Measurement: Sensors 27 (2023) 100793

Contents lists available at ScienceDirect

Measurement: Sensors

journal homepage: www.sciencedirect.com/journal/measurement-sensors

An effective fraud detection using competitive swarm optimization based deep neural network

T Karthikeyan [a,*], M Govindarajan [a], V Vijayakumar [b]

[a] Department of Computer Science & Engg., Annamalai University, Annamalai Nagar, India
[b] Department of Computer Science & Engg., Sri Manakula Vinayagar Engineering College, Puducherry, India

ARTICLE INFO

Keywords:
Fraud activity
Optimization
Deep learning
Classification
Online transaction
Neural network
And credit card

ABSTRACT

Frauds have no persistent patterns. They constantly alter their behaviour, necessitating the use of unsupervised learning. Fraudsters gain access to latest technology that enables them to commit scams via internet transactions. Consumers' habitual behaviour is assumed by fraudsters, and fraud trends very quickly. Fraud detection systems must utilize unsupervised learning to identify online payments since some fraudsters initially use online channels before switching towards other methods. The goals of this research are to construct a deep convolutional neural network model to detect anomalies from regular patterns produced by competitive swarm optimization, with a particular emphasis on fraud scenarios that cannot be detected using prior records or supervised learning (CSO). An unsupervised learning method called the suggested CSO-DCNN employs ReLu by making the inputs and outputs equal. The CSO-DCNN is classifies the fraud activity using real-time and other available dataset that is compared with the existing algorithms. The experimental results shows that the suggested CSO-DCNNobtains98.20%, 99.77%, 95.23% of accuracy for credit card, insurance and Mortgage Data set.

1. Introduction

The current state of traditional commerce is changing in response of virtual businesses and the internet. E-commerce has greater value now that there is a worldwide market, more freedom, and more competition. E-commerce also makes it simpler and more accessible to innovate in the payment and banking sectors. E-commerce makes life simple for users, whether they are consumers or business owners. It is important in the more competitive global economy [1]. People are moving away from conventional markets and toward the growing global market. It has a variety of restrictions in addition to providing customers with conveniences. For e-commerce or the digital market, online payment is essential [2].

The requirement for the financial and banking sectors has grown as the size of the global market has expanded. With the use of connected phones like a laptop, mobile phone, desktop, PDA, etc., online payments enable you to make transactions from any place [3]. The elimination of traditional commerce's constraints is the primary driver of the expansion of electronic payments. The user may physically visit the bank to complete the transaction without having to wait in a big line. It offers a number of advantages, such as speedier transactions that may be completed in a matter of minutes without having to visit a bank or wait in line [4].

There are two methods of execution both online and offline electronic payments. Virtual payment may be recognised as payment processing. Account owner name, PIN, card number, expiration date, and other sensitive details are needed for online payments [5]. Physical payment can be identified as offline payment. Cardholder presence and PIN are needed for offline payments. The first item needed for online payment fraud is a cardholder's credit card number. These frauds can be carried out in a variety of ways. Credit - card fraud is frequently committed using techniques including phishing, identity theft, skimming, using lost or stolen cards, card cloning, etc. [6].

In addition to these techniques, there are other mechanisms that enable credit card frauds, such as malware or keyloggers that can steal credit card information during an online transaction, or scanners which are used to read your credit card information. While online payment does not require signature or PIN number of your card, it makes process easier. Most of the websites are stealing card details and selling them to third party, number of the fraudsters are available on dark web so difficulty to trap [7,8]. For online payment there are various payment mechanisms available in market.

* Corresponding author.
E-mail addresses: karthi4cse@gmail.com (T. Karthikeyan), govind_aucse@yahoo.com (M. Govindarajan), vijayakumarv@smvec.ac.in (V. Vijayakumar).

https://doi.org/10.1016/j.measen.2023.100793
Received 31 December 2022; Received in revised form 20 March 2023; Accepted 12 May 2023
Available online 13 May 2023

Users are using different types of payment mechanisms as per their need and choice. Various types of payment mechanisms are Debit card, credit card, and online banking, e-wallet, etc. One more reason still around 60% of Internet users are preferring COD i.e. Cash on Delivery payment mechanism, just because of the risk factor in Online Payment [9,10]. CC is a popular, widely used, and promising digital payment method for online purchasing. As banks are giving customers credits for transactions that last a certain amount of time. Customer can choose to pay using CC with ease. Today, more people are using online payment methods [11]. The credit card is used both online and offline. As it gives facilities to users, end number of customers are there for CC [12].

The Reserve Bank is attempting to strengthen security in light of the increasing incidence of cybercrimes. In particular, a digital trade is seeing significant increase nowadays. The central banks agenda is to enhancing the security against cyber risk [13]. This is to confirm nonstop security beside the changing forms of internet-based security extortions. Indian market provides a massive E-Commerce. Currently, it provides facility to every age of groups. Where they can easily perform online shopping. It provides more benefits to today's generation compare with offline or traditional shopping techniques. Ecommerce offers huge number of choices, facilities, and discounts and offers [14].

The main contribution of this paper is.

- To study the problem of fraud in banks and its resolution by the hybrid deep learning techniques.
- We present an analysis of the bank fraud problem and its different forms and propose for each form, the variant of DCNNs that can be used for its resolution and the necessary adaptations.
- We also propose a hybridization of competitive swarm optimization and Deep Convolutional neural network (CSO – DCNN) methods to enhance the detection of fraudulent transactions. A system summarizing the use of the proposed solutions is designed and compared in this work.

The rest of the paper is organized as follows: we first present the various forms of fraud in banks as well as the indices used to discover it, then we discuss the types of deep learning solutions that can be used. In the third section, we discuss about the existing approach of Deep Convolutional neural network and its limitations in the bank transactions. In Fourth section, discuss the use of hybrid CSO-DCNN methods to fulfill the requirements for detecting each type of scam. The suggested methods are tested on bank datasets in the fifth part to demonstrate their reliability. The essay is concluded with a summation and future views in the sixth part.

1.1. Related works

One main advantage to ecommerce market in present days is the use of unlimited Internet access. Today rural are area are also using large numbers access, low price of data packs, compilations within service providing companies, offers, affordable smart phones, etc. are the main reason behind the growth of Digital Market. One more reason is India comes to the second highest position in terms of population. This is another reason behind Digitalization is increasing, covering Global market, etc.

As mentioned above, India is a large marketplace and there are no boundaries of opportunities. This is the reason India has covered the Global market by gaining the attention of international companies. Because of more facilities and offers, user's interest is increased for online shopping [15]. The fraud detection is considered as a necessary process in the digital and this article focusses on the classification of fraud activity. An effective optimization based classification approach is developed for classification of normal and fraud transaction that is competitive swarm based deep convolutional neural network.

The credit card fraud classification is attained by sequence and static learners [16]. The process of classification is accomplished with the deep learning based generalized adversarial network [17]. Support vector machines, recurrent neural networks, and auto emitter based Restricted Boltzmann Machines (AERBM) are deep learning- and machine learning techniques, respectively, which are utilized to classify credit card malfunction [18–21].

Using a hybrid data resampling approach and an artificial neural ensemble classifier, an unique method for identifying credit card malfunction [31]. The ensemble classifier in the adaptive boosting (AdaBoost) technique is built utilizing a long short-term memory (LSTM) network as the basis trainer. In the meanwhile, hybrid resampling is carried out using the edited nearest neighbour (SMOTE-ENN) method and the artificial minority oversampling methodology. The utility of the suggested strategy is portrayed by publicly accessible databases of actual card transactions. Machine learning (ML)-based methods databases containing real card transaction history have been described in recent study, however their detection scores still need to be improved because a measure of the imbalanced data in a particular dataset. A few methods have generated remarkable outcome on diverse datasets [32, 33].

The classification approaches used in the previously deployed system suffers from the problem of being unable to comprehend lengthy statements. The prevalence of errors could lead to incorrect categorization. The existing methods are ineffective due to the high computational expense and poor precision. There is a difficulty with both the quantity of data and the quality of the instruction. By taking these constraints into consideration, it is possible to develop an effective hybrid model for deep learning that is built on competitive swarm optimization.

2. Existing approach

One of the most cutting-edge artificial intelligence methods for tackling computer vision problems is the Deep Neural Network (Deep CNN). As a feed-forward ann model, the Deep CNN formed a class of deep learning and was used in various agricultural picture categorization efforts. The convolutional layer, which is efficient to Deep CNN, utilize filters to extract the information from the input images. A large number of training examples is need to enhance Deep CNN's performance. One of the key benefits of using Deep CNN for photo categorization is that it reduces the requirement for feature engineering. Deep CNN has numerous layers, each of which has several convolutions. They produce a range of interpretations of the training examples in the earlier, more thorough layers, moving up to more complex ones in the different levels. The pooling layers, which initially operate as feature extraction from the convolutional layers' effectiveness as feature representation, are then used to lower the complexity of the training data.

In order to provide more discriminative features, the convolutional layers extract multiple lower level characteristics. The convolution layer are also fundamental parts of Deep CNN. Feature engineering, a specific behaviour aspect of deep learning, represents a significant advancement over standard machine learning. The downsampling procedure is carried out along the spatial by the pooling layer. It promotes having small parameter choices. The maxpooling approach was used in the convolution layers of the proposed models. Max pooling achieves better than

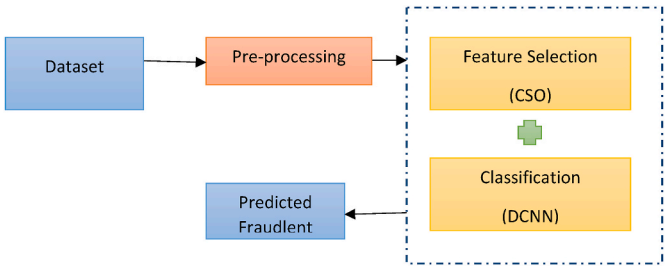

Fig. 1. Block Diagram for proposed method.

Table 1
Dataset description credit card fraud dataset.

Description	Credit Card Fraud
Number of Instances	58016
Number of Attributes	40
Number Class	2
Number of Positive Samples	57879
Number of Negative Samples	137

Table 2
Dataset description mortgage fraud dataset.

Description	Mortgage Fraud
Number of Instances	48674
Number of Attributes	11
Number Class	2
Number of Positive Samples	48066
Number of Negative Samples	608

average pooling in the proposed Deep Classification algorithm. Dropout is a crucial layer that explains removing objects from the network. It is an overfitting reduction regularisation approach. With dropout values ranging from 0.2 to 0.8, the suggested model was trained and compared. Utilizing the results from the convolutional and pooling layers, the dense layer then conducts the classification.

Deep CNN is a very iterative process, and to get the optimal model, many models must be trained. Gradient descent is a fundamental optimization approach that executes the gradient steps while using all training examples for each stage. It is also known as batch gradient descent. Gradient descent is challenging to accomplish with a large training set [15].

3. Proposed methodology

The proposed method is unique hybrid combination of deep learning method CSO – DCNN (in which feature selection CSO and classification is DCNN) for detecting financial fraudulent detection. This hybrid feature selection and classification provides effective output when comparing to the other methods.

Table 3
Dataset description insurance fraud dataset.

Description	Insurance Fraud
Number of Instances	44326
Number of Attributes	39
Number Class	2
Number of Positive Samples	43814
Number of Negative Samples	512

3.1. Feature selection using competitive swarm optimization

The proficient competitive swarm optimization (CSO) solves the massive optimization model, and the process of learning is derived from the selected competitors. The particle in the population is dynamically separated into two parts in each iteration, and bilateral competitiveness is achieved amongst particle in each group [21,22]. The initiation of particle of G in an N-dimensional space. The fitness function evaluates the characteristics by verifying the fitness value, and the greatest potential responses are obtained. For the computation of the fitness function, the terms are chosen at random, and a feasible solution is generated for the terms that were chosen at random. The MAD value is estimated in equqtion (1) as follows,

$$MAD\left(ve_{xyg}\right) = \sum_{p=1}^{fe}\sum_{q=1}^{sg}\frac{\left|v_{xyg} - \overline{v_{xyg}}\right|}{g} \quad (1)$$

Where fe denotes the number of the feature, g denotes the reviews retrieved from the user tweet, ve_{xyg} denotes the result of the swarm intelligence path vector, and sg denotes the number of the segmented group containing x features. The primary goal of employing MAD as a fitness function is to determine the score of each individual part as well as the distance between them. In the particular region of search, the solution value chooses the fitness value with the highest MAD.

The concepts that satisfy the fitness function are referred to as winners, while the alternative terms are referred to as losers. The winners of the competition are instantly moved on to the following iteration, while losers will gain to what the champions have discovered. Swarms that are given a position at random and are designated losers will adjust their position by changing their velocity. By figuring out where the winners are, the losers may determine their position in relation to the winner. Meanwhile, the position and velocity will be described in equations (2) and (3) as follow:

$$s_{1o}^{t+1} = RG_1^t S_1^t + RG_2^t\left(p_{wi}^t - p_{lo}^t\right) + \varnothing RG_3^t\left(p^{-t} - p_{lo}^t\right) \quad (2)$$

$$p_{lo}^{t+1} = p_{lo}^t + s_{lo}^{t+1} \quad (3)$$

The winning swarm is denoted by p_wiˆt, the loser swarm by p_loˆt, the mean position by∅ for pˆ(-t), and the influences are regulated by for p. The iteration is denoted by t, the random vectors are denoted by RG_1ˆt,RG_2ˆt, ⟦RG_3ˆt, and these values lie in the range of [0, 1]n [27].

The velocity and position will be updated as a consequence of the values in the loser, resulting in a new position. Equation (1) is used to evaluate the function value for the freshly created swarms, and the swarm with the highest fitness is transmitted to the winner. Because of the new selection swarms, the CSO will experience dispersion, which would be a significant disadvantage of the CSO and occurs instantly after a certain generation. The below Fig. 1 shows the block diagram of proposed method followed that algorithm of CSO is also provided. The

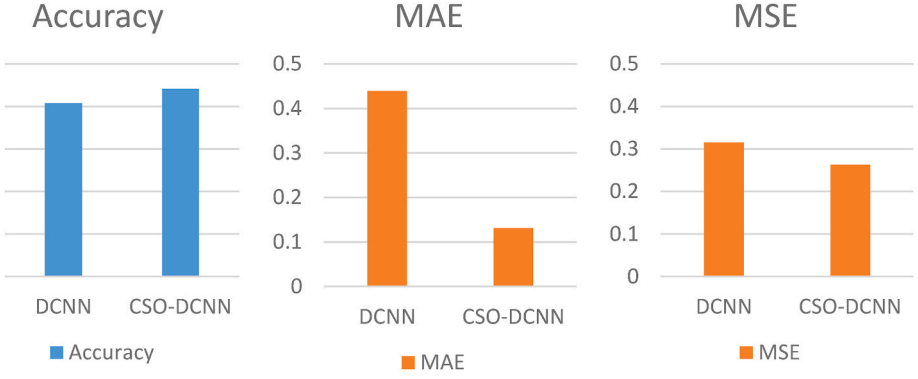

Fig. 2. Comparison of performance of credit card fraud.

advantage of the proposed method is it offers a high level of accuracy, and it allows one to retrieve more time and determine the precision, to decrease fraudulent use of credit cards as well as to forecast potential fraudulent activity, even in massive data sets, to alleviate some of the budgetary concerns through more efficient categorization and to lower the risk of deception when making internet payments.

Algorithm 1. Proposed CSO for feature selection

Algorithm 1. Proposed CSO for feature selection

Initialization of swarms with N dimension

t=0

randomly initiate the swarm

while term_con not satisfied do

estimation of mean absolute distance //fitness function calculation

$$MAD(ve_{xyg}) = \sum_{p=1}^{fe}\sum_{q=1}^{sg}\frac{\left|v_{xyg} - \overline{v_{xyg}}\right|}{g}$$

while MAD≠ ∅ then

randomly elects the swarm from N dimension of swarm

if the fitness value is satisfied

the swarms are winners

else

the swarms are losers

end if

The position of the swarm is updated by

$$GF_{xy}^{d}(\mathrm{t}) = \sum_{y\in K_{BEST}, y\neq x} p_{wij}^{t}.GF_{xy}^{d}(\mathrm{t})$$

Update position and velocity

$$s_{1o}^{t+1} = RG_1^t S_1^t + RG_2^t(p_{wi}^t - p_{lo}^t) + \emptyset RG_3^t(p^{-t} - p_{lo}^t)$$

$$p_{lo}^{t+1} = p_{lo}^t + s_{lo}^{t+1}$$

end while

best features are retrieved;

t=t+1;

end while

3.2. Classification using deep convolutional neural network

Deep Convolutional Neural Network (DCNN) using Rectified Linear Unit (ReLu) as a learning rate is used to help classify the best features [24,25]. The feature extraction and classification phases make up DCNN. Convolution and pooling layer are included in the feature learning step. The fully connected and softmax layer are present during the classification step. The learning of picture features is facilitated by the Deep CNN, and categorization is straightforward.

Convolution Layer: The aggregate of the output from this layer's multiple filters as they pass over the data input is done using the component by component multiplication approach. The result of this layer is then calculated as the input's responsive rate [26]. The succeeding layer uses the normalized cumulative value as an input component. The main topic is dragged to complete the data of the other patterns in the convolutional layer's output. In the convolution layer, zero padding, stride, and filter size are provided for each operation. The Rectified Linear Unit (ReLU) functions as an activation function, speeding rate of convergence of the probabilistic ascent gradients. The construction of ReLU is simple, and it takes advantage of thresholding, in which the activation function's outcome is transferred to zero. If it obtains a negative number, it delivers zero, and if it accepts a positive benefit, it delivers to t. The ReLU (AFn) is given in equation (4) as,

$$AFn = \max(0, t) \tag{4}$$

The gradient processterminates learning when the AFn value attains zero and the leaky ReLU is started in that instance. Its function is specifiedin equation (5) as follows,

Table 4
Comparison of performance for credit card fraud.

Method	Accuracy (%)	MAE	MSE
Existing-DCNN	91.42	0.439	0.315
Proposed-CSO-DCNN	98.20	0.131	0.263

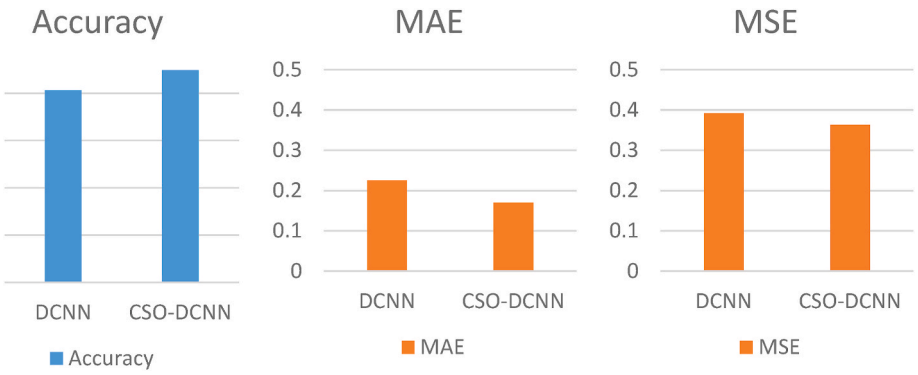

Fig. 3. Comparison of performance for mortgage fraud.

Table 5
Comparison of performance for mortgage fraud.

Method	Accuracy (%)	MAE	MSE
Existing-DCNN	91.32	0.225	0.392
Proposed-CSO-DCNN	99.77	0.170	0.363

Table 6
Comparison of performance for insurance fraud.

Method	Accuracy	MAE	MSE
Existing-DCNN	90.17	0.417	0.327
Proposed-CSO-DCNN	95.23	0.313	0.219

$$AF_l = \begin{cases} t & t > 0 \\ o \times t & t \le 0 \end{cases}, o \tag{5}$$

where o is designated as the preset parameter, and the value 0.01 is given.

Pooling Layer: The output dimensionality is decreased by the pooling layer, and most well-known max pooling technique is used to display the maximum pooling filter value. Max pooling is a workable method that significantly minimises the size of the input sample [23]. The summation and merging procedures are accomplished via the maximum pooling methodology.

Fully Connected Layer: The convolutional layer result reveals that this layer integrates non-linear data from high-range properties. This layer learns the non-linear functionality in that region.

Softmax Layer: This stage involves categorization, and the output units—a normalized exponential quantity of output data—use the softmax function. This denotes that the output frequency and functionality are distinct. In addition, the progressive pixel value improves the likelihood to the highest degree possible. The softmax is equated in (6) as,

$$SOp_x \frac{e^{z_x}}{\sum_{x=1}^{M} e^{z_x}} \tag{6}$$

where zx is the output count before the softmax, sopx is the output of a softmax for output count x, and M is the total output node count In this layer, the classifiers are classified.

4. Result and discussion

The effectiveness of the suggested technique is examined in this part using three distinct datasets [28,29and30]. Tables 1–3 provide a description of the dataset.

4.1. Dataset description

4.2. Metric unit

4.2.1. Accuracy

The degree to which a certain value closely resembles situations that

Table 7
Comparison of Performance for different approaches.

Credit Card Fraud Dataset		
References	Methods	Accuracy (%)
Proposed	CSO + DCNN	98.20
Esraa Faisal Malik et al. (2022) [34]	Adaboost + LGBM	82.00
Ajeet Singh&Anurag Jain (2021) [35]	CFS + RF + firefly	96.23
Maria Nancy et al. (2020) [36]	CNN-KNN	98

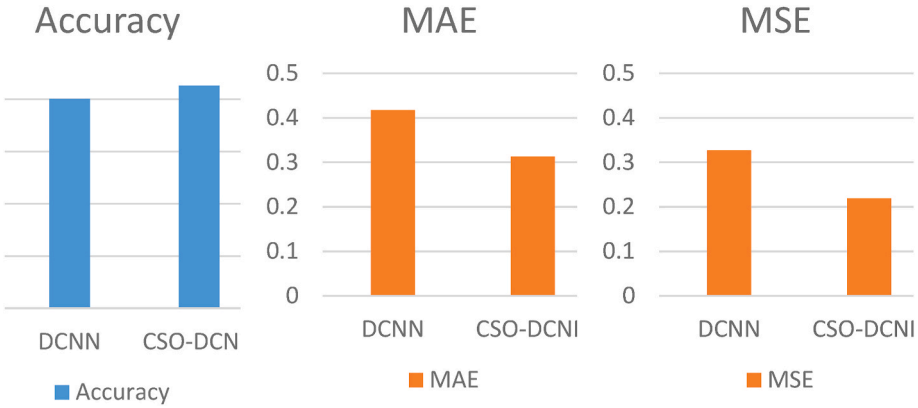

Fig. 4. Comparison of performance for insurance fraud.

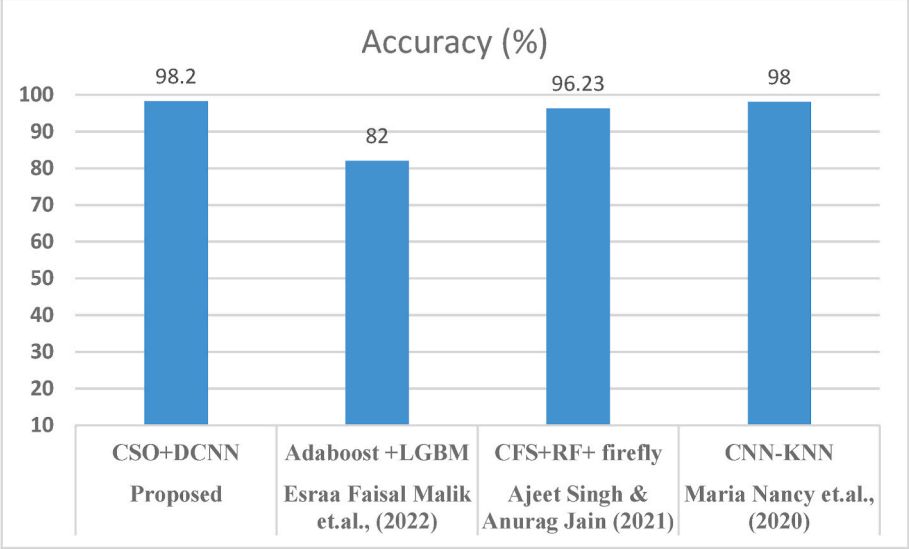

Fig. 5. Comparison of Performance forCredit Card Fraud Dataset.

Table 8
Comparison of Performance for different approaches.

Mortage Fraud Dataset		
References	Methods	Accuracy (%)
Proposed	CSO + DCNN	99.77
Theja.G.S et al. (2020) [37]	CF + XGBoost	98
Yao Zhi Xu et al. (2019) [38]	DT + LR	72
Bing Chu et al., 2018 [39]	RF + CNN	92

have been classified is known as accuracy. The characterization of methodological problems and quantitative biased is accuracy. Furthermore, it is the agreement (TP and TN values added together) among the number of evaluated classes along with the estimation's suitability for the actual value. The presence of least precision is which causes the disparity between the resultant and real resultant values. It is the proportion of correctly identified terms over all instances that have been considered. It's calculated as (7),

$$Accuracy = \frac{TP + True\ Negative\ (TN)}{TP + TN + FP + FN} \tag{7}$$

4.2.2. MAE

Absolute mean error (MAE) The MAE calculates the average size of prediction errors without taking into account their direction. It measures accuracy for continuous variables.

4.2.3. MSE

The mean squared error (MSE) of a regression line shows how close it is to a collection of points. By doubling the lengths between the endpoints and the regression line, it is able to do this (these lengths are the "errors"). Any unfavourable signals must be eliminated using squaring. Significant differences are also given greater weight. It is known as the mean squared error since you are averaging a number of errors. The MSE decreases as the forecast gets better.

5. Result and discussion

Fig. 2 and Table 4 illustrates the comparison of accuracy, MAE, and MSE for the credit card fraud detection.

Fig. 2 shows a thorough analysis of the CSO-DCNN model's accuracy, mean absolute error, and mean squared error in comparison to the

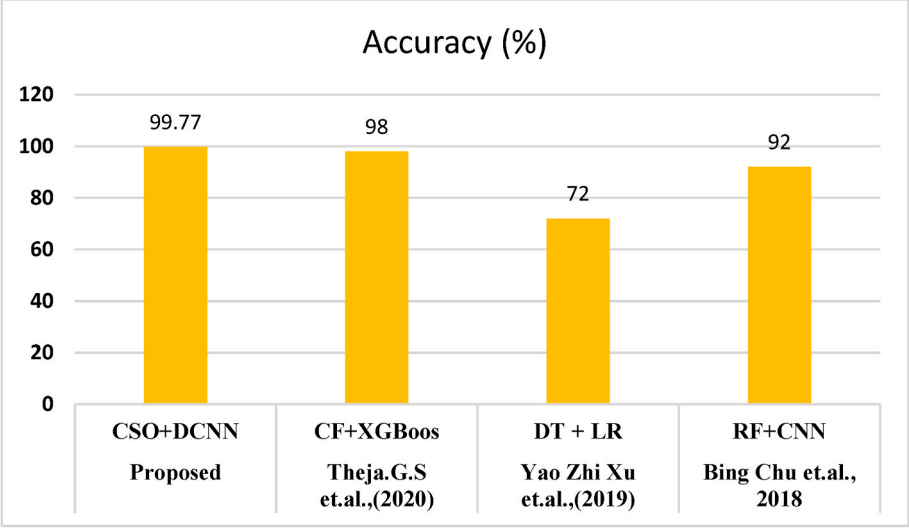

Fig. 6. Comparison of performance for mortage fraud dataset.

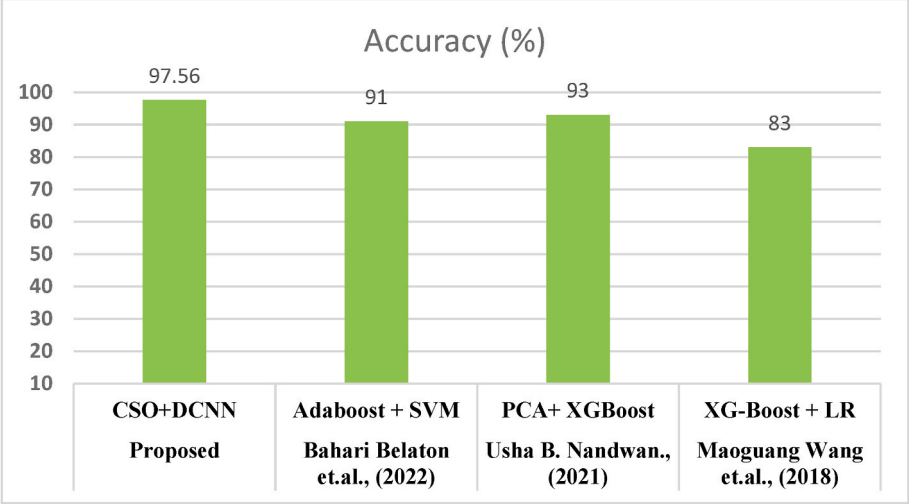

Fig. 7. Comparison of performance for insurance fraud dataset.

Table 9
Comparison of Performance for different approaches.

Insurance Fraud Dataset		
References	Methods	Accuracy (%)
Proposed	CSO + DCNN	97.56
Bahari Belaton et al. (2022) [40]	Adaboost + SVM	91
Usha B. Nandwan., (2021) [41]	PCA + XGBoost	93
Maoguang Wang et al. (2018) [42]	XG-Boost + LR	83

Table 10
Comparison of Performance for different approaches.

Credit Card Fraud Dataset		
References	Methods	MAE
Proposed	CSO + DCNN	0.131
Hasoon et al. (2022) [43]	DNN + LSTM	0.137
Meran et al. (2020) [44]	LSTM + NN	0.181
HAsaniuo et al. (2016) [45]	PSO + KNN	0.224

DCNN model that is currently in use on the Credit card Dataset. The resulting figures showed that the CSO-DCNN model had enhanced accuracy of 98.20%, a lowered MAE of 0.131, and a lower MSE of 0.263, compared to the DCNN model's slightly decreased accuracy of 91.42%, increased MAE of 0.439, and lower MSE of 0.315.

Fig. 3 and Table 5 illustrates the comparison of accuracy, MAE, and MSE for the mortgage fraud.

Fig. 3 shows an in-depth comparison of the accuracy, mean absolute error, and mean squared error of the CSO-DCNN model with the current DCNN model on the Mortgage Fraud Dataset. The results showed that the CSO-DCNN model had improved accuracy of 99.77%, decreased MAE of 0.170, and MSE of 0.363, whereas the DCNN model had slightly decreased accuracy of 91.32%, increased MAE of 0.225, and MSE of 0.392.

Fig. 4 and Table 6 illustrates the comparison of accuracy, MAE, and MSE for the insurance fraud.

Fig. 4 shows a thorough analysis of the accuracy, mean absolute error, and mean squared error of the CSO-DCNN model in comparison to the DCNN model that is currently in use for the Insurance Fraud Dataset. The results showed that the CSO-DCNN model had better accuracy of 95.23%, a lowered MAE of 0.313, and a lower MSE of 0.219, compared to the DCNN model's slightly decreased accuracy of 90.17%, increased MAE of 0.417, and higher MSE of 0.327.

Fig. 5 and Table 7 illustrates the comparison of accuracy, for the Credit Card Fraud Dataset.

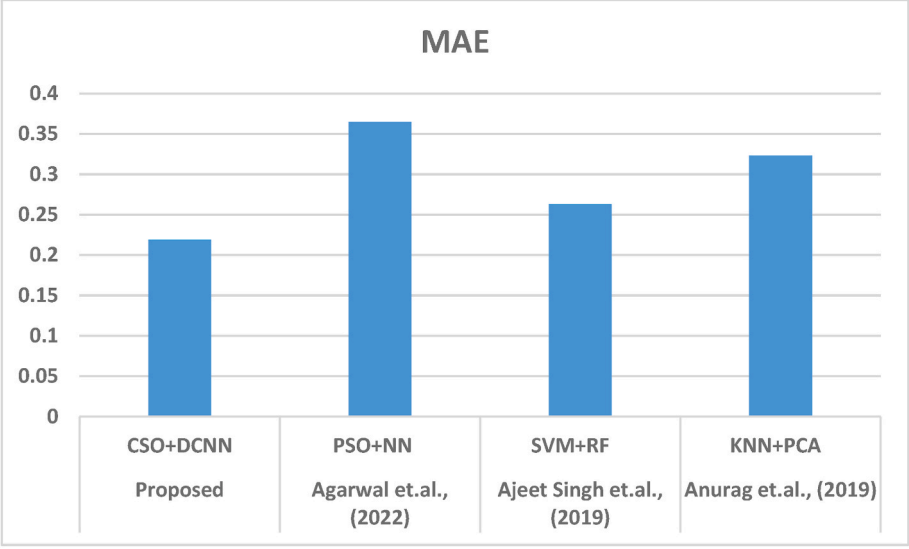

Fig. 8. Comparison of performance for credit card fraud dataset.

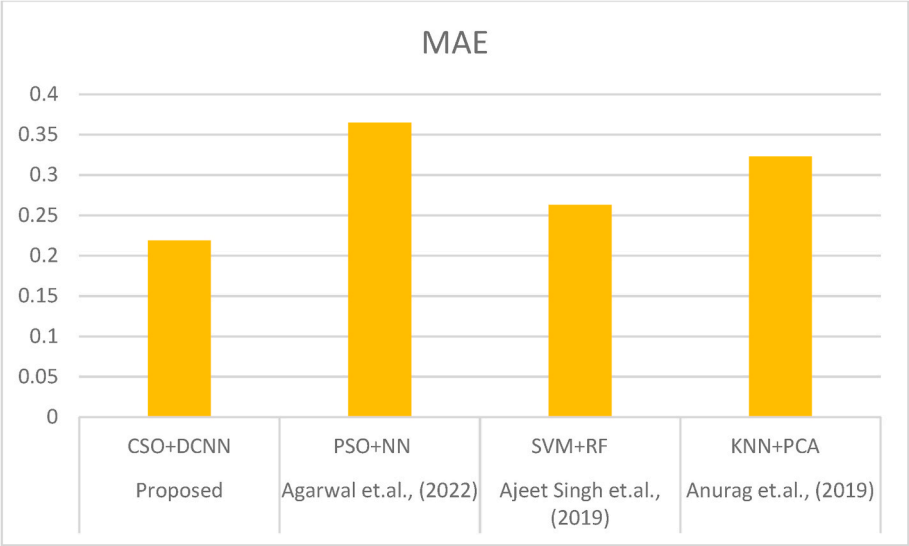

Fig. 9. Comparison of performance for mortage fraud dataset.

Table 11
Comparison of Performance for different approaches.

Mortage Fraud Dataset		
References	Methods	MAE
Proposed	CSO + DCNN	0.170
Oikonomidis et al. (2022) [46]	CNN + XGBoost	0.263
Yan Li et al. (2019) [47]	LSTM + CNN	0.181
Lin T et al. (2017) [48]	NN + LSTM	0.192

Fig. 5 shows a thorough comparison of the CSO-DCNN model's accuracy with the pre-existing models on Credit Card Fraud Dataset. The calculated figures showed that the accuracy of the CSO-DCNN model was improved to 98.20%, whereas the accuracy of the CNN-KNN model, Adaboost + LGBM, CFS + RF + firefly, and CFS + RF + firefly was slightly decreased to 82.00%, 96.23%, and 98%, respectively.

Fig. 6 and Table 8 illustrates the comparison of accuracy, for the Mortage Fraud Dataset.

In Fig. 6, a thorough evaluation of the CSO-DCNN model's accuracy in comparison to other models on Mortage Fraud Dataset is shown. The resulting figures showed that the accuracy of the CSO-DCNN model has increased to 99.77%, while the accuracy of the CF + XGBoost, DT + LR, and RF + CNN models has somewhat decreased to 98%, 72%, and 92%, respectively.

Fig. 7 and Table 9 illustrates the comparison of accuracy for the Insurance Fraud Dataset.

Table 12
Comparison of Performance for different approaches.

Insurance Fraud Dataset		
References	Methods	MAE
Proposed	CSO + DCNN	0.313
Agarwal et al. (2022) [49]	PSO + NN	0.335
Ajeet Singh et al. (2019) [50]	SVM + RF	0.412
Anurag et al. (2019) [51]	KNN + PCA	0.498

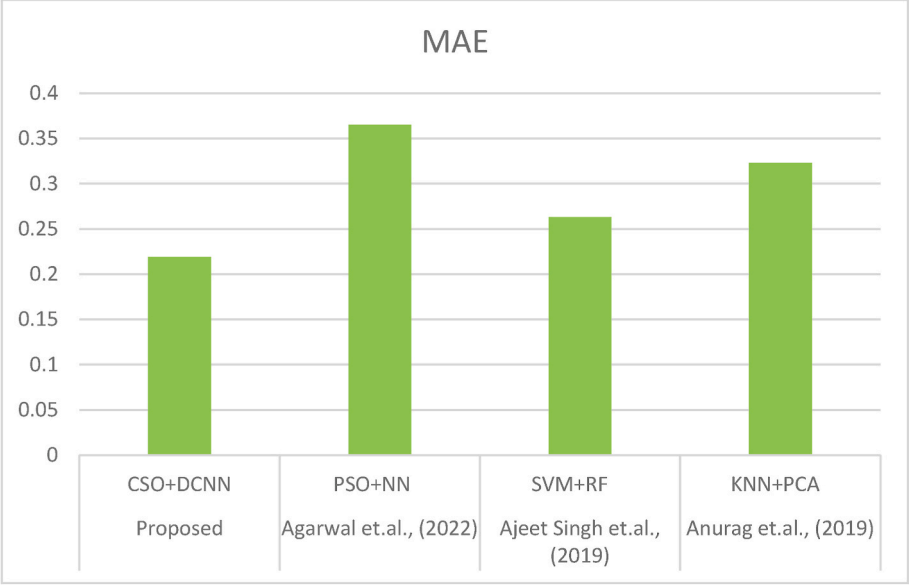

Fig. 10. Comparison of performance for insurance fraud dataset.

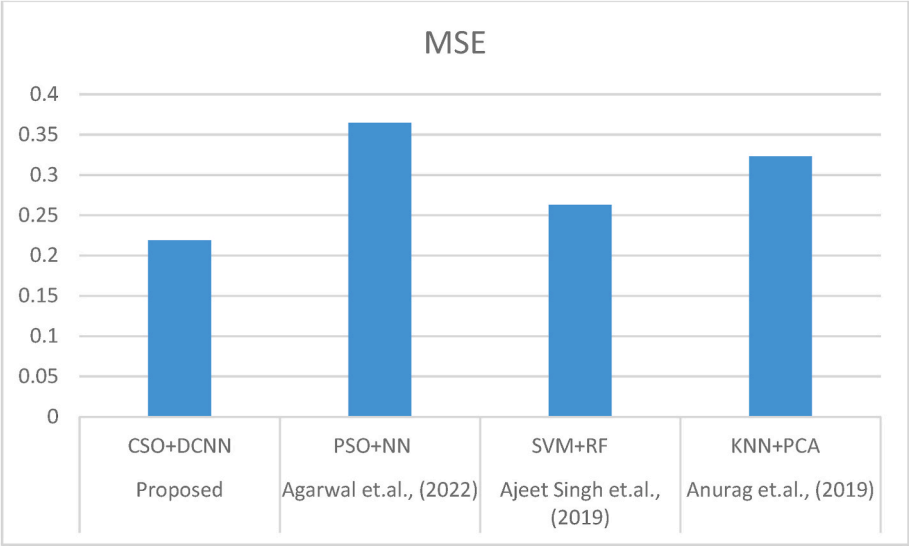

Fig. 11. Comparison of performance for credit card fraud dataset.

Table 13
Comparison of Performance for different approaches.

Credit Card Fraud Dataset		
References	Methods	MSE
Proposed	CSO + DCNN	0.263
Hasoon et al. (2022) [43]	DNN + LSTM	0.413
Meran et al. (2020) [44]	LSTM + NN	0.351
HAsaniuo et al. (2016) [45]	PSO + KNN	0.780

Table 14
Comparison of Performance for different approaches.

Mortage Fraud Dataset		
References	Methods	MSE
Proposed	CSO + DCNN	0.363
Oikonomidis et al. (2022) [46]	CNN + XGBoost	0.374
Yan Li et al. (2019) [47]	LSTM + CNN	0.412
Lin T et al. (2017) [48]	NN + LSTM	0.382

Fig. 7 shows a thorough comparison of the CSO-DCNN model's accuracy with the other models on Insurance Fraud Dataset. The final data showed that the accuracy of the CSO-DCNN model was increased to 97.56%, whereas the accuracy of the Adaboost + SVM, PCA + XGBoost, and XG-Boost + LR models was somewhat decreased to 91%, 93%, and 83%, respectively.

Fig. 8 and Table 10 illustrates the comparison of MAE for the Credit Card Fraud Dataset.

Fig. 8 compares the Mean Absolute Error of the CSO-DCNN model to the previous models on Credit Card Fraud Dataset in detail. The resulting results showed that the DNN + LSTM, LSTM + NN, and PSO + KNN models all had somewhat higher Mean Absolute Error values, ranging from 0.137 to 0.224, while the CSO-DCNN model had a lowered Mean Absolute Error of 0.131.

Fig. 9 and Table 11 illustrates the comparison of MAE for the Mortage Fraud Dataset.

A detailed investigation on Mean Absolute Error of the CSO - DCNN model with the existing models on the Mortage Fraud Dataset is displayed in Fig. 9. The resultant values highlighted that the CSO-DCNN model has attained decreased Mean Absolute Error of 0.170 while the

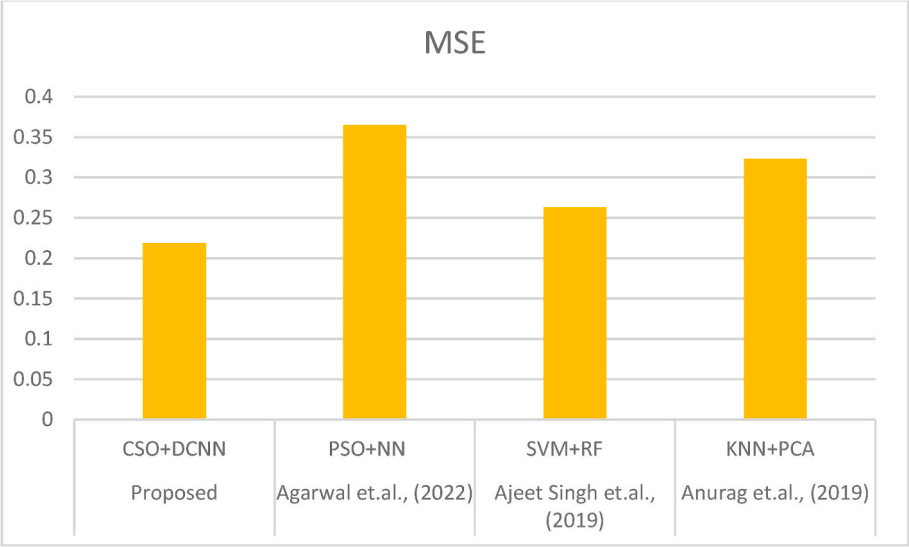

Fig. 12. Comparison of performance for mortage fraud dataset.

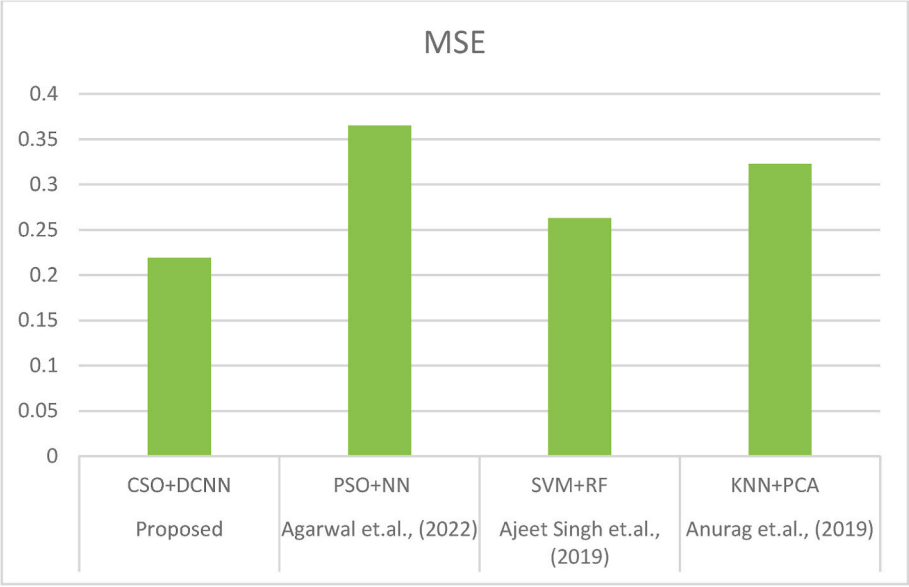

Fig. 13. Comparison of performance for insurance fraud dataset.

Table 15
Comparison of Performance for different approaches.

Insurance Fraud Dataset		
References	Methods	MSE
Proposed	CSO + DCNN	0.219
Agarwal et al. (2022) [49]	PSO + NN	0.365
Ajeet Singh et al. (2019) [50]	SVM + RF	0.263
Anurag et al. (2019) [51]	KNN + PCA	0.323

CNN + XGBoost, LSTM + CNN, and NN + LSTM model has gained slightly increased Mean Absolute Error value of 0.263, 0.181 and 0.192 respectively.

Fig. 10 and Table 12 illustrates the comparison of MAE for the Insurance Fraud Dataset.

Fig. 10 compares the Mean Absolute Error of the CSO-DCNN model to the previous models on Insurance Fraud Dataset in detail. The resulting results showed that the PSO + NN, SVM + RF, and KNN + PCA models all had somewhat higher Mean Absolute Error values than the CSO-DCNN model, which had a lowered Mean Absolute Error of 0.313.

Fig. 11 and Table 13 illustrates the comparison of MSE for the Credit Card Fraud Dataset.

Fig. 11 presents a thorough analysis of the Mean Squared Error of the CSO-DCNN model in comparison to the pre-existing models on Credit Card Fraud Dataset. The calculated results showed that the DNN + LSTM, LSTM + NN, and PSO + KNN models all had somewhat higher Mean Squared Error values than the CSO-DCNN model (0.263 vs. 0.413, 0.351, and 0.780, respectively).

Fig. 12 and Table 14 illustrates the comparison of MSE for the Mortage Fraud Dataset.

Fig. 12 presents a thorough analysis of the Mean Squared Error of the CSO-DCNN model in comparison to the pre-existing models on Mortage Fraud Dataset. The resulting results showed that the CNN + XGBoost, LSTM + CNN, and NN + LSTM models all had somewhat higher Mean Squared Error values than the CSO-DCNN model (0.363 vs. 0.374, 0.412, and 0.382, respectively).

Fig. 13 and Table 15 illustrates the comparison of MSE for the Insurance Fraud Dataset.

A detailed investigation on Mean Squared Error of the CSO - DCNN model with the existing models on the Insurance Fraud Dataset is displayed in Fig. 13. The resultant values highlighted that the CSO-DCNN model has attained decreased Mean Squared Error of 0.219 while the PSO + NN, SVM + RF, and KNN + PCA model has gained slightly increased Mean Squared Error value of 0.365, 0.263 and 0.323 respectively. In this data set proposed model CSO-DCNN has negative rate of Mean Squared Error.

6. Conclusion

Online payments are significant in today's global computing world since they employ simply the user credentials from the credit card to complete an application and then charge money. As a result, it's critical to develop the best approach for detecting the greatest number of frauds in online systems. The performance of the proposed and existing approaches is investigated using accuracy, MAE, and MSE. The performance outcome indicates the efficiency of the proposed technique and it outperforms the existing approaches. The proposed approach achieves highest accuracy for three data sets over existing techniques namely DCNN. Attainment of highest accuracy shows that the suggested method is successful in identification of cyber-attacks. The error rate is minimal in the proposed approach over existing techniques. In the work that will be done in the future, the primary focus will be on including additional datasets that contain numerically sensitive characteristics and implementing ensemble prepossessing techniques to address the problem of uneven data.

CRediT authorship contribution statement

Karthikeyan T: Conceptualization, Methodology. **Govindarajan M:** Validation. **Vijayakumar V:** Comments evaluated, Training and testing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

No data was used for the research described in the article.

References

[1] A. Abdallah, M.A. Maarof, A. Zainal, Fraud detection system: a survey, J. Netw. Comput. Appl. 68 (2016) 90–113.

[2] J. West, M. Bhattacharya, Intelligent financial fraud detection: a comprehensive review, Comput. Secur. 57 (2016) 47–66.

[3] V. Van Vlasselaer, T. Eliassi-Rad, L. Akoglu, M. Snoeck, B. Baesens, Gotcha! Network-based fraud detection for social security fraud, Manag. Sci. 63 (9) (2017) 3090–3110.

[4] J.O. Awoyemi, A.O. Adetunmbi, S.A. Oluwadare, Credit card fraud detection using machine learning techniques: a comparative analysis, in: 2017 International Conference *on Computing Networking and Informatics (ICCNI)*, IEEE, 2017, October, pp. 1–9.

[5] S.Y. Huang, C.C. Lin, A.A. Chiu, D.C. Yen, Fraud detection using fraud triangle risk factors, Inf. Syst. Front 19 (6) (2017) 1343–1356.

[6] A. Chouiekh, E.H.I.E. Haj, Convnets for fraud detection analysis, Proc. Comput. Sci. 127 (2018) 133–138.

[7] A. Dal Pozzolo, G. Boracchi, O. Caelen, C. Alippi, G. Bontempi, Credit card fraud detection: a realistic modeling and a novel learning strategy, IEEE Transact. Neural Networks Learn. Syst. 29 (8) (2017) 3784–3797.

[8] B. Baesens, V. Van Vlasselaer, W. Verbeke, Fraud Analytics Using Descriptive, Predictive, and Social Network Techniques: a Guide to Data Science for Fraud Detection, John Wiley & Sons, 2015.

[9] A.C. Bahnsen, D. Aouada, A. Stojanovic, B. Ottersten, Feature engineering strategies for credit card fraud detection, Expert Syst. Appl. 51 (2016) 134–142.

[10] T.F. Kummer, K. Singh, P. Best, The effectiveness of fraud detection instruments in not-for-profit organizations, Manag. Audit J. 30 (4) (2015) 435–455.

[11] S.K. Majhi, Fuzzy clustering algorithm based on modified whale optimization algorithm for automobile insurance fraud detection, Evolutionary intelligence 14 (1) (2021) 35–46.

[12] D. Cheng, S. Xiang, C. Shang, Y. Zhang, F. Yang, L. Zhang, Spatio-temporal attention-based neural network for credit card fraud detection, in: Proceedings of the AAAI Conf*erence on Artificial Intelligence* vol. 34, 2020, April, pp. 362–369, 1.

[13] A. Zakaryazad, E. Duman, A profit-driven Artificial Neural Network (ANN) with applications to fraud detection and direct marketing, Neurocomputing 175 (2016) 121–131.

[14] A.S. Bekirev, V.V. Klimov, M.V. Kuzin, B.A. Shchukin, Payment card fraud detection using neural network committee and clustering, Opt. Mem. Neural Network. 24 (3) (2015) 193–200.

[15] S. Georgieva, M. Markova, V. Pavlov, Using neural network for credit card fraud detection, in: AIP Conference Proceedings vol. 2159, AIP Publishing LLC, 2019, October, 030013, 1.

[16] J. Jurgovsky, M. Granitzer, K. Ziegler, S. Calabretto, P.E. Portier, L. He-Guelton, O. Caelen, Sequence classification for credit-card fraud detection, Expert Syst. Appl. 100 (2018) 234–245.

[17] U. Fiore, A. De Santis, F. Perla, P. Zanetti, F. Palmieri, Using generative adversarial networks for improving classification effectiveness in credit card fraud detection, Inf. Sci. 479 (2019) 448–455.

[18] A. Pumsirirat, L. Yan, Credit card fraud detection using deep learning based on auto-encoder and restricted Boltzmann machine, Int. J. Adv. Comput. Sci. Appl. 9 (1) (2018) 18–25.

[19] S. Wang, C. Liu, X. Gao, H. Qu, W. Xu, Session-based fraud detection in online e-commerce transactions using recurrent neural networks, in: Joint European Conference on M*achine Learning and Knowledge Discovery in Databases*, Springer, Cham, 2017, September, pp. 241–252.

[20] V.B. Nipane, P.S. Kalinge, D. Vidhate, K. War, B.P. Deshpande, Fraudulent detection in credit card system using SVM & decision tree, International Journal of Scientific Development and Research (IDSDR) 1 (5) (2016) 590–594.

[21] R. Cheng, Y. Jin, A competitive swarm optimizer for large scale optimization, IEEE Trans. Cybern. 45 (2) (2014) 191–204.

[22] Y. Li, Z.H. Zhan, S. Lin, J. Zhang, X. Luo, Competitive and cooperative particle swarm optimization with information sharing mechanism for global optimization problems, Inf. Sci. 293 (2015) 370–382.

[23] J. Yosinski, J. Clune, Y. Bengio, H. Lipson, How Transferable Are Features in Deep Neural Networks?, 2014 *arXiv preprint arXiv:1411.1792*.

[24] R. Miikkulainen, J. Liang, E. Meyerson, A. Rawal, D. Fink, O. Francon, B. Hodjat, Evolving deep neural networks, in: Artificial Intelligence in the Age of Neural Networks and Brain Computing, Academic Press, 2019, pp. 293–312.

[25] J. He, L. Li, J. Xu, C. Zheng, Relu Deep Neural Networks and Linear Finite Elements, 2018 *arXiv preprint arXiv:1807.03973*.

[26] Credit card fraud detection, Accessed: Nov. 9, 2020. [Online]. Available: https://www.kaggle.com/mlg-ulb/creditcardfraud.

[27] IEEE-CIS fraud detection, Accessed: Nov. 9, 2020. [Online]. Available: https://www.kaggle.com/c/ieee-fraud-detection/data.

[28] Credit Card Fraud Detection. Accessed: Nov. 9, 2020. [Online]. Available:https://www.kaggle.com/datasets/mlg-ulb/creditcardfraud.

[29] Credit Card Fraud Detection. Accessed: Nov. 9, 2020. [Online]. Available:https://www.kaggle.com/roshansharma/insurance-claim.

[30] Credit Card Fraud Detection. Accessed: Nov. 9, 2020. [Online]. Available:https://www.kaggle.com/code/arjunjoshua/predicting-fraud-in-financial-payment-services/data.

[31] E. Esenogho, I.D. Mienye, T.G. Swart, K. Aruleba, G. Obaido, A neural network ensemble with feature engineering for Improved Credit Card Fraud Detection, IEEE Access 10 (2022) 16400–16407.

[32] A. Alharbi, M. Alshammari, O.D. Okon, A. Alabrah, H.T. Rauf, H. Alyami, T. Meraj, A novel text2IMG mechanism of credit card fraud detection: a deep learning approach, Electronics 11 (5) (2022) 756.

[33] K.G. Dastidar, M. Granitzer, W. Siblini, The importance of future information in credit card fraud detection, in: International Conference *on Artificial Intelligence and Statistics*, PMLR, 2022, May, pp. 10067–10077.

[34] Ajeet Singh &Anurag Jain, Hybrid bio-inspired model for fraud detection with correlation based feature selectiom, J. Discrete Math. Sci. Cryptogr. 24 (2021, Sep) 1365–1374.

[35] A. Maria Nancya, G. Senthil Kumar, S. Veena, N.A. Vinoth, Fraud detection in credit card transaction using hybrid model, AIP Conf. Proc. 2277 (2020), 130010.

[36] Esraa Faisal Malik, Khai Wah Khaw, Bahari Belaton, Wai Peng Wong, XinYing Chew, Credit card fraud detection using a new hybrid machine learning architecture, Mathematics (2022) 1480.

[37] Bing Zhu, Wenchuan Yang, Huaxuan Wang, Yuan Yuan, A hybrid deep learning model for consumer credit scoring, in: International Conference on Artificial Intelligence and Big Data, 2018.

[38] Yao-Zhi Xu. Jian-Lin Zhang, Ying Hua, Lin-Yue Wang, Dynamic credit risk evaluation method for E-commerce sellers based on a hybrid artificial intelligence model. Sustainability, 11(19), 5521.

[39] S. ThejasG, S.S. SuryaDheeshjith, N. Iyengar, Sunitha, Prajwal Badrinath, A hybrid and effective learning approach for Click Fraud detection, Machine Learning With Applications 3 (2020), 10016.

[40] Esraa Faisal Malik, Khai Wah Khaw, Bahari Belaton, Wai Peng Wong, XinYing Chew, Credit card fraud detection using a new hybrid machine learning architecture, Mathematics 10 (2022) 1480.

[41] Maoguang Wang, Jiayu Yu, Zijian Ji, Credit fraud risk detection based on XGBoost-LR hybrid model, in: Proceedings of the 18th International Conference on Electronic Business, 2018, pp. 336–343.

[42] Usha B. Nandwani, Archana Deshmukh, Deepti Gangurde, Anuj Satavas, Personal loan fraud detection based on hybrid supervised and unsupervised learning, International Journal for Research in Engineering Application & Management (IJREAM) 7 (2021) 1–6.

[43] T.S. Meron, Crime analysis and prediction using hybrid deep learning algorithms, Journal of Electrical Engineering Innovations 25 (2) (2021) 1131–1139.

[44] Safwan Hasoon, Muzahem Al-Hashimi, Hybrid deep neural network and long short term memory network for predicting of sunspot time series, Int. J. Math. Comput. Sci. 17 (3) (2022) 955–967.

[45] M. Hasanluo, F. Soleimanian Gharehchopogh, Software cost estimation by a new hybrid model of particle swarm optimization and K-nearest neighbor algorithms, J. Electr. Comput. Eng. Innovat. 4 (1) (2016) 49–56.

[46] LIi Yan, Wei Dai, Bitcoin price forecasting method based on CNN-LSTM hybrid neural network model, J. Eng. 2020 (13) (2019) 344–347.

[47] Tao Lin, Tian Guo, Karl Aberer, Hybrid neural networks for learning the trend in time series, in: Proceedings of the Twenty-Sixth International Joint Conference on Artificial Intelligence, 2020, pp. 2273–2280.

[48] Alexandros Oikonomidis, Cagatay Catal, Hybrid deep learning-based models for crop yield prediction, Appl. Artif. Intell. 36 (1) (2022) 456–467.

[49] George S Atsalakis and Kimon P Valavanis. Forecasting stock market short-term trends using a neuro-fuzzy based methodology, Expert Syst. Appl., 36 (7),10696–10707.

[50] V.E. Cabrera, D. Solís, G.A. Baigorria, D. Letson, Managing climate variability in agricultural analysis, in: John A. Long, S. David (Eds.), Ocean Circulation and El Niño: New Research, Wells 163–79, Nova Science Publishers, Hauppauge, NY, 2009.

[51] Y. LeCun, Y. Bengio, G. Hinton, Deep learining, Nature 571 (7553) (2015) 436–444.

Mr. T. Karthikeyan is working as an Assistant Professor in the Department of Computer Science and Engineering, Christ Institute of Technology, Puducherry. He has completed his M.E in Computer Science and Engineering in Sathyabama University, Chennai. He has got more than 10 years of teaching experience. He has published nearly five research articles both in national and international journals. His research is being focusing on Deep Learning. He has also participated in more than 10 seminars and workshops and trying to work in projects for welfare for the society.

Dr. M. Govindarajan is currently an Associate Professor in the Department of Computer Science and Engineering, Annamalai University, Tamil Nadu, India. He received the B.E, M.E and Ph.D Degree in Computer Science and Engineering from Annamalai University, Tamil Nadu, India in 2001, 2005 and 2010 respectively. He did his post-doctoral research in the Department of Computing, Faculty of Engineering and Physical Sciences, University of Surrey, Guildford, Surrey, United Kingdom in 2011 and at CSIR Centre for Mathematical Modelling and Computer Simulation, Bangalore in 2013. He has presented and published more than 140 papers at Conferences and Journals and also received best paper awards. He has delivered invited talks at various national and international conferences. His current research interests include Data Mining and its applications, Web Mining, Text Mining, and Sentiment Mining. He was the recipient of the Achievement Award for the field in the Conference in Bio-Engineering, Computer Science, Knowledge Mining (2006), Prague, Czech Republic. He received Career Award for Young Teachers (2006), All India Council for Technical Education, New Delhi and Young Scientist International Travel Award (2012), Department of Science and Technology, Government of India, New Delhi. He is a Young Scientists awardee under Fast Track Scheme (2013), Department of Science and Technology, Government of India, New Delhi and also granted Young Scientist Fellowship (2013),Tamil Nadu State Council for Science and Technology, Government of Tamil Nadu, Chennai. He also received the Senior Scientist International Travel Award (2016), Department of Science and Technology, Government of India, New Delhi. He has completed two major projects as principal investigator and has produced four Ph.Ds and also applied patent in the area of data mining. He has published four books and ten book chapters in national and international levels. He has visited countries like Czech Republic, Austria, Thailand, United Kingdom (twice), Malaysia, U.S.A (twice), and Singapore. He is an active Member of various professional bodies and Editorial Board Member of various conferences and journals.

Dr. V. Vijayakumar has been working as an Associate Professor in the Department of Computer Science and Engineering, Sri Manakula Vinayagar Engineering College, Puducherry. He Completed his doctorate from Pondicherry Central University. He has got more than 9 years of experience in teaching. He has published 11 research articles in International Journals. And his study focuses on load balancing and effective data dissemination in VANET. Additionally, he attended more than 25 advanced technology conferences and workshops.