IJITCE

ISSN 2347- 3657

International Journal of Information Technology & Computer Engineering

Email : ijitce.editor@gmail.com or editor@ijitce.com

International Journal of Information Technology & Computer Engineering

ISSN 2347–3657

Volume 11, Issue 4, 2023

A Hybrid Deep Learning and Data Resampling Approach to Detect Credit Card Fraud

Mrs.Sujatha Godavarthi[1]., K.Aasritha[2]

1 Assistant Professor, Department of CSE, Malla Reddy College of Engineering for Women., Maisammaguda., Medchal., TS, India

2, B.Tech CSE (20RG1A0529),

Malla Reddy College of Engineering for Women., Maisammaguda., Medchal., TS, India

Abstract: Credit cards play an essential role in today's digital economy, and their usage has recently grown tremendously, accompanied by a corresponding increase in credit card fraud. Machine learning (ML) algorithms have been utilized for credit card fraud detection. However, the dynamic shopping patterns of credit card holders and the class imbalance problem have made it difficult for ML classifiers to achieve optimal performance. In order to solve this problem, this paper proposes a robust deep-learning approach that consists of long short-term memory (LSTM) and gated recurrent unit (GRU) neural networks as base learners in a stacking ensemble framework, with a multilayer perceptron (MLP) as the meta-learner. Meanwhile, the hybrid synthetic minority oversampling technique and edited nearest neighbor (SMOTE-ENN) method is employed to balance the class distribution in the dataset. The experimental results showed that combining the proposed deep learning ensemble with the SMOTE-ENN method achieved a sensitivity and specificity of 1.000 and 0.997, respectively, which is superior to other widely used ML classifiers and methods in the literature.Next we introduce advanced ensemble models, including Stacking and Voting Classifiers, evaluating them on both original and SMOTE-ENN datasets. Additionally, a Flask framework with SQLite integration enables user signup, signin, and testing for enhanced project functionality and user interaction.

Index terms -*Credit card, deep learning, ensemble learning, fraud detection, machine learning, neural network.*

1. INTRODUCTION

Information technology advancements have significantly impacted the financial sector, leading to the broad adoption of electronic commerce (e-commerce) platforms. Also, the recent outbreak of the novel coronavirus (COVID-19) pandemic has further shown the need for a more digital world and further expanded the e-commerce industry [1], [2]. One of the major issues associated with modern e-commerce is the high cases of credit card fraud [3]. Also, in the last decade, there has been an increase in credit card fraud, which is a huge burden on financial institutions [4]. The increased credit card fraud rate is associated with the expansion of e-commerce and increased online transactions. Therefore, credit card fraud detection (CCFD) is crucial for financial companies to avoid losses.

Artificial intelligence (AI) and machine learning applications in the financial sector can produce excellent results for companies, such as improved efficiency, reduced operational cost, and enhanced

customer satisfaction [5]. Several ML-based systems have been developed to detect credit card fraud. For example, Malik et al. [6] studied the use of hybrid models in CCFD. The hybrid models were achieved by combining a variety of ML algorithms, including extreme gradient boosting (XGBoost), random forest, adaptive boosting (AdaBoost), and light gradient boosting machine (LGBM). The experimental results indicated that the hybrid model based on AdaBoost and LGBM obtained the best classification performance. In a similar research work, Alfaiz and Fati [7] conducted a performance evaluation of ML classifiers and data resampling techniques for detecting credit card fraud. The classifiers used in the study include LGBM, XGBoost, random forest, categorical boosting (CatBoost), logistic regression, and naïve Bayes. The results indicated that the CatBoost classifier integrated with a k-nearest neighbor-based undersampling technique performed better than the other methods.

Meanwhile, building robust machine learning-based CCFD models has remained a challenge for some reasons. Firstly, conventional classifiers make predictions based on thetransaction details only, such as amount, transaction country, and transaction type, ignoring the sequence of transactions that defines the clients' shopping behaviour, which is useful in identifying appropriate fraud patterns [8], [9]. Secondly, credit card fraud datasets are highly imbalanced since genuine transactions significantly outnumber fraudulent transactions [10]. Imbalance classification is a predictive modelling problem where there is an uneven distribution of samples across the classes [11]. The class that makes up a large proportion of the dataset is called the majority class, while the class with a smaller proportion is called the minority class. Imbalance classification is a challenge because most ML algorithms were designed with the assumption of an even class distribution. Therefore, using imbalanced data such as the credit card dataset results in models with poor classification performance, especially for the minority class, i.e., fraudulent transactions. Furthermore, correctly identifying the minority class samples is of utmost importance in imbalance classification problems [12].

Deep learning (DL) and ensemble learning have recently dominated the ML field [13], [14], [15], [16], achieving excellent prediction performances in complex problems, and they could be applied to solve the challenges in credit card fraud detection. Deep learning, a subset of machine learning, is mainly a neural network with multiple layers [17]. Deep learning models using recurrent neural networks (RNN) have been employed for different sequential modelling-based ML tasks [18], [19], [20]. For example, Shen et al. [21] noted that algorithms that utilize sequential modelling, such as RNNs, usually perform better than conventional ML models. Meanwhile, simple RNN-based models are prone to the vanishing gradient problem, a situation where the RNN is unable to propagate relevant gradient information from the model's output end back to the layers near the input end [22]. However, LSTM and GRU-based RNNs were proposed to solve the vanishing gradient problem and have shown good performances in different sequence classification tasks [8], [23], [24].

2. LITERATURE SURVEY

the Countless research works of deep neural networks (DNNs) in the task of credit card fraud detection have

focused on improving the accuracy of point predictions and mitigating unwanted biases by building different network architectures or learning models [1]. Quantifying uncertainty accompanied by point estimation is essential because it mitigates model unfairness and permits practitioners to develop trustworthy systems which abstain from suboptimal decisions due to low confidence. Explicitly, assessing uncertainties associated with DNNs predictions is critical in real-world card fraud detection settings for characteristic reasons, including (a) fraudsters constantly change their strategies, and accordingly, DNNs encounter observations that are not generated by the same process as the training distribution, (b) owing to the time-consuming process, very few transactions are timely checked by professional experts to update DNNs [8,23,24]. Therefore, this study proposes three uncertainty quantification (UQ) techniques named Monte Carlo dropout, ensemble, and ensemble Monte Carlo dropout for card fraud detection applied on transaction data. Moreover, to evaluate the predictive uncertainty estimates, UQ confusion matrix and several performance metrics are utilized. Through experimental results, we show that the ensemble is more effective in capturing uncertainty corresponding to generated predictions. Additionally, we demonstrate that the proposed UQ methods provide extra insight to the point predictions, leading to elevate the fraud prevention process.

Credit card fraud is becoming a serious and growing problem as a result of the emergence of innovative technologies and communication methods, such as contactless payment. In this article, [2] we present an in-depth review of cutting-edge research on detecting and predicting fraudulent credit card transactions conducted from 2015 to 2021 inclusive. The selection of 40 relevant articles is reviewed and categorized according to the topics covered (class imbalance problem, feature engineering, etc.) and the machine learning technology used (modelling traditional and deep learning). Our study shows a limited investigation to date into deep learning, revealing that more research is required to address the challenges associated with detecting credit card fraud through the use of new technologies such as big data analytics, large-scale machine learning[13], [14], [15], [16], and cloud computing. Raising current research issues and highlighting future research directions, our study provides a useful source to guide academic and industrial researchers in evaluating financial fraud detection systems and designing robust solutions.

With the development of e-commerce, fraud behaviors have been becoming one of the biggest threats to the e-commerce business. [3] Fraud behaviors seriously damage the ranking system of e-commerce platforms and adversely influence the shopping experience of users. It is of great practical value to detect fraud behaviors on e-commerce platforms. However, the task is non-trivial, since the adversarial action taken by fraudsters. Existing fraud detection systems used in the e-commerce industry easily suffer from performance decay and can not adapt to the upgrade of fraud patterns, as they take already known fraud behaviors as supervision information to detect other suspicious behaviors. In this article, we propose a competitive graph neural networks (CGNN)-based fraud detection system (eFraudCom) to detect fraud behaviors at one of the largest e-commerce platforms, “Taobao”1. In the eFraudCom system, (1) the competitive graph neural

networks (CGNN) as the core part of eFraudCom can classify behaviors of users directly by modeling the distributions of normal and fraud behaviors separately; (2) some normal behaviors will be utilized as weak supervision information to guide the CGNN to build the profile for normal behaviors that are more stable than fraud behaviors [31,32]. The algorithm dependency on fraud behaviors will be eliminated, which enables eFraudCom to detect fraud behaviors in presence of the new fraud patterns; (3) the mutual information regularization term can maximize the separability between normal and fraud behaviors to further improve CGNN. eFraudCom is implemented into a prototype system and the performance of the system is evaluated by extensive experiments. The experiments on two Taobao and two public datasets demonstrate that the proposed deep framework CGNN is superior to other baselines in detecting fraud behaviors. A case study on Taobao datasets verifies that CGNN is still robust when the fraud patterns have been upgraded.

The problem of imbalanced datasets is a significant concern when creating reliable credit card fraud (CCF) detection systems. In this work, we study and evaluate recent advances in machine learning (ML) algorithms and deep reinforcement learning (DRL) used for CCF detection systems, including fraud and non-fraud labels. Based on two resampling approaches, SMOTE and ADASYN are used to resample the imbalanced CCF dataset. [4] ML algorithms are, then, applied to this balanced dataset to establish CCF detection systems. Next, DRL is employed to create detection systems based on the imbalanced CCF dataset. The diverse classification metrics are indicated to thoroughly evaluate the performance of these ML and DRL models. Through empirical experiments, we identify the reliable degree of ML models based on two resampling approaches and DRL models for CCF detection. When SMOTE and ADASYN are used to resampling original CCF datasets before training/test split, the ML models show very high outcomes of above 99% accuracy. However, when these techniques are employed to resample for only the training CCF datasets, these ML models [4] show lower results, particularly in terms of logistic regression with 1.81% precision and 3.55% F1 score for using ADASYN. Our work reveals the DRL model is ineffective and achieves low performance, with only 34.8% accuracy.

The negative effect of financial crimes on financial institutions has grown dramatically over the years. To detect crimes such as credit card fraud, several single and hybrid machine learning approaches have been used. However, these approaches have significant limitations as no further investigation on different hybrid algorithms for a given dataset were studied. This research [6] proposes and investigates seven hybrid machine learning models to detect fraudulent activities with a real word dataset. The developed hybrid models consisted of two phases, state-of-the-art machine learning algorithms were used first to detect credit card fraud, then, hybrid methods were constructed based on the best single algorithm from the first phase. Our findings indicated that the hybrid model Adaboost + LGBM is the champion model as it displayed the highest performance. Future studies should focus on studying different types of hybridization and algorithms in the credit card domain.

3. METHODOLOGY

i) Proposed Work:

The proposed system introduces a powerful solution for credit card fraud detection, harnessing the capabilities of deep learning ensembles. It combines long short-term memory (LSTM) and gated recurrent unit (GRU) neural networks as base learners in a stacking ensemble, with a multilayer perceptron (MLP) serving as the meta-learner. This approach effectively tackles the challenges of dynamic shopping patterns and class imbalance in credit card fraud detection. To mitigate class imbalance, the system employs the hybrid Synthetic Minority Oversampling Technique and Edited Nearest Neighbor (SMOTE-ENN) method. Experimental results demonstrate its superior sensitivity and specificity compared to conventional machine learning methods, making it a compelling choice for real-time fraud detection. The proposed system is compared with AdaBoost, Random Forest, MLP, LSTM, GRU models[8], [23], [24]. And then we incorporate advanced ensemble techniques such as Stacking Classifier, comprising Random Forest and MLP, and a Voting Classifier combining AdaBoost and RandomForest. These models are evaluated on both the original and SMOTE-ENN enhanced datasets. Furthermore, a Flask framework with SQLite integration has been developed, facilitating user signup, signin, and testing functionalities. This extension enhances the project's robustness, providing a comprehensive evaluation of diverse classifiers and incorporating a user-friendly interface for seamless interaction and testing.

ii) System Architecture:

The system begins by collecting credit card transaction data, which includes information on both normal and potentially fraudulent transactions. The collected data undergoes preprocessing, which involves tasks like data cleaning, handling missing values, and data transformation to ensure data quality. To address class imbalance, data sampling techniques are applied. This includes oversampling the minority class (fraudulent transactions) using methods like SMOTE-ENN[27], [28], [29]., which generates synthetic samples, and possibly undersampling the majority class to balance the dataset. Feature selection methods are employed to identify the most relevant attributes or features for fraud detection. This reduces dimensionality and focuses on the data attributes that contribute the most to the classification. The selected features are used as input for ML and DL classifiers. These classifiers are trained on the preprocessed and sampled data to learn patterns that distinguish between normal and fraudulent transactions. The system incorporates a validation phase to assess the performance of the trained classifiers. This typically involves using a separate validation dataset to evaluate the model's ability to generalize. The performance of the classifiers is evaluated using metrics such as accuracy, precision, recall, F1 score, ROC curve, AUC, sensitivity, and specificity. This evaluation is conducted for both normal and fraudulent transactions to measure the system's effectiveness. Based on the evaluation, the system generates results indicating the classification of new credit card transactions as either normal or potentially fraudulent.

International Journal of Information Technology & Computer Engineering

ISSN 2347–3657

Volume 11, Issue 4, 2023

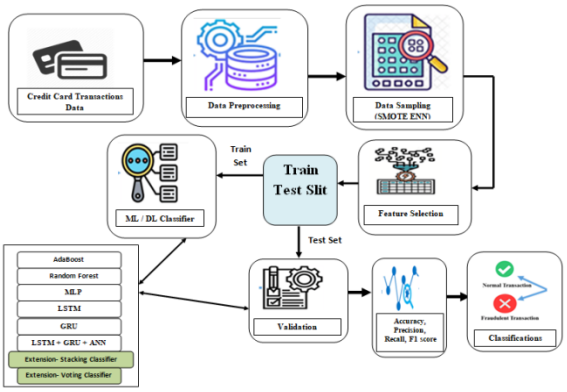

Fig 1 System Architecture

iii) Dataset collection:

The study utilizes a dataset available on Kaggle and employs data augmentation techniques to address the Problem using card fraud data, employ exploratory data analysis, and feature correlation analysis to better understand the dataset. These techniques help reveal data distributions, outliers, and relationships between variables, aiding in subsequent data processing and model building.We have used Credit Card Fraud Detection dataset taken from Kaggle to train machine learning algorithms [17]. The dataset originally had various transaction-related features, like "Amount," "Time," and "V1" to "V28." Details about the original features were kept confidential to safeguard sensitive information.

V1	V2	V3	V4	V5	V6	V7	V8	V9	V10	...	V23	V
-0.611712	-0.769705	-0.149759	-0.224877	2.028577	-2.019887	0.292491	-0.523020	0.358468	0.070050	...	0.380739	0.0234
-0.814682	1.319219	1.329415	0.027273	-0.284871	-0.653985	0.321552	0.435975	-0.704298	-0.600684	...	0.090660	0.4011
-0.318193	1.118618	0.969864	-0.127052	0.569563	-0.532484	0.706252	-0.064966	-0.463271	-0.528357	...	-0.123884	-0.4956
-1.328271	1.018378	1.775426	-1.574193	-0.117696	-0.457733	0.681867	-0.031641	0.383872	0.334853	...	-0.239197	0.0099
1.276712	0.617120	-0.578014	0.879173	0.061706	-1.472002	0.373692	-0.287204	-0.084482	-0.696578	...	-0.076738	0.2587

32 columns

Fig 2 Dataset

iv) Data Processing:

Data processing involves transforming raw data into valuable information for businesses. Generally, data scientists process data, which includes collecting, organizing, cleaning, verifying, analyzing, and converting it into readable formats such as graphs or documents. Data processing can be done using three methods i.e., manual, mechanical, and electronic. The aim is to increase the value of information and facilitate decision-making. This enables businesses to improve their operations and make timely strategic decisions. Automated data processing solutions, such as computer software programming, play a significant role in this. It can help turn large amounts of data, including big data, into meaningful insights for quality management and decision-making.

v) Feature selection:

Feature selection is the process of isolating the most consistent, non-redundant, and relevant features to use in model construction. Methodically reducing the size of datasets is important as the size and variety of datasets continue to grow. The main goal of feature selection is to improve the performance of a predictive model and reduce the computational cost of modeling.

Feature selection, one of the main components of feature engineering, is the process of selecting the most important features to input in machine learning algorithms. Feature selection techniques are employed to reduce the number of input variables by eliminating redundant or irrelevant features and narrowing down the set of features to those most relevant to the machine learning model. The main benefits of performing feature selection in advance,

ISSN 2347–3657

Volume 11, Issue 4, 2023

rather than letting the machine learning model figure out which features are most important.

vi) Algorithms:

AdaBoost, or Adaptive Boosting, is a machine learning algorithm that enhances classification accuracy by combining multiple simple models. It starts with a basic model, like a one-level decision tree, and iteratively trains new models while giving more importance to the data points that the previous models misclassified. By combining these models, AdaBoost creates a powerful ensemble that can make accurate predictions, making it valuable in your project for improving credit card fraud detection by learning from the mistakes of previous models and boosting overall performance [36].

```
from sklearn.ensemble import AdaBoostClassifier

# instantiate the model
ada = AdaBoostClassifier(n_estimators=100, random_state=0)

ada.fit(X_train, y_train)

y_pred = ada.predict(X_test)
```

Fig 3 Adaboost

Random Forest is an ensemble learning method that combines multiple decision trees to make predictions. It works by training a collection of decision trees on random subsets of the data and then averaging their predictions. This ensemble approach enhances accuracy, reduces overfitting, and provides robust performance for both classification and regression tasks.

```
from sklearn.ensemble import RandomForestClassifier

# instantiate the model
forest = RandomForestClassifier(max_depth=2, random_state=0)

forest.fit(X_train, y_train)

y_pred = forest.predict(X_test)
```

Fig 4 Random forest

The **Multilayer Perceptron** (MLP) is a type of artificial neural network used in this project for credit card fraud detection. It comprises multiple layers of interconnected neurons that process data and learn complex patterns. During training, the MLP adjusts its internal parameters to minimize prediction errors. This adaptability and its ability to capture non-linear relationships in data make the MLP an effective tool for identifying fraudulent credit card transactions.

Fig 5 MLP

LSTMs are designed to overcome the limitations of traditional RNNs when working with sequential data. They are capable of learning and remembering over long sequences, making them well-suited for various tasks like natural language processing, speech recognition, time series analysis, and more. [9] LSTMs utilize a system of cells, gates, and states to capture and propagate information over time, allowing them to model complex dependencies and patterns in sequential data effectively.

```
inputs1=Input((1,11))
att_in=LSTM(50,return_sequences=True,dropout=0.3,recurrent_dropout=0.2)(inputs1)
att_in_1=LSTM(50,return_sequences=True,dropout=0.3,recurrent_dropout=0.2)(att_in
att_out=attention()(att_in_1)
outputs1=Dense(1,activation='sigmoid',trainable=True)(att_out)
model1=Model(inputs1,outputs1)
```

Fig 6 LSTN

International Journal of Information Technology & Computer Engineering

ISSN 2347–3657

Volume 11, Issue 4, 2023

The **Stacking Classifier** is a machine learning technique that combines the predictive abilities of multiple base classifiers to create a more powerful and accurate model. In your provided code, two base classifiers, Random Forest and Multilayer Perceptron (MLP), are used within the Stacking Classifier framework. The final prediction is determined by the Light Gradient Boosting Machine (LGBM) classifier. By leveraging the diverse strengths of these classifiers, the Stacking Classifier aims to improve overall prediction performance. This ensemble approach can be valuable for addressing complex datasets and challenging classification tasks by amalgamating the knowledge from different base classifiers.

```
estimators = [('rf', RandomForestClassifier(n_estimators=10)),('mlp', MLPClassifier(random_state=1, ma
clf1 = StackingClassifier(estimators=estimators, final_estimator=LGBMClassifier(n_estimators=10))

clf1.fit(X_train,y_train)
y_pred = clf1.predict(X_test)
```

Fig 7 Stacking classifier

The**Gated Recurrent Unit (GRU)** is a recurrent neural network (RNN) architecture that excels at processing sequential data. It shares similarities with the Long Short-Term Memory (LSTM) model but is designed for more efficient computation. [8] GRU's strength lies in its ability to capture dependencies and patterns in sequences while being computationally lighter. It achieves this through a gating mechanism that controls the flow of information, allowing it to retain important details and discard less relevant information. GRU is widely used in applications like natural language processing, time series analysis, and speech recognition, where handling sequential data is crucial. Its simplicity and effectiveness make it a popular choice for various machine learning tasks.

```
inputs1=Input((1,11))
att_in=GRU(50,return_sequences=True,dropout=0.3,recurrent_dropout=0.2)(inputs1)
att_in_1=GRU(50,return_sequences=True,dropout=0.3,recurrent_dropout=0.2)(att_in)
att_out=attention()(att_in_1)
outputs1=Dense(1,activation='sigmoid',trainable=True)(att_out)
model1=Model(inputs1,outputs1)
```

Fig 8 GRU

In this project, a powerful ensemble model is crafted by combining Long Short-Term Memory (LSTM), Gated Recurrent Unit (GRU), and an Artificial Neural Network (ANN) Multilayer Perceptron (MLP). LSTM and GRU, two types of recurrent neural networks (RNNs), excel at understanding sequences and their dependencies, with LSTM being proficient at long-range connections and GRU providing computational efficiency [8], [23], [24]. The addition of MLP as the meta-learner enhances the ensemble's capacity to learn intricate patterns in credit card transaction data. This combination, known for its ability to capture both short-term and long-term dependencies, significantly boosts the accuracy and effectiveness of fraud detection in the project.

```
inputs1=Input((1,11))
att_in=LSTM(50,return_sequences=True,dropout=0.3,recurrent_dropout=0.2)(inputs1)
att_in_1=GRU(50,return_sequences=True,dropout=0.3,recurrent_dropout=0.2)(att_in)
att_out=attention()(att_in_1)
outputs = Dense(8, activation='relu')(att_out)
outputs = Dense(6, activation='relu')(att_out)
outputs1=Dense(1,activation='sigmoid',trainable=True)(att_out)
model1=Model(inputs1,outputs1)
```

Fig 9 LSTM + GRU + ANN

The Soft **Voting Classifier algorithm** is a part of ensemble learning in machine learning. In this approach, it combines the predictions from multiple individual classifiers to make a final prediction. Instead of assigning equal weight to each classifier, it takes into account the probability estimates assigned by each classifier for different classes. The algorithm then combines these probability estimates, effectively giving more weight to the classifiers that are more confident in their predictions. This results in a more refined and accurate final prediction. In the context of

International Journal of Information Technology & Computer Engineering

ISSN 2347–3657

Volume 11, Issue 4, 2023

credit card fraud detection, using a Soft Voting Classifier with diverse base classifiers like AdaBoost and Random Forest can improve the system's performance by leveraging the strengths of different models.

```
eclf2 = VotingClassifier(estimators=[('ad', clf1), ('rf', clf2)], voting='soft')
eclf2.fit(X_train, y_train)
y_pred = eclf2.predict(X_test)
```

Fig 10 Voting classifier

4. EXPERIMENTAL RESULTS

Precision: Precision evaluates the fraction of correctly classified instances or samples among the ones classified as positives. Thus, the formula to calculate the precision is given by:

Precision = True positives/ (True positives + False positives) = TP/(TP + FP)

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

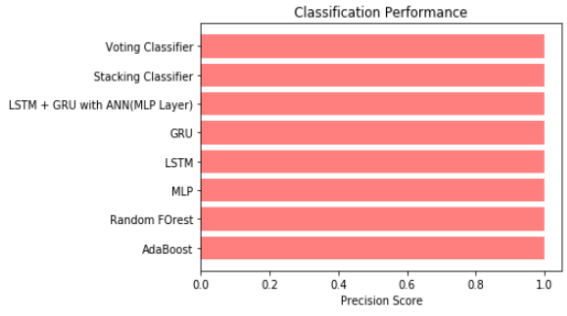

Fig 11 Precision comparison graph

Recall: Recall is a metric in machine learning that measures the ability of a model to identify all relevant instances of a particular class. It is the ratio of correctly predicted positive observations to the total actual positives, providing insights into a model's completeness in capturing instances of a given class.

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

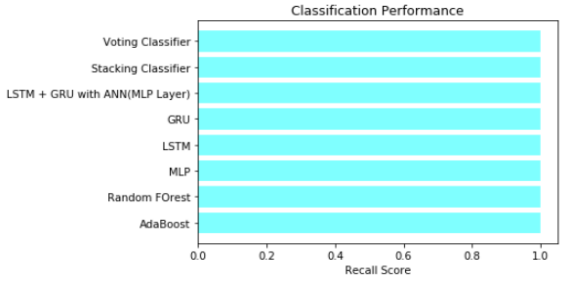

Fig 12 Recall comparison graph

Accuracy: Accuracy is the proportion of correct predictions in a classification task, measuring the overall correctness of a model's predictions.

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

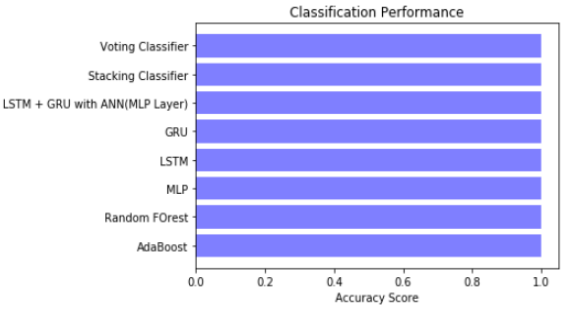

Fig 13 Accuracy graph

F1 Score: The F1 Score is the harmonic mean of precision and recall, offering a balanced measure that

International Journal of Information Technology & Computer Engineering

ISSN 2347–3657

Volume 11, Issue 4, 2023

considers both false positives and false negatives, making it suitable for imbalanced datasets.

$$\text{F1 Score} = 2 * \frac{\text{Recall} \times \text{Precision}}{\text{Recall} + \text{Precision}} * 100$$

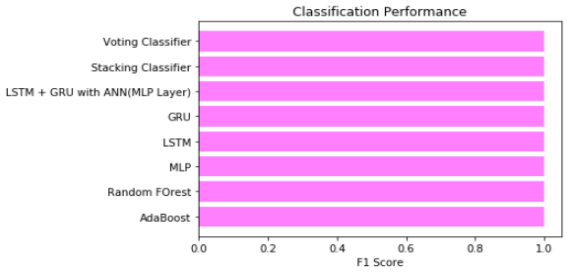

Fig 14 F1Score

	ML Model	Accuracy	Precision	Recall	F1-Score
0	AdaBoost	0.999	0.999	0.999	0.999
1	Random Forest	0.999	0.999	0.999	0.999
2	MLP	0.999	0.999	0.999	0.999
3	LSTM	0.998	1.000	0.998	0.999
4	GRU	0.998	1.000	0.998	0.999
5	LSTM + GRU with ANN(MLP Layer)	0.998	1.000	0.998	0.999
6	**Extension- Stacking Classifier**	**1.000**	**0.999**	**0.999**	**0.999**
7	**Extension- Voting Classifier**	**1.000**	**1.000**	**1.000**	**1.000**

Fig 11 Performance Evaluation original dataset

	ML Model	Accuracy	Precision	Recall	F1-Score
0	AdaBoost	0.952	0.953	0.952	0.952
1	Random Forest	0.931	0.938	0.931	0.931
2	MLP	0.999	0.999	0.999	0.999
3	LSTM	0.499	1.000	0.499	0.666
4	GRU	0.499	1.000	0.499	0.666
5	LSTM + GRU with ANN(MLP Layer)	0.499	1.000	0.499	0.666
6	**Extension- Stacking Classifier**	1.000	1.000	1.000	1.000
7	**Extension- Voting Classifier**	1.000	1.000	1.000	1.000

Fig 12 Performance Evaluation SMOTE-ENN dataset

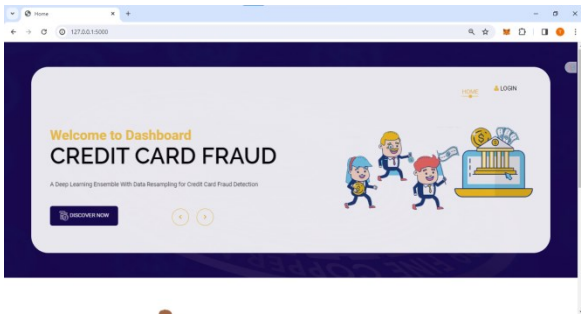

Fig 13 Home page

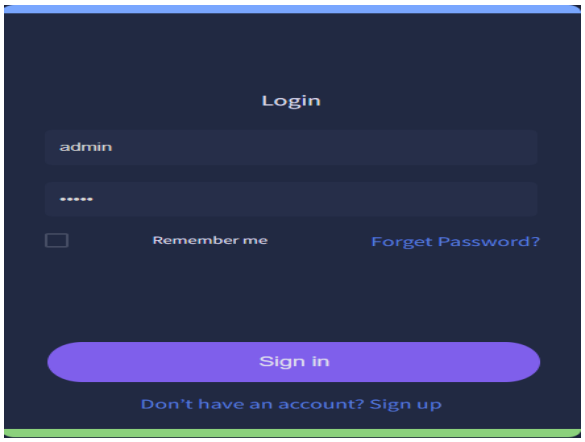

Fig 14 Login page

Fig 15User input

International Journal of Information Technology & Computer Engineering

ISSN 2347–3657

Volume 11, Issue 4, 2023

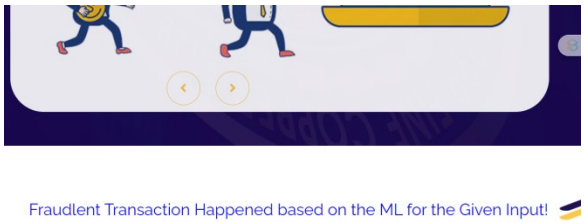

Fig 16 Predict result for given input

5. CONCLUSION

The project successfully addresses the growing challenge of credit card fraud detection in the digital era, providing a crucial solution as reliance on digital transactions continues to rise globally. Utilizing various data sampling and scaling techniques, the project ensures the dataset's optimal condition for machine learning models, reflecting the importance of meticulous data organization in enhancing model performance. Building and assessing diverse models, including AdaBoost, Random Forest, MLP, LSTM, GRU, and LSTM + GRU + MLP, revealed their effectiveness[8], [23], [24].. The subsequent introduction of voting and stacking classifiers as an extension to the project, with the Voting Classifier outperforming others, showcased improved accuracy. The incorporation of ensemble methods significantly elevated the accuracy and robustness of the fraud detection system. By emphasizing teamwork among models, the project achieved outstanding results, highlighting the potential for further advancements in the field. The integration of a user-friendly front-end interface using the Flask framework, coupled with user authentication, underscores the project's commitment to accessibility and ease of use. This approach ensures the system's practicality for users, allowing convenient interaction for input and classification of fraudulent transactions [10].

6. FUTURE SCOPE

Future research can explore enhancing model diversity by combining LSTM with various other classifiers, including random forest, logistic regression, or SVM, to further improve credit card fraud detection accuracy [34]. Conducting feature importance analysis in upcoming studies can help identify the most critical variables in credit card fraud detection, aiding in the development of more effective and efficient detection methods. Future research might delve into risk factor analysis to understand the underlying elements contributing to credit card fraud. This understanding can inform the development of more robust detection methods. Improvements to the proposed deep learning ensemble approach could involve investigating different model architectures, optimization techniques, and hyperparameter tuning methods to refine the system's performance. The proposed approach's applicability can be extended to encompass other fraud detection domains beyond credit card fraud, such as insurance fraud or online transaction fraud, contributing to a broader range of fraud prevention solutions. Additionally, exploring real-time implementation and deployment possibilities can provide immediate fraud detection and prevention in financial transactions.

REFERENCES

[1] M. Habibpour, H. Gharoun, M. Mehdipour, A. Tajally, H. Asgharnezhad, A. Shamsi, A. Khosravi, M. Shafie-Khah, S. Nahavandi, and J. P. S. Catalao, ‘‘Uncertainty-aware credit card fraud detection using deep learning,’’ 2021, arXiv:2107.13508.

[2] A. Cherif, A. Badhib, H. Ammar, S. Alshehri, M. Kalkatawi, and A. Imine, ‘‘Credit card fraud detection in the era of disruptive technologies: A systematic review,’’ J. King Saud Univ. Comput. Inf. Sci., vol. 35, no. 1, pp. 145–174, Jan. 2023, doi: 10.1016/j.jksuci.2022.11.008.

[3] G. Zhang, Z. Li, J. Huang, J. Wu, C. Zhou, and J. Yang, ‘‘eFraudCom: An e-commerce fraud detection system via competitive graph neural networks,’’ ACM Trans. Inf. Syst., vol. 40, no. 3, pp. 1–27, Mar. 2022, doi: 10.1145/3474379.

[4] T. K. Dang, T. C. Tran, L. M. Tuan, and M. V. Tiep, ‘‘Machine learning based on resampling approaches and deep reinforcement learning for credit card fraud detection systems,’’ Appl. Sci., vol. 11, no. 21, p. 10004, Oct. 2021, doi: 10.3390/app112110004.

[5] J. Chaquet-Ulldemolins, F.-J. Gimeno-Blanes, S. Moral-Rubio, S. MuñozRomero, and J.-L. Rojo-Álvarez, ‘‘On the black-box challenge for fraud detection using machine learning (I): Linear models and informative feature selection,’’ Appl. Sci., vol. 12, no. 7, p. 3328, Mar. 2022, doi: 10.3390/app12073328.

[6] E. F. Malik, K. W. Khaw, B. Belaton, W. P. Wong, and X. Chew, ‘‘Credit card fraud detection using a new hybrid machine learning architecture,’’ Mathematics, vol. 10, no. 9, p. 1480, Apr. 2022, doi: 10.3390/math10091480.

[7] N. S. Alfaiz and S. M. Fati, ‘‘Enhanced credit card fraud detection model using machine learning,’’ Electronics, vol. 11, no. 4, p. 662, Feb. 2022, doi: 10.3390/electronics11040662.

[8] I. Benchaji, S. Douzi, B. El Ouahidi, and J. Jaafari, ‘‘Enhanced credit card fraud detection based on attention mechanism and LSTM deep model,’’ J. Big Data, vol. 8, no. 1, p. 151, Dec. 2021, doi: 10.1186/s40537-021- 00541-8.

[9] E. Esenogho, I. D. Mienye, T. G. Swart, K. Aruleba, and G. Obaido, ‘‘A neural network ensemble with feature engineering for improved credit card fraud detection,’’ IEEE Access, vol. 10, pp. 16400–16407, 2022, doi: 10.1109/ACCESS.2022.3148298.

[10] E. Btoush, X. Zhou, R. Gururaian, K. Chan, and X. Tao, ‘‘A survey on credit card fraud detection techniques in banking industry for cyber security,’’ in Proc. 8th Int. Conf. Behav. Social Comput. (BESC), Oct. 2021, pp. 1–7, doi: 10.1109/BESC53957.2021.9635559.

[11] I. D. Mienye and Y. Sun, ‘‘Performance analysis of cost-sensitive learning methods with application to imbalanced medical data,’’ Inform. Med. Unlocked, vol. 25, Jan. 2021, Art. no. 100690, doi: 10.1016/j.imu.2021.100690.

[12] S. A. Ebiaredoh-Mienye, T. G. Swart, E. Esenogho, and I. D. Mienye, ‘‘A machine learning method with filter-based feature selection for improved prediction of chronic kidney disease,’’ Bioengineering, vol. 9, no. 8, p. 350, Jul. 2022, doi: 10.3390/bioengineering9080350.

[13] C. Ho, Z. Zhao, X. F. Chen, J. Sauer, S. A. Saraf, R. Jialdasani, K. Taghipour, A. Sathe, L.-Y. Khor, K.-H. Lim, and W.-Q. Leow, ‘‘A promising deep learning-assistive algorithm for histopathological screening of colorectal cancer,’’

Sci. Rep., vol. 12, no. 1, pp. 1–9, Feb. 2022, doi: 10.1038/s41598-022-06264-x.

[14] P. Goel, R. Jain, A. Nayyar, S. Singhal, and M. Srivastava, ‘‘Sarcasm detection using deep learning and ensemble learning,’’ Multimedia Tools Appl., vol. 81, no. 30, pp. 43229–43252, Dec. 2022, doi: 10.1007/s11042- 022-12930-z.

[15] I. D. Mienye, P. Kenneth Ainah, I. D. Emmanuel, and E. Esenogho, ‘‘Sparse noise minimization in image classification using genetic algorithm and DenseNet,’’ in Proc. Conf. Inf. Commun. Technol. Soc. (ICTAS), Mar. 2021, pp. 103–108, doi: 10.1109/ICTAS50802.2021.9395014

[16] R. T. Aruleba, T. A. Adekiya, N. Ayawei, G. Obaido, K. Aruleba, I. D. Mienye, I. Aruleba, and B. Ogbuokiri, ‘‘COVID-19 diagnosis: A review of rapid antigen, RT-PCR and artificial intelligence methods,’’ Bioengineering, vol. 9, no. 4, p. 153, Apr. 2022, doi: 10.3390/bioengineering9040153.

[17] G. Nguyen, S. Dlugolinsky, M. Bobák, V. Tran, L. L. García, I. Heredia, P. Malík, and L. Hluchý, ‘‘Machine learning and deep learning frameworks and libraries for large-scale data mining: A survey,’’ Artif. Intell. Rev., vol. 52, no. 1, pp. 77–124, Jun. 2019. [Online]. Available: http://link.springer.com/10.1007/s10462-018-09679-z

[18] S. O. Alhumoud and A. A. Al Wazrah, ‘‘Arabic sentiment analysis using recurrent neural networks: A review,’’ Artif. Intell. Rev., vol. 55, no. 1, pp. 707–748, Jan. 2022, doi: 10.1007/s10462- 021-09989-9.

[19] Z. Zhong, Y. Gao, Y. Zheng, B. Zheng, and I. Sato, ‘‘Real-world video deblurring: A benchmark dataset and an efficient recurrent neural network,’’ Int. J. Comput. Vis., vol. 131, no. 1, pp. 284–301, Jan. 2023, doi: 10.1007/s11263-022-01705-6.

[20] J. Van Gompel, D. Spina, and C. Develder, ‘‘Satellite based fault diagnosis of photovoltaic systems using recurrent neural networks,’’ Appl. Energy, vol. 305, Jan. 2022, Art. no. 117874, doi: 10.1016/j.apenergy.2021.117874.

[21] F. Shen, X. Zhao, G. Kou, and F. E. Alsaadi, ‘‘A new deep learning ensemble credit risk evaluation model with an improved synthetic minority oversampling technique,’’ Appl. Soft Comput., vol. 98, Jan. 2021, Art. no. 106852, doi: 10.1016/j.asoc.2020.106852.

[22] A. Tsantekidis, N. Passalis, and A. Tefas, ‘‘Chapter 5—Recurrent neural networks,’’ in Deep Learning for Robot Perception and Cognition, A. Iosifidis and A. Tefas, Eds. New York, NY, USA: Academic, 2022, pp. 101–115, doi: 10.1016/B978-0-32-385787-1.00010-5.

[23] Y. Xie, G. Liu, C. Yan, C. Jiang, M. Zhou, and M. Li, ‘‘Learning transactional behavioral representations for credit card fraud detection,’’ IEEE Trans. Neural Netw. Learn. Syst., early access, Oct. 5, 2022, doi: 10.1109/TNNLS.2022.3208967.

[24] Y.-C. Wei, Y.-X. Lai, and M.-E. Wu, ‘‘An evaluation of deep learning models for chargeback fraud detection in online games,’’ Cluster Comput., vol. 26, pp. 927–943, Jul. 2022, doi: 10.1007/s10586-022-03674-4.

[25] S. Mishra, K. Shaw, D. Mishra, S. Patil, K. Kotecha, S. Kumar, and S. Bajaj, ‘‘Improving the accuracy of ensemble machine learning classification models using a novel bit-fusion algorithm for

healthcare AI systems,'' Frontiers Public Health, vol. 10, May 2022, Art. no. 858282. [Online]. Available: https://www.frontiersin. org/articles/10.3389/fpubh.2022.858282

[26] I. D. Mienye, G. Obaido, K. Aruleba, and O. A. Dada, ''Enhanced prediction of chronic kidney disease using feature selection and boosted classifiers,'' in Intelligent Systems Design and Applications. Cham, Switzerland: Springer, 2022, pp. 527–537, doi: 10.1007/978-3-030-96308- 8_49.

[27] N. L. Fitriyani, M. Syafrudin, G. Alfian, C.-K. Yang, J. Rhee, and S. M. Ulyah, ''Chronic disease prediction model using integration of DBSCAN, SMOTE-ENN, and random forest,'' in Proc. ASU Int. Conf. Emerg. Technol. Sustainability Intell. Syst. (ICETSIS), Jun. 2022, pp. 289–294, doi: 10.1109/ICETSIS55481.2022.9888806.

[28] J. Yang and J. Guan, ''A heart disease prediction model based on feature optimization and smote-Xgboost algorithm,'' Information, vol. 13, no. 10, p. 475, Oct. 2022, doi: 10.3390/info13100475.

[29] D. S. Sisodia, N. K. Reddy, and S. Bhandari, ''Performance evaluation of class balancing techniques for credit card fraud detection,'' in Proc. IEEE Int. Conf. Power, Control, Signals Instrum. Eng. (ICPCSI), Sep. 2017, pp. 2747–2752, doi: 10.1109/ICPCSI.2017.8392219.

[30] H. Guan, Y. Zhang, M. Xian, H. D. Cheng, and X. Tang, ''SMOTE-WENN: Solving class imbalance and small sample problems by oversampling and distance scaling,'' Int. J. Speech Technol., vol. 51, no. 3, pp. 1394–1409, Mar. 2021, doi: 10.1007/s10489-020-01852-8.

[31] L. Ni, J. Li, H. Xu, X. Wang, and J. Zhang, ''Fraud feature boosting mechanism and spiral oversampling balancing technique for credit card fraud detection,'' IEEE Trans. Computat. Social Syst., early access, Feb. 13, 2023, doi: 10.1109/TCSS.2023.3242149.

[32] H. Fanai and H. Abbasimehr, ''A novel combined approach based on deep autoencoder and deep classifiers for credit card fraud detection,'' Exp. Syst. Appl., vol. 217, May 2023, Art. no. 119562, doi: 10.1016/j.eswa.2023.119562.

[33] F. Itoo, Meenakshi, and S. Singh, ''Comparison and analysis of logistic regression, Naïve Bayes and KNN machine learning algorithms for credit card fraud detection,'' Int. J. Inf. Technol., vol. 13, no. 4, pp. 1503–1511, Aug. 2021, doi: 10.1007/s41870-020-00430-y.

[34] S. K. Saddam Hussain, E. Sai Charan Reddy, K. G. Akshay, and T. Akanksha, ''Fraud detection in credit card transactions using SVM and random forest algorithms,'' in Proc. 5th Int. Conf. I-SMAC (IoT Social, Mobile, Analytics Cloud) (I-SMAC), Nov. 2021, pp. 1013–1017, doi: 10.1109/I-SMAC52330.2021.9640631.

[35] I. D. Mienye, Y. Sun, Z. Wang, ''Prediction performance of improved decision tree-based algorithms: A review,'' Proc. Manuf., vol. 35, pp. 698–703, Jan. 2019, doi: 10.1016/j.promfg.2019.06.011.

[36] K. Randhawa, C. K. Loo, M. Seera, C. P. Lim, and A. K. Nandi, ''Credit card fraud detection using AdaBoost and majority voting,'' IEEE Access, vol. 6, pp. 14277–14284, 2018, doi: 10.1109/ACCESS. 2018.2806420.

ISSN 2347–3657

Volume 11, Issue 4, 2023

[37] T.-H. Lin and J.-R. Jiang, ‘‘Credit card fraud detection with autoencoder and probabilistic random forest,’’ Mathematics, vol. 9, no. 21, p. 2683, Oct. 2021, doi: 10.3390/math9212683.

[38] A. Rb and S. K. Kr, ‘‘Credit card fraud detection using artificial neural network,’’ Global Transitions Proc., vol. 2, no. 1, pp. 35–41, Jun. 2021, doi: 10.1016/j.gltp.2021.01.006.

[39] S. C. Dubey, K. S. Mundhe, and A. A. Kadam, ‘‘Credit card fraud detection using artificial neural network and BackPropagation,’’ in Proc. 4th Int. Conf. Intell. Comput. Control Syst. (ICICCS), May 2020, pp. 268–273, doi: 10.1109/ICICCS48265.2020.9120957.

[40] O. N. Akande, S. Misra, H. B. Akande, J. Oluranti, and R. Damasevicius, ‘‘A supervised approach to credit card fraud detection using an artificial neural network,’’ in Applied Informatics. Cham, Switzerland: Springer, 2021, pp. 13–25, doi: 10.1007/978-3-030-89654-6_2.