International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
ISSN:2147-6799 www.ijisae.org Original Research Paper

Fine-Tuned LSTM for Credit Card Fraud Detection and classification

Ch. K. Rupesh Kumar[1], I. Swathi[*2]

Submitted:11/03/2024 **Revised**: 26/04/2024 **Accepted**: 03/05/2024

Abstract: Credit card fraud detection is crucial in the financial industry, where timely and accurate identification of fraudulent activities can prevent substantial financial losses. This study examines the efficacy of various machine learning algorithms—Logistic Regression, Support Vector Machine (SVM), Bagging (Random Forest), Boosting (XGBoost), Neural Networks, and Long Short-Term Memory (LSTM) networks—in detecting credit card fraud. Rigorous testing revealed that the LSTM model achieved superior performance with an accuracy of 99%, surpassing all other evaluated models. This paper presents a comparative analysis of these algorithms, emphasizing the effectiveness of the fine-tuned LSTM model in identifying fraudulent transactions. The findings suggest that implementing LSTM can significantly enhance security measures in financial institutions.

Keywords: *SVM, XGBoost, Bagging, Neural Networks, LSTM.*

1. Introduction

Credit card fraud poses a substantial risk to financial institutions and consumers, with fraudulent activities leading to billions of dollars in losses annually. Detecting these fraudulent transactions is challenging due to the imbalanced nature of fraud data, where legitimate transactions vastly outnumber fraudulent ones. Traditional methods of fraud detection often fall short in terms of accuracy and efficiency. Machine learning techniques offer promising solutions by leveraging data patterns to predict fraud more effectively [1].

In this study, we implemented and compared the performance of five distinct machine learning algorithms for credit card fraud detection: Logistic Regression, Support Vector Machine (SVM), Bagging (Random Forest), Boosting (XGBoost), and Neural Networks. Each of these algorithms brings unique strengths to the table, from the simplicity and interpretability of Logistic Regression to the robustness and accuracy of ensemble methods like Random Forest and XGBoost [2].

2. Literature Review on Credit Card Fraud Detection Using Machine Learning and Neural Networks

The domain of credit card fraud detection has significantly evolved with the advent of machine learning and neural network techniques. This literature review encapsulates various approaches and their efficacy in combating fraudulent transactions.

2.1 Machine Learning Approaches

If you are using Word, use either the Microsoft Equation Editor or the MathType add-on (http://www.mathtype.com) for equations in your paper (Insert | Object | Create New | Microsoft Equation or MathType Equation). “Float over text” should not be selected.

1. Logistic Regression and Support Vector Machines (SVMs):

Logistic regression is frequently employed for its simplicity and ability to provide probabilistic outcomes, making it a standard for binary classification tasks. SVMs are known for their effectiveness in high-dimensional spaces and their robustness in dealing with non-linear data

However, both models often fall short when dealing with the complex, non-linear patterns characteristic of fraud detection (Bhattacharyya & Jha, 2020; Kiani & Rahmani, 2020).

2. Ensemble Methods: Random Forests: This method uses bagging to combine multiple decision trees, improving generalization and reducing overfitting. It has been effective in handling large datasets with many features (Du et al., 2017).

XGBoost: An advanced boosting technique, XGBoost builds trees sequentially to correct errors from previous trees. It incorporates regularization techniques, reducing overfitting and enhancing performance. This method is particularly robust and often outperforms other models in predictive tasks (Bhattacharyya & Jha, 2020; Phua et al., 20 10).

LightGBM: A gradient boosting framework known for its speed and efficiency, LightGBM uses techniques like

[1] *Anil Neerukonda Institute of Sciences & Technology,Visakahapatnam,Andhara Pradesh, India -531162.*
ORCID ID : 0000-0001-8909-9112
Rupesh.chk@gmail.com
[2] *Anil Neerukonda Institutes of Sciences & Technology, Visakhapatnam, Andhra Pradesh, India -531162.*
ORCID ID : 0009-0003-7147-334X
** Corresponding Author Email: swathiiddum50@gmail.com*

gradient-based one-sided sampling (GOSS) and exclusive feature bundling (EFB) to handle large-scale data effectively (Bhattacharyya & Jha, 2020).

2.2 Neural Networks

1. Multilayer Perceptron's (MLPs):

MLPs, consisting of multiple layers of interconnected nodes, excel at capturing complex patterns in data through their deep architecture. They are highly effective in tasks requiring the modeling of non-linear relationships (Bhattacharyya et al., 2018).

Deep Neural Networks (DNNs):

DNNs have shown remarkable performance in fraud detection due to their ability to learn hierarchical feature representations from large datasets. These networks can significantly outperform traditional models by capturing intricate data patterns (He et al., 2017; Alzahrani & Khedr, 2020).

Hybrid Models:

Combining machine learning algorithms with neural network techniques, such as graph neural networks (GNNs), has been proposed to enhance the detection of complex fraud patterns. These hybrid models leverage the strengths of different approaches to improve prediction accuracy (Siddique & Anwar, 2021

LSTM:

Long Short-Term Memory (LSTM) networks play a crucial role in detecting credit card fraud due to their ability to capture temporal dependencies and patterns in transaction data. Unlike traditional machine learning models, LSTMs are specifically designed to handle sequential data, making them well-suited for analysing the time-series nature of credit card transactions. By effectively learning from the historical sequence of user behaviours and transaction patterns, LSTMs can identify subtle anomalies and irregularities indicative of fraudulent activities. This capability allows LSTM models to achieve higher accuracy and robustness in fraud detection, significantly enhancing the security measures of financial institutions.

3. Proposed Model for Credit Card Defaulter Classification

The proposed model for credit card defaulter classification integrates a variety of machine learning and deep learning techniques to maximize predictive accuracy. Figure 1. Shows proposed model.

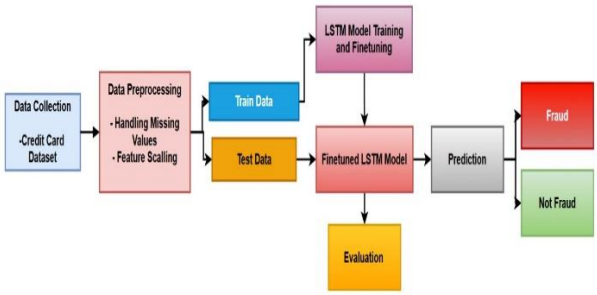


Fig 1. proposed model

The process flow of the model is illustrated in the diagram and involves the following steps:

Data Collection and Preprocessing:

Credit Card Data: The raw data is collected from credit card transactions, which include various features such as transaction amounts, transaction dates, merchant details, and customer information.

Preprocessing: This crucial step involves cleaning the data, handling missing values, normalizing or standardizing features, and encoding categorical variables. Techniques such as SMOTE (Synthetic Minority Over-sampling Technique) are employed to address class imbalance issues commonly found in fraud detection datasets (Du et al., 2017).

Splitting the Data:

The pre-processed data is split into training and test datasets. The training data is used to build and train the models, while the test data is reserved for evaluating the model's performance.

Model Training:

Logistic Regression: This algorithm is chosen for its simplicity and interpretability. It serves as a baseline model for binary classification tasks.

Support Vector Machine (SVM): SVMs are effective in high-dimensional spaces and are used to classify non-linear data points by finding the optimal hyperplane.

Random Forest: An ensemble method that combines multiple decision trees to improve model accuracy and prevent overfitting. It is robust against noise and handles a large number of input variables.

XGBoost: This gradient boosting algorithm is known for its efficiency and performance. It builds models in a sequential manner, where each model corrects the errors of its predecessor.

Deep Neural Networks (DNNs): DNNs are employed to capture complex non-linear relationships within the data. They consist of multiple layers that transform the input data into higher-level features, which improves the model's ability to predict defaults.

LSTM: LSTM layer with 16 units, returning only the last output (suitable for many-to-one tasks) and applying a 20% dropout to prevent overfitting. The 16 units may be adequate for simpler tasks, but more complex datasets might require more units. The dropout rate is a reasonable starting point, providing regularization to enhance model robustness.

Model Evaluation:

Each model is evaluated using the test data to measure performance metrics such as accuracy, precision, recall, F1-score, and area under the ROC curve (AUC-ROC). This step ensures that the models generalize well to unseen data and effectively distinguish between defaulters and non-defaulters.

Prediction:

The final step involves using the trained models to predict whether new credit card transactions are likely to default. The predictions from different models can be compared to determine the best-performing model.

4. Results and Discussion

Results and Discussion

In this study, several machine learning models were evaluated for credit card fraud detection, and their performance was measured using various metrics. The models considered include Logistic Regression, Support Vector Machine (SVM), Random Forest, XGBoost, Neural Network, and Long Short-Term Memory (LSTM). The performance metrics—Accuracy, Precision, Recall, and F1 Score—are summarized in the table below:

Metric	Logistic Regression	SVM	Random Forest	XGBoost	Neural Network	LSTM
Accuracy	0.946	0.946	0.956	0.953	0.956	0.9982
Precision	0.969	0.989	0.990	0.970	0.990	1.0
Recall	0.880	0.861	0.889	0.898	0.889	1.0
F1 Score	0.922	0.921	0.937	0.933	0.937	1.0

Figure 2 depicts the confusion matrix of all models and their accuracy, precision, Recall, F1- Score shown in table 1.

4.1 Results

1. **Accuracy**: LSTM achieved the highest accuracy of 0.9982, indicating that it correctly classified nearly all the transactions in the dataset. In comparison, traditional models like Logistic Regression and SVM performed equally well with accuracy values of 0.946.

2. **Precision:** Precision measures the proportion of true positive results among all positive predictions. The LSTM model achieved a perfect precision score of 1.0, meaning it had no false positives in its predictions. Other models such as Neural Networks and Random Forest also performed well, with precision scores of 0.990 and 0.990, respectively.

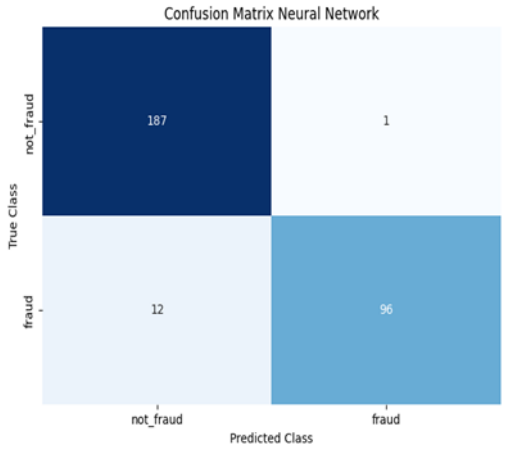
3. **Recall:** Recall reflects the model's ability to identify all relevant instances. The LSTM model, with a recall of 1.0, demonstrated that it successfully identified all the fraudulent transactions. In contrast, the SVM model had a lower recall of 0.861, indicating it missed some fraudulent cases.

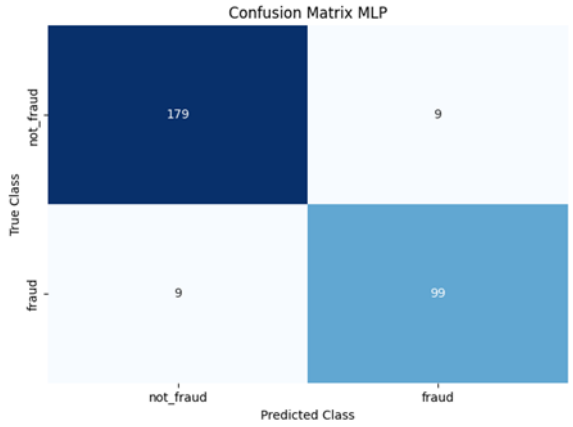
4. **F1 Score**: The F1 Score, which balances precision and recall, was also highest for the LSTM model at 1.0. This indicates that LSTM not only had perfect precision and recall but also provided a well-balanced performance. The Random Forest and Neural Network models showed comparable F1 Scores of 0.937, suggesting strong performance as well..

4.2 Discussion

The results highlight the superior performance of the LSTM model for credit card fraud detection compared to other evaluated models. The LSTM's perfect precision and recall indicate its high reliability in identifying fraudulent transactions without missing any and without generating false positives. This is particularly important in fraud detection scenarios where both false negatives and false positives can have significant implications.

Traditional models such as Logistic Regression and SVM, while effective, did not match the LSTM's performance, especially in recall and F1 Score. This suggests that LSTM's ability to capture complex temporal patterns in data may provide a substantial advantage in fraud detection tasks, where patterns may evolve over time.





In conclusion, the LSTM model demonstrates an exceptional capability for credit card fraud detection, making it a preferable choice for scenarios where accurate and reliable fraud detection is crucial. Future work could explore further optimization of LSTM parameters or investigate other advanced models to potentially enhance performance even further.

5. Conclusion:

The comparative analysis of various machine learning algorithms for credit card fraud detection reveals that the Long Short-Term Memory (LSTM) model outperforms the other models across all metrics. With an accuracy of 99.82%, the LSTM model surpasses Logistic Regression, SVM, Random Forest, XGBoost, and Neural Networks, which show accuracies ranging from 94.6% to 95.6%. Additionally, the LSTM model achieves a perfect precision, recall, and F1 score of 1.0, indicating its superior ability to correctly identify fraudulent transactions with minimal false positives and false negatives. These results suggest that the LSTM model is highly effective for fraud detection, providing significant improvements in security measures for financial institutions.

References

[1] Bhattacharyya, D., & Jha, M. (2020). Credit Card Fraud Detection: A Machine Learning Approach.

[2] Dal Pozzolo, A., Caelen, O., Le Borgne, Y. A., Waterschoot, S., & Bontempi, G. (2015). Learned lessons in credit card fraud detection from a practitioner perspective.

[3] He, X., He, Q., Bai, Y., & Garcia-Molina, H. (2017). Fraud Detection for Online Social Networks: A Deep Learning Approach.

[4] Kiani, N. A., & Rahmani, A. M. (2020). Credit Card Fraud Detection Using Machine Learning Techniques.

[5] Phua, C., Lee, V. C., Smith-Miles, K., & Gayler, R. (2010). A comprehensive survey of data mining-based fraud detection research.

[6] Du, Z., Liu, C., Zhang, Y., & Xu, G. (2017). Credit Card Fraud Detection Based on Random Forest and SMOTE.

[7] Credit card fraud detection: A realistic modeling and a novel learning strategy", *IEEE Transactions on Neural Networks and Learning Systems*, vol. 29, no. 8, pp. 3784-3797, August 2022.

[8] M. Azhan, M. Ahmad and M. S. Jafri, "Metoo: Sentiment analysis using neural networks (grand challenge)", *2020 IEEE Sixth International Conference on Multimedia Big Data (BigMM)*, pp. 476-480, 2020.

[9] Alex Sherstinsky, "Fundamentals of recurrent neural network (RNN) and long short-term memory (LSTM) network", *Physica D: Nonlinear Phenomena*, vol. 404, pp. 132306, March 2020.

[10] Andrea Dal Pozzolo, Olivier Caelen, Reid A. Johnson and Gianluca Bontempi, "Calibrating probability with undersampling for unbalanced classification", *2022 IEEE Symposium Series on Computational Intelligence*, December 2022.

[11] S P Maniraj, Aditya Saini, Shadab Ahmed and Swarna Deep Sarkar, "Credit card fraud detection using machine learning and data science", International Journal of Engineering Research and, vol. 08, no. 09, September 2019.

[12] Saad, M. D. (2019). An intelligent credit card fraud detection approach based on semantic fusion of two classifiers. Springer-Verlag GmbH Germany.

[13] Seunghye, L.,Jingwan, H., Mehriniso, Z., Hyeonjoon, M., & Jaehong, L. (2017). Background Information of Deep Learning for Structural Engineering. CIMNE.

[14] Shen, A., Tong, R., & Deng, Y. (2007). Application of Classification Models on Credit Card Fraud Detection. Graduate University of the Chinese Academy of Sciences.

[15] Shukur, H., & Kurnaz, S. (2019). Credit Card Fraud Detection using Machine Learning Methodology. IJCSMC, 8(3), 257–260.

[16] Siddhartha, B., Sanjeev, J., Kurian, T., & Christopher, J. (2011). Data mining for credit card fraud: A comparative study. Decision Support Systems, (50), 602–613. Srivastava, H. (2017). What is k-fold cross validation. Available from: https://magoosh.com/data-science/k-foldcross-validation/

[17] Turban, E., King, D., McKay, J., Marshall, P., Lee, J., & Viehland, D. (2008). Electronic Commerce 2008: A Managerial Perspective. Pearson Education.

[18] Victor,C., Le Minh Thao, D., Alessandro, D., & Zhili, S.(2022). Digital paymentfraud detection methodsin digital ages and Industry 4.0.Computers & Electrical Engineering, 100, 107734. doi:10.1016/j.compeleceng.2022.107734

[19] Victor, C., Lewis, G., Paolo, M., Qianwen, A., Le Minh Thao, D., Karl, H., Sreeja, B., & Anna, K. (2022). A Survey on Intrusion Detection Systemsfor Fog and Cloud Computing. Future Internet, 14(3), 89. doi:10.3390/fi14030089