International Journal of
Intelligent Engineering & Systems

http://www.inass.org/

Hybrid Sampling and Similarity Attention Layer in Bidirectional Long Short Term Memory in Credit Card Fraud Detection

Valliammal Narayan[1]* **Sudhamathy Ganapathisamy**[1]

[1]*Avinashilingam Institute for Home Science and Higher Education for Women, India*
* Corresponding author's Email: valliammal_cs@avinuty.ac.in

Abstract: Machine learning methods are widely applied in credit card fraud detection to improve its efficiency and automate the process. The existing methods in credit card fraud detection have limitations of imbalance data problem. In this research, Hybrid Sampling (HS) - Similarity Attention Layer (SAL) – Bidirectional Long Short Term Memory (BiLSTM) is proposed to improve the classification performance of credit card fraud. Two datasets such as European data and Revolutionary analytics are applied to evaluate the SAL-BiLSTM model. Hybrid sampling of SMOTE-ENN model over-sample the minority class and under-sample majority class that helps to reduce difference between original data and generated data. The SAL is introduced to measure similarity of sequence of data to provide weight to unique features and reduces overfitting problem in classification. The proposed method has the advantage of removing the overlapped samples in the majority class and increasing the instances in the minority class. The SAL method in the proposed method helps to focus on unique features in the datasets to improve the classification performance. The BiLSTM model performs forward and backward analysis to find the relevant features for classification. The proposed HS-SAL-BiLSTM has a 99.2 % recall value and the existing RF-SMOTE-Support Vector Machine (SVM) has a 97.7 % recall value.

Keywords: Bidirectional long short term memory, Credit card fraud detection, Hybrid sampling, Similarity attention layer, Support vector machine.

1. Introduction

Recently, credit card frauds in online payment have been increased dramatically pushing e-commerce organizations and banks to apply automatic fraud detection systems based on machine learning techniques on transaction logs [1]. Supervised binary classification system trained on the sample dataset which provides a promising solution to distinguish fraudulent and non-fraudulent instances to identify the illicit transaction. Dataset present for fault detection consists of highly imbalance class [2]. Data-driven rules or expert-driven rules or both types of rules combination are used to find the features in a fraud detection system. Fraud discovery of specific scenarios is identified using fraud investigators in expert-driven rules. Expert driven rules learn fraudulent patterns and detect new incoming transactions of data-stream [3]. Machine learning techniques such as Support Vector Machines, decision trees, logistic regression, rule-induction techniques, and Artificial Neural Networks were used for classification. Several methods are combined or standalone methods were used to form a hybrid model for detection [4]. Imbalance classification has minority class consisting of a small number of data instances compared to data instances in majority class in the dataset. This problem is defined as data skewed distribution and extremely imbalance dataset. The ratio of criminal or fraudulent activities is considerably smaller than genuine and legitimate ones [5].

Every transaction was manually verified to detect fraud is infeasible for credit card issuers due to huge credit card transactions. Statistical and machine learning methods are used for the automatic detection of fraudulent transactions [6]. A credit card transaction dataset is highly unbalanced, where fraudulent transactions instances are less than 0.1 %. The supervised learning methods face big challenges for unbalanced datasets and cost function is needed for adjustment or good sampling procedure. The unsupervised methods have high false alarm rates due to the relation of a legitimate transaction and unusual behavior [7, 8]. Deep learning methods are powerful learning techniques and achieved promising results in various fields such as image processing and pattern recognition. Long Short Term Memory (LSTM) model is applied to detect credit card fraud in supervised learning category and sequence classification [9, 10]. The objectives and contributions of the research are discussed as below

1. Hybrid sampling technique of SMOTE-ENN were applied to balance the data instances in the dataset. SMOTE method over-sample the minority class in the dataset and ENN technique under-sample the majority class in dataset.
2. Similarity Attention Learning helps to focus on unique features related to the class to reduce overfitting and solve imbalance data problem. The SAL measure the similarity of sequence of data in Bi-LSTM model to find the importance of data.
3. The SMOTE-ENN model effectively solves the imbalance data problem, SAL method reduces overfitting problem and Bi-LSTM model provides efficient classification performance. The hybrid sampling-SAL-BiLSTM model provides efficient performance in credit card fraud detection compared to existing model.

The organization of the paper is given as follows: credit card fraud detection of recent research was reviewed in Section 2 and the explanation of the proposed method is given in Section 3. The simulation setup is given in Section 4. The result is given in Section 5 and the conclusion of this research work is given in Section 6.

2. Literature review

Credit Card Fraud is involved in stealing money using credit card information and it leads to loss of money from victims. Credit card fraud detection was carried out by various researchers and some of the notable papers were reviewed in this section.

Rtayli and Enneya [11] applied hybrid method of Recursive Feature Elimination (RFE), GridSearchCV and Synthetic Minority Oversampling Technique (SMOTE). The RFE method involves finding the relevant features for the prediction process. The GridSearchCV method was applied for the hyperparameter optimization and SMOTE method was applied for the sampling of data to overcome the imbalance data problem. The SVM classifier was applied for the sampling data and GridSearchCV is applied to select the parameter for a classifier. The three datasets of European data, PaySim data and Data set 03 were used to evaluate the developed method. The oversampling method involves in bias the model and the Grid search method selects the parameter in the specific value. The RFE reduce relevant features and SMOTE data is deviated from the original data.

Xie [12] applied a heterogeneous ensemble learning model based on data distribution to overcome the limitation of data imbalance problem. The two real credit card datasets were used to test the performance of the developed method. The KNN and K-means methods were applied to find the distribution of the majority class and reduce the information. The RMDD undersampling method is applied to analyze the distribution of the majority class and reduces the information loss within the majority class. The RMDD and ensemble learning method provides higher performance in a highly imbalanced dataset. The KNN method is sensitive to the outlier in the data distribution and the k-means method has random initialization.

Xiao and Jiao, [13] applied Multiple Instance Learning (MIL) and the self-training LSTM model for credit card detection. The MIL method consists of the Affinity Propagation (AF) clustering method for the learning process. The real-world dataset and simulated dataset were used to evaluate the MIL-LSTM model in credit card fraud detection. The PaySim simulator generates the data and evaluates the hiding label in the developed method. The developed method shows higher learning in the few labeled time series dataset. The MIL-LSTM has lower efficiency in the large-scale dataset and learning efficiency is low. The LSTM model has vanishing gradient problem that affects model efficiency in classification.

Seera [14] applied 13 statistical and machine learning methods for payment card fraud detection. The developed method was tested on both publicly available and real transaction datasets. The genetic algorithm identified aggregated features was evaluated based on a statistical hypothesis test. This provides better discriminative power than original features in the fraud detection process. The Bayesian method, tree based method, neural network and regression methods were used for the credit card fraud detection. The developed method handles imbalance dataset and provides considerable performance. The feature selection of genetic algorithm has the limitation of lower convergence and is easily trapped into local optima.

Carcillo [15] combined supervised and unsupervised methods for the credit card fraud detection process. Unsupervised outlier scores at various levels of granuality were used for the detection of credit card fraud. The outlier score was measured in the developed method to eliminate the outliers in the dataset. Principal Component Analysis (PCA) was applied for the dimensional reduction in the feature selection. Independent variable is less interpretable in the feature selection and this tends to information loss in the system.

Trisanto [16] proposed modified Focal loss for imbalance XGBoost to improve the ability of focal loss and this is used to provide weight to the class that are often misinterpreted. The modified focal loss method is evaluated using credit card fraud dataset and compared with existing method. The W-loss is used in the focal loss to use imbalance parameter and tuning hyper-parameter. The model has overfitting problem due to increases the weight values to input data.

Trisanto [17] proposed two stage feature reduction technique for selection of optimal features from the dataset. Random undersampling and Instance Hardness Threshold sampling were applied to deal with imbalance data problem. The two stage feature reduction technique was evaluated using ULB credit card fraud detection dataset for classification. The under-sampling method provides boost to the recall and MCC score. The developed method has limitation of outlier and overfitting problem in the classification.

3. Proposed method

The credit card fraud of two datasets were used in this research to evaluate HS-SAL-BiLSTM model. The hybrid sampling method of SMOTE and ENN was applied to reduce the overlapped majority class and increase the minority class. The SAL is introduced in the proposed method to focus on unique features to improve the classification performance. The BiLSTM is applied with sampled data and weight values to perform credit card fraud detection. The overview of the proposed HS-SAL-BiLSTM in credit card fraud detection is shown in Fig. 1.

3.1 Sampling method

Data balancing or data sampling is a common method applied for imbalance datasets and machine learning that consists of three kinds, under-sampling, over-sampling and hybrid sampling. The minority class data samples are generated in over-sampling and data samples in the majority class are eliminated by under-sampling. The hybrid method combines the under-sampling and over-sampling methods.

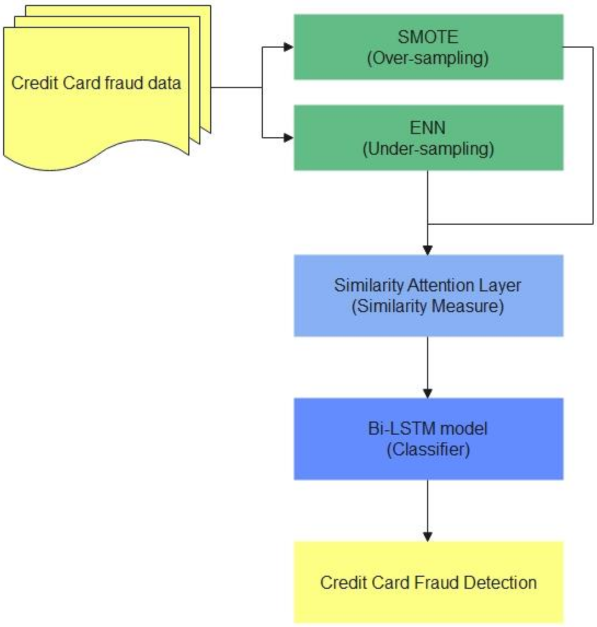

Figure. 1 The proposed HS-SAL-BiLSTM model in credit card fraud detection

This research applies a hybrid SMOTE-ENN model to balance data instances in the credit card fraud detection dataset. Generally, Edited Nearest Neighbor (ENN) [18] method is applied to remove the irrelevant overlapping samples to balance the dataset and SMOTE [19] method over-samples the minority class to balance the training dataset. The minority and majority classes of a dataset are subjected to the presence and absence of credit card fraud. The SMOTE technique is applied to randomly generate new samples based on Neural Networks of minority class samples to increase the number of minority classes. Then ENN model is applied to eliminate the overlapping samples in the dataset.

The distribution pattern from original samples is eliminated with overlapped samples and creates new artificial samples based on the SMOTE-ENN method. The distribution data attributes of the training dataset are implemented with the SMOTE-ENN model.

3.2 Similarity attention layer

Opinion terms and aspect terms of high correlation and contextual information are not related to a sentiment polarity of a sentence. Context words are not equally distributed in sentence semantics content and the self-attention technique is used to extract relevant words based on a higher weight to

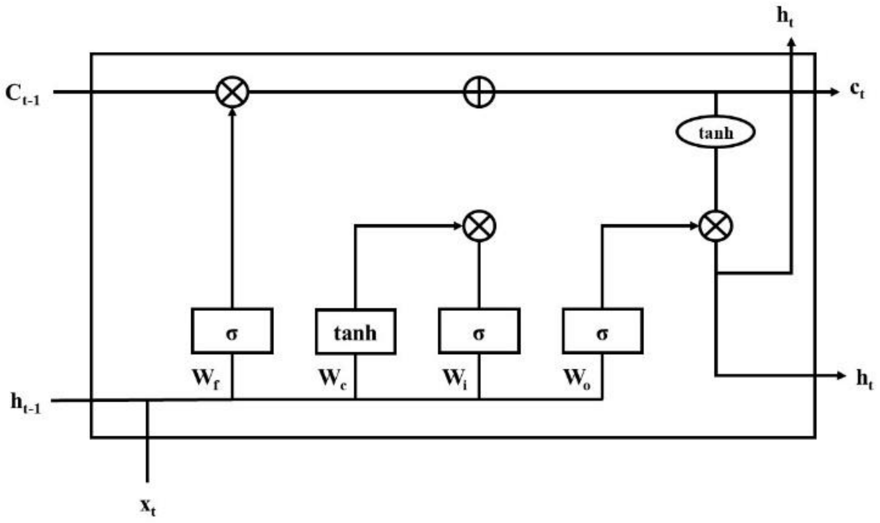

Figure. 2 The LSTM cell architecture

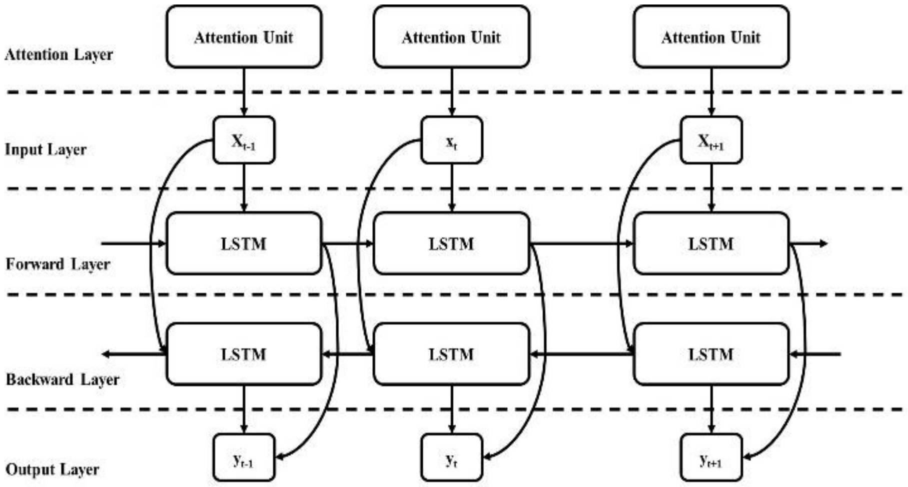

Figure. 3 Attention layer and BiLSTM model

improve the importance in the self-attention mechanism.

A hidden vector h_t is produced by the BiLSTM neural networks. A simple Multi-layer perceptron provides input to the hidden vector h_t to get a new hidden representation u_t. The words importance symbolize a weight value to calculate a context vector of a word level u_w and for h_t as u_t. A high dimensional representation of context vector u_w to measure the importance of different words in sentences, which is based on joint learning and random initialize in the training process. The hidden vector h_t the weighted mean is measured based on a softmax function that measures similarity between u_t and u_w using function M. The formula for each step is given in Eqs. (1) to (3).

$$u_t = \tanh(W_w h_t + b_w) \quad (1)$$

$$\partial_t = \frac{\exp(u_t^T u_w)}{\sum_t \exp(u_t^T u_w)} \quad (2)$$

$$s = \sum_t \partial_t h_t \quad (3)$$

3.3 Bidirectional long short term memory

The LSTM model stores the useful information for the long term based on forgetting gate and cell. Classification of credit card fraud requires both recent data and also historical data. The self-feedback method of hidden layer handles long-term dependence problems in the LSTM model and the LSTM model uses three gates and memory cells to store information in long-term features [20]. The Bi-LSTM cell structure is shown in Fig. 2.

The x_t denotes input data at time t in LSTM cell, h_{t-1} denotes previous moment, c_t denotes memory cell value, and h_{t-1} denotes the output. The LSTM unit calculation process is explained in each step.

The weight matrix of the candidate memory cell is represented as W_c, the bias is represented as b_c, and candidate memory cell is represented as $\widetilde{c_t}$, as shown in Eq. (4).

$$\widetilde{c_t} = \tanh(W_c.[h_{t-1}, x_t] + b_c) \tag{4}$$

The output unit of LSTM model h_t is calculated in Eq. (5).

$$h_t = o_t * \tanh(c_t) \tag{5}$$

The weight matrix of the output gate is represented as W_0, the bias is represented as b_0, output gate controls the state value of the memory cell, and the output gate o_t is calculated, as given in Eq. (6).

$$o_t = \sigma(W_0.[h_{t-1}, x_t] + b_0) \tag{6}$$

The last LSTM unit state value is represented as c_{t-1}, and current moment memory cell c_t is measured, as given in Eq. (7).

$$c_t = f_t * c_{t-1} + i_t * \widetilde{c_t} \tag{7}$$

Where "*" denotes dot product. The candidate and last cell state value of memory cell update control input and forget gate.

The weight matrix of forget gate is represented as W_f, the bias is represented as b_f, forget gate controls historical data update in-state value of memory cell, and forget gate f_t is measured, as given in Eq. (8).

$$f_t = \sigma(W_i.[h_{t-1}, x_t] + b_f) \tag{8}$$

The σ represents the sigmoid function, W_i denotes the weight matrix, b_i denotes the bias, current input data of memory cell state value controlled by input gate i_t, as shown in Eq. (9).

$$i_t = \sigma(W_i.[h_{t-1}, x_t] + b_i) \tag{9}$$

LSTM model read, reset, update using control gates and memory cells to keep long-time information. Internal parameters of sharing mechanism in LSTM model control dimension's settings of weight matrix is output dimensions.

The sequence of each token is used for learning in two LSTMs in the Bi-LSTM model based on past and future token context. The sequence is used for learning in LSTM from left to right and other from right to left, i.e., forward and backward manner. The function $\vec{h}$ of a hidden unit on a hidden forward layer at each time step t that is used to compute hidden state h_{t-1} of the previous step and current input step x_t. The current input step x_t and hidden state $\overleftarrow{h}_{t+1}$ of future step are used to compute a hidden backward layer. The $\overleftarrow{h_t}$ and $\vec{h_t}$ represent backward and forward context representation that concatenates into a long vector. The teacher given target signals of combined outputs are used for classification. The Bi-LSTM model overview is shown in Fig. 3.

4. Simulation results

Credit Card Fraud detection involves applying machine learning methods to detect fraud automatically. The dataset information, metrics, system requirement and parameter settings were given in this section.

Datasets: Credit Card Fraud data [21] and Revolution Analytics [22] datasets were used to evaluate the proposed HS – SAL – BiLSTM method performance. Credit Card Fraud data consists of European transaction information that consists of 30 attributes. Revolution Analytics consists of 7 features and 1 target variable. The Revolution Analytics consists of 5.96 % fraud data in the datasets for detection.

System Requirement: The system of Intel i7 processor, 6 GB of a graphics card, 16 GB RAM and Windows 10 64-bit OS. The proposed and existing methods were trained and tested in the same environment and same dataset.

Metrics: Accuracy metrics were used to analyze the overall performance of the model in fraud detection. Precision and Recall metrics were used to measure the class-wise performance of the detection model. The formulas for accuracy, precision, and recall are given in Eqs. (10) to (14), respectively.

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \times 100 \tag{10}$$

$$Precision = \frac{TP}{TP+FP} \times 100 \tag{11}$$

$$Recall = \frac{TP}{TP+FN} \times 100 \tag{12}$$

$$Specificity = \frac{TN}{TN+FP} \times 100 \tag{13}$$

$$F - measure = 2 \times \frac{Precision \times Recall}{Precision + Recall} \tag{14}$$

Parameter Settings: The BiLSTM model batch size is 64, the number of epochs is 50, dropout rate is 0.2, and learning rate is 0.1.

Table 1. Performance analysis of sampling method

Sampling Method	Methods	Accuracy (%)	Precision (%)	Recall (%)	Specificity (%)	F-Measure (%)
Without SMOTE	RF	53	69	64	63	66.41
	SVM	56	58	52	54	54.84
	LSTM	58	51	64	67	56.77
	BiLSTM	61	67	66	69	66.50
SMOTE	RF	60	65	74	73	69.21
	SVM	73	60	71	74	65.04
	LSTM	74	73	68	71	70.41
	BiLSTM	74	61	66	76	63.40
ENN	RF	72	75	77	78	75.99
	SVM	76	71	74	76	72.47
	LSTM	76	79	71	73	74.79
	BiLSTM	80	74	75	77	74.50
Hybrid Sampling	RF	89	91	93	94	91.99
	SVM	90	88	92	91	89.96
	LSTM	91	90	90	92	90.00
	BiLSTM	96.3	97.1	97.3	97.4	97.20

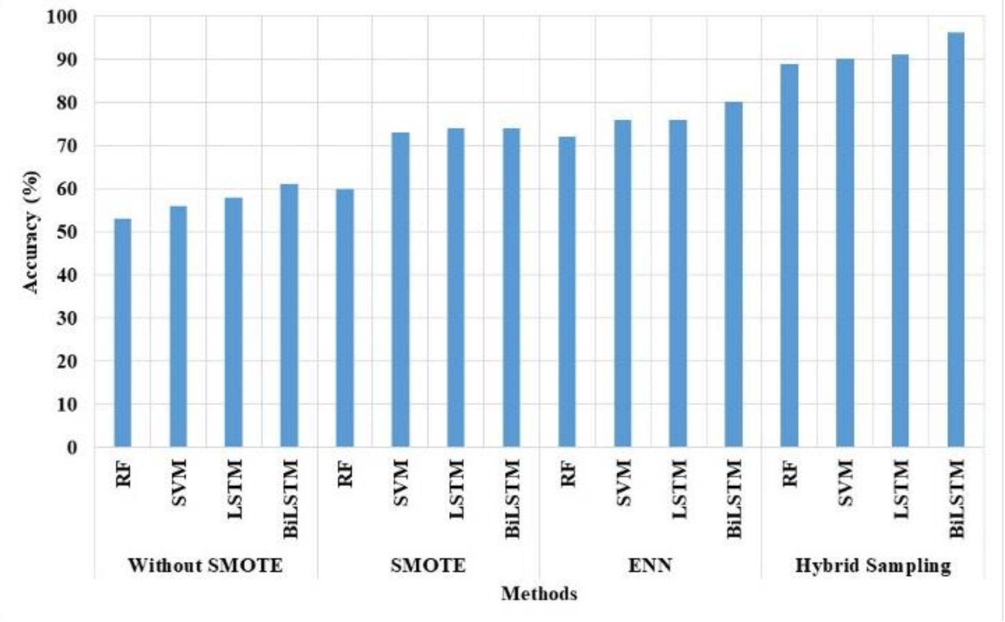

Figure. 4 Accuracy of the sampling methods

5. Results

Credit card fraud detection based on transaction information is required in the bank to reduce the fraud and loss in the bank. Various existing methods were applied for credit card fraud detection methods and have the limitation of imbalance data. In this research, HS – SAL – BiLSTM model is proposed to increase the efficiency of credit card fraud detection.

The proposed hybrid sampling method is compared with the single sampling methods such as SMOTE and ENN, as shown in Table 1. The classifiers such as Support Vector Machine (SVM), Random Forest (RF), LSTM and BiLSTM models were used for performance analysis. This shows that the hybrid sampling method has higher performance than individual sampling methods in terms of accuracy, precision, and recall. The hybrid method involves in increases the minority class and decreasing in majority class without high deviation in the dataset. The hybrid method is also involved in reducing the overlapping samples in the dataset. The Hybrid Sampling method improves the precision and recall that shows the model provides higher performance in class-wise detection.

The accuracy of the hybrid sampling method and individual sampling methods is shown in Fig. 4. The hybrid sampling method provides higher accuracy compared to the individual sampling method. The hybrid sampling method improves the

Table 2. Performance analysis of attention layer

Sampling Method	Methods	Accuracy (%)	Precision (%)	Recall (%)	Specificity (%)	F-Measure (%)
Without Attention Layer	RF	53	69	64	63	66.41
	SVM	56	58	52	53	54.84
	LSTM	58	51	64	65	56.77
	BiLSTM	61	67	66	68	66.50
Similarity Attention Layer	RF	85	92	86	85	88.90
	SVM	85	88	95	92	91.37
	LSTM	87	87	90	90	88.47
	BiLSTM	93	88	90	91	88.99
HS - SAL - BiLSTM	RF	90	95	89	89	91.90
	SVM	93	94	90	92	91.96
	LSTM	95	94	92	94	92.99
	BiLSTM	99.1	99.3	99.2	97.4	99.25

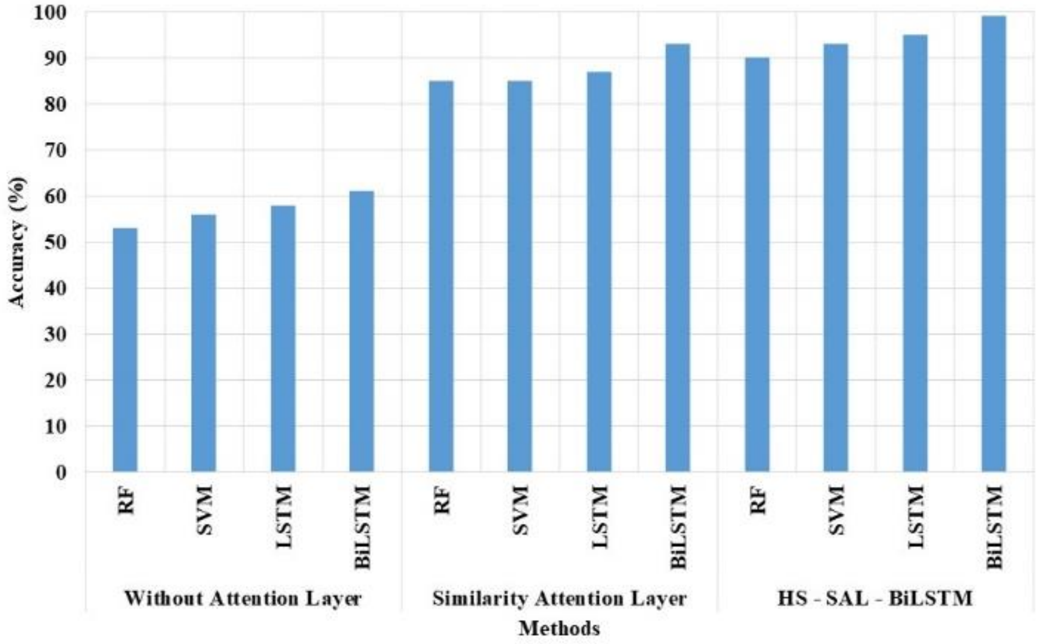

Figure. 5 Accuracy of attention layer

performance of the classifier in credit card fraud detection. The SVM model has lower performance in class wise classification that affects overall performance. The BiLSTM model has the advantage of examining the data in the forward and backward manner to improve classification performance.

The proposed HS–SAL–BiLSTM model is compared with SAL method without sampling and standard classifiers, as shown in Table 2. The HS - SAL - BiLSTM method has higher efficiency in terms of Accuracy, Precision, and Recall model. Table 2 shows that the similarity attention layer model has improved the performance of fraud detection. The similarity attention layer measures the similarity of the sequence of the data in the BiLSTM network to find the importance of the input. This process helps to find the unique features in the given input dataset and apply them for the classification. The HS - SAL - BiLSTM method has the advantage of balancing the dataset and finding the unique features in the datasets for improving the detection performance.

Table 3. Performance of hybrid method

Methods	Accuracy (%)	Precision (%)	Recall (%)
HS - BiLSTM	96.3	97.1	97.3
SAL - BiLSTM	93	88	90
HS - SAL - BiLSTM	99.1	99.3	99.2

The accuracy of the proposed method and attention layer for credit card fraud detection is shown in Fig. 5. This shows that the attention layer improves the overall efficiency of the proposed method. The BiLSTM model provides higher performance in the classifier due to its advantages of analyzing the data in a forward and reverse manner. The SVM model has lower efficiency in handling the

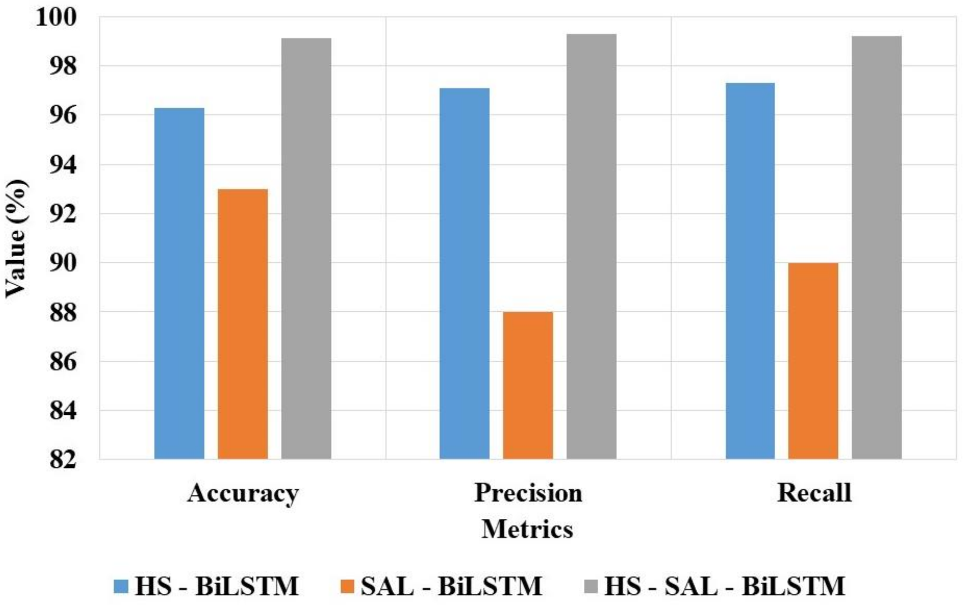

Figure. 6 Performance of the hybrid method

Table 4. Comparative analysis of proposed method

Methods	Accuracy (%)	Recall (%)	Precision (%)	MCC (%)
HS - SAL - BiLSTM	99.1	99.2	99.3	99.4
RFE -SMOTE - SVM [11]	97.2	97.7	96.4	96.8
KNN - K Means [12]	95.3	95.1	95.4	95.3
Self-training LSTM [13]	93.2	92.4	93.7	93.4
GA - SVM [14]	87.2	84.7	88.1	89.3
PCA [15]	85.2	83.1	87.1	87.5

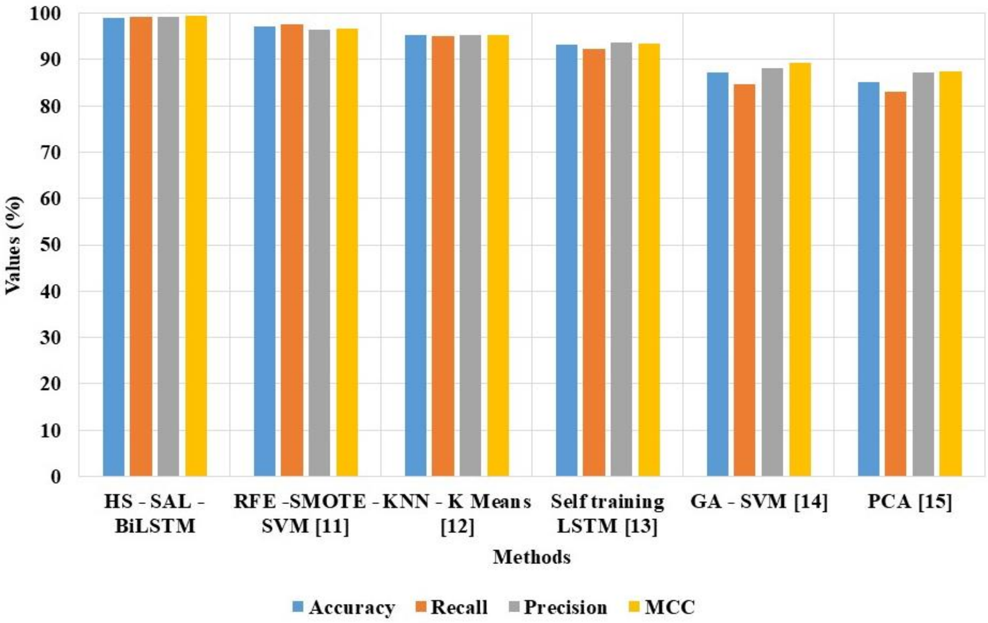

Figure. 7 Comparative analysis of proposed method

imbalance data due to feature measure values are similar. The attention layer has lower performance in class wise classification due to imbalance data. The proposed method overcome this problem based on balancing the dataset.

The hybrid method is compared with similarity attention layer and hybrid sampling, as shown in Table 3 and Fig. 6. The proposed method has the advantage of balancing the dataset and focusing on unique features to perform classification. The hybrid

Received: March 27, 2022. Revised: July 6, 2022.

method has higher performance than an individual method in terms of accuracy, precision, and recall.

5.1 Comparative analysis

The HS-SAL-BiLSTM method is compared with existing research in credit card fraud detection in terms of accuracy and recall.

The proposed HS-SAL-BiLSTM method is compared with existing research models in credit card fraud detection, as shown in Table 4 and Fig. 7. This shows that the proposed HS-SAL-BiLSTM model has higher performance in terms of accuracy and recall. The accuracy shows the overall performance and recall shows the class wise performance. The LSTM and KNN model has lower efficiency in handling higher imbalance datasets. The existing SVM method has lower efficiency in handling the imbalance data. The PCA model reduces the number of features in the dataset and affects the performance of the model.

6. Conclusion

Credit Card Fraud Detection based on machine learning methods has the limitation of imbalance data problem. This research proposes the HS-SAL-BiLSTM method to improve the performance of credit card fraud detection. The HS method combines the SMOTE and ENN model to decrease the instances in the majority class and increase the instances in the minority class. The SAL helps to focus on the unique features for BiLSTM that help to improve the classification performance. The HS-SAL method improves the class wise classification performance that improved the precision and recall value. The proposed HS-SAL-BiLSTM method has 99.1 % accuracy, 99.2 % recall, and the existing RFE-SMOTE-SVM method has 97.2 % accuracy and 97.7 % recall. The future work of the proposed model can involve improving the multi-class classification performance.

Conflicts of Interest

The authors declare no conflict of interest.

Author Contributions

The paper conceptualization, methodology, software, validation, formal analysis, investigation, resources, data curation, writing—original draft preparation, writing—review and editing, visualization, have been done by 1st author. The supervision and project administration, have been done by 2nd author.

References

[1] U. Fiore, A. D. Santis, F. Perla, P. Zanetti, and F. Palmieri, "Using generative adversarial networks for improving classification effectiveness in credit card fraud detection", *Information Sciences*, Vol. 479, pp. 448-455, 2019.

[2] F. Carcillo, A. D. Pozzolo, Y. A. L. Borgne, O. Caelen, Y. Mazzer, and G. Bontempi, "Scarff: a scalable framework for streaming credit card fraud detection with spark", *Information Fusion*, Vol. 41, pp. 182-194, 2018.

[3] J. Jurgovsky, M. Granitzer, K. Ziegler, S. Calabretto, P. E. Portier, L. H. Guelton, and O. Caelen, "Sequence classification for credit-card fraud detection", *Expert Systems with Applications*, Vol. 100, pp. 234-245, 2018.

[4] K. Randhawa, C. K. Loo, M. Seera, C. P. Lim, and A. K. Nandi, "Credit card fraud detection using AdaBoost and majority voting", *IEEE Access*, Vol. 6, pp. 14277-14284, 2018.

[5] S. Makki, Z. Assaghir, Y. Taher, R. Haque, M. S. Hacid, and H. Zeineddine, "An experimental study with imbalanced classification approaches for credit card fraud detection", *IEEE Access*, Vol. 7, pp. 93010-93022, 2019.

[6] X. Zhang, Y. Han, W. Xu, and Q. Wang, "HOBA: A novel feature engineering methodology for credit card fraud detection with a deep learning architecture", *Information Sciences*, Vol. 557, pp. 302-316, 2021.

[7] H. Zhu, G. Liu, M. Zhou, Y. Xie, A. Abusorrah, and Q. Kang, "Optimizing Weighted Extreme Learning Machines for imbalanced classification and application to credit card fraud detection", *Neurocomputing*, Vol. 407, pp. 50-62, 2020.

[8] Y. Lucas, P. E. Portier, L. Laporte, L. H. Guelton, O. Caelen, M. Granitzer, and S. Calabretto, "Towards automated feature engineering for credit card fraud detection using multi-perspective HMMs", *Future Generation Computer Systems*, Vol. 102, pp. 393-402, 2020.

[9] A. A. Taha and S. J. Malebary, "An intelligent approach to credit card fraud detection using an optimized light gradient boosting machine", *IEEE Access*, Vol. 8, pp. 25579-25587, 2020.

[10] A. G. D. Sá, A. C. Pereira, and G. L. Pappa, "A customized classification algorithm for credit card fraud detection", *Engineering Applications of Artificial Intelligence*, Vol. 72, pp. 21-29, 2018.

[11] N. Rtayli and N. Enneya, "Enhanced credit card fraud detection based on SVM-recursive feature

elimination and hyper-parameters optimization", *Journal of Information Security and Applications*, Vol. 55, p. 102596, 2020.

[12] Y. Xie, A. Li, L. Gao, and Z. Liu, "A Heterogeneous Ensemble Learning Model Based on Data Distribution for Credit Card Fraud Detection", *Wireless Communications and Mobile Computing*, 2021.

[13] Z. Xiao and J. Jiao, "Explainable Fraud Detection for Few Labeled Time Series Data", *Security and Communication Networks*, 2021.

[14] M. Seera, C. P. Lim, A. Kumar, L. Dhamotharan, and K. H. Tan, "An intelligent payment card fraud detection system", *Annals of Operations Research*, pp. 1-23, 2021.

[15] F. Carcillo, Y. A. L. Borgne, O. Caelen, Y. Kessaci, F. Oblé, and G. Bontempi, "Combining unsupervised and supervised learning in credit card fraud detection 2021", *Information Sciences*, Vol. 557, pp. 317-331.

[16] D. Trisanto, N. Rismawati, M. F. Mulya, and F. I. Kurniadi, "Modified Focal Loss in Imbalanced XGBoost for Credit Card Fraud Detection", *International Journal of Intelligent Engineering and Systems*, Vol. 14, No. 4, pp. 350-358, 2021, doi: 10.22266/ijies2021.0831.31.

[17] D. Trisanto, N. Rismawati, M. F. Mulya, and F. I. Kurniadi, "Effectiveness undersampling method and feature reduction in credit card fraud detection", *International Journal of Intelligent Engineering and Systems*, Vol. 13, No. 2, pp. 173-181, 2020, doi: 10.22266/ijies2020.0430.17.

[18] G. Douzas and F. Bacao, "Geometric SMOTE a geometrically enhanced drop-in replacement for SMOTE. Information sciences", Vol. 501, pp. 118-135, 2019.

[19] L. I. Kuncheva, P. Yousefi, and J. Almeida, "Edited nearest neighbour for selecting keyframe summaries of egocentric videos", *Journal of Visual Communication and Image Representation*, Vol. 52, pp. 118-130, 2018.

[20] T. Chen, R. Xu, Y. He, and X. Wang, "Improving sentiment analysis via sentence type classification using BiLSTM-CRF and CNN", *Expert Systems with Applications*, Vol. 72, pp. 221-230, 2017.

[21] "Credit card Fraud data - dataset by raghu543 | data.world.", https://data.world/raghu543/credit-card-fraud-data(accessed Dec. 09, 2021).

[22] "Index of /datasets/.", https://packages.revolutionanalytics.com/datasets/ (accessed Jan. 18, 2020).