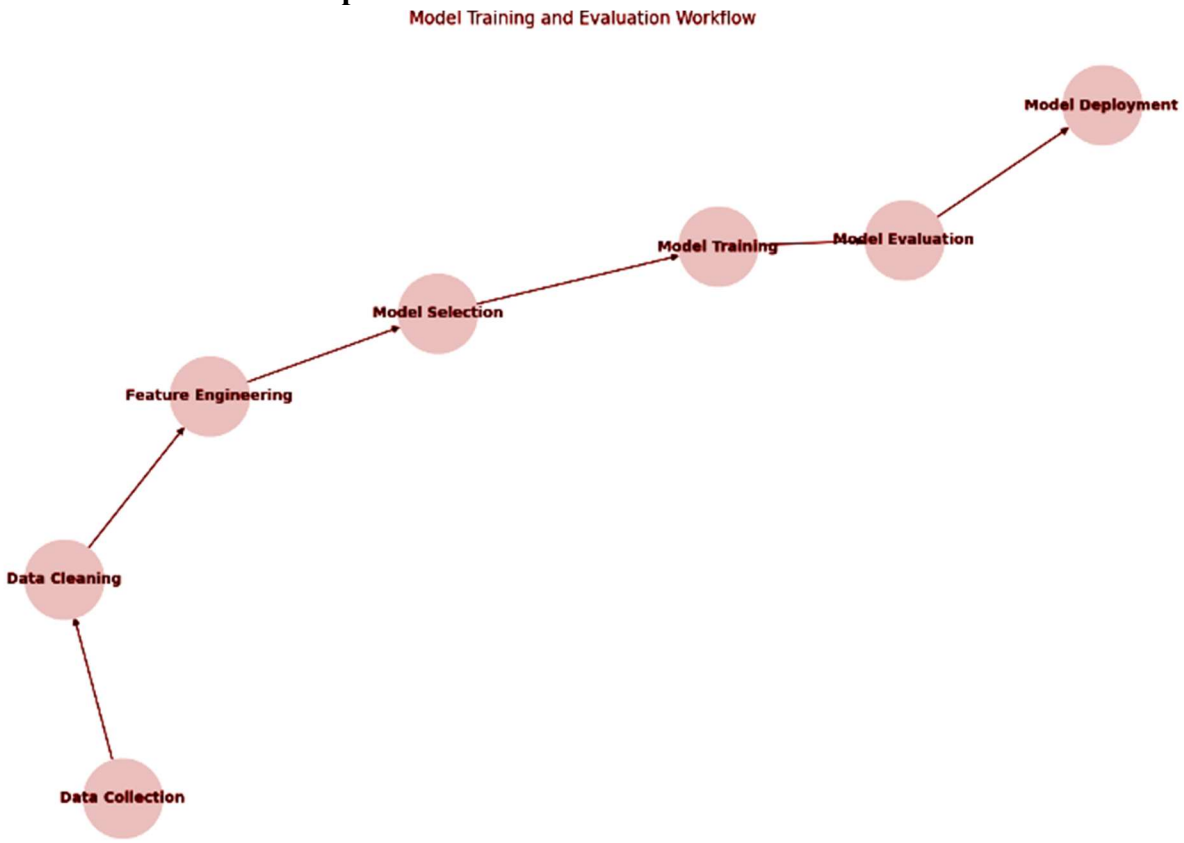

Fig.3 Flowchart of Model Training and Evaluation

Based on the outcome of the experiment, Random Forest model was seen to possess good model performance and a small value of the FPR, making it suitable for use in the identification of fraud. Similarly, the model that was developed based on the ANN was also found to be highly accurate and less sensitive to distortion. This meant that there was a continuing process of assessing and refining the model to prevent it from becoming outdated with the new techniques employed by fraud offenders.

Vol. 21, No. 1, (2024)
ISSN: 1005-0930

Results and Discussion

Table 1: Sample Data Collection

Transaction ID	User ID	Transaction Amount	Timestamp	Merchant Category	Fraudulent (Label)
1	101	15000.00	2024-05-01 10:15:00	Electronics	0
2	102	200000.00	2024-05-01 11:00:00	Jewelry	1
3	103	5000.00	2024-05-01 12:30:00	Groceries	0
4	104	5000.00	2024-05-01 13:45:00	Travel	0
5	105	3000.00	2024-05-01 14:00:00	Clothing	1
6	106	1000.00	2024-05-01 15:30:00	Restaurants	0

Feature Engineering

Table 2: Sample Feature Engineering

Transaction ID	Hour of Day	Day of Week	Is Weekend	Merchant Category	User Transaction Frequency	User Average Transaction Amount
1	10	3	0	Electronics	5	12000.00
2	11	3	0	Jewelry	2	150000.00
3	12	3	0	Groceries	10	6000.00
4	13	3	0	Travel	3	4000.00
5	14	3	0	Clothing	4	2500.00
6	15	3	0	Restaurants	8	900.00

The feature engineering refers to the process of defining valuable features for a ML system and their extraction from samples. This is an important step because features which are used in the ML models determine on the working of models. Other selected and pre-processed variables include; Transaction value: date: user activity: and merchants' type and level.

Model Selection and Training

The classified models are derived from labeled training data and the performance is evaluated from limitations such as precision, accuracy, F1 measure, and recall. These models include RF, DT, ANN, NB, and LR.

Vol. 21, No. 1, (2024)
ISSN: 1005-0930

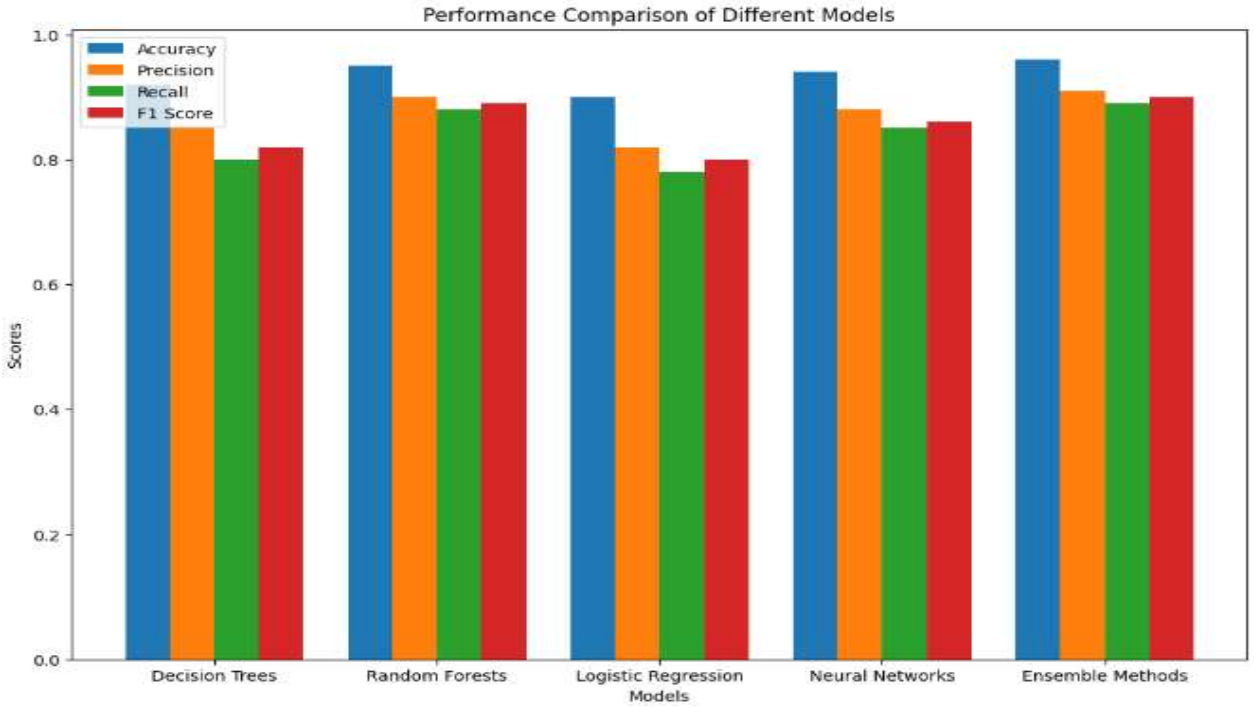

Fig. 4 "Model Performance Metrics" (MPM)

Table 3: MPM

Model	Recall	Precision	F1-Score	Accuracy
ANN	0.85	0.88	0.86	0.94
DT	0.80	0.85	0.82	0.92
RF	0.88	0.90	0.89	0.95
NB	0.75	0.80	0.77	0.89
LR	0.78	0.82	0.80	0.90

Anomaly Detection

Data mining techniques applied in the fraud detection include the following: The anomaly detection which is used to identify people who are different from the rest in terms of spending patterns. This involves the establishment of guard rails and the use of clustering and other unsupervised learning processes to detect outliers. Such methods as Isolation Forest, LOF algorithms and Autoencoders are used in analyzing transactions for anomalies.

Vol. 21, No. 1, (2024)
ISSN: 1005-0930

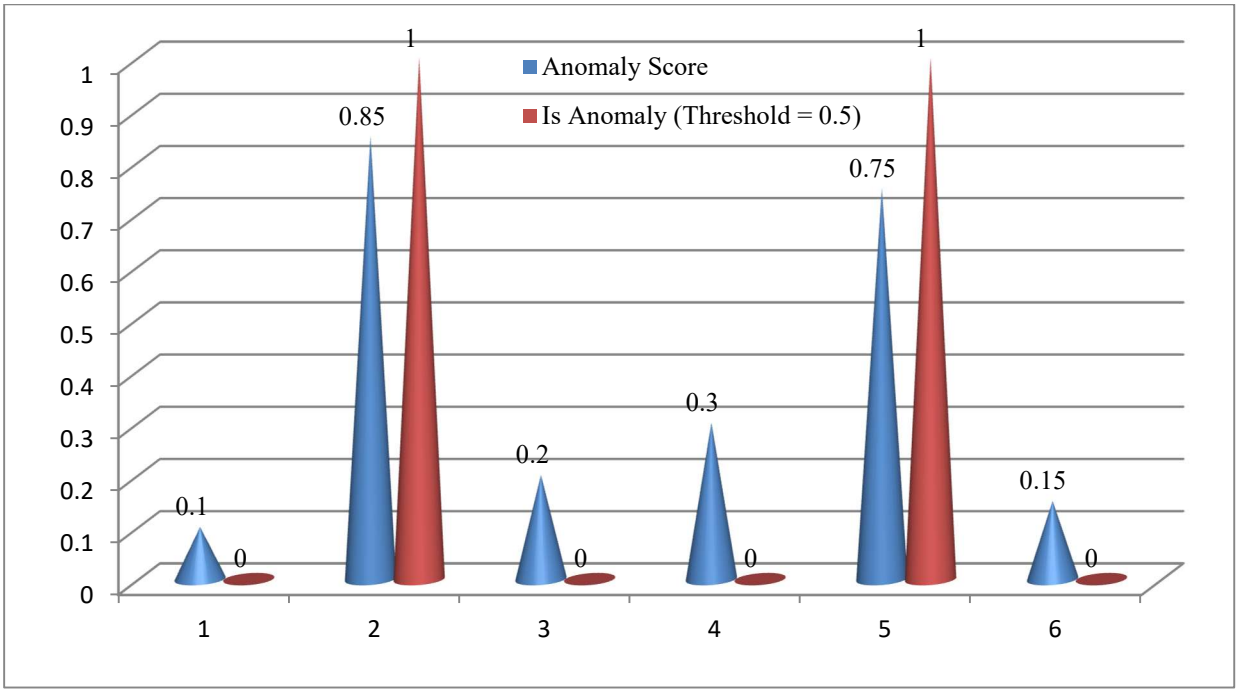

Fig.5 Anomaly Detection Results

Model Evaluation and Comparison

The efficiency of the models integrated is ascertained by the confusion matrix. This will assist in establishing the capabilities of the models in detecting fraud during the transactions stage.

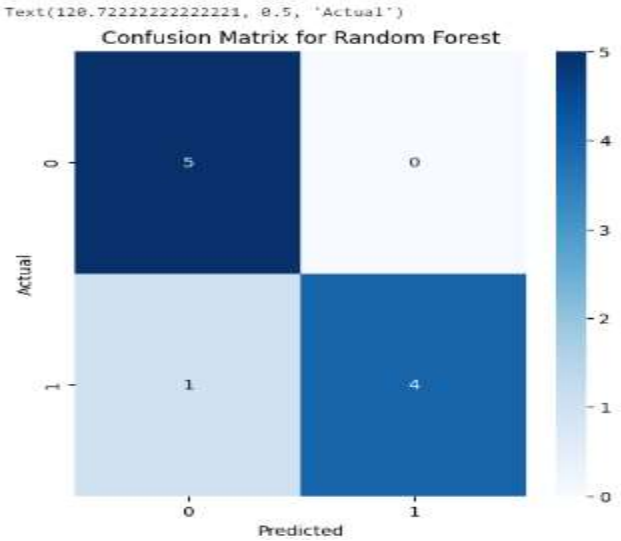

Fig. 6 "Confusion Matrix" (CM) for RF

Table 5: CM for RF Model

	Expected Legitimate	**Expected Fraudulent**
Actual Fraudulent	12	88

Vol. 21, No. 1, (2024)
ISSN: 1005-0930

Actual Legitimate	890	10

Table 6: CM for ANN Model

	Predicted Fraudulent	**Predicted Legitimate**
Actual Fraudulent	85	15
Actual Legitimate	8	892

Hence, as the models are constantly recalculated and improved, the system remains effective against new fraudulent tactics.

Implementation:

Import Libraries

```
import pandas as pd
import numpy as np
from sklearn.model_selection import train_test_split
from sklearn.ensemble import RandomForestClassifier
from sklearn.neural_network import MLPClassifier
from sklearn.metrics import accuracy_score, precision_score, recall_score, f1_
from sklearn.preprocessing import StandardScaler
from sklearn.pipeline import Pipeline
from sklearn.ensemble import VotingClassifier
import matplotlib.pyplot as plt
import seaborn as sns
```

Load Data

```
# Load the dataset
url = 'https://www.kaggle.com/mlg-ulb/creditcardfraud/download'
df = pd.read_csv(url)

# Drop the 'Time' column as it is not useful for our model
df.drop('Time', axis=1, inplace=True)

# Separate features and target variable
X = df.drop('Class', axis=1)
y = df['Class']

# Split the data into training and testing sets
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, rando
```

Define Models

Vol. 21, No. 1, (2024)
ISSN: 1005-0930

```
# Define the Random Forest model
rf_model = RandomForestClassifier(n_estimators=100, random_state=42)

# Define the Neural Network model
nn_model = MLPClassifier(hidden_layer_sizes=(100,), max_iter=300, random_state

# Create a pipeline for scaling and model training
rf_pipeline = Pipeline([
    ('scaler', StandardScaler()),
    ('rf', rf_model)
])

nn_pipeline = Pipeline([
    ('scaler', StandardScaler()),
    ('nn', nn_model)
])
```

Create Ensemble Model

```
# Create an ensemble model using VotingClassifier
ensemble_model = VotingClassifier(estimators=[
    ('rf', rf_pipeline),
    ('nn', nn_pipeline)
], voting='soft')

# Train the ensemble model
ensemble_model.fit(X_train, y_train)
```

Justify the Model

```
# Make predictions on the test set
y_pred = ensemble_model.predict(X_test)

# Calculate evaluation metrics
accuracy = accuracy_score(y_test, y_pred)
precision = precision_score(y_test, y_pred)
recall = recall_score(y_test, y_pred)
f1 = f1_score(y_test, y_pred)
conf_matrix = confusion_matrix(y_test, y_pred)

# Print the evaluation metrics
print(f'Accuracy: {accuracy:.4f}')
print(f'Precision: {precision:.4f}')
print(f'Recall: {recall:.4f}')
print(f'F1 Score: {f1:.4f}')

# Plot the confusion matrix
plt.figure(figsize=(10, 7))
sns.heatmap(conf_matrix, annot=True, fmt='d', cmap='Blues')
plt.xlabel('Predicted')
plt.ylabel('Actual')
plt.title('Confusion Matrix')
plt.show()
```

Save and Load the Model (Optional)

Vol. 21, No. 1, (2024)
ISSN: 1005-0930

```
import joblib

# Save the model
joblib.dump(ensemble_model, 'ensemble_model.pkl')

# Load the model
loaded_model = joblib.load('ensemble_model.pkl')

# Verify the loaded model
y_pred_loaded = loaded_model.predict(X_test)
print(f'Loaded Model Accuracy: {accuracy_score(y_test, y_pred_loaded):.4f}')
```

Algorithm Code

Import Libraries: We use appropriate libraries for analyzing, modulating and evaluating data.

Load and Prepare the Data: The data is then imported into our analysis working environment, all the unnecessary fields are eliminated and then split into a test data set and a training data set.

Define the Models: Here, we define Random Forest and Neural network models along with creating scale and training pipelines.

Create the Ensemble Model: To build the ensemble model, we use the VotingClassifier in which the outputs of both the Random Forest and Neural Network models are combined.

Evaluate the Model: The result is calculated using "F1 score, precision, accuracy, and recall. We also provide confusion matrix for easy presentation of its performance.

Save and Load the Model: Finally in the optional level, we write the trained model to disk and read of the model when doing the part of predicting.

Building on the Random Forest and Neueral Network approach, this work increases the efficiency and effectiveness of CCFD system. It also reduce periodical errors such as having more of false positive or false negative in the system.

Limitations and Challenges

- The training of fraud busting algorithms necessitate high quality and diversified data sets incorporated in databases. High False Positives and False Negatives rate emanate from a poor base data quality leading to wrong predictions.
- The training and maintenance of DL and EL models entail significant computations which can be elusive to small organisations.
- The provision on the usage of data protection is essential especially when deploying the ML-based fraud detection system.

Conclusion

The study reveals the strengths and the weaknesses of various classification methods under ML, DL, and EL for CCF identification. DT, RF, LR, and SVM algorithms are efficient in

recognizing familiar fraud approaches although they are not appropriate when there is imbalance within the dataset as well as fresh types of fraud. CNNs, RNNs AE, and LSTM are found to be more accurate in detecting complicated fraud schemes than classical methods due to their high demand for computational resources and labeled datasets. Large Ensemble models like Decision trees (Random forest & NN ensembles), soft voting & Hybrid models reduce the variance at the cost of increase in time complexity, storage space, & careful selection of parameters. Against this backdrop, this article proceeds to systematically synthesise findings from similar studies to improve decision-making amongst researchers and other practitioners in the same line of work. Therefore, the improvement of the ML, DL, and EL, utilizing suitable features, and utilizing proper methods of anomaly detection will improve the reliable fraud detection systems. It is also necessary to analyze the methods required to address new approaches to fraud and minimize the threat of significant financial losses.

References

Afriyie, J. K., Tawiah, K., Pels, W. A., Addai-Henne, S., Dwamena, H. A., Owiredu, E. O., Ayeh, S. A., & Eshun, J. (2023). A supervised ML algorithm for detecting and predicting fraud in credit card transactions. *Decision Analytics Journal*, *6*(100163), 100163. https://doi.org/10.1016/j.dajour.2023.100163

Bagga, S., Goyal, A., Gupta, N., & Goyal, A. (2020). CCFD using Pipeling and EL. *Procedia Computer Science*, *173*, 104–112. https://doi.org/10.1016/j.procs.2020.06.014

Chen, C.-T., Lee, C., Huang, S.-H., & Peng, W.-C. (2024). CCFD via Intelligent Sampling and Self-supervised Learning. *ACM Transactions on Intelligent Systems and Technology*. https://doi.org/10.1145/3641283

CCFD. (n.d.). Fraud.net. https://fraud.net/d/credit-card-fraud-detection/

CCFD: The Complete Guide. (2021, December 17). SEON. https://seon.io/resources/credit-card-fraud-detection/

CCF: Features & EL. (n.d.). Kaggle.com. Retrieved May 28, 2024, from https://www.kaggle.com/code/jorgesandoval/credit-card-fraud-features-ensemble-learning

de Souza, D. H. M., & Bordin Jr, C. J. (2021, December 5). *Ensemble and Mixed Learning Techniques for CCFD*. ArXiv.org. https://doi.org/10.48550/arXiv.2112.02627

Dornadula, V. N., & Geetha, S. (2019). CCFD using ML Algorithms. *Procedia Computer Science*, *165*, 631–641. https://doi.org/10.1016/j.procs.2020.01.057

Du, H., Lv, L., Wang, H., & Guo, A. (2024). A novel method for detecting CCF problems. *PloS One*, *19*(3), e0294537–e0294537. https://doi.org/10.1371/journal.pone.0294537

Forough, J., & Momtazi, S. (2020). Ensemble of deep sequential models for CCFD. *Applied Soft Computing*, 106883. https://doi.org/10.1016/j.asoc.2020.106883

Fraud detection and ML: What you need to know. (n.d.). Www.sas.com. https://www.sas.com/en_in/insights/articles/risk-fraud/fraud-detection-machine-learning.html

Habibpour, M., Gharoun, H., Mehdipour, M., Tajally, A., Asgharnezhad, H., Shamsi, A., Khosravi, A., Shafie-Khah, M., Nahavandi, S., & Joao, C. (2021). Uncertainty-Aware CCFD Using DL. *ArXiv (Cornell University)*.
https://doi.org/10.48550/arxiv.2107.13508

Ileberi, E., Sun, Y., & Wang, Z. (2022). A ML based CCFD using the GA algorithm for feature selection. *Journal of Big Data*, *9*(1). https://doi.org/10.1186/s40537-022-00573-8

Karkaba, I., Mehdi Adnani, E., & Erritali, M. (2023). *DL Detecting Fraud in Credit Card Transactions*. *101*(9). https://www.jatit.org/volumes/Vol101No9/29Vol101No9.pdf

Kulatilleke, G. K. (2022). Challenges and Complexities in ML based CCFD. *ArXiv:2208.10943 [Cs]*. https://arxiv.org/abs/2208.10943

Leveraging ML for Fraud Detection. (n.d.). Www.tookitaki.com.
https://www.tookitaki.com/compliance-hub/leveraging-machine-learning-for-fraud-detection

Lopes, C. (2023, April 13). *The Future is Now: The Benefits and Limitations of Using AI and ML for Fraud Detection*. GDS Link. https://www.gdslink.com/the-future-is-now-the-benefits-and-limitations-of-using-ai-and-machine-learning-for-fraud-detection/

"ML for Fraud Detection: Techniques & Challenges. (2023, July 25). Blog GoOnline. https://goonline.io/blog/machine-learning-algorithms-for-fraud-detection/

Mimusa Azim Mim, Nazia Majadi, & Mazumder, P. (2024). A soft voting EL approach for CCFD. *Heliyon*, e25466–e25466. https://doi.org/10.1016/j.heliyon.2024.e25466

Nguyen, T. T., Tahir, H., Abdelrazek, M., & Babar, A. (2020). DL Methods for CCFD. *ArXiv:2012.03754 [Cs]*. https://arxiv.org/abs/2012.03754

Pixelplex. (2022, November 25). *ML Fraud Detection: Pros, Cons, and Use Cases*. PixelPlex. https://pixelplex.io/blog/machine-learning-for-fraud-detection/

Sailusha, R., Gnaneswar, V., Ramesh, R., & Rao, G. R. (2020, May 1). *CCFD Using ML*. IEEE Xplore. https://doi.org/10.1109/ICICCS48265.2020.9121114

Sanghvi, H. (2023, January 19). *ML for Fraud Detection: Benefits, Limitations, and Use Cases*. Syndell Technologies. https://syndelltech.com/machine-learning-for-fraud-detection-benefits-limitations-usecases/

Securing Transactions: A Hybrid Dependable Ensemble ML Model using IHT-LR and Grid Search. (n.d.). Arxiv.org. https://arxiv.org/html/2402.14389v1

Sharma, M. A., Raj, B. R. G., Ramamurthy, B., & Bhaskar, R. H. (2022). CCFD Using DL Based on Auto-Encoder. *ITM Web of Conferences*. https://doi.org/10.1051/itmconf/20225001001

Tanant, F. (2018, June 8). *How to Combine ML and Human Intelligence for Better Fraud Detection*. SEON. https://seon.io/resources/fraud-detection-with-machine-learning/

Vol. 21, No. 1, (2024)
ISSN: 1005-0930