# DEEP LEARNING BASED HYBRID APPROACH OF DETECTING FRAUDULENT TRANSACTIONS

**[1]MIN JONG CHEON, [2]DONG HEE LEE, [3]HAN SEON JOO, OOK LEE**

[1]Department of Information Systems, Hanyang University, 222 Wangshimni-ro, Seongdong-gu, 04673 Seoul, South Korea
[2]Department of Information Systems, Hanyang University, 222 Wangshimni-ro, Seongdong-gu, 04673 Seoul, South Korea
[3]Department of Philosophy, Catholic University of Korea, 43, Jibong-ro, Bucheon-si, Gyeonggi-do, South Korea
*Department of Information Systems, Hanyang University, 222 Wangshimni-ro, Seongdong-gu, 04673 Seoul, South Korea

E-mail: [1]jmj2316@hanyang.ac.kr, [2]ryu03153@hanyang.ac.kr , [3]jooking012@catholic.ac.kr
*ooklee@hanyang.ac.kr

## ABSTRACT

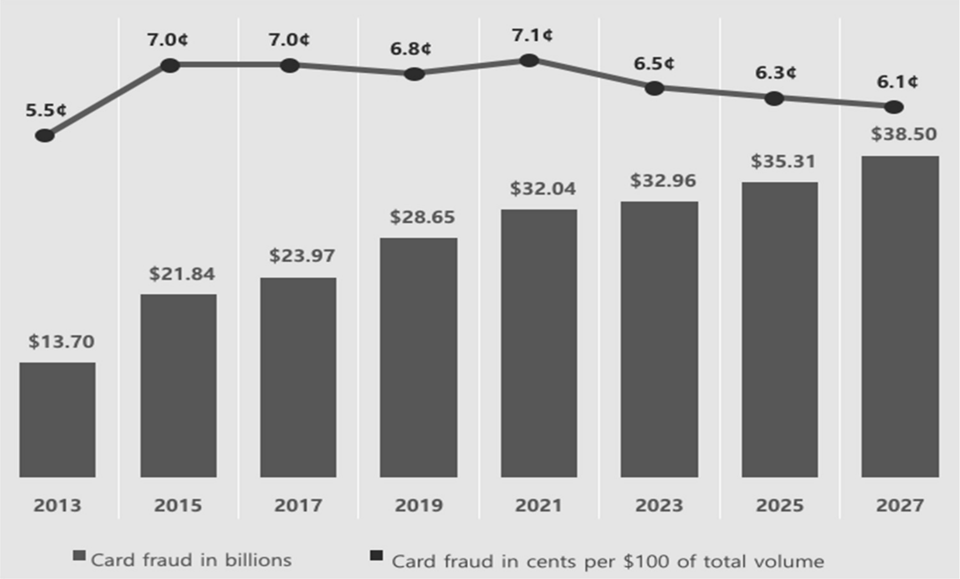
As daily transactions made with credit cards have been increasing, fraudulent transactions have also continuously increased. Therefore, the importance of detecting anomalous transactions has kept rising. The given dataset, from Kaggle, consists of imbalanced data, 99.83% of normal data and 0.17% of fraud data. Therefore, in order to solve this imbalance problem, we decided to construct a fraud detecting algorithm. Through constructing a new model with a hybrid approach of deep learning and machine learning, which is composed of a Bi-LSTM-Autoencoder and Isolation Forest, we successfully detected fraudulent transactions in the given dataset. This proposed model yielded an 87% detection rate of fraudulent transactions. Compared to other models (Isolation Forest, Local Outlier, and LSTM-Autoencoder), which show 79%, 3% and 82% detection rates, respectively, our proposed model attained the highest rate. On the contrary, when evaluated by accuracy score, our proposed model did not show a higher score. Even though our model has a similar accuracy score compared to other models and does not implement the Variational Autoencoder for feature selection, this model could potentially be utilized as an effective process to detect fraudulent transactions, especially with the number of global cases increasing along with the need for productivity, quicker detection.

**Keywords:** *Artificial Intelligence, Machine Learning, Deep Learning, EEG, Olfactory Impairment, Diagnosis*

## 1. INTRODUCTION

### 1.1 Background

Given the many daily transactions made by credit cards, the risk of fraudulent transactions has continuously increased. Losses caused by this type of fraud affect not only financial companies but also customers. According to the Nilson Report, worldwide losses from card fraud have continued to rise. Gross fraud losses to issuers, merchants, and acquirers of card transactions from merchants, as well as acquirers of card transactions from ATMs reached $28.65 billion in 2019, up 2.9% from $27.85 billion in 2018. Card-based payment systems worldwide experienced gross fraud losses equal to 6.78¢ for every $100 of total volume in 2019. By 2025, total payment card volume worldwide is projected to be $56.182 trillion, with gross card fraud globally expected to rise to $35.31 billion[1]. This is a very relevant problem that demands the attention of machine learning and data science communities where the solution to this problem can be automated. This problem is particularly challenging from the perspective of learning, as it is characterized by various factors such as class imbalance. The number of fraudulent transactions is much less than that of valid transactions [2]. To relieve this problem, this paper builds an efficient fraud detection model using anomaly detection. Detecting fraudulent transactions is an essential issue in anomaly detection. There are various models for detecting anomalies including SVM, logistic regression, decision tree, etc. Among the many anomaly detection methods, this research will apply the Isolation Forest algorithm, which is one of the most studied branches in anomaly detection.



*<Figure 1> Graph Of Worldwide Card Fraud*

### 1.2 Objectives

As we try to resolve the fraudulent problem mentioned above, where the dataset is imbalanced, we need to consider methods to solve the imbalance problem. Two solutions can be applied to this issue. The first is augmenting the given dataset and classifying the fraud[3]. Another option is to use a fraud detecting algorithm[4]. We focused on fraud detection and tried several fraud detecting methods from machine learning models. The most representative models are Local Outlier and Isolation Forest. The objective was to find the most optimal fraud detecting method for the dataset by utilizing these methods. As several studies were already conducted with these models, we focused on developing a new model with higher performance. We believed that feature extraction from the dataset was vital because of the large and imbalanced dataset. Thus, we constructed a hybrid approach of deep learning and machine learning, consisting of a Bidirectional Long Short-Term Memory Autoencoder (Bi-LSTM-Autoencoder) and Isolation Forest. When our suggested model detected fraud within the given datasets, synthetic data about the fraud were created through oversampling methods, the performance was evaluated through accuracy and a probability to detect a fraudulent transaction in the identified dataset. We combined them with the original input dataset then, compared the result of our model to different machine learning models such as Isolation Forest, Local Outlier and Long Short-Term Memory Autoencoder (LSTM-Autoencoder). Furthermore, as the accuracy score is not appropriate for the unbalanced dataset, we suggest the other evaluation score for the fraudulent detection.

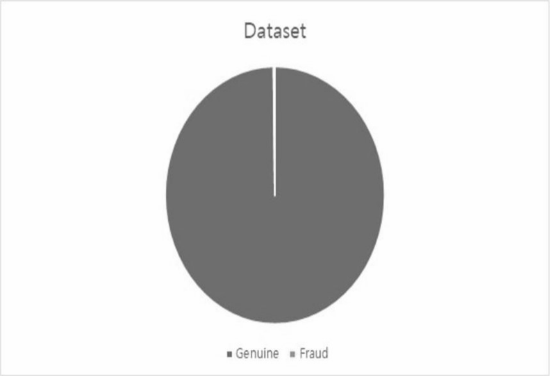
## 2. RELATED WORK

John and Naaz used local outlier factor and Isolation Forest to get the higher accuracy and they obtained the accuracy of 97% by outlier factor and 76% by Isolation Forest[4]. Mittal and Tyagi implemented both supervised learning and unsupervised learning algorithms to compare the performance on an imbalanced dataset. The results show that in an imbalanced dataset, an unsupervised learning algorithm, such as Isolation Forest and Local Outlier are better than supervised learning ones, such as Logistic Regression or Random Forest [5]. Pumsirirat and Yan built a model of a Restricted Boltzmann Machine (RBM) and Autoencoder (AE) based on Keras to classify the target and they achieved an AUC score of 0.96[6]. Varmedja et al. utilized a Synthetic Minority Over-sampling Technique (SMOTE) to solve an imbalanced dataset and Logistic Regression, Random Forest, Naive Bayes and multi-layer perceptron for classifying the target column. The results show accuracy of 97.64%, 99.96%, 99.23% and 99.93%, respectively [7]. Dhankhad et al. conducted research about applying supervised machine learning algorithms for credit card fraudulent transaction detection. Random Forest, Stacking Classifier, XGB Classifier and other machine learning models were utilized and the highest precision score was 95% from Random Forest [8].

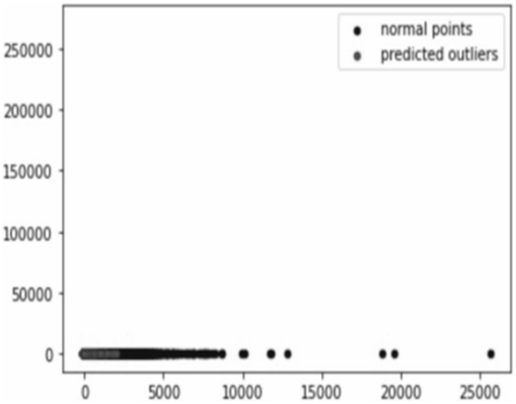
## 3. MATERIALS AND METHODS

### 3.1 Dataset Description

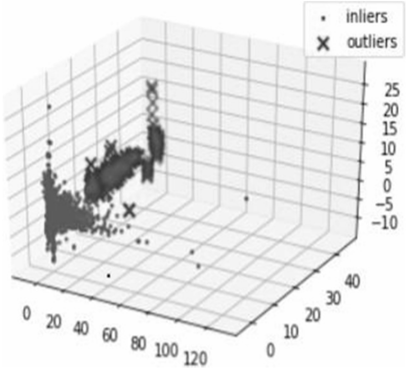
Our dataset is from Kaggle, and the dataset is available from https://www.kaggle.com/mlg-ulb/creditcardfraud[9]. The dataset is from transactions of a European bank in September 2013 that are CSV format. It contains 492 (0.17%) fraud transactions and 284,807 (99.83%) genuine transactions. The dataset is composed of 31 numerical features and V1, V2... V28 are from the Principal Component Analysis (PCA). Among the features, "Time" represents the time between the first transaction and other transactions. "Amount" indicates the total amount of transactions from the credit card. The target column is the "Class" that contains only 1 representing fraud transactions, and 0 for genuine transactions. Therefore, the binary classification should be conducted for treating the dataset.



*<Figure 2> Proportion Of Genuine And Fraud Data From Dataset*



<Figure 3> Visualizing dataset through 2D PCA



*<Figure 4> Visualizing Dataset Through 3D PCA*

### 3.2 Train and Test Dataset

With this dataset, 70% are used as training sets, and the remaining 30% as test sets through the train_test_split function in scikit-learn library. After splitting the dataset, we scaled each dataset through a standard scaler, which is also from scikit-learn library. The standard scaler is calculated by (1):

$$Z = (x - u) / s \quad (1)$$

where x denotes sample data, u denotes the mean of the training samples, and s denotes the standard deviation of the training samples.

### 3.3 LSTM

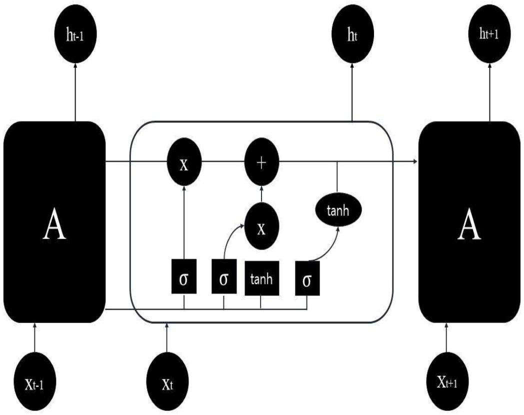
Long short-term memory (LSTM) is a recurrent neural network (RNN) which belongs to deep learning algorithms. Deep neural network (DNN) basically consists of a one-way network which implies that the input data went through the neural network once. On the contrary, RNN has a difference in overall network structure. Unlike the DNN, the output from the nodes in RNN becomes an input for the same nodes, and in this respect, it is

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

called “recurrent”. However, as RNN models have the drawback of a vanishing gradient problem, LSTM appears to overcome the disadvantage by including a ‘memory cell’. It can preserve information for long periods of time. The overall architecture of LSTM includes an input gate, an output gate, and a “forget” gate.

Firstly, the input data passes through the cell state and receives a sigmoid layer in sequence to decide whether to preserve the information or forget it through (2) and (3). In the second place, the tanh layer generates a $\tilde{C}_t$, in order to update the cell state (4). Thirdly, a new vector is created through (5), and by multiplying ft, the forget gates of LSTM determines whether to pass or update the information through the previous stage. Furthermore, $i_t * \tilde{C}_t$ is added and Ct gets updated. Moreover, the states are determined by output gates drawn on the previous cell states through (6). Lastly, the final output can be obtained through a discriminative passage of information with (7)[10].

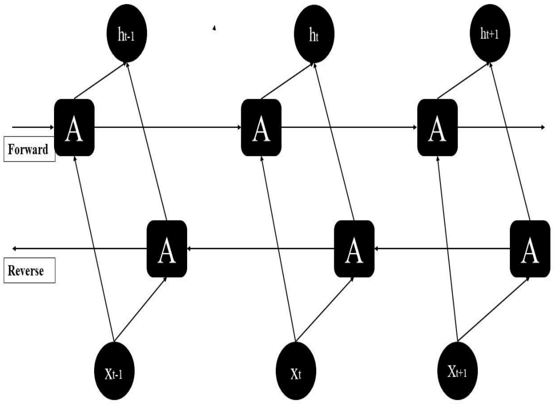
$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \tag{2}$$
$$i_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \tag{3}$$
$$\tilde{C}_t = \tanh(W_C \cdot [h_{t-1}, x_t] + b_C) \tag{4}$$
$$C_t = f_t * C_{t-1} + i_t * \tilde{C}_t \tag{5}$$
$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \tag{6}$$
$$h_t = o_t * \tanh(C_t) \tag{7}$$



<Figure 5> Overall architecture of LSTM

### 3.3 Bi - LSTM

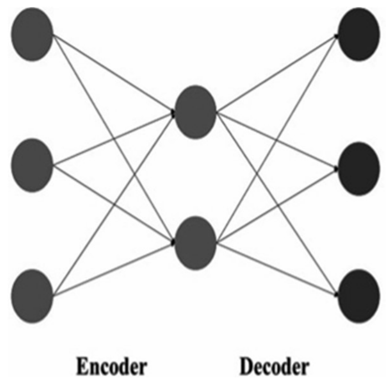
Unlike the DNN, the hidden layers of RNN and LSTM can preserve past information. However, as the calculation in these models is conducted through one way direction, the result of the output gates usually affects the previous cells. This is a main disadvantage of RNN and LSTM. In order to solve this issue, Bi–LSTM trains the model in two ways, both forward and reverse directions. This difference makes Bi-LSTM conduct end-to-end learning, which allows the model to train whole parameters while minimizing loss from the output [11].



*<Figure 6> Overall Architecture Of Bi-LSTM*

### 3.4 Autoencoder

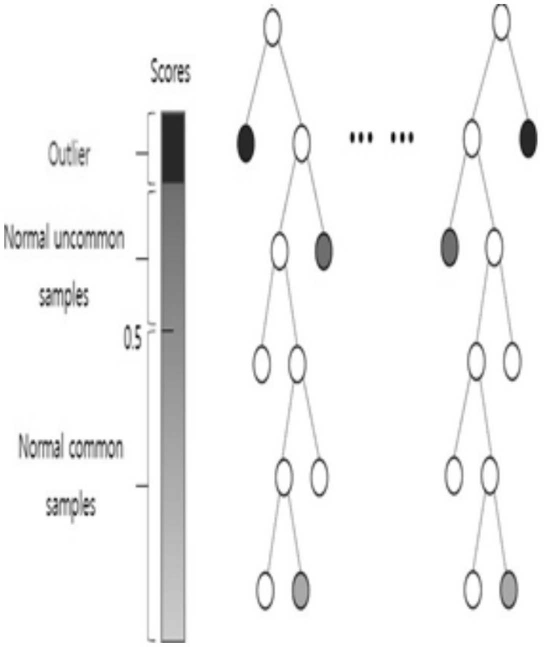
Autoencoder has a structure of feed-forward neural network that consists of the same input and out dimension. Autoencoder simply copies the input to the output, as shown in Figure 7, Setting a limitation to the neural network can make this simple neural network more complicated. Representative constraint is to limit the number of neurons between n input and output. This constraint has two advantages. Firstly, it would prevent autoencoders from simply copying inputs to the output. Secondly, it would learn how to represent data more effectively. This simple autoencoder is defined as an undercomplete autoencoder, which is a symbolic model of the autoencoder. The autoencoder incorporates two separate parts, which are the encoder and decoder. The encoder converts inputs into inner representations. The decoder transforms inner representations into outputs [12].



*<Figure 7> Overall Architecture Of Autoencoder*

### 3.5 Isolation Forest

Anomaly detection through unsupervised learning mostly uses training data to calculate the distance or density between the observations to define the normal category. Then it calculates the value from the test data to determine that as an anomaly if it is not included into the normal cate- gory. However. there are two disadvantages here. As the model is optimized for normal points, the anomaly detection performance is poor. Furthermore, it is difficult to apply density or observation calculations to large data sets or high- dimensional data because of their high computational cost. Isolation Forest is one of the unsupervised anomaly detection algorithms, which is often used to detect outliers in a dataset. As the name suggests, it is implemented based on a tree, which randomly splits data to isolate all observations. The Isolation Anomaly detection through unsupervised learning mostly uses training data to calculate the distance or density between the observations to define the normal category. Then it calculates the value from the test data to determine that as an anomaly if it is not included into the normal cate- gory. However. there are two disadvantages here. As the model is optimized for normal points, the anomaly detection performance is poor. Furthermore, it is difficult to apply density or observation calculations to large data sets or high-dimensional data because of their high computational cost. Isolation Forest is one of the unsupervised anomaly detection algorithms, which is often used to detect outliers in a dataset. As the name suggests, it is implemented based on a tree, which randomly splits data to isolate all observations. The Isolation Forest algorithm randomly selects dimensions to divide the space by any criterion. For normal data x0 1 inside a cluster, a large number of space splits must be performed to leave only one point in space and in complete isolation, but outlier data x1, far from the cluster, can only be isolated with a small number of space splits. In other words, normal data is treated near the terminal node of the tree with a large path length and outlier data is treated near the root node of the tree which has a small path length [13].

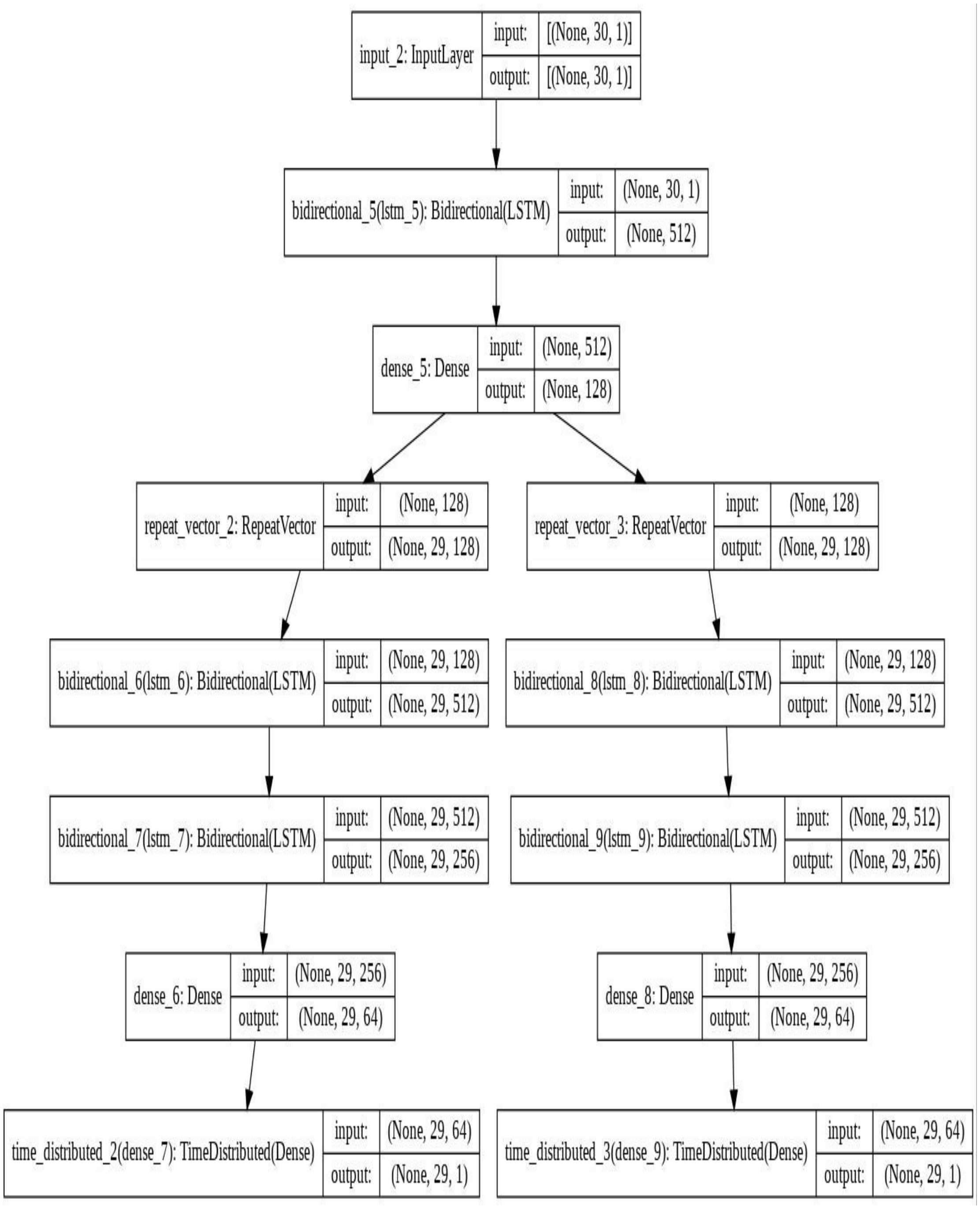


*<Figure 8> Overall Architecture Of Isolation Forest*

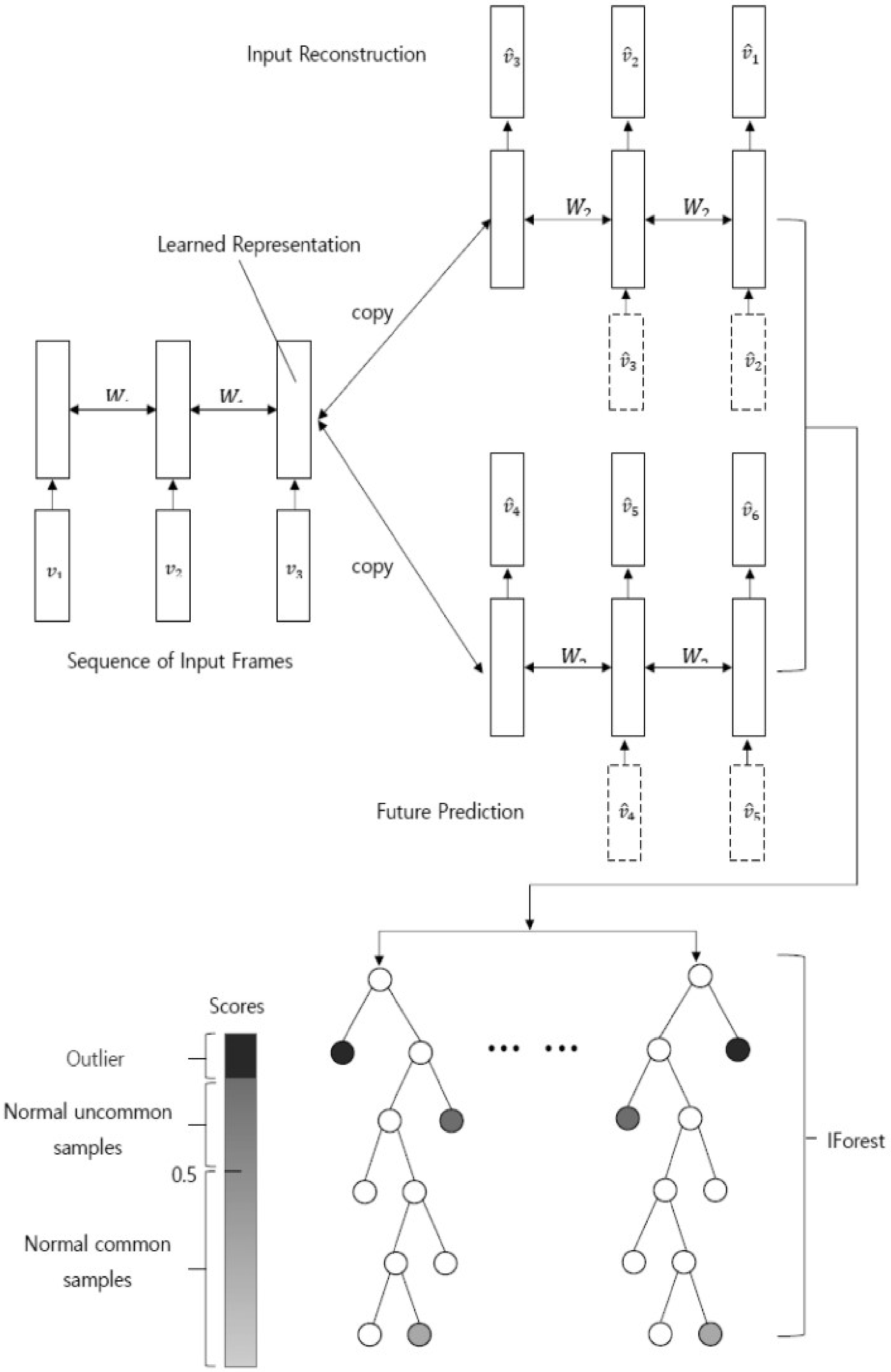
### 3.6 Proposed Model

Our proposed model consists of Bi-LSTM-Autoencoder and Isolation Forest. As previous research has proved, the autoencoder shows high efficiency in feature extraction from a dataset [14]. As our dataset is a time series one, we decided to use the Bi-LSTM-Autoencoder for feature extraction because Bi-LSTM shows higher performance than LSTM. Our suggested autoencoder consists of 2 parts, which are reconstruction and prediction. For the reconstruction, as our encoder is based on LSTM, the input shape should be three dimensional. Prediction is a structure for time-series prediction and the input sequence, and it constructs data to learn one point ahead by leaving the current time point (t) and output point (t+1). For compiling the encoder, the Adam optimizer was used and MSE for the loss function of the given autoencoder. We got predicted values from the constructed encoder and then used them as input data for the Isolation Forest. Then, Isolation Forest detects the anomaly data and shows the performance through accuracy and a probability to detect a fraudulent transaction.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195



*<Figure 9> Overall Summary Of Proposed Model*



*<Figure 10> Overall Architecture Of Proposed Model*

## 4. Result

In our experiment, we examined three other models. These models were Local Outlier, Isolation Forest, and LSTM-Autoencoder. On account of the experimental procedure, we were able to determine the most appropriate models for the fraud detection. The performance of the classifier was measured by an accuracy score, which is the simple ratio of the correctly predicted observations to total observations. The performance of each algorithm was evaluated by accuracy and a probability to detect a fraudulent transaction. Furthermore, a confusion matrix was provided for each model. <Figure 11>, <Figure 12>, <Figure 13>, <Figure 14> showed how the fraudulent transaction data were classified. <Figure 15> showed the accuracy score between the models, and most of them achieved about 97 %. However, as the data are too much imbalanced, the accuracy score is less meaningful. Therefore, we evaluated our models by probability of detecting fraudulent transaction, and in Figure <16>, our proposed model showed an 87% detection rate of fraudulent transactions. Other models such as Isolation Forest, Local Outlier, and LSTM-Autoencoder yielded 79%, 3% and 82%

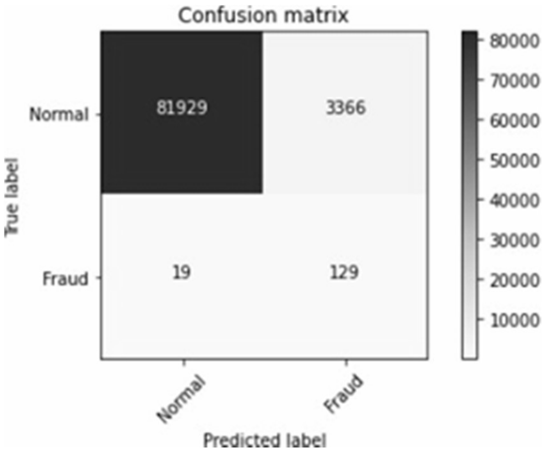
| | Predicted : NO | Predicted : YES |
|---|---|---|
| Actual : NO | FN | TP |
| Actual : YES | TN | FP |

*<Table 1> Evaluation Matrix In Machine Learning*

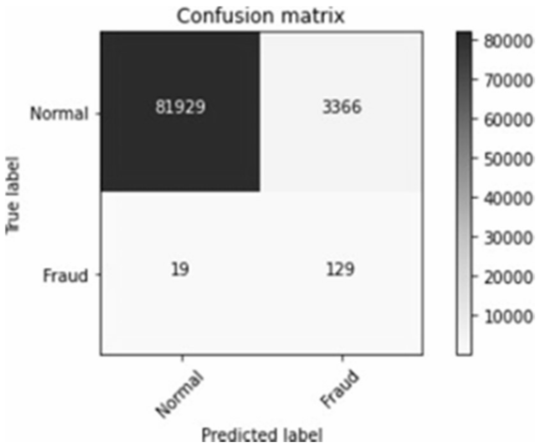
TP, TN, FN, FP denote as follows :

- True Positives(TP) : Transaction data where the true label is positive and which are correctly predicted to be positive
- False Positives(FP) : Transaction data where the true label is negative and which are correctly predicted to be positive
- True Negatives(TN) : Transaction data where the true label is negative and which are correctly predicted to be negative
- False Negatives(FN) : Transaction data where the true label is positive and which are correctly predicted to be negative

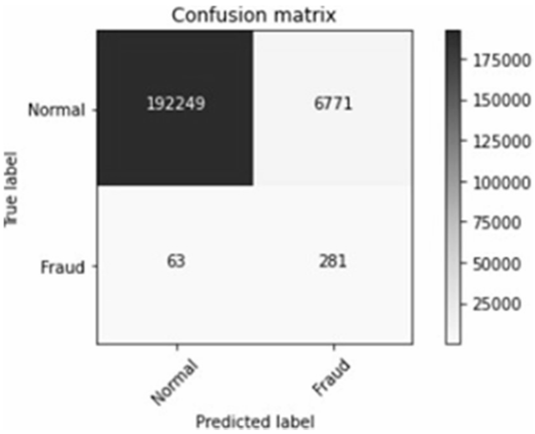
$$\text{Accuracy :}\ \frac{TP+TN}{TP+FN+TN+FP} \qquad (8)$$



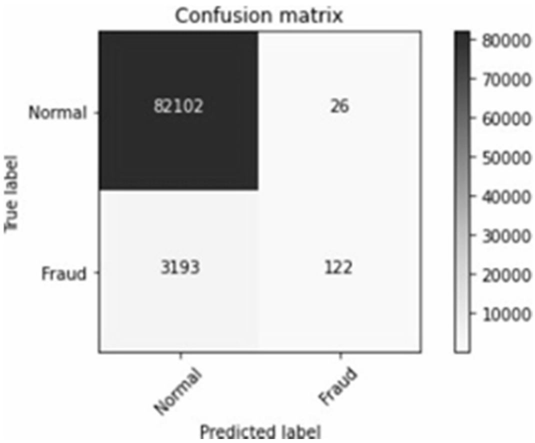
<Figure 11> Confusion matrix of proposed model



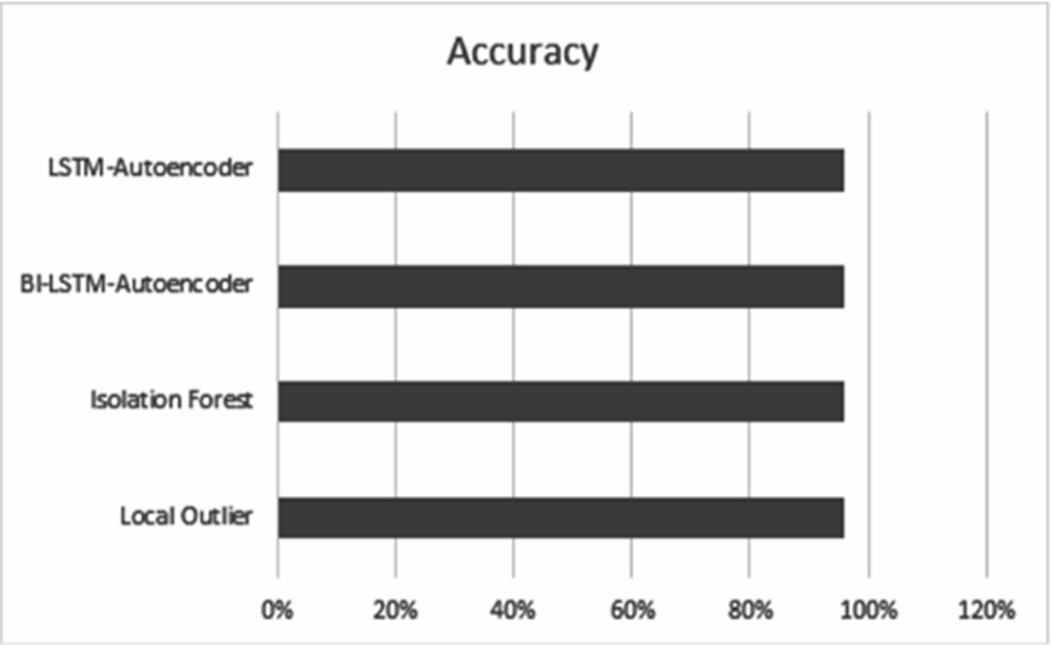
*<Figure 12> Confusion Matrix Of LSTM-Autoencoder*



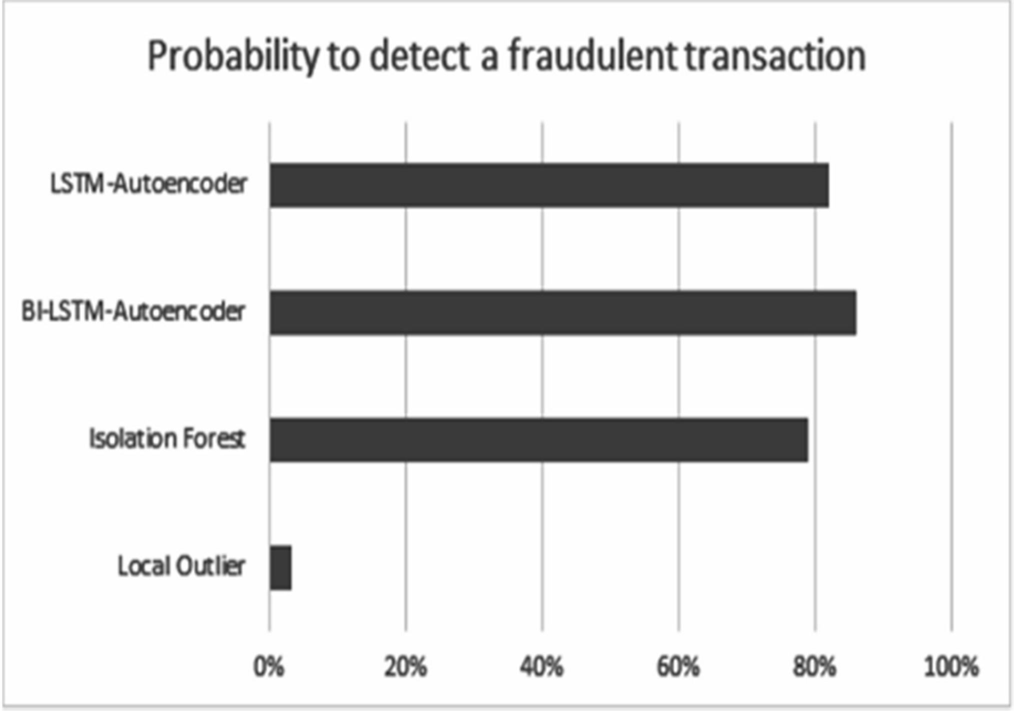
<Figure 13> Confusion Matrix Of Isolation Forest



<Figure 14> Confusion Matrix Of Local Outlier



<Figure 15> Accuracy Of Each Model



*<Figure 16> Probability Of Detecting Fraud Transactions*

## 5. DISCUSSION

### 5.1 Limitation

One critical limitation to note is that the accuracy of our model shows similar accuracy when compared to different models such as LSTM – Autoencoder, Bi-LSTM – Autoencoder, Isolation Forest, Local Outlier. Furthermore, even though Variational Autoencoder (VAE) and extended isolation forest is a renowned algorithm for fraud detection and feature extraction, we did not utilize this model for our dataset[15].

### 5.2 Principal Finding

One of the principal findings our proposed model showed is the highest probability for detecting fraud (87%) in the given dataset. Furthermore, our model successfully combined Bi-LSTM-Autoencoder and Isolation Forest. From the result, we concluded that extracting features from the autoencoder has superior results than solely using Isolation Forest or Local Outlier algorithms. In addition, our proposed model was superior to LSTM-Autoencoder, which means that bidirectional LSTM is more efficient than LSTM. Lastly, unlike the related works which provide the accuracy score for the evaluation, we suggest the probability of finding fraudulent one. From this result, further research could apply our models to different datasets for fraud detection, especially when the dataset is time series one.

## 6. Conclusion

The main goal of our experiment was to develop a hybrid model which is composed of the Bi-LSTM-Autoencoder and Isolation Forest from the unbalanced dataset. We succeeded in discovering the fraudulent transactions on the given dataset and also proposed a new evaluation method which makes our research worth while compared to the previous researches. Although there are some limitations that still need to be addressed, with this novel approach, our proposed model could still be applied to fraud detection with a high probability of detecting fraud transactions. The number of fraud cases is increasing rapidly from day to day, our suggested model could be applied to different datasets in different fields, especially when it is time series one, as Bi- LSTM based models show higher efficiency on time series dataset.

## REFERENCE

[1] Card fraud Losses Reach $28.65 Billion [Internet]. [cited 2021Mar1]. Available from: https://nilsonreport.com/mention/1313/1link/

[2] S P Maniraj, Aditya Saini, Shadab Ahmed, Swarna Deep Sarkar. Credit Card Fraud Detection using Machine Learning and Data Science. International Journal of Engineering Research and. 2019;08(09).

[3] Moon J, Jung S, Park S, Hwang E. Conditional Tabular GAN-Based Two-Stage Data Generation Scheme for Short-Term Load Forecasting. IEEE Access. 2020;8:205327–39.

[4] John H, Naaz S. Credit Card Fraud Detection using Local Outlier Factor and Isolation Forest. International Journal of Computer Sciences and Engineering. 2019;7(4):1060–4.

[5] Tyagi S, Mittal S. Sampling Approaches for Imbalanced Data Classification Problem in Machine Learning. Lecture Notes in Electrical Engineering. 2019;:209–21.

[6] Pumsirirat A, Yan L. Credit Card FraudDetection using Deep Learning based on Auto-Encoder and Restricted Boltzmann Machine. International Journal of Advanced Computer Science and Applications. 2018;9(1).

[7] Varmedja D, Karanovic M, Sladojevic S, Arsenovic M, Anderla A. Credit Card Fraud Detection - Machine Learning methods. 2019 18th International Symposium INFOTEH-JAHORINA (INFOTEH). 2019;

[8] Dhankhad S, Mohammed E, Far B. Supervised Machine Learning Algorithms for Credit Card Fraudulent Transaction Detection: A Comparative Study. 2018 IEEE International Conference on Information Reuse and Integration (IRI). 2018;

[9] ULB MLG-. Credit Card Fraud Detection [Internet]. Kaggle. 2018 [cited 2021Jun1]. Available from: https://www.kaggle.com/mlg-ulb/creditcardfraud

[10] Hochreiter S, Schmidhuber J. Long Short-Term Memory. Neural Computation. 1997;9(8):1735–80.

[11] Graves A, Schmidhuber J. Framewise phoneme classification with bidirectional LSTM and other neural network architectures. Neural Networks. 2005;18(5-6):602–10.

[12] Chen J, Sathe S, Aggarwal C, Turaga D. Outlier Detection with Autoencoder Ensembles. Proceedings of the 2017 SIAM International Conference on Data Mining. 2017;:90–8. 90–98.

[13] Liu FT, Ting KM, Zhou Z-H. Isolation Forest. 2008 Eighth IEEE International Conference on Data Mining. 2008;

[14] Yao R, Liu C, Zhang L, Peng P. Unsupervised Anomaly Detection Using Variational Auto-Encoder based Feature Extraction. 2019 IEEE International Conference on Prognostics and Health Management (ICPHM). 2019;

[15] Kawachi Y, Koizumi Y, Harada N. Complementary Set Variational Autoencoder for Supervised Anomaly Detection. 2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). 2018;