MSW MANAGEMENT -Multidisciplinary, Scientific Work and Management Journal
ISSN: 1053-7899
Vol. 36 Issue 1s, 2026, Pages: 3399-3402

EXPENSE TRACKER WITH CREDIT CARD FRAUD DETECTION

Mrs. Vinitha Mary J[1*],
Assistant Professor, Department of Information Technology,
Manakula Vinayagar Institute of Technology, Puducherry, India.
Email:jvinitha11@gmail.com
Ramana.S[2],
UG Student, Department of Information Technology,
Manakula Vinayagar Institute of Technology, Puducherry, India
Mohammed Arshad.A[33],
UG Student, Department of Information Technology,
Manakula Vinayagar Institute of Technology, Puducherry, India
Hariharan.S[4],
UG Student, Department of Information Technology,
Manakula Vinayagar Institute of Technology, Puducherry, India
Dendulkar.R[5],
UG Student, Department of Information Technology,
Manakula Vinayagar Institute of Technology, Puducherry, India

Abstract
Credit card fraud detection is an important aspect of financial security in Artificial Intelligence (AI) and focuses on using deep learning techniques to detect anomalies. It is critically important to efficiently and accurately identify suspicious transactions in light of the increasing level of sophistication with which fraud is committed. From a deep learning perspective, the subdomain labelled deep learning with imbalanced classification is particularly challenging as fraud cases are rare events where there are a number of legitimate transactions. The imbalance has an enormous impact on the classification performance of machine learning models. Current techniques for credit card fraud detection rely on machine learning algorithms; however, such algorithms all have issues with imbalanced noisy data. The most common classifiers, Random Forests, XGBoost, and LightGBM, are typically used in conjunction with resampling techniques to aid in detection. None of the existing systems leverage temporal transaction identification which is particularly damaging to performance in cases of fraudulent activity. In addition, there is the challenge of imbalanced data where fraud cases are outnumbered by latitude transactions significantly. There is also the issue of noise present in the data which means the accuracy of each classifier is reduced.
The proposed system presents a new way forwards - a hybrid SMOTE-ENN resampling technique with a stacking ensemble model. The goal, as in all fraud detection problems, is to improve performance.

Artificial Intelligence (AI), SMOTE-ENN, LSTM, Random Forest, MLP, XGBoost, LightGBM

I. INTRODUCTION:

DEEPLEARNING- Deep learning is a type of machine learning that is essentially a neural network (NN) with three or more layers. These neural networks attempt to simulate the human brain's behavior allowing it to "learn" from large amounts of data. A neural network with a single layer can still make approximate predictions, but adding hidden layers can optimize performance. In deep learning, a computer model learns classification tasks from images, text or sound. Deep learning models can achieve state-of-the-art accuracies, sometimes exceeding human performance and can automatically learn based on data. When models are trained using a large amount of labeled data based on neural network architectures with many layers, it presents a new paradigm based on learning and improving on its own, studying computer algorithms. However, the technological advances of Big Data analytics have enabled larger and more sophisticated neural networks allowing computers to view, learn, and interpret complex scenarios faster than out of human capabilities.Deep learning has augmented vital innovations in the field of image classification, language translation, speech recognition. Overall, it could be operated for any pattern recognition task and decisions can be taken without human input.Deep learning is one form of machine learning. A machine learning workflow starts with relevant features being manually extracted from images. The features were used to create a model that categorizes the objects in the image. In deep learning, relevant features are automatically extracted from images. Furthermore, deep learning does "end-to-end learning", where the network is given raw data and a task, such as classification, and it learns how to do the task automatically. An additional key difference is that deep learning scales with data and shallow learning converges. A major advantage of deep learning networks is that they still continue to improve as your data size increases.
The other major distinction is that deep learning algorithms scale with data while shallow learning converges. The big benefit of deep learning networks is that they typically continue to get better as the data size grows. Deep learning has matured alongside the digital age, which has experienced a data outburst in multiple formats and from every part of the world. This massive data - Big Data - is created from social media, internet search engines, e-commerce, online cinema and so on. This vast amount of data has never been more available and in comparison can be shared through fintech platforms like cloud computing. Deep learning is a machine learning model that enables computers to do what is natural for a human being - learning from examples. Deep learning is a central technology to the operation and function of driverless cars, seeing a stop sign or the difference between a pedestrian and a lamppost. it is essential to voice control in consumer devices like phones, tablets, TVs, etc. Deep learning is getting lots of hype right now and it is deserved. It is delivering results that were not possible before. In deep learning, a computer model learns to conduct classification tasks directly from images, text

II.LITERATURE REVIEW:

Electronic financial transactions have become commonplace in the business world, whether online or in person, throughout the world. If we look at online purchases, we can see that fraud and instance of default payments has grown equally as fast as the use of credit cards online. As the access to customers and the natural tendency for operators to assume an account is legitimate increases, so have the rapid observations of fraudulent payments that attribute towards serious monetary losses to identify stolen items for sale. An identification has been made that some form of machine learning models, evaluating and utilizing their own ensembles, has been sort after by researchers to identify anomalies of credit card transaction data. Identifying anomalies in their data is more challenging due to overlapping classes and an imbalanced class distribution. Therefore, the anomaly detection of the minority class samples is much lower as learning algorithms can typically be biased towards the majority class samples.Through my study of the CCAD model, it suggests that there is value of utilising the hybrid model for anomalies in credit card fraud detection when introducing challenges such as imbalanced data and overlapping features of the dataset. The CCAD model employed a stacked approach that integrates classifiers and significantly improves the total accuracy the model achieved secured in relation to typical techniques Conventional Approaches. As the model stacked the classifiers, The ensemble method allows the classification to capture more complex patterns

MSW MANAGEMENT -Multidisciplinary, Scientific Work and Management Journal
ISSN: 1053-7899
Vol. 36 Issue 1s, 2026, Pages: 3399-3402

with more observations which the model can safely use as a good benchmark for fraudulent detection worth using. Conclusively, I would argue that the volumes say that the CCAD model provides good value to use the notorial outcomes.[1]

As e-commerce has advanced and technology has developed, the use of credit cards has exploded in both online and offline transactions. This has led to a substantial increase in the number of fraudulent transactions on a daily basis. Consequently, across all sectors and industries, organizations and financial institutions are losing billions of dollars every year to credit card fraud.

Because both legitimate and fraudulent transactions are distributed globally, differentiating between the two can be challenging. Moreover, relatively small percentage of transactions are fraudulent in relation to the total, making it a problem with class imbalance. Therefore, it is vital to have a robust fraud-detection method to maintain the trustworthiness of the payment system.[2]

Fraud related to credit card transactions is a demanding security issue financial institutions globally have to contend with. The dynamic setting of fraud characteristics, coupled with the class imbalance and complete separation issues present in fraud data, makes it hard to predict which transactions will be, used fraudulently, and to put in place real-time fraud detection systems.

The objective of their study is to develop a new real-time fraud detection framework that can be adapted for online, real-time implementation, and help mitigate the issues of non-stationary changes in transaction and fraud characteristics, class imbalance, and complete separation.

They proposed a new approach for addressing the impact of non-stationary changes in transaction patterns for fraud, which also increases efficiency in previously proposed model training, especially considering the enormous size of datasets to be analyzed.[3]

Proposed model signifies important improvements to contend with this problem in real-time solutions.

In their model, they took advantage of Deep Convolutional Neural Networks (DeepConvNet) and a group of optimization approaches. Optimization algorithms refer to a group of computational/mathematical approaches used to find the best solution or best set of solutions to the problem. They provided a comparison of proven effective and validated optimization algorithms: Stochastic Gradient Descent (Sgd), Adaptive Gradient (Adagrad), Adaptive Moment Estimation (Adam), and Root Mean Squared Propagation (Rmsprop).

They applied their optimization algorithms to the Deep Convolutional Neural Network (DeepConvNet) for this specific problem statement, which is credit card fraud detection (CCFD). After consideration of the problem nature, the properties of the objective function, and computational aspects, they found that all four algorithm optimization techniques are applicable for our CCFD task. However, based on the results of our experiment, they demonstrate that Rms prop produce a tremendous 99.93%accuracy compared to others.[4]

As e-commerce has advanced and technology has developed, the use of credit cards has exploded in both online and offline transactions. This has led to a substantial increase in the number of fraudulent transactions on a daily basis. Consequently, across all sectors and industries, organizations and financial institutions are losing billions of dollars every year to credit card fraud.

Because both legitimate and fraudulent transactions are distributed globally, differentiating between the two can be challenging. Moreover, relatively small percentage of transactions are fraudulent in relation to the total, making it a problem with class imbalance. Therefore, it is vital to have a robust fraud-detection method to maintain the trustworthiness of the payment system. [5]

III.PROPOSED SYSTEM:

The proposed credit card fraud detection system provides numerous key benefits that, when implemented, will broaden its ability to detect fraud. The hybrid SMOTE- ENN method deals with class imbalance and allows machine learning models to be trained on a balanced dataset which now includes sufficient instances of fraudulent transactions. With more accurate models it will also reduce the number of false negatives (where fraud is present but not detected). The use of LSTM networks means there are effective temporal patterns in transaction data and the modelling system will be able to detect poorly distinguishing patterns of evolving fraudulent behaviour. The operational architecture introduces a stacking ensemble model that fuses various classifiers that were previously used, thus taking positive attribution of predictive power to Random Forest and MLP, which improves their prediction accuracy and model robustness. Overall this strategy leads to improved fraud detection rates and increases customer confidence for credit card transactions.

The proposed detection system for credit card fraud has many compelling benefits that improve its accuracy and reliability. The use of the hybrid SMOTE- ENN approach tackles class imbalance effectively, thereby allowing machine learning models to train on a balanced dataset that included a proportional number of fraudulent transactions, which moreover allows for improved performance and a decrease in false negatives. The inclusion of LSTM networks also allows the system to model the temporal aspects of transaction data allowing it to identify changing patterns of fraud. The stacking ensemble model takes the strongest classifiers, for example Random Forest and MLP, to then enhance prediction accuracy and robustness. This overall approach improves process improvement with respect to fraud detection as it improves the likelihood of identification of fraud occurring and increasing consumer trust in the security of their credit card transaction.

IV.ARCHITECTURE DIAGRAM

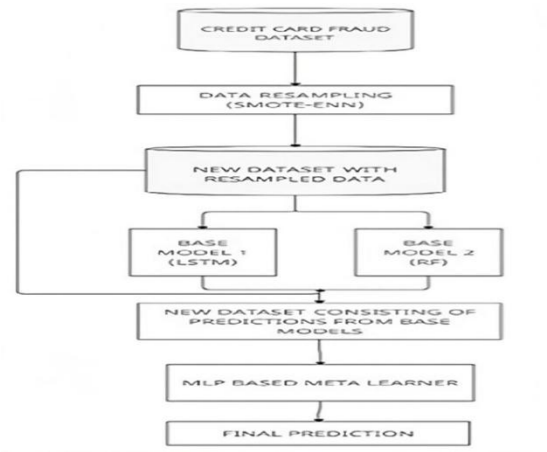

Figure 3.1 Architectural Workflow of Credit card fraud detection.

The machine learning pipeline for detecting credit card fraud begins with a dataset specifically designed for this purpose. Given the challenges of imbalanced data, where fraudulent transactions are significantly fewer than non-fraudulent ones, a data resampling technique is employed to address this issue. The pipeline utilizes the SMOTE-ENN method, whichcombinesSyntheticMinorityOver-samplingTechniqu e(SMOTE) to generate synthetic samples of the minority class (fraudulent transactions) and Edited Nearest Neighbors (ENN) to filter out noise from the

MSW MANAGEMENT -Multidisciplinary, Scientific Work and Management Journal
ISSN: 1053-7899
Vol. 36 Issue 1s, 2026, Pages: 3399-3402

majority class. This results in a new, balanced dataset. Following this,twobasemodels—LongShort-TermMemory(LSTM)andRandomForest(RF)—process the resampled data, making independent predictions. The outputs of these models are then aggregated to form a new dataset. This data set serves as input for a MetaLearner, specifically a Multi-Layer Perceptron (MLP), which synthesizes the predictions from the base models to enhance accuracy. By leveraging the distinct strengths ofboth the LSTM and RF models.

V.RESULT AND DISCUSSION

The suggested credit card fraud detection solution is built as an advanced machine learning framework that exploits multiple models in the form of a stacked ensemble. This is accomplished with Long Short-Term Memory (LSTM) networks and Random Forest (RF) models as base learners, and combined through a Multi-Layer Perceptron (MLP) model as the meta-learner. This structure captures a variety of properties of the transaction data. In particular, LSTM models are great in the analysis of time-series data and understanding patterns that unfold over time is crucial for the detection of fraud, as fraudulent behaviour usually unfolds over time. Random Forest incorporates a strong classification mechanism through an ensemble of decision trees, which generally handle structure data well. The MLP meta-learner looks to combine the classification of both the LSTM and Random Forest models by learning from both the models to predict the final classifications. Finally, it is also important to note the role played by the SMOTE-ENN technique in the preprocessing of the dataset. SMOTE (Synthetic Minority Over-sampling Technique) is used to create synthetic examples of fraud occurrences, which are scarce in reality. ENN (Edited Nearest Neighbors) is used to clean any noise from the data. In summary, we are confident that this procedure will improve the actual fraud detection performance because the model is being trained on a better-balanced dataset using the SMOTE-ENN technique, clearly separating fraudulent class instances, and allowing for higher overall performance for fraud detection.

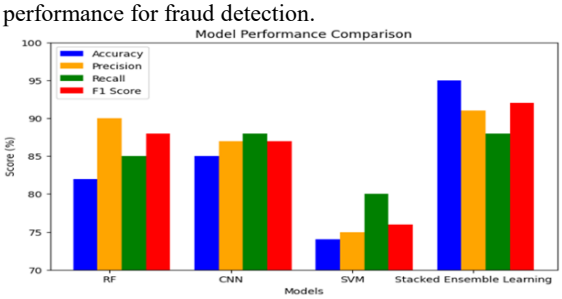

Figure:5.1 :Model Performance Comparison.

2)Comparative Analysis: The researchers evaluated their proposed stacked ensemble model against a few traditional machine learning models (Support Vector Machine (SVM), Convolutional Neural Networks (CNN), and Random Forest (RF) separately). Based on four key metrics (accuracy, precision, recall, F1-score), the performance of the stacked ensemble LSTM + MLP models was the best as compared to the other models. This confirms the advantage of combining time-based models (like LSTMs) with classifiers based on features (such as RF) and having a meta learner (MLP) that optimises the final decision. Though RF and CNN performed somewhat well (indicating they could detect some amount of fraud), they lagged behind because of their individual limitations of using sequential or imbalanced data (as mentioned above). SVM performed the worse out of all the models, again due to imbalanced data, but there is also an intrinsic characteristic of SVM being sensitive to the nature of the high dimensionality of transaction data with complex patterns.

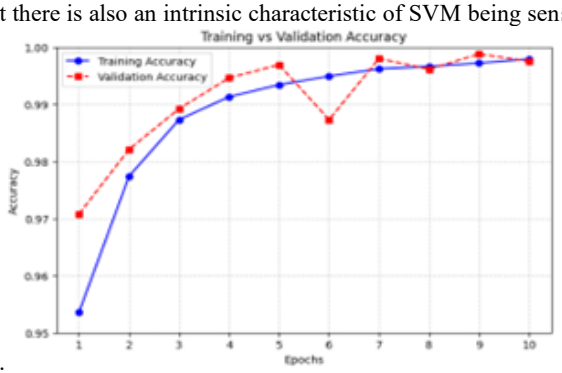

Figure:5.2 :Training vs Validation Accuracy

3)Visualization: In addition, the report provides numerous visualizations (e.g., accuracy curves, loss plots, comparison bar graphs) to illustrate their findings Figure 7.3), which enhance their overall presentation, and have helped summarize all performance improvements from the stacked ensemble. For example, the accuracy and loss plots detail the training performance of the model over epochs, while the comparison graph can be used to highlight how the proposed models (LSTM + MLP) outperformed on every evaluation metric. The visual comparisons clearly demonstrate that the LSTM + MLP approach was superior to other approaches, which establishes confidence in the claims made in the discussion. The inclusion of visual tools definitely aids in interpretation of the results and validates the findings from the numerical data.

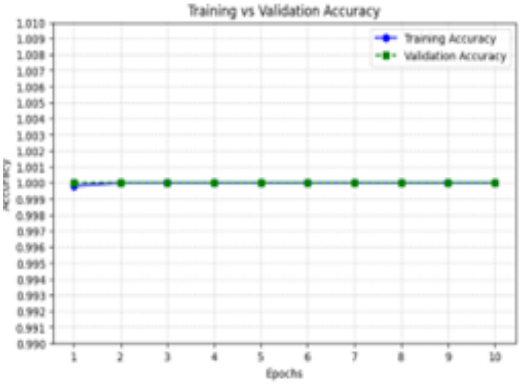

Figure:5.3: Visualization Accuracy

MSW MANAGEMENT -Multidisciplinary, Scientific Work and Management Journal
ISSN: 1053-7899
Vol. 36 Issue 1s, 2026, Pages: 3399-3402

4)Discussion Highlights: The discussion section emphasizes that integrating the functionality of a LSTM with the classification approach of a Random Forest provides the model with two ways to classify transactions, as both sequential behavior and other characteristics made up of features create a holistic understanding of user transactions. This is important since detecting fraud means identifying the characteristics of the individual transaction (for example, amount, merchant) and the sequence of previous transactions (for example, unusual purchases). The meta-learning layer (MLP) acts as a 'decision referee', improving final prediction accuracy by virtue of discovering the errors and strengths made by the base models. Foremost among any potential improvements was the successful reduction of false positives - legitimate transactions identified as fraud. Reducing false positives is critical for protecting merchants from attack to preserve user trust, and save merchants the hassle or expenses of investigating genuine consumers' activities. On the other hand, if recalls are still relatively high, even with a false positive rate, reducing false highs still improves transaction security through increased fraud sensitivity. The balance between true positives and true negatives is essential in a real-world application, where the socio-economic inconvenience of under-detection and over-detection can cause substantial and unwanted repercussions.

Figure:4.4: Loss over Epochs

PREDICTED OUTPUT: When a user uploads their transactional data into the application, the system analyzes each piece of input and determines if the transaction is legitimate or fraudulent. This determination is made by the crux of the system - a stacked ensemble made up of Long Short-Term Memory (LSTM) networks and Random Forest (RF) classifiers, with a Multi-Layer Perceptron (MLP) serving as the meta-learner. Together, they allow the system to detect time-based procedures and categorical anomalies, which greatly enhances its ability to identify fraud.he system provides a confidence score with each classification, indicating the model's level of confidence in its predicted classification. For example, a transaction could be identified as "Fraudulent" with a confidence score of 92%, giving users further understanding of the seriousness or reliability of the detection. Understanding the scoring provides the user with a means of making informed decisions, especially in cases where the data is almost a) borderline, etc. To better help the user receive a holistic understanding of the data, the system also provided visual aids. These would include pie charts showing how expenses are broken down into various categories, bar graphs that depict activity in monthly spending areas, and line charts that visualize the user's behavioral anomalies over time. Fraudulent transactions are also visually indicated or marked in some cases to help the user recognize the areas where issues could be flagged for additional investigation and examine areas of concern or an unusual spike in usage.Finally, the system provides the ability to download a full report in CSV format. The report contains all the transaction data that was analyzed, predicted labels (fraud User True or fraud User False), the associated confidence scores and important metadata like time and amount. This feature is helpful for users who want to keep a record of the transaction data or share the outcome with customers or auditors.The predicted output as a whole is designed to not only detect fraud but also to give users insight into their spending behavior. In doing so, it encourages a level of transparency or trustworthiness to engage users and provide early alerts to help people stay on top of their finances. This is accomplished by providing meaningful fraud detection paired with a user-friendly layer for presentation.

V. CONCLUSION AND FUTURE WORK

The project “Expense Tracker with Credit Card Fraud Detection” addresses important issues related to financial security in today’s digital world. The methods required to provide effective fraud detection have shortcomings when developed independently, leading to underperforming models characterized by poor fraud detection (increasingly not detecting fraud), increased false positives (whether true or false), and an inability to adapt to the changeworld of fraud. Addressing these limitations in fraud detection is possible through the proposed system that integrates the SMOTE-ENN hybrid data resampling method with a powerful Stacked Ensemble Learning framework. With the base learners of the models as Long Short-Term Memory (LSTM) networks and Random Forest classifiers situated in a multi-layer perception (MLP) meta-learner, the method relies upon sequential and structured data to provide a reliable and accurate fraud detection prediction. While the data was preprocessed into a balanced dataset through SMOTE-ENN for training, perceived accuracy of the predictions was a result of both precise identification of fraud and a rejection of marked clean transactions as potential fraud. The experimental results show significant increases, across all metrics of performance (accuracy, recall, F1-score), and decreases in losses for the models. In summary, the proposed system presents for the reader an effective means for real-world fraud detection that is reliable, and scalable and efficient while allowing users to track their expenses and provide security against unauthorized transactions.

Reference

[1. Md AmirulIslam, Md AshrafUddin, Sunil Aryal, Giovanni Stea, "An ensemble learning approach for anomaly detection in credit card data with imbalanced and overlapped classes," *Available online 7 October 2023*, 2214-2126/© 2023 The Author(s). Published by Elsevier Ltd.

[2. Mimusa Azim Mim, Nazia Majadi, Peal Mazumder, "A soft voting ensemble learning approach for credit card fraud detection," *Available online 2024*, 2405-8440/© 2024 The Author(s). Published by Elsevier Ltd.

[3. Georgios Charizanos, Haydar Demirhan, Duygu İçen , "An online fuzzy fraud detection frame work for credit card transactions ,"*Available online 3May2024* ,0957-4174/©2024The Author(s). Published by Elsevier Ltd.

[4. Chandana Gouri Tekkalia, Karthika Natarajan, "Assessing CNN’s Performance with Multiple Optimization Functions for Credit Card Fraud Detection," *Available online 2024*,The Authors. Published by Elsevier B.V. Peer-review under responsibility of the scientific committee of the International Conference on Machine Learning and Data Engineering (ICMLDE 2023).

[5. Zorion, Priyanshu Katiyar, Lakshay Sachan, Rithik Chhabra, Vishal Pandey, Dr. Hoor Fatima, "Credit Card Financial Fraud Detection Using Deep Learning," *Available at SSRN*, November 10, 2023.

[6. E.F.Malik,K. W. Khaw, B.Belaton, W.P.Wong,X .Chew, "Credit card fraud detection using any why brid machine learning architecture," *Mathematics* ,vol. 10, no.9,p.1480,Apr. 2022, doi: 10.3390/math10091480.

[7. N. S. Alfaiz, S. M. Fati, "Enhanced credit card fraud detection model using machine learning," *Electronics*, vol. 11, no. 4, p. 662, Feb. 2022, doi: 10.3390/electronics 11040662

[8. I.Benchaji,S.Douzi,B.ElOuahidi,J.Jaafari,"Enhancedcredit card fraud detection based on attention mechanism and LSTM deep model," *J.BigData*,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, G. Obaido, "A neural network ensemble with feature engineering for improved credit card fraud detection,"*IEEEAccess*,vol. 10, pp. 16400–16407, 2022, doi: 10.1109/ACCESS.2022.3148298.

[10. E.Btoush, X. Zhou, R. Gururaian, K. Chan, 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.