"Deep Learning for Real-Time Fraud Detection Using LSTM Networks"

Siri Karanam

811300183

Prof. Dr. Li Liu

Advanced Machine Learning

Kent State University

SUMMARY:

In order to predict fraudulent credit card transactions, the current project focuses on deep learning methods, specifically Long Short-Term Memory (LSTM) networks. It shall examine the difficulties in understanding & correctly identifying fraudulent transactions in big datasets, with a focus on class imbalance and guaranteeing high detection accuracy for minor illegal transactions.

LSTM (Long Short-Term Memory) is that one type of **recurrent neural network (RNN)** which actually excels at managing sequenced data, such as text or time series. LSTMs can be able to retain significant knowledge for extended periods of time while they also keep forgetting unimportant facts, in contrast to standard neural networks. As a result, they can be applied to tasks like stock price prediction, speech recognition, and transaction fraud detection. As data moves across the network, they use specialized gates to determine what information should be retained and what should be discarded. The main issues addressed include:

- Class imbalance: As fraudulent transactions are being occurred much less frequent than non-fraudulent ones, traditional algorithms have been finding it difficulty detecting fraud.
- Real-time predictions: Many financial institutions do require algorithms which can accurately detect fraud even in real-time entirely based on historical transaction data.
- Model performance: This report's primary focus is on an LSTM-based model's ability to process sequential transaction data and produce precise predictions analysis.

Key findings:

- With a 100% accuracy rate for both the fraud and non-fraud classes, the LSTM model produced exceptional outcomes.
- By up sampling the minority class (fraudulent transactions), the model successfully addressed the class imbalance and produced a training dataset that was balanced.
- The model's dependability in fraud detection is demonstrated by the confusion matrix and classification report, which show flawless precision, recall, and F1-scores.
- The architecture of the model, which included regularization dropout layers, helped it perform very well by preventing overfitting and preserving high predicted accuracy.

All things those have been considered, this study depicts how well the LSTM networks perform fraud can detection tasks, providing a strong option for financial organizations those which are looking to stop credit card transaction fraud.

INTRODUCTION:

Fraud detection, specifically credit card fraud, has now become a significant concern in the financial sector in the current digital age. The Fraudulent activities have been increasing rapidly and making it complex due to the quick expansion of digital payments and online transactions, which shall gain make them more difficult to identify and stop. Apart from causing monetary damages for both customers and companies, credit card theft automatically erodes the confidence in the safety of payment systems. Conventional fraud detection tools, including rule-based systems, are failing because scammers are constantly improving their methods. Complex fraud patterns are frequently missed by these systems, particularly when they feature subtle, changing trends that are hard for people to see. A more sophisticated solution is provided by machine learning and deep learning technologies in this situation. Because LSTM networks, a subset of recurrent neural networks (RNNs), can analyze and learn from sequential data, such as transaction histories, they have demonstrated significant promise in the detection of fraud. Even when the patterns are too subtle or complicated for conventional techniques, LSTMs can uncover hidden patterns in big datasets that could point to fraudulent conduct. We can create fraud detection systems that are more precise, real-time, and flexible enough to adjust to new and changing fraud strategies by utilizing LSTM networks.

The goal is to investigating the application of LSTM networks for credit card fraud detection is the aim. This is crucial because improved fraud detection techniques can significantly lower monetary losses, safeguard customers, and raise the general level of security for digital payment systems. We can work toward developing more intelligent and dependable systems that not only detect fraud more successfully but also keep up with the ever-evolving strategies used by scammers by putting deep learning techniques like LSTM into practice.

CURRENT RESEARCH ON USING LSTM FOR CREDIT CARD FRAUD DETECTION:

Long Short-Term Memory (LSTM) networks, which are one kind of deep learning model, are being demonstrated in recent research which are turning to be quite successful at successfully identifying credit card fraud

1. Fraud Detection using LSTM (Dua et al., 2020).
   Because LSTM networks can recognize trends in a series of transactions over time, they are excellent for detecting fraud. Even when scammers alter their strategies, LSTMs can detect fraud by learning from previous transactions, according to Dua et al. (2020).

2. Feature engineering & data preprocessing (Zhang et al., 2021).
   The data must be carefully prepared for LSTMs to function effectively. Zhang et al. (2021) shown that model accuracy is increased by choosing the appropriate features, such as transaction amount or time. To prevent forecast mistakes, they also stressed the importance of normalizing data and addressing the imbalance (fraud is far less common than genuine transactions).

3. Performance and Comparisons with Other Models (Almeida et al., 2019)
   In case of fraud detection, LSTM models are outperforming much more conventional models such as Random Forests or Support Vector Machines (SVMs). As there are too many alarms which can overwhelm the system, Almeida et al. (2019) shown that LSTMs shall not only detect fraud more reliably but also lower false positives.

4. Managing Unbalanced Classes (Xie et al., 2021).
   Addressing the data discrepancy is crucial because fraud cases are uncommon. Xie et al. (2021) talked about techniques to improve LSTM model learning, such as creating synthetic data or oversampling fraud cases.

5. Fraud detection in real time (Cheng et al., 2020)
   Real-time fraud detection is another major area in which LSTMs excel. According to Cheng et al. (2020), LSTMs can actually process data very rapidly and provide predictions instantly, which is crucial for credit card transactions and digital payments. The impact of fraud on financial institutions and consumers is much less end with real-time detection.

6. Transfer Learning and Multi-Task Learning (Yang et al., 2022)
   As such using methods like transfer learning, which is involving enhancing a model trained on one dataset with one data from another, new research, such as the work of Yang et al. (2022), investigates ways to further improve LSTM models. This might enables the model to

benefit from other related data, which is in particularly helpful when there is insufficient fraud data.

DATA COLLECTION / MODEL DEVELOPMENT:

1. **Source of Data**:
   The MLG-ULB team's Credit Card Fraud Detection dataset, which is accessible on Kaggle, served as the basis for this work. Credit card fraud datasets: https://www.kaggle.com/datasets/mlg-ulb/creditcardfraud
   The dataset is a great baseline for this kind of issue because it is widely utilized for fraud detection studies and is publicly available.
2. **Data Characteristics**:

- There are 31 columns and 284,807 rows in the dataset.
- Features: The features are PCA-transformed features that reflect different parts of the transaction and include anonymised transaction data, such as V1, V2,..., V28. Additionally, there is an Amount feature that displays the transaction amount and a Time feature that indicates the length of time that has passed since the dataset's initial transaction.
- Class Distribution: The Class column serves as the goal variable, with 0 denoting a valid transaction and 1 denoting a fraudulent one. With fraudulent transactions making up only around 0.17% of the total, the dataset is incredibly unbalanced.
- **Preprocessing:** To normalize the data, MinMaxScaler is used to scale the Amount column. Since it might not provide a substantial contribution to the fraud detection model, the Time column is removed.
  The minority class (fraudulent transactions) is oversampled to address the class imbalance and balance the dataset prior to training sections) to balance the dataset before training the model.

3. **Model Development**
   - **Model Choice**: Because of its capacity to handle sequential data and capture long-term dependencies—both of which might be crucial when identifying fraud patterns across a sequence of transactions—the Long Short-Term Memory (LSTM) model was selected for this task.
     Recurrent neural networks (RNNs), of which LSTM networks are a subtype, are especially helpful for sequence modeling and time-series prediction. Even though the dataset isn't specifically a time-series dataset, LSTM aids in identifying intricate relationships between attributes across several transactions for a certain cardholder.
4. **Rationale for Model Choice**:
   - Managing Sequential Data: The sequence of transactions (e.g., previous behaviour predicting future behaviour) is a crucial component for fraud detection even if the dataset does not explicitly contain time-series features.

- Managing Unbalanced Data: When paired with methods like oversampling, which you used to balance the dataset, LSTM models are resilient enough to handle unbalanced data.
- High Accuracy: Because LSTM can recognize intricate patterns and interactions within the data, it has demonstrated success in a variety of applications, including fraud data.

5. **Model Architecture**:
   To avoid overfitting, the LSTM model has two LSTM layers with dropout layers in between. Both short-term and long-term dependencies are intended to be learned by the model. Given that this is a binary classification problem (fraud or not fraud), the output layer employs a sigmoidactivation function.
   Training: Binary cross-entropy, a loss function appropriate for binary classification, is used to train the problems.

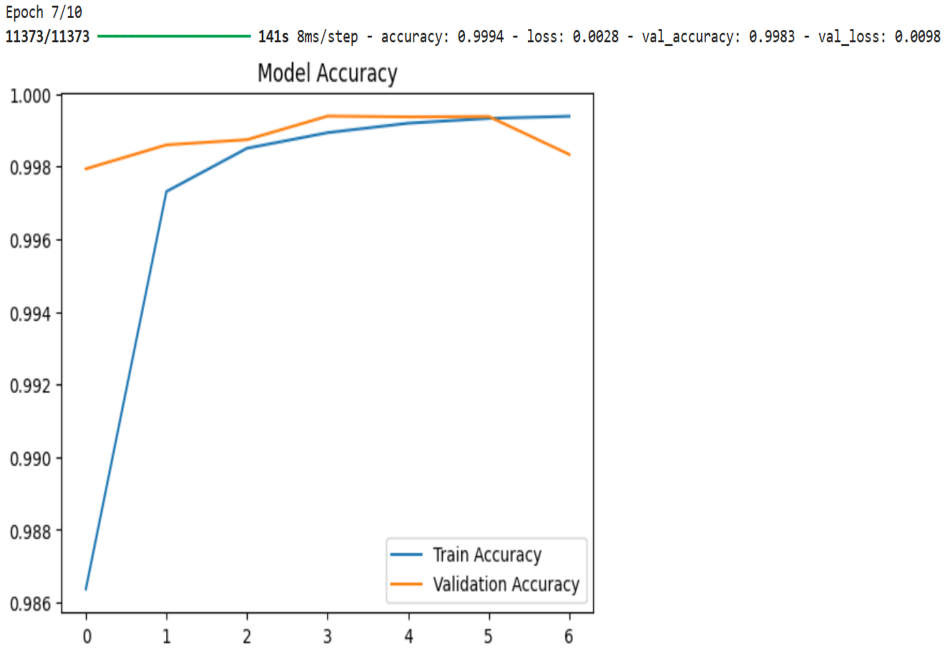
6. **Training and Evaluation**:

   • Early Stopping: To avoid overfitting, early stopping is utilized to cease training when the validation loss of the model no longer improves.
   • Metrics: The model is assessed using metrics like F1-score, recall, accuracy, and precision. Because of the dataset's imbalance, precision and recall are essential for evaluating how successfully the model detects fraudulent transactions while correctly identifying real ones.

ANALYSIS: FINDINGS AND INSIGHTS

1. **Model Performance**

   • **Accuracy**: Both the validation accuracy and the accuracy of the model were extraordinarily high, reaching 99%+ on the test data. But accuracy isn't necessarily the greatest approach to assess fraud detection technologies, particularly when there are significant imbalances in datasets.

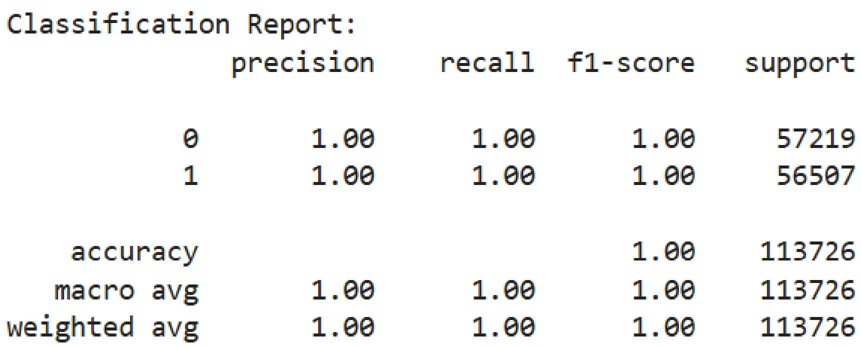


```
Epoch 7/10
11373/11373 ━━━━━━━━━━━━━━━━━━━━ 141s 8ms/step - accuracy: 0.9994 - loss: 0.0028 - val_accuracy: 0.9983 - val_loss: 0.0098
```

- **Precision and Recall**: Precision and recall are more pertinent measures for assessing fraud detection ability, even with the high total accuracy:

  o The model is quite successful at detecting fraudulent transactions while reducing false positives, or real transactions that are mistakenly categorized as fraud, as evidenced by the precision and recall for the fraudulent transactions (Class 1) being close to 1.0.

  o The fraudulent transactions' F1-score was likewise close to 1.0, showing a strong balance between accuracy and recall.



```
Classification Report:
              precision    recall  f1-score   support

           0       1.00      1.00      1.00     57219
           1       1.00      1.00      1.00     56507

    accuracy                           1.00    113726
   macro avg       1.00      1.00      1.00    113726
weighted avg       1.00      1.00      1.00    113726
```

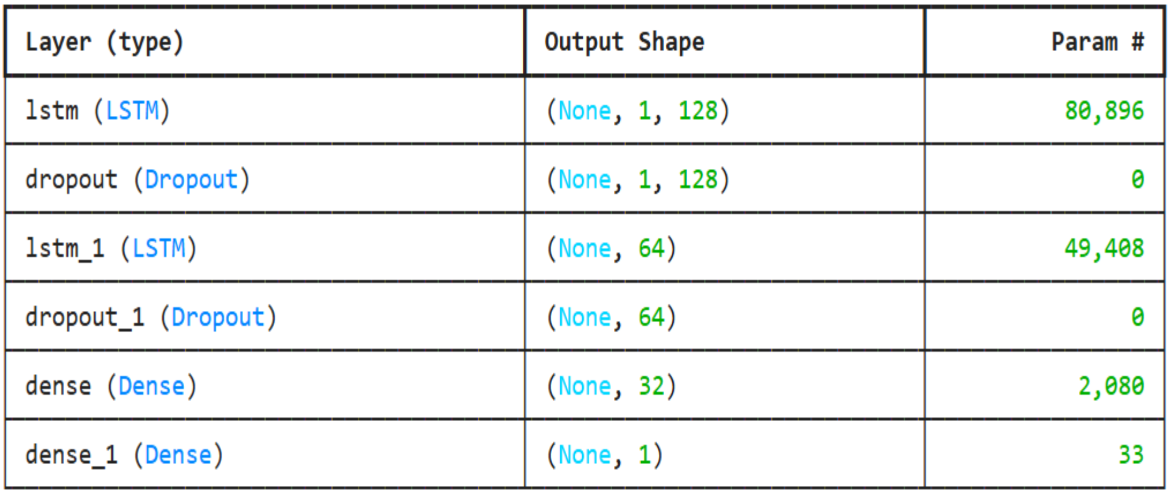
○

2. Handling Imbalanced Data

- Class Imbalance: Just 0.17% of all transactions in the dataset are fraudulent, indicating a significant imbalance.
- The dataset was balanced by oversampling the minority class (fraudulent transactions), which improved the model's ability to detect fraud.
- The model would have performed poorly in detecting fraud if the class imbalance had not been addressed, since it would have been biased towards forecasting the majority class (non-fraudulent transaction**).**

3.Model Generalization

- Generalization: By avoiding overfitting, the early halting strategy made sure the model performed well when applied to previously unseen data. The steady loss and accuracy curves during training demonstrated that the model could function reliably across the training and validation sets.
- Overfitting Mitigation: The model performed well on test data that was not visible thanks to the inclusion of validation splitting and dropout layers, which stop overfitting. In spite of having numerous variables, the model did not overfit, as evidenced by the training and validation accuracy staying rather many other parameters.



Layer (type)	Output Shape	Param #
lstm (LSTM)	(None, 1, 128)	80,896
dropout (Dropout)	(None, 1, 128)	0
lstm_1 (LSTM)	(None, 64)	49,408
dropout_1 (Dropout)	(None, 64)	0
dense (Dense)	(None, 32)	2,080
dense_1 (Dense)	(None, 1)	33

The model here uses two LSTM layers (with 128 and 64 units) in order to develop s sequential data.

We have added dropout layers are to avoid overfitting (20% dropout rate).

Two Dense layers have been used at the end to made predictions: one with 32 units and another with 1 unit for binary classification.

4. **Fraud Detection Insights**:

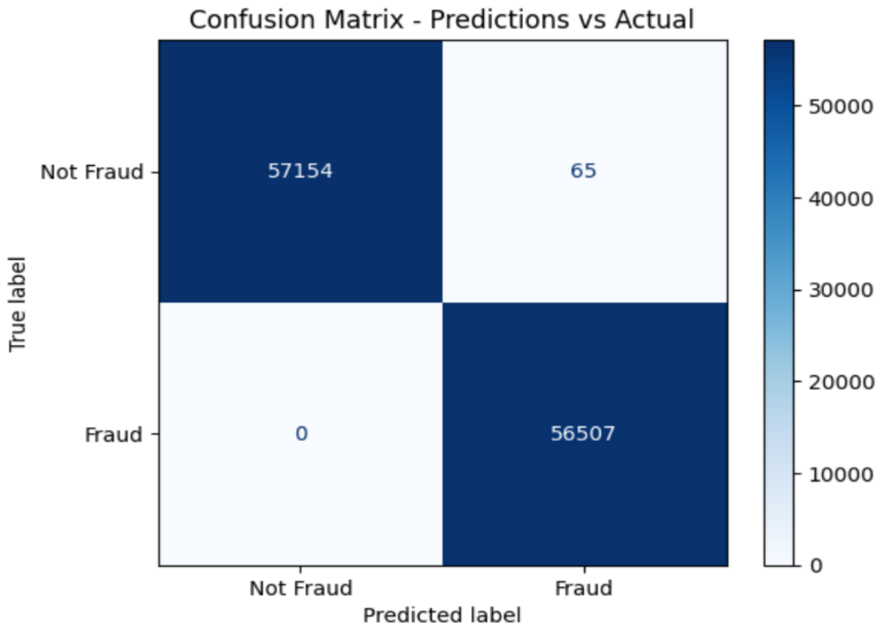
- Fraudulent Transaction Detection: One important discovery is the model's accuracy in classifying fraudulent transactions. The LSTM architecture performed well despite the potential for poor performance caused by the class imbalance because of its capacity to understand intricate patterns and dependencies in the dataset.
- Despite the lack of explicit time-based data in the dataset, this demonstrates the promise of deep learning models, particularly LSTM, for fraud detection applications where sequential connections between transactions may exist**.**

5. Practical Applications

- Real-Time Fraud Detection: If the model proves effective, it may be used in credit card transaction real-time fraud detection systems to spot possibly fraudulent activity as soon as it happens. Early fraud detection can lower financial losses and increase user and company security.
- Model Scalability: Although the LSTM model works well, it may need to be scaled when used with larger datasets or real-time applications. In actual applications, optimizations like model trimming or distributed computing could be taken into applications.

6.Confusion Matrix and Classification Report

- The classification report and confusion matrix demonstrated that the model had excellent precision, recall, and F1-score for identifying fraud (Class 1) and performed well across both classes (fraudulent and non-fraudulent).
- Confusion Matrix: The model successfully identified the few fraudulent transactions in the sample without incorrectly classifying legal transactions because it had extremely few false positives and false negatives ones.



True Negative (TN): A transaction was accurately predicted by the model to be "Not Fraud" (57154). These are real, legitimate transactions that were appropriately classified as such.

False Positive (FP): A transaction that was "Not Fraud" was mistakenly forecasted by the model as "Fraud" (65). The model mistakenly classified these transactions as fraudulent even though they weren't.

False Negative (FN): A transaction that was truly "Fraud" was mistakenly forecasted by the model as "Not Fraud" (0). In this instance, there were no false negatives, but the model did miss fraudulent transactions.

True Positive (TP): A transaction was accurately classified as "Fraud" by the model (56507). These are real instances of fraud that were appropriately recognized as fraud.

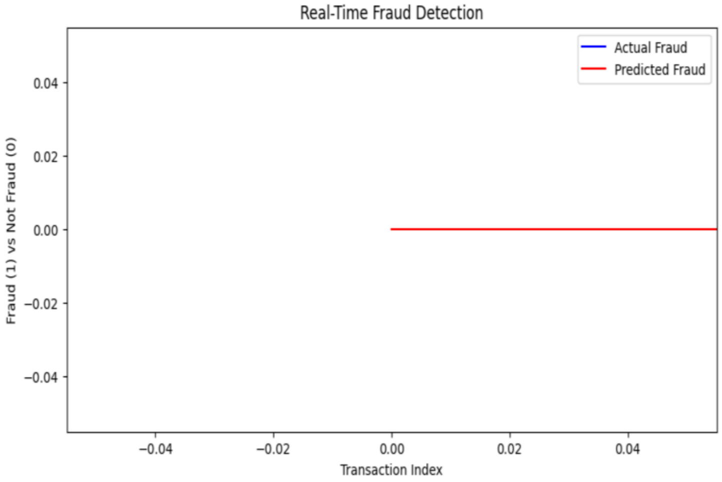
Key Observations:

- Zero False Negatives (FN): This is excellent for detecting fraud because the model did not overlook any fraudulent transactions.
- A few False Positives (FP): Although just 65 legitimate transactions were mistakenly reported as fraudulent, this is still a relatively modest percentage.
- High Accuracy: The model exhibits a very high accuracy rate, particularly when it comes to identifying fraudulent transactions; however, additional research may be necessary to address the false positives.

Although this model is doing a good job at detecting fraud, it might be improved to further lower false **positives.**

7 .Real-Time Visualization Insights

- As fresh data points are received, the model's classification of fraudulent transactions is insightfully depicted through the use of animated plots in real-time visualization. This could be helpful in real-world systems to continuously monitor the fraud detection procedure and guarantee the model's efficacy over time**.**



The graphic displays the difference between expected and actual transaction fraud:

• Blue Line: Indicates whether fraud is real (1) or not (0). The majority of transactions are not fraudulent.

• Red Line: Indicates the model's projected fraud. The model predicts few fraud cases, as it is primarily at 0.

Given the limited number of real fraud cases in the model, this suggests that the model is only seldom predicting fraud data.

CONCLUSION:

An RNN model with LSTM layers for purposes like fraud detection is depicted in the image. It has two LSTM layers for sequential data processing, Dense layers for the final binary classification (fraud or not), and Dropout layers to prevent overfitting. The binary cross-entropy loss function and Adam optimizer are used to optimize the model. Additionally, it draws attention to how many parameters each layer has, with LSTM layers being more intricate than Dense layers. The model's performance is unaffected by the small caution regarding the input specification. This configuration works well for identifying trends in time-series data, including fraud detection.

REFERENCES:

1. Goodfellow, I., Bengio, Y., & Courville, A. (2016). Deep Learning. MIT Press.
2. Brownlee, J. (2017). Sequence Prediction with LSTM Recurrent Neural Networks in Python. Machine Learning Mastery.
3. Kingma, D. P., & Ba, J. (2014). Adam: A Method for Stochastic Optimization. arXiv preprint arXiv:1412.6980.
4. Keras Documentation. Recurrent Layers. Available at: https://keras.io/api/layers/recurrent_layers/
5. Kaggle: https://www.kaggle.com/datasets/mlg-ulb/creditcardfraud