ISSN : 2455-9679
SJIF Impact Factor : 6.008

IJRTSM
INTERNATIONAL JOURNAL OF RECENT TECHNOLOGY SCIENCE & MANAGEMENT

"OPTIMIZING CREDIT CARD FRAUD TRANSACTIONS IDENTIFICATION AND CLASSIFICATION IN BANKING INDUSTRY USING MACHINE LEARNING ALGORITHMS"

Honie Kali [1]

[1] *Independent Researcher, honieresearch@gmail.com*

ABSTRACT

The banking industry is increasingly confronted with the problem of credit card theft. The ability to identify fraudulent transactions is crucial for credit card companies to prevent consumers from being charged for items that were not really purchased. There was a time when several classifiers and machine learning algorithms were used to detect fraudulent transactions. By comparing Logistic Regression (LR), Convolutional Neural Networks (CNNs), and Long Short-Term Memory (LSTM) networks, this paper proposes the best framework for detecting fraud. In addition to data balancing utilizing SMOTE to handle class imbalance, the approach incorporates data pretreatment processes such as managing missing values, removing outliers, and min-max normalization. To improve model efficiency, feature selection approaches are used. Important evaluation metrics, including F1-score, AUC, recall, and accuracy, are satisfied by the LSTM model, which attains an accuracy of 92.54%. Through the identification of temporal correlations in transaction sequences, LSTM improves fraud detection while reducing false positives and negatives, as shown in experiments. This study improves financial security by providing a robust, data-driven framework for fraud detection suitable for real-world banking applications.

Key Words: *Credit card transaction, Banking, European dataset, Artificial intelligence, SMOTE, Machine Learning*.

I. INTRODUCTION

The banking industry is one of the most volatile and risk-prone sectors, facing numerous challenges, including the surge in digital transactions, which has caused a noticeable increase in credit card fraud [1]. Cashless transactions are progressively replacing traditional cash payments as technology advances, providing customers with a simple option to undertake financial operations using contactless and internet payment systems [2]. However, this shift has also provided opportunities for fraudulent activities, where scammers exploit vulnerabilities to carry out illegal activities and steal credit card information [3][4][5]. The rising default rate brought on by the growing number of credit card users has made measures to prevent fraud even more difficult.

The issue of financial fraud affects the financial industry in a significant way and day-to-day living. Fraud may impact people's cost of living, undermine economies, and erode trust in business [6]. The manual methods used in traditional procedures, such as audits, were ineffective and unreliable because of the complexity of the issue [7].

The illicit use of credit or debit cards to gain funds, products, or services is known as credit card fraud [8]. Financial fraud has far-reaching consequences that erode consumer confidence and destabilize economies. It includes identity theft, money laundering, credit card fraud, and tax evasion. Traditionally, banks and financial institutions used manual audits and rule-based techniques to discover fraud. However, these methods are reactive, time-consuming, and lack flexibility in adapting to evolving fraud techniques [9][10]. With the rapid advancements in technology, financial institutions are shifting towards proactive, AI-driven approaches to more effectively identify fraudulent transactions.

Computers may learn from previous transactions and increase forecast accuracy without explicit programming thanks

to ML, a kind of AI [11][12]. Systems for detecting fraud based on machine learning can find hidden patterns in big datasets, enhancing fraud detection and classification capabilities. By leveraging ML techniques, banks and e-commerce platforms can develop automated fraud detection models that minimize false positives and improve real-time transaction monitoring [13]. To increase the precision of fraud detecting and lower financial losses in the banking sector, this study employs ML approaches to optimize recognizing and classification of fraudulent credit card transactions.

A. Significance and Contribution

This technology is significant because it uses cutting-edge ML methods to address the banking industry's growing problem of spotting fraudulent credit card purchases. This method combines an LSTM for fraud categorization and compares it with CNN and LR in order to improve decision-making and make fraud detection more accurate while lowering false positives. Using data balancing strategies like SMOTE guarantees a more accurate detection model even when there are class imbalances, which eventually leads to safer financial transactions and lower losses. The main contributions are:

- Using the European Credit Card Database for Identifying Bank Credit Card Fraud Transactions.
- Utilizes SMOTE for data balancing, tackling the issue of class disparity that commonly comes up during fraud detection.
- Evaluates how well LSTM performs in comparison to CNN and LR to determine the value of several different ML methods for detecting fraud.
- Evaluates the model's performance using a number of measures, including F1-score, AUC, recall, accuracy, precision, and loss.

B. Justification and Novelty

This research is justified by the growing prevalence and sophistication of credit card fraud, underscoring the need for sophisticated fraud detection tools that can manage high transaction volumes. The study offers a fresh strategy for improving ML-based fraud identification in the banking industry, using the European credit card dataset to increase classification accuracy. The incorporation of an LSTM for fraud categorization is what makes this study innovative. By incorporating feature selection, hyperparameter tuning, and evaluating performance with various metrics, the study offers a robust, adaptive, and efficient solution for eventually identifying credit card fraud contributes to increased financial transaction security and dependability.

C. Structure of the Paper

The study is set up as follows: Section II presents pertinent studies on credit card fraud transactions. Section III explains in great depth the tools and methods utilized. The experimental results of the suggested system are shown in Section IV. Section V concludes the investigation and provides an overview of its findings.

II. LITERATURE REVIEW

A few review articles on ML methods used by banks to handle credit card fraud transactions are included in this area. The Table I highlights the paper, methods, dataset, key findings, and limitations/future work.

Sahu and Sahu (2023) examine many ML methods and show that the KNN approach works best, with an accuracy rate of 99.9%. This offers a comprehensive analysis of how well different ML techniques work, which is beneficial for creating apps that detect credit card fraud. The NB approach also attains the lowest accuracy of 98.6%, as this article shows [14].

Yuhes Raajha et al. 2023 provide an overview of the newest developments in credit card theft and real-time fraud detection methods. There were credit card transactions that were checked for fraud using machine learning methods like SVM, LR, RF, and FSVM on a collection of those transactions. It was tested using different classification methods to see how precise, sensitive, specific, and accurate the credit card fraud detection system was. FSVM did better than the other algorithms, as shown by the comparison test, which showed a 98.61% success rate [15].

Geetha et al. (2023) article discusses the use of ML methods to develop APIs that identify security flaws. Their API typically identifies harmful URLs and fraudulent credit card transactions and hosts them online, saving users from

ISSN : 2455-9679
SJIF Impact Factor : 6.008

having to download software files locally. All results from earlier models, which provided a general accuracy of 70% to 90%, are consistent with this research [16].

Ahmed and Saini (2023) six supervised ML algorithms NB, SVM, RF, KNN, LR, and Boost presented in this study are utilized to create a classification model capable of accurately detecting this kind of fraud. When it comes to accuracy, SVMs are the best. Findings from comparing all of these algorithms' capabilities to detect Financial transactions that are both criminal and genuine [17].

Singh et al. (2022) article aims at the newest developments and uses of machine learning to find credit card fraud. This study looks at how accurate four different machine learning methods are. Has a track record of 99.87% accuracy. Researchers have shown that the Cat boost algorithm is the best way to spot credit card scams. The dataset that is used to find credit card fraud is open to everyone through Kaggle [18].

The accuracy of traditional techniques for identifying credit card fraud is problematic, adaptability, and are unable to minimize false positives while maintaining high fraud detection rates. Many existing machine learning models struggle to balance sensitivity and specificity, often resulting in either undetected fraudulent transactions or an excessive number of flagged legitimate transactions. Furthermore, dealing with imbalanced datasets is still a difficult problem that affects the whole effectiveness of fraud detection systems. To get over these restrictions, a brand-new LSTM network-based fraud detection paradigm is suggested. Fraud identification was improved by using this approach, which combines the advanced feature selection, synthetic data balancing techniques, namely, SMOTE, and a DL-based classification approach.

Table I Summary of background study for credit card fraud transactions in banking industry using machine learning

Authors	Methods	Dataset	Key Findings	Limitations & Future Gaps
Sahu and Sahu (2023)	KNN, various ML algorithms	Kaggle	KNN achieved the highest accuracy of 99.9%; the lowest was 98.6% for Naive Bayes.	Dataset details and generalizability of the results not discussed; performance on imbalanced data unknown.
Yuhes Raajha et al. (2023)	FSVM, RF, LR, SVM	Credit card user transaction dataset	FSVM outperformed others with an accuracy of 98.61%.	Real-time performance and scalability of FSVM not elaborated; lacks discussion on class imbalance handling.
Geetha et al. (2023)	ML algorithms with API deployment	Previous models' datasets	API detects malicious URLs and fraudulent transactions with 70%-90% accuracy. Hosted online to avoid local installation.	Lower accuracy range suggests need for improved feature engineering and model tuning.
Ahmed and Saini (2023)	Naive Bayes, SVM, RF, KNN, LR, XGBoost	Kaggle	SVM was the most reliable with the highest accuracy.	Dataset type not mentioned; lacks comparison on execution time and computational cost.
Singh et al. (2022)	CatBoost, and three other unnamed algorithms	Kaggle credit card fraud dataset	CatBoost achieved 99.87% accuracy and was the best performer.	Specific ML methods not all named; lacks detailed explanation of preprocessing and feature selection.

III. METHODOLOGY

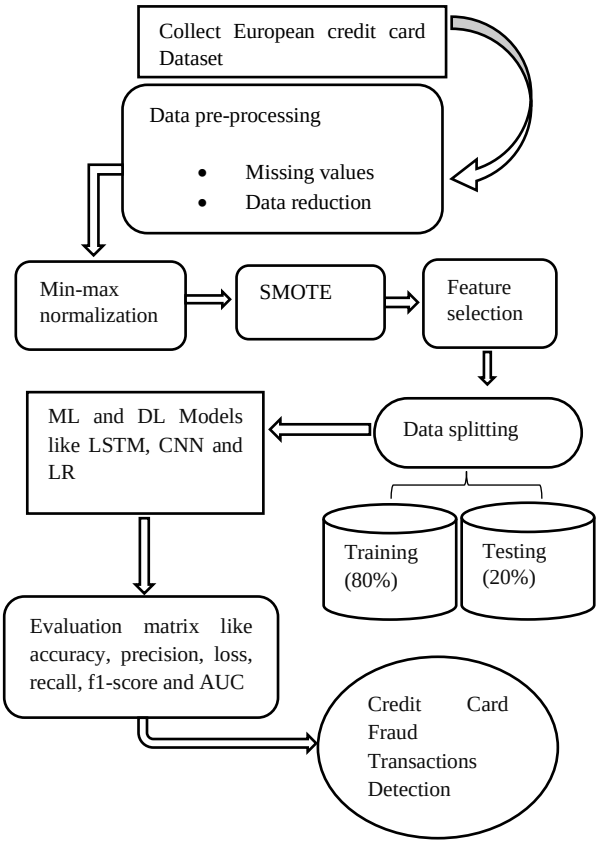
Financial institutions can increase their capacity to identify and categories credit card fraud transactions by utilizing ML algorithms. Preprocessing procedures, including eliminating outliers, decreasing data, and managing missing values, are used to manage data of interest from the dataset of credit cards in Europe. In the event of a class imbalance, they use SMOTE in conjunction with min-max normalization to balance the data. Finding the most important attributes using feature selection methods gives you the best model performance. Next, an 80:20 training-testing ratio is used on the dataset. In this, various models like CNN and LR are employed to classify fraud, along with the suggested model, LSTM, to mention a few. In order to increase predictive power during model training and optimization, hyperparameter tweaking is utilized. Recall, F1-score, loss, and area under the curve (AUC) are some of the metrics used by banks to judge how well their scam detection and classification systems work. Figure 1 shows the flowchart that is used to spot activities that involve credit card fraud.

A. Data Collection

The Kaggle platform was used to get the European credit card information. Out of the 284,807 events in the dataset,

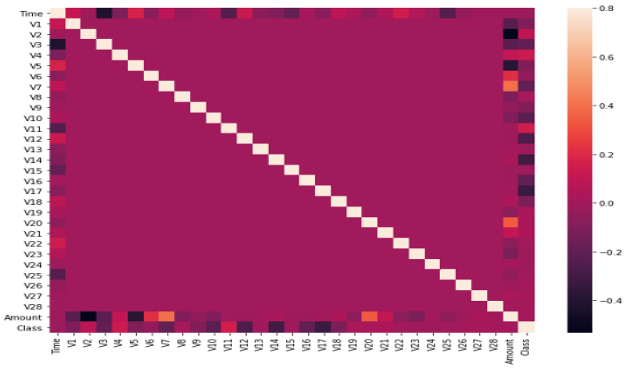
ISSN : 2455-9679
SJIF Impact Factor : 6.008

only 492 have been confirmed as fake. The other 284,315 have been confirmed as real. There is a notable imbalance in the dataset. The dataset has 31 characteristics, such as 'Amount' for transaction amount, 'Time' for transaction time, and 28 more parameters designated V1 through V28. Every transaction's attribute is determined by its class, with a binary value of 0 denoting a valid transaction and 1 denoting an illegal one. The data visualization graphics are provided below:



Flowchart for Credit Card Fraud Transactions Identification

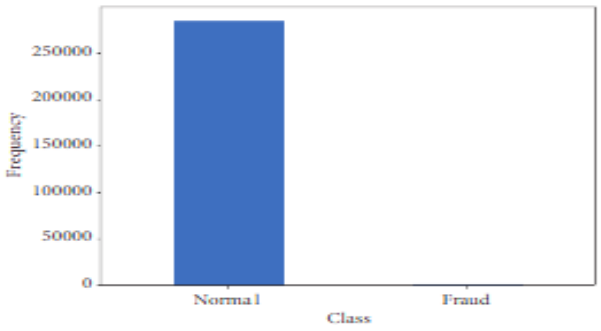
The following steps of the proposed methodology are explained in below:



Correlation Matrix

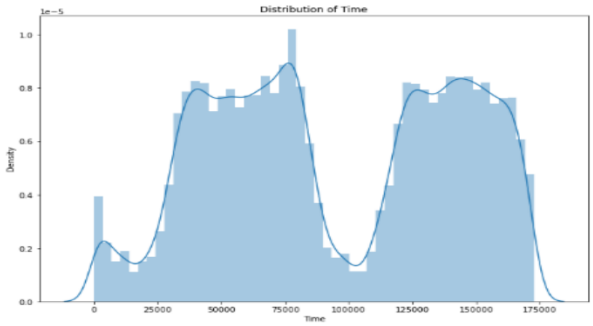
Figure 2 illustrates a correlation matrix showing the links between attributes for the European credit card dataset. The heatmap indicates correlation strength (-0.4 to -0.8), with diagonal values showing perfect self-correlation. Notably, the "Class" variable, representing fraud, correlates with features like "V1" and "V28," suggesting their significance in fraud detection. The "Amount" variable also shows some correlation with "Class" and other features, indicating its potential relevance.

ISSN : 2455-9679
SJIF Impact Factor : 6.008



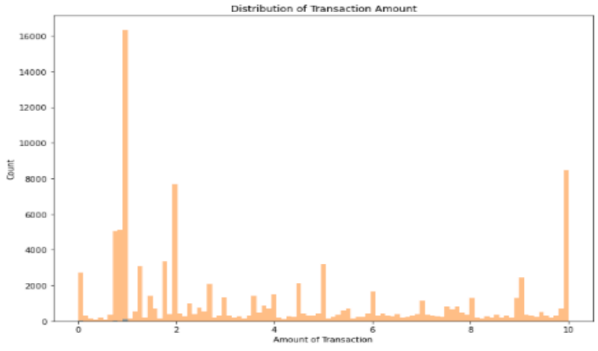
Distribution of European Credit card Dataset

This bar graph, Figure 3, depicts the frequency distribution of two classes, "Normal" and "Fraud," with "Normal" exhibiting a significantly higher frequency, exceeding 250,000, while "Fraud" shows a frequency close to zero, highlighting a severe class imbalance in the dataset.



Transaction Time of European Dataset

The temporal distribution's low peak value is seen in Figure 4 due to the notable distinction between daytime and nighttime transactions. The density plot's x-axis displays the transaction time, while the y-axis displays the attribute density.



Transaction Amount of European Dataset

Figure 5 shows the total amount of money that was exchanged. Most transactions are tiny, and only a small percentage approach the maximum transaction value.

B Data Preprocessing

The pre-processing effort will be given special attention because real-world banking data is frequently noisy. This includes resolving missing values as well as shrinking and optimizing the dataset for subsequent model deployment. The most pertinent pre-processing techniques found in the literature will be combined and used in this study, even if certain pre-processing procedures are predicated on knowledge of the data and context. The pre-processing of European

ISSN : 2455-9679
SJIF Impact Factor : 6.008

credit card data is provided below:

- **Missing values:** Missing value handling is the initial stage in data cleansing. The term "missing values" describes the voluntary or involuntary removal of data from a record. Although identifying and encoding missing data is the first stage, resolving the missing values is the second.
- **Data reduction:** It is crucial to optimize the feature and, in this example, lower the number of unique values for categorical variables once the data has been cleared of missing values and other potential biases. To put related observations into a single group (cluster), clustering was done.
- **Outlier Removal:** The majority of the numerical variables in the section were found to be severely skewed with substantial kurtosis based on the preliminary investigation. The cube root approach was used to handle negatively skewed features, whereas the log(x+1) transformation was used to treat positively skewed variables. Log(x+1) will be utilized in place of log(x) to preserve the zero in the data and prevent any issues with missing values

C. Min-Max Normalization

In real-world datasets, most properties will vary in terms of magnitude, range, and unit. When one characteristic is larger than the others, it becomes problematic since it will inevitably take precedence over the others. Therefore, in order to match classification algorithms and remove the influence of different quantitative units, raw data needs be scaled. Therefore, the features in this study were rescaled between 0 and 1 using the Min Max scaler approach. Because it employs statistical methods that do not alter the variance of the data, this method has the advantage of being resilient against outliers Equation (1).

$$x' = x - \frac{\min(x)}{\max(x)} - \min(x) \qquad (1)$$

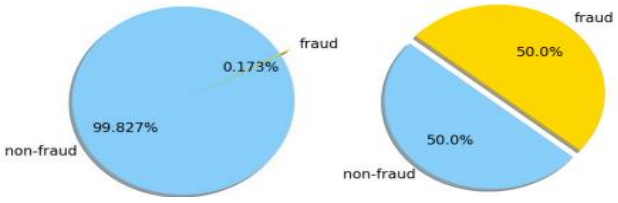
Features are represented by the variable X, where Xmin is the feature's lowest value and Xmax is its greatest value.

D. Feature Selection

The examination of cardholders' spending patterns serves as the foundation for identifying credit card fraud. An appropriate set of variables that reflects the distinct behavior of a credit card is used to assess this spending profile. The characteristics of both fraudulent and authorized transactions are often evolving. It is necessary to select the best features that substantially differentiate the two profiles in order to classify credit card transactions. The components of the card use profile and the techniques used affect how well credit card fraud detection systems function. These variables are derived from past transactions and a credit card's transaction history.

E. Balancing with SMOTE

When categorization categories are not distributed approximately equally, ML algorithms struggle to learn. In order to effectively train the model, some sort of balancing must be done because the provided data is extremely unbalanced. Adjusting the class distribution is often accomplished either by sampling the dominant class too thoroughly or the minority class too little, or a mix of the two. When applied to unbalanced datasets, the well-liked Synthetic Minority Oversampling Technique (SMOTE) has shown promise. To enhance random oversampling, SMOTE was suggested as a technique (Figure 6).



Class Distribution Before and After Sampling

F. Data Splitting

In the European credit cards dataset is separated into two categories, 20% will go towards testing, while 80% will go towards training.

ISSN : 2455-9679
SJIF Impact Factor : 6.008

G. Classification with LSTM

A sophisticated kind of RNN called the LSTM network is order for it to understand long-term associations. The vanishing gradient problem had to be resolved. Three gates make up an LSTM cell: a forget gate (ft), an output gate (ot), and an input gate (it). Ct is used to indicate cell state. The following Equations (2 to 6) characterize the operations that take place within an LSTM cell:

$$f_t = \sigma\left(W_f[h_{t-1}, x_t] + b_f\right) \quad (2)$$

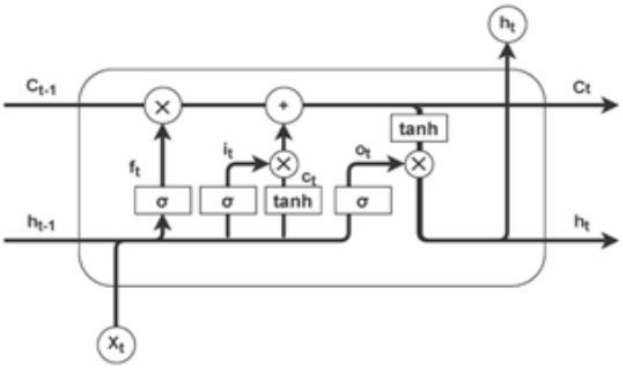
$$i_t = \sigma(W_i[h_{t-1}, x_t] + b_i) \quad (3)$$

$$o_t = \sigma(W_o[h_{t-1}, x_t] + b_o) \quad (4)$$

$$C_t = f_t * C_{t-1} + i_t * tanh(W_C[h_{t-1}, x_t] + b_C) \quad (5)$$

$$h_t = o_t * \tanh(C_t) \quad (6)$$

where σ and tanh are activation functions, b terms are biases, and W matrices are weights. Time-series forecasting and natural language processing are two applications where LSTMs excel in learning long-term dependencies. They are perfect for complicated sequential patterns where crucial information may be separated by several time steps, as seen in Figure 7, because of the ability to retain and update knowledge over long periods of time.



LSTM Architecture

H. Evaluation Metrics

A performance matrix that contrasted the actual observations with the efficacy of the selected models was evaluated using the model's predictions. Among the measures that were part of the performance matrix were AUC, F1-score, recall, accuracy, and precision. Each class's metrics were calculated as follows: TP is a measure of the quantity of correctly identified positive cases, whereas the number of correctly categorized negative examples is denoted by TN. These are known as FP and FN, where FPs represent the number of cases that are mistakenly categorized as positive, and FNs represent the instances that are mistakenly labelled as negative. The generally used criteria for classifying in this inquiry, there will be fraudulent credit card transactions used, are then explained in the following paragraphs:

1) Accuracy

Finding the accuracy (ACC) is as simple as dividing the total number of predictions or entries in the sample by the number of correct forecasts. The Equation (7):

$$Accuracy = \frac{TP+TN}{TP+Fp+TN+FN} \quad (7)$$

ISSN : 2455-9679
SJIF Impact Factor : 6.008

2) **Precision**

One way to figure out how accurate a model is to divide the number of true positives (TP) by the total number of predicted positives (TP + FP). To rephrase, the ability of the model to actually implement the optimistic predictions it generates. It is provided by Equation (8):

$$Precision = \frac{TP}{TP+FP} \quad (8)$$

3) Recall

A statistic called recall is used to determine how well the ML model can locate every instance of the class that is positive. It is computed by dividing up all of the favorable findings. that actually happened by the number of accurately anticipated positive observations. It is given by Equation (9):

$$Recall = \frac{TP}{TP+FN} \quad (9)$$

4) F1-score

In generally speaking, Recall and accuracy are combined to create the F1-score. To determine the F1-score, the following formula is used Equation (10):

$$F1 - Score = \frac{2(Precision*Recall)}{Precision+Recall} \quad (10)$$

5) ROC

The ROC curve is a measure that is linked to the AUC. The ROC curve is a good way to compare the pros and cons of TPR and FPR for a threshold-value-based predictor. It is the threshold number that changes both TPRs and FPRs. Set the x-axis to FPR and the y-axis to TPR to see the AUC, or area under the ROC curve. Higher AUC values indicate superior classifiers. Classifiers are considered no-skill if their AUC value is 0.5. It is deemed perfect when the classifier's AUC is 1.

The machine and DL models are determined using these matrices..

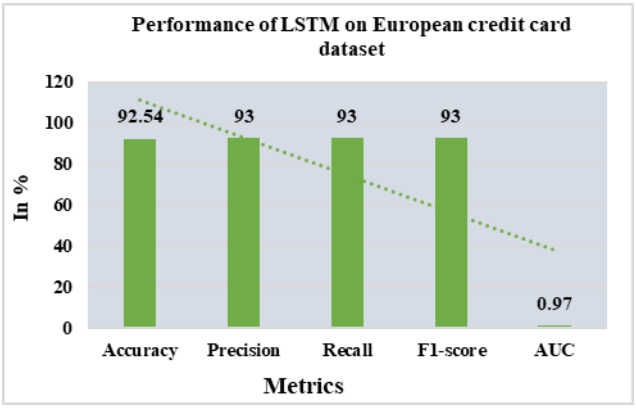
IV. RESULT & DISCUSSION

The evaluation machine for finding credit card fraud has an i5 8th generation CPU, a 240 GB SSD, 8 GB of DDR4 RAM, and is set up to use the Python v3 language. The 1050 H CPU type has a clock speed of 2.6 GHz to 5.0 GHz with turbo boost and a RAM frequency of 2565 MHz, making it the fastest for testing and training models. This part shows the results from ML and DL models that banks use to find credit card transactions. The F1 score, recall, accuracy, and precision are all performance measures. Table II displays how well the LSTM model worked with credit card data from Europe.

TABLE I. ML AND DL MODELS ON THE EUROPEAN CREDIT CARDS DATASET FOR CREDIT CARD FRAUD DETECTION

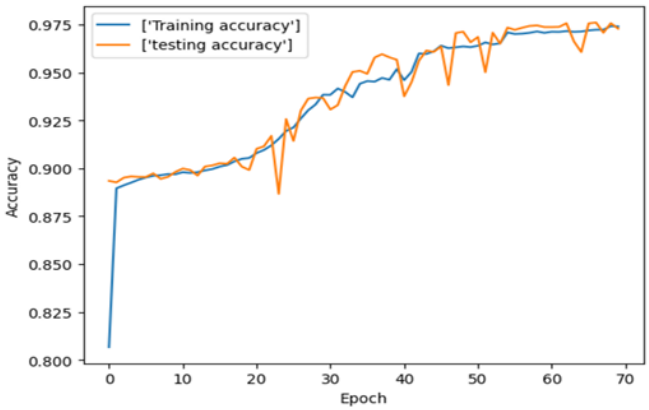
TABLE II.

Performance Measures	LSTM
Accuracy	92.54
Precision	93
Recall	93
F1-score	93
AUC	0.97



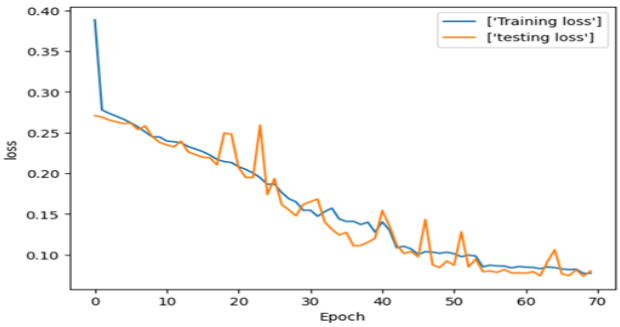
Bar Graph for LSTM Model Performance

Table II displays the LSTM model's performance characteristics for identifying the European dataset on credit card fraud. The model demonstrates a strong compromise between successfully detecting fraudulent transactions and lowering false positives, with F1-scores were 92.54%, 93%, and recall, and accuracy, respectively. Furthermore, a high discrimination capability between fraud and regular transactions is indicated by the AUC score of 0.97. A bar graph comparing several ML and DL models in Figure 8 shows how well LSTM performs in identifying fraudulent transactions, surpassing conventional machine learning models in important assessment measures.



The Accuracy Graph for the European Dataset

Figure 9 illustrates the accuracy progression over 70 training epochs for the European dataset. The graph depicts both training and testing accuracy, with the training curve (blue) demonstrating a steady increase from approximately 80% to nearly 97.5%. The testing accuracy (orange) follows a similar trend, exhibiting slight fluctuations but converging closely with the training curve in later epochs. The overall performance suggests effective model generalization with minimal overfitting.

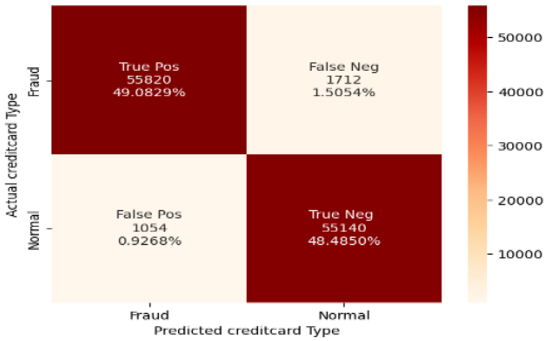


The loss function graph for the European dataset

Figure 10 shows the 70-epoch European dataset's training and testing loss curves. While the testing loss (orange) has a

ISSN : 2455-9679
SJIF Impact Factor : 6.008

similar trend with minor oscillations, the training loss (blue) begins at around 0.38 and progressively declines. Both curves show good model optimization as they converge to a low loss value close to 0.1. The model appears to generalize effectively without experiencing severe overfitting, as seen by the constancy of testing and training loss.



The confusion matrix for the European dataset

Figure 11 illustrates the European dataset's confusion matrix for classifying fraud using credit cards using the LSTM model. The matrix shows that the model correctly identifies 55,820 fraud cases TP and 55,140 normal transactions TN, achieving a balanced classification. However, 1,712 fraud cases are misclassified as normal FN, while 1,054 normal transactions are incorrectly classified as fraud FP. The distribution of values indicates strong classification performance, with the majority of predictions falling in the correct categories.

A. Comparative Analysis and Discussion

This section includes a comparative analysis of European credit card datasets for detecting diabetes. It compares ML and DL models like LSTM, CNN [19], and LR [20]. Performance metrics, including F1-score, recall, accuracy, and precision, are displayed in the results based on Table III.

There is a comparison of European credit card datasets for finding diabetes in this part. It checks out ML and DL models like CNN [19], LR [20], and LSTM. The results based on Table III show performance measures such as F1-score, recall, accuracy, and precision.

TABLE III. COMPARISON OF MACHINE LEARNING AND DEEP LEARNING MODELS ON A SET OF EUROPEAN CREDIT CARDS TO FIND STOLEN CARDS

Performance Measures	LSTM	CNN	Logistic Regression
Accuracy	92.54	89.60	54.86
Precision	93	88.60	38.36
Recall	93	76.90	58.33
F1-score	93	82.34	53.13

A comparison of ML and DL models Table III displays the credit card dataset used for credit card transaction identification. CNN and LR achieve 89.60% and 54.86% accuracy, respectively, but the LSTM network does better with 92.54% accuracy. LSTM gets the highest precision (93%), recall (93%), and F1 score (93%) too, so it is robust enough to detect the fraudulent transaction. CNN follows this up with 76.9% recall, 88.60% precision, and 82.34% F1-score, as well as a recall performance that is less precise. It turns out that LR does not work very well compared to others, with an F1-score of 53.13%, 58.33% recall, 54.86% accuracy, and 38.36% precision indicate that it has trouble processing complex transaction patterns. This is due to the fact that deep learning models (LSTM in particular) are effective in improving the precision and dependability of detecting fraud.

The suggested LSTM model produces extremely few high-quality results, detecting credit card fraud with 92.54% accuracy and capturing the temporal correlations of transaction data, together with false positives and false negatives. In practical financial security applications, it ensures excellent efficiency and resilience for large-scale fraud detection. LSTM performs superior than models such as CNN and LR in terms of F1-score, accuracy, and recall, proving that it is a more reliable method for identifying fraudulent transactions. Its capacity to learn intricate transaction patterns makes

LSTM a stable and adaptive one in choosing whether to accept credit transactions or deny them as fraudulent credit card transactions.

V. CONCLUSION & FUTURE WORK

A prevalent issue that costs banks, credit card companies, and consumers money is credit card fraud. The threat to financial institutions is growing (CCF). Fraudsters will constantly come up with novel methods for committing fraud. The resilient classifier handles fraud shifting nature. Accurately identifying which cases should be reported as fraud and which a fraud detection system shouldn't be primarily responsible for, LSTM is used to classify and identify credit card fraud. At 92.54% accuracy, the LSTM model did better than more common ML models like LR and CNN. Evaluating performance. they tested the model using recall, accuracy, precision, F1 score, and AUC as parameters. Fraud detection is made easier when temporal correlations in the data are captured by the LSTM model. This enhances banking apps' financial security by lowering FP and FN.

This study leads to several areas for future research despite its promising results. Secondly, additional improvement of detection accuracy can be achieved by using hybrid DL models that integrate DL with such mechanisms as LSTM and Attention or Transformers. Moreover, extra sophisticated feature engineering methods like graph-based analysis and anomaly detection strategies to aid in fraud classification might be tried to lower the quantity of false positives. Thirdly, real-time fraud detection and model deployment in banking systems could be explored to assess performance in dynamic environments. Lastly, applying federated learning techniques to ensure data privacy while improving fraud detection across multiple financial institutions could be a valuable direction for further research. Future research can help create more reliable, effective, and secure fraud detection frameworks in the banking sector by addressing these factors

REFERENCES

[1] S. Arora, S. Bindra, S. Singh, and V. Kumar, “Materials Today : Proceedings Prediction of credit card defaults through data analysis and machine learning techniques,” Mater. Today Proc., no. xxxx, 2021, doi: 10.1016/j.matpr.2021.04.588.

[2] Y. Bing Chu, Z. Min Lim, B. Keane, P. Hao Kong, A. Rafat Elkilany, and O. Hisham Abusetta, “Credit Card Fraud Detection on Original European Credit Card Holder Dataset Using Ensemble Machine Learning Technique,” J. Cyber Secur., vol. 5, no. 0, pp. 33–46, 2023, doi: 10.32604/jcs.2023.045422.

[3] P. Raghavan and N. El Gayar, “Fraud Detection using Machine Learning and Deep Learning,” in Proceedings of 2019 International Conference on Computational Intelligence and Knowledge Economy, ICCIKE 2019, 2019. doi: 10.1109/ICCIKE47802.2019.9004231.

[4] J. Kumar Chaudhary, S. Tyagi, H. Prapan Sharma, S. Vaseem Akram, D. R. Sisodia, and D. Kapila, “Machine Learning Model-Based Financial Market Sentiment Prediction and Application,” in 2023 3rd International Conference on Advance Computing and Innovative Technologies in Engineering (ICACITE), IEEE, May 2023, pp. 1456–1459. doi: 10.1109/ICACITE57410.2023.10183344.

[5] Suhag Pandya, “A Machine and Deep Learning Framework for Robust Health Insurance Fraud Detection and Prevention,” Int. J. Adv. Res. Sci. Commun. Technol., pp. 1332–1342, Jul. 2023, doi: 10.48175/IJARSCT-14000U.

[6] J. West and M. Bhattacharya, “Intelligent financial fraud detection: A comprehensive review,” Computers and Security. 2016. doi: 10.1016/j.cose.2015.09.005.

[7] T. H. Lin and J. R. Jiang, “Credit card fraud detection with autoencoder and probabilistic random forest,” Mathematics, 2021, doi: 10.3390/math9212683.

[8] S. Tyagi, “Analyzing Machine Learning Models for Credit Scoring with Explainable AI and Optimizing Investment Decisions,” Am. Int. J. Bus. Manag., vol. 5, no. 01, pp. 5–19, 2022.

[9] 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, 2022, doi: 10.3390/math10091480.

[10] S. Tyagi, T. Jindal, S. H. Krishna, S. M. Hassen, S. K. Shukla, and C. Kaur, “Comparative Analysis of Artificial Intelligence and its Powered Technologies Applications in the Finance Sector,” in Proceedings of 5th International Conference on Contemporary Computing and Informatics, IC3I 2022, 2022. doi: 10.1109/IC3I56241.2022.10073077.

[11] A. Immadisetty, “Machine Learning for Real-Time Anomaly Detection,” Int. J. Multidiscip. Res., vol. 6, no. 6, 2022.

[12] V. Pillai, “Anomaly Detection for Innovators: Transforming Data into Breakthroughs,” Lib. Media Priv. Ltd., 2022.

[13] E. Ileberi, Y. Sun, and Z. Wang, “A Machine Learning Based Credit Card Fraud Detection Using The GA Algorithm For Feature Selection,” J. Big Data, vol. 9, no. 24, Dec. 2022, doi: 10.1186/s40537-022-00573-8.

[14] S. Sahu and N. Sahu, “Analysis of Credit Card Fraud Transaction Detection using Machine Learning Algorithms,” in Proceedings of

International Conference on Contemporary Computing and Informatics, IC3I 2023, 2023. doi: 10.1109/IC3I59117.2023.10397696.

[15] R. M. Yuhes Raajha, A. Kavin, D. Rajkumar, R. Reshma, R. Santhosh, and N. Mekala, "An Analytical Approach to Fraudulent Credit Card Transaction Detection using Various Machine Learning Algorithms," in Proceedings of the 2023 2nd International Conference on Electronics and Renewable Systems, ICEARS 2023, 2023. doi: 10.1109/ICEARS56392.2023.10085157.

[16] S. Geetha, Y. Mohammed Khan, R. Sujay, S. P. Yoganand, and R. B, "Fraudulent URL and Credit Card Transaction Detection System Using Machine Learning," in 2023 International Conference on Advances in Electronics, Communication, Computing and Intelligent Information Systems (ICAECIS), 2023, pp. 709–714. doi: 10.1109/ICAECIS58353.2023.10170677.

[17] A. N. Ahmed and R. Saini, "Detection of Credit Card Fraudulent Transactions Utilizing Machine Learning Algorithms," in 2023 2nd International Conference for Innovation in Technology, INOCON 2023, 2023. doi: 10.1109/INOCON57975.2023.10101137.

[18] A. Singh, A. Singh, A. Aggarwal, and A. Chauhan, "Design and Implementation of Different Machine Learning Algorithms for Credit Card Fraud Detection," in International Conference on Electrical, Computer, Communications and Mechatronics Engineering, ICECCME 2022, 2022. doi: 10.1109/ICECCME55909.2022.9988588.

[19] T. Berhane, T. Melese, A. Walelign, and A. Mohammed, "A Hybrid Convolutional Neural Network and Support Vector Machine-Based Credit Card Fraud Detection Model," Math. Probl. Eng., 2023, doi: 10.1155/2023/8134627.

[20] J. O. Awoyemi, A. O. Adetunmbi, and S. A. Oluwadare, "Credit card fraud detection using machine learning techniques: A comparative analysis," in 2017 International Conference on Computing Networking and Informatics (ICCNI), IEEE, Oct. 2017, pp. 1–9. doi: 10.1109/ICCNI.2017.8123782.