AIMS Mathematics, 8(10): 23200–23217.
DOI: 10.3934/math.20231180
Received: 04 June 2023
Revised: 12 July 2023
Accepted: 13 July 2023
Published: 20 July 2023

http://www.aimspress.com/journal/Math

Research article

A novel deep learning-based hybrid Harris hawks with sine cosine approach for credit card fraud detection

Altyeb Taha*

Department of Information Technology, Faculty of Computing and Information Technology in Rabigh, King Abdulaziz University, Jeddah 21911, Saudi Arabia

* **Correspondence:** Email: aaataha@kau.edu.sa; Tel: +966552971512.

Abstract: Credit cards have become an integral part of the modern financial landscape, and their use is essential for individuals and businesses. This has resulted in a significant increase in their usage in recent years, especially with the growing popularity of online payments. Unfortunately, this increase in credit card use has also led to a corresponding rise in credit card fraud, posing a serious threat to financial security and privacy. Therefore, this research introduces a novel deep learning-based hybrid Harris hawks with sine cosine method for credit card fraud detection system (HASC-DLCCFD). The aim of the presented HASC-DLCCFD approach is to identify fraudulent credit card transactions. The suggested HASC-DLCCFD scheme introduces a HASC technique for feature selection, by combining Harris hawks optimization (HHO) with the sine cosine algorithm (SCA). For the purpose of identifying credit card fraud, an architecture of a convolutional neural network combined with long short-term memory (CNN–LSTM) is utilized in this study. Finally, the adaptive moment estimation (Adam) algorithm is utilized as a hyperparameter optimizer of the CNN-LSTM model. The performance of the suggested HASC-DLCCFD approach was experimentally evaluated using a publicly available database. The results demonstrate that the suggested HASC-DLCCFD approach outperforms other current techniques and achieved the highest accuracy of 99.5%.

Keywords: neural network; deep learning; hyperparameter optimization; features selection; fraud detection; credit card
Mathematics Subject Classification: 68M25

1. Introduction

The method by which we pay for products and services has been drastically transformed due to the advancement of online business transactions and electronic payment techniques. In recent decades, transactions made with credit cards have significantly increased, which has unfortunately attracted the attention of criminals [1,2]. Credit card fraud takes place when someone gains access to credit card information or uses fake cards to carry out unauthorized transactions, resulting in huge financial losses and unauthorized access to the financial information of legitimate cardholders [3,4]. The growth of e-commerce and the increase in online transactions are both factors contributing to the rising rate of credit card fraud. The financial industry has been significantly impacted by fraudulent credit card transactions. Based on the research conducted in [5], in 2018, the total amount lost due to credit card fraud was close to $27.85 billion, indicating an increase of 16.2% in comparison to the $23.97 billion loss recorded in 2017. These losses are expected to increase further, reaching 35 billion USD by 2023. Credit card fraud detection (CCFD) is essential for financial businesses to prevent losses [6].

Artificial intelligence (AI) applications in the finance sector can bring significant benefits to businesses, such as improved productivity, reduced operating costs and increased customer satisfaction. Many machine learning (ML) techniques have been developed to effectively detect fraudulent credit card activities. For instance, Malik et al. [7] investigated the application of hybrid models in identifying credit card fraud; these approaches were created by merging a number of ML algorithms, such as light gradient boosting machines (LGBM), extreme gradient boosting (XGBoost), random forest (RF) and adaptive boosting (AdaBoost). The findings of the experiments suggested that the hybrid method built using AdaBoost and LGBM had the greatest classification performance. In our previous study [8], we used an optimized light gradient boosting machine (OLightGBM) for detecting fraudulent credit card transactions; we compared our findings to the results achieved by other approaches. Our suggested method (OLightGBM) attained the best accuracy and outperformed the other approaches.

For several reasons, developing advanced CCFD models using machine learning is still challenging. First, fraud is concealed. Just a small percentage of transactions are fraudulent, resulting in a significant imbalance of class distribution [9,10]. As a result, there are limited instances of fraudulent behavior to train the machine learning models, which might lead to a large proportion of false positives (normal transactions mislabeled as fraudulent). The class that comprises a greater ratio of the dataset is known as the majority class, whereas the class with a lower ratio is known as the minority class. Because the majority of ML algorithms were developed on the presumption of balanced class distribution, classification based on imbalance classes is challenging [11]. Second, credit card fraud is dynamic and evolves continuously. Criminals quickly adapt to new security measures and detection strategies [12]. As new methods of fraud continue to surface at a rapid pace, it becomes increasingly important for fraud detection systems to be flexible and responsive. These systems must quickly adapt and upgrade their functionalities to stay ahead of emerging fraudulent techniques, ensuring effective detection and prevention measures are in place. By overcoming these challenges, machine learning can provide an important layer of protection against credit card fraud in the digital age.

Deep learning is a sophisticated AI approach that is widely used in a variety of applications. Deep learning has a high capability for learning from big datasets, enables unsupervised learning and is capable of great generalization. It is more powerful and capable of handling more sophisticated applications than shallow models, such as indoor object identification [13], tiredness detection [14] and forecasting issues [15].

The HASC-DLCCFD system, which is based on deep learning and hybrid Harris hawks with sine and cosine, is proposed in this study. The HASC-DLCCFD approach develops a HASC strategy for selection of significant features with the CNN-LSTM model for the identification of fraudulent credit card transactions. Finally, the adaptive moment estimation (Adam) algorithm is utilized to optimize the CNN-LSTM method's hyperparameters. The performance of the suggested HASC-DLCCFD approach is evaluated using a publicly available credit card fraud dataset. The following list summarizes some of the study's important contributions.

- This research presents a new method for identifying credit card fraud, called the HASC-DLCCFD system. It incorporates three key components: HASC for the selection of significant features, CNN-LSTM for accurate classification and Adam for hyperparameter optimization. Based on the researcher's knowledge, this HASC-DLCCFD approach for detecting credit card fraud has not yet been suggested by any authors in the literature.
- The HASC method has been designed by integrating the HHO algorithm with SCA method for the most effective feature selection.
- Adam has been introduced in this research with a CNN-LSTM method for identifying credit card fraud.
- The suggested model's prediction performance for test data is improved by tuning the hyperparameter of CNN-LSTM model utilizing the Adam algorithm and cross-validation.
- When compared to numerous advanced methodologies, the practical experiments indicate that the suggested approach is superior.

This research is structured as follows: An overview of related works is presented in section 2. Section 3 presents a description of the suggested model. The detailed results discussion is given in Section 4. The paper's conclusion is presented in Section 5.

2. Related work

The rising significance of identifying fraudulent credit card transactions has caused a rise in research endeavors in the field. This section introduces an overview of notable research conducted in this area. For a more extensive analysis, further extensive reviews are available in [16–18].

In our previous study [8], we introduced an efficient method for fraud detection in credit card transactions by utilizing an optimized light gradient boosting machine (OLightGBM). Our suggested method integrates a Bayesian-based hyperparameter optimization method to logically fine-tune the parameters of a light gradient boosting machine (LightGBM). To validate the efficacy of OLightGBM in identifying credit card fraud, we conducted tests based on two publicly available datasets containing both genuine and fraudulent transactions. When compared to alternative approaches on these datasets, our proposed method demonstrated superior performance, achieving the highest accuracy of 98.40%, F1-score of 56.95%, precision of 97.34% and area under the curve (AUC) of 92.88%.

Sudha et al. [19] suggested an approach for detecting fraudulent credit card transactions based on the features of the transaction using support vector machine (SVM) and RF methods. The extracted operational features were used as inputs for the SVM and random forest; their results illustrate that SVM obtained the best accuracy at 98%. In [20], Wang and Zhao proposed a modified logistic regression (LR) model to identify fraudulent credit card transactions. First, they employed the Synthetic minority oversampling technique (SMOTE) algorithm to balance the dataset and eliminate extraneous details. Second, they trained the LR model, determined the optimum model parameters

using grid CV search and assessed how well they performed. They compared LR with a number of machine learning algorithms, such as K-nearest neighbors (KNN), decision tree (DT) and SVM to confirm the usefulness of LR. The results from the experiment demonstrated that LR attained an accuracy rating of 94%, while KNN, DT and SVM obtained 91%, 93% and 92%, respectively. Afriyie et al. [21] conducted a study where they examined the achievement of three distinct machine learning algorithms, namely, DT, RF and LR, in classifying and predicting credit card fraud. The authors compared the performance of these models in identifying fraudulent credit card transactions, and they found that RF attained the best accuracy of 96%.

The authors in [22] suggested a credit card fraud identification approach utilizing an optimized back propagation (BP) network based on whale method. They optimized the weights of the BP network by employing the whale swarm optimization method. The dataset used was from Kaggle. There were 492 fraudulent transactions among the 284,807 total transactions. Positives (fraud) represented 0.172% of all transactions, making the dataset unbalanced. Their method achieved an accuracy of 96.40%. The researchers in [23] suggested a hybrid approach using a deep learning model for identifying credit card fraud. To improve the achievement of their deep learning model, they used a variety of methods, such as memory compression, features engineering and mixed and precision methods. The IEEE-CIS fraud dataset, which contains almost a million credit card transactions, was contributed by Vesta Corporation and used to train and test the model. Their approach obtained an accuracy of 95.80%. In [24], Zhang et al. developed a system for detecting fraud that combines a deep learning design with a sophisticated feature engineering process using homogeneity-oriented behavior analysis (HOBA). They conducted a comparative analysis based on an actual dataset from one of the biggest financial banks in China to examine how well the suggested framework works. The test findings showed that their suggested technique is a practical and successful mechanism for detecting credit card fraud. Their method obtained an accuracy of 98.25 %. The authors in [25] addressed the challenging task of credit card fraud detection by proposing a novel model that improves upon long short-term memory (LSTM) with a time-aware gate. The model aims to accurately capture fraudulent patterns by learning representations based on historical transactions of users. Key components of the model include a current-historical attention module that establishes connections between current and historical transactional behaviors, enabling the capture of behavioral periodicity, and an interaction module that learns comprehensive and rational behavioral representations. Extensive experiments conducted on a large real-world transaction dataset and a public dataset validate the effectiveness of the proposed method. The results demonstrate a clear distinction between legitimate and fraudulent behaviors and superior fraud detection performance compared to state-of-the-art methods. The authors in [26] proposed a novel model for credit card fraud detection that extracts transactional behaviors and learns new representations. The model incorporates time-aware gates to capture long- and short-term transactional habits, a time-aware-attention module to extract behavioral information and an interaction module for enhanced representations. Experimental results on real-world and public datasets show that the proposed method effectively distinguishes fraudulent behaviors and outperforms existing approaches in credit card fraud detection.

The study [27] tackles the complex issue of class imbalance with overlap in detecting credit fraud transactions. Fraudsters deliberately craft fraudulent transactions to closely resemble legitimate ones, leading to a significant overlap between fraudulent and genuine data, making them challenging to differentiate. To address this, the authors present a unique hybrid approach that employs a divide-and-conquer strategy. It involves training an anomaly detection model on minority samples to eliminate

outliers and a substantial portion of majority samples. Consequently, an overlapping subset is created, characterized by reduced interference. This subset is then effectively distinguished using a non-linear classifier. To evaluate the quality of the overlapping subset, the authors introduce a novel assessment criterion called dynamic weighted entropy (DWE). Extensive experiments demonstrate the outstanding performance of the proposed method.

3. The suggested approach

In this research, the author has developed a new HASC-DLCCFD method for the detection of credit card fraud. It includes three phases, the selection of features using the HASC method, CNN-LSTM based credit card fraud identification and Adam based parameter optimization. Figure 1 depicts the complete process of the suggested HASC-DLCCFD method.

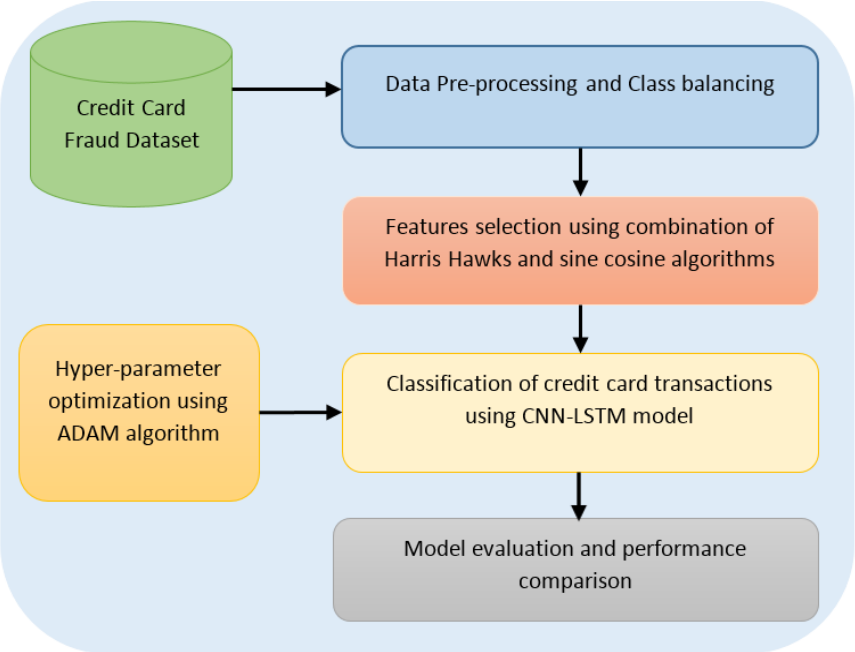


Figure 1. The complete process of the suggested HASC-DLCCFD method.

3.1. Design of the hybrid approach HASC for selecting significant features

At the initial phase, the HASC-DLCCFD approach designs a new HASC method for the selection of feature subset [28]. It is anticipated that the solution's caliber will be enhanced, along with the refinement of its convergence pattern. Moreover, the utilization of a hybrid mechanism can lead to the creation of an exceptionally efficient search by incorporating frequent jumps within the search space, thus evading challenges posed by local optima. Consequently, this approach generates a multitude of varied solutions. The HASC technique is structured in a hierarchical manner. An individual created by the HHO at the top layer is enhanced by the SCA at the lowest layer. The topmost layer consists of M HHO search agents, corresponding to the M group count in the bottommost layer. The N population is formed by all of the groups in the lowest layer. In the early stage of the updating process, the SCA is implemented at the lowest layer to determine the novel location. Subsequently, the individual's position

is improved at the top layer based on the obtained optimal solution. As a result, new equations are generated to depict the stages of exploitation and exploration. The implementation of the exploration stage in the HASC technique utilizes the following mathematical expression:

$$Y_{t+1}^{i} = \begin{cases} y_{rand} - r_2 \left| y_{rand} - 2r_2 \left[y_t + r_8 \sin(r_9) \times \left| r_{10} p_t^i - y_t \right| \right] \right|, c \geq 0.5 \delta a r_{11} < 0.5 \\ y_{rand} - r_2 \left| y_{rand} - 2r_2 \left[y_t + r_8 \cos(r_9) \times \left| r_{10} p_t^i - y_t \right| \right] \right|, c \geq 0.5 \delta r_{11} < 0.5 \\ y_{prey} - y_m - r_3 \left[lb_t + r_4 [ub_t - lb_t] \right], c < 0.5 \end{cases} \quad (1)$$

In formula (1), $\boldsymbol{Y_{t+1}^{i}}$ indicates the location of the t^{th} individual in the uppermost layer who matches the i^{th} search element in the lowermost layer. The location of the searching agent at the t[th] topmost layer is represented by $\boldsymbol{y_t}$. The variable t indicates the current number of iterations.

The variable $\boldsymbol{y_{prey}} = p_i^t$ represents the improved position achieved during the current iteration. The parameters c and r_2 , r_3 , r_4 , r_{11} are random variables. Meanwhile, y_m, ub and lb respectively refer to the average, upper boundary and lower boundary.

$$r_8 = 2 - t \left(\frac{2}{T} \right)$$

$$r_9 = 2_\pi . rand() \quad (2)$$

$$r_{10} = 2. \ rand()$$

The abovementioned besieging strategy utilizes the exploitation phase. Besieging hawks employ this method to target the prey with the least energy during their escape. This is indicated by the conditions r ≥ 0.5 and E < 0.5. The hybrid approach presented incorporates these methods in the following manner.

$$Y_{t+1}^{i} \begin{cases} y_{prey} - E \left| y_{prey} - 2r_2 \left[y_t + r_8 \sin(r_9) \times \left| r_{10} p_i^t - y_t \right| \right] \right|, r_{11} < 0.5 \\ y_{prey} - E \left| y_{prey} - 2r_2 \left[y_t + r_8 \cos(r_9) \times \left| r_{10} p_i^t - y_t \right| \right] \right|, r_{11} \geq 0.5 \end{cases} \quad (3)$$

$$E = 2E_0 \left(1 - \frac{t}{T} \right), t = \{1,2,3, \cdots, T\} \quad (4)$$

When the prey's energy is significantly reduced, the siege becomes intense with frequent rapid descents. This siege is specified by the conditions r < 0.5 and E < 0.5.

$$Y_{t+1}^{i} = \begin{cases} Z \ if \ F(Z) < F(y_t) \ \delta \\ y_t = \begin{cases} y_t + r_8 \sin(r_9) \times \left| r_{10} p_i^t - y_t \right|, r_{11} < 0.5 \\ y_t + r_8 \cos(r_9) \times \left| r_{10} p_i^t - y_t \right|, r_{11} \geq 0.5 \end{cases} \\ X \ if \ F(y_t) \ \delta \\ y_t = \begin{cases} y_t + r_8 \sin(r_9) \times \left| r_{10} p_i^t - y_t \right|, r_{11} < 0.5 \\ y_t + r_8 \cos(r_9) \times \left| r_{10} p_i^t - y_t \right|, r_{11} \geq 0.5 \end{cases} \end{cases} \quad (5)$$

where Z= X + S × LF(D), X = y_{prey} – E | $J_{y_{prey}} - y_m$|,

S denotes random vector of 1 × D, D represents the dimension, and r_7 represents the random parameter.

$$J = 2(1 - r_7) \quad (6)$$

$$LF(D) = \frac{\beta \times u}{|v|^{\frac{1}{\sigma}}} \times 0.01 \quad (7)$$

$$\beta = \left(\frac{sin\left(\frac{\pi\sigma}{2}\right) \times \Gamma\,(1+\sigma)}{\Gamma\left(\frac{1+\sigma}{2}\right) \times \sigma \times 2\left(\frac{\sigma-1}{2}\right)} \right) \tag{8}$$

Mild siege occurs when the hawks take action if r ≥ 0.5 and E ≥ 0.5.

$$Y_{t+1}^{i} = \begin{cases} y_{prey} - [y_t + r_8 \sin(r_9) \times |r_{10} p_i^t - y_t|] - E|y_{prey} - \\ \quad 2_{r2}[y_t + r_8 \sin(r_9) \times |r_{10} p_i^t - y_t|]|, r_{11} < 0.5 \\ y_{prey} - [y_t + r_8\, cos(r_9) \times |r_{10} p_i^t - y_t|] - E|y_{prey} - \\ \quad 2_{r2}[y_t + r_8\, cos(r_9) \times |r_{10} p_i^t - y_t|]|, r_{11} \geq 0.5 \end{cases} \tag{9}$$

$$Y_{t+1}^{i} = \begin{cases} Z\ if\ F(Z) < F(y_t) \delta\ y_t = \\ \quad \begin{cases} y_t + r_8 \sin(r_9) \times |r_{10} p_i^t - y_t|, r_{11} < 0.5 \\ y_t + r_8 \cos(r_9) \times |r_{10} p_i^t - y_t|, r_{11} \geq 0.5 \end{cases} \\ X\ if\ F(X) < F(y_t) \delta\ y_t = \\ \begin{cases} y_t + r_8 \sin(r_9) \times |r_{10} p_i^t - y_t|, r_{11} < 0.5 \\ y_t + r_8 \cos(r_9) \times |r_{10} p_i^t - y_t|, r_{11} \geq 0.5 \end{cases} \end{cases} \tag{10}$$

where

Z = X + S × LF(D)

X = $y_{prey} - E\,|J_{y_{prey}} - y_t|$

The method proposed integrates the objective into a unified equation, allowing the current weight to consider all significant objectives [28]. A fitness function is utilized in this approach, which combines the objectives of FS in the following manner.

$$Fitness\,(X) = \alpha\,.E(X) + \beta \times \left(1 - \frac{|R|}{|N|}\right) \tag{11}$$

Equation (11) defines the fitness of a subset X, denoted as Fitness (X). The term E(X) represents the frequency of mistakes made by the classifier using the features chosen in subset X. The quantities |R| and |N| indicate the numbers of chosen and genuine features, respectively. The parameter α belongs to the range [0, 1], and β is defined as (1−α). Here, α and β determine the respective weights assigned to the classifier mistake and reduction ratio.

Utilizing Harris hawks with sine cosine method to select the significant features for credit card fraud detection can achieve superior results compared to traditional methods due to its unique advantages. First, it leverages the collaborative hunting behavior of Harris hawks, which inspire an optimization algorithm that promotes effective feature selection. This approach enhances the method's ability to identify relevant features accurately. Second, the sine cosine algorithm provides an efficient search mechanism that strikes a balance between exploration and exploitation, leading to better convergence towards optimal solutions. This fusion of nature-inspired optimization and intelligent search algorithms contributes to the excellent performance of the Harris hawks with sine cosine method for credit card fraud detection, making it a promising approach in this domain.

3.2. Optimal CNN-LSTM based credit card fraud identification

The integration of CNN and LSTM leads to a greater enhancement in classification accuracy as

this fusion incorporates both the localized (regional) details within the features and the distant relationships between them [29]. The CNN-LSTM method is utilized for an accurate identification of the fraudulent credit card transactions. The CNN is a type of feedforward neural network, and it possesses remarkable feature extraction capabilities and demonstrates excellent performance across numerous applications [30, 31]. Figure 2 illustrates the convolutional layer and pooling layer, which together make up the majority of the CNN's fundamental design. The objective of the CNN is to create multiple filters that employ a layer-by-layer process of convolution and pooling on input data in order to extract valuable information.

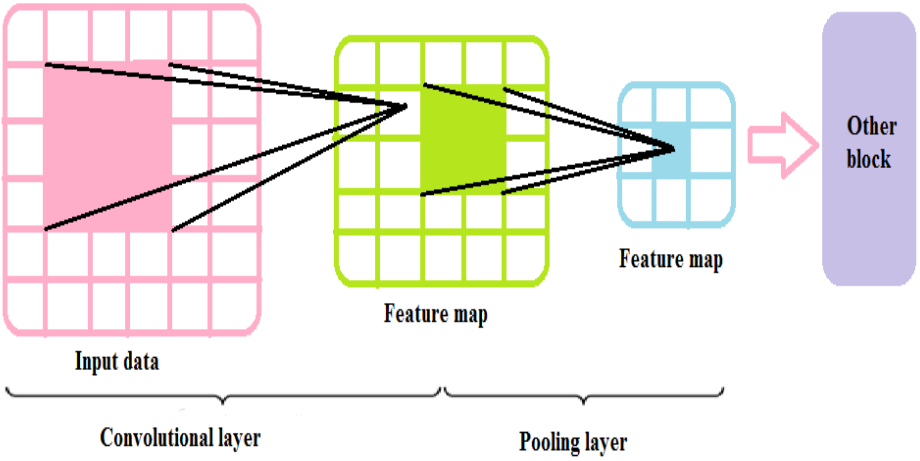


Figure 2. The fundamental structure of a standard CNN.

Convolution kernels, which are small windows, are present in large quantities in the convolutional layer. A convolution kernel is used to convolve the feature maps from the preceding layer, and an activation function creates the output feature. The model's performance can be enhanced by the newly created features, which are often more valuable than the input data's original features. Following is an explanation of how the convolutional layer works:

$$m_j^l = a\left(\sum_{i\in M_j} m_i^{l-1} * k_{ij}^l + b_j^l\right) \tag{12}$$

where m_j^l denotes the j^{th} feature map generated by the l^{th} layer. M_j indicates the chosen input maps. The weights connecting the i^{th} input map to the j^{th} output map are represented as k_{ij}^l. The "*" symbol signifies the convolution process. The term b_j^l indicates the bias associated with the convolution kernel. The function $a(\cdot)$ denotes an activation function, like the rectified linear unit (ReLU). This function allows the feature maps to express themselves nonlinearly, enhancing their ability to represent features effectively.

After performing the convolution operation, the characteristics of the initial data are obtained. However, these features often have a high dimension, leading to practical issues due to their associated costs. To address this problem, a common approach is to incorporate a pooling layer after the convolution layer. This layer performs an important part in accelerating the convergence of the network by reducing the dimensionality of the extracted features. The pooling layer acts as a technique for subsampling, selecting specific values from the convolutional features and generating matrices with smaller dimensions. It operates similarly to the convolution layer, employing a small sliding window to handle the convoluted attributes and produce a fresh output value. Consequently, the output of the

pooling layer can be seen as a compressed representation of the attributes derived from the convolution layer. The options for pooling processes are three: maximum, minimum and average pooling. The mathematical expression for the pooling layer's functionality can be represented by Eq (13).

$$m_j^l = a\left(\zeta_j^l mp\left(m_i^{l-1}\right) + b_j^l\right) \tag{13}$$

where m_j^l denotes the function for performing max pooling, which is a sub-sampling technique. ζ_j^l denotes the bias term. The pooling process guarantees that the convolutional neural network (CNN) acquires a relatively robust feature representation. The output value of the pooling layer remains unchanged even if there are slight differences in the input data. This is because minor variations in the input data do not affect the outcome of the pooling layer.

LSTM is an enhanced version of a recurrent neural network that incorporates specialized cells to preserve long-term memory. It presents a gating method to manage the state of these cells [31]. LSTM has found extensive use across different domains, including but not limited to natural language processing, weather prediction and autonomous vehicle technology. The structure of the LSTM model is explained in Figure 3. The gating mechanism is composed of an input gate, forget gate and output gate. By utilizing historical data x, LSTM generates predictions for the output sequence $y=(y_1, y_2, \ldots, y_d)$, where d denotes the duration for which the prediction is made. The process of managing and updating information follows a series of steps. Initially, the input gate defines the amount of recently acquired knowledge that can be preserved in the cell state and calculates the candidate potential worth $\hat{C}_t$ which can potentially be included in the cell state.

$$i_t = \sigma\left(W_i.[h_{t-1}, x_t] + b_i\right) \tag{14}$$

$$\hat{C}_t = \tanh(W_c\,.\,[h_{t-1}, x_t] + b_c) \tag{15}$$

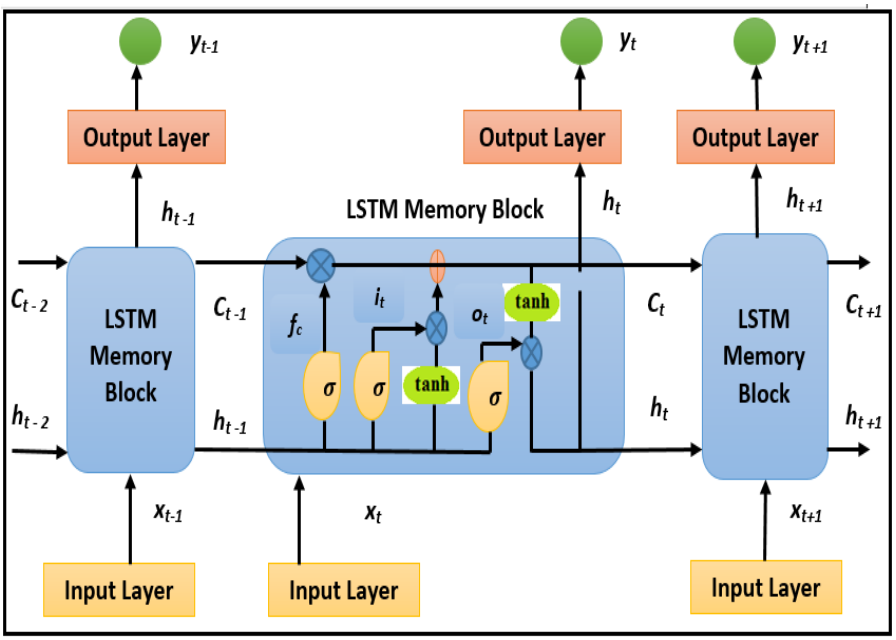


Figure 3. The composition of the memory block in LSTM.

Following that, the forget gate performs an essential part in specifying the amount of information that needs to be discarded or forgotten.

$$f_t = \sigma\left(W_f \,.\, [h_{t-1}, x_t] + b_f\right) \tag{16}$$

The calculation of the cell state in this block, denoted as Ct, involves discarding certain information from the prior cell state Ct−1 and incorporating the cell state candidate of this block, C^t.

$$C_t = f_t \odot C_{t-1} + \hat{C}_t \tag{17}$$

Ultimately, the quantity of information passed to the next memory block is dictated by the output gate. Consequently, the ultimate result is determined based on this factor.

$$o_t = \sigma\left(W_o \,.\, [h_{t-1}, x_t] + b_o\right) \tag{18}$$

$$h_t = o_t \odot \tanh\left(C_t\right) \tag{19}$$

$$y_t = \Phi\left(W_y h_t + b_y\right) \tag{20}$$

In this given context, the following symbols and notations are used: h_t denotes hidden layer state, while i_t, f_t and o_t denote the states of the input gate, forget gate and output gate, respectively. W_i, W_f, W_c, W_o and W_y refer to matrices of weights that are associated with specific operations. Additionally, b_i, b_f, b_c, b_o and b_y represent vectors of bias corresponding to these operations. The sigmoid function is denoted by σ, and the hyperbolic tangent function is represented by *tanh()*. The element-wise product of vectors is symbolized by $\odot$. Furthermore, Φ denotes the activation function applied to the network's output.

3.3. Performance evaluation metrics

The suggested method is assessed using several measures including accuracy, area under the curve (AUC), recall, precision and F1 measure. These metrics are commonly used in the evaluation of classification models and provide a detailed assessment of the model's achievement.

Accuracy: Accuracy is the ratio of cases that are properly categorized to all instances. It is computed using the following equation:

$$Accuracy = \frac{TP+TN}{FP+FN+TP+TN}. \tag{21}$$

Precision: Precision is the ratio of true positive occurrences to all positive occurrences classified by the model. It is computed using the following equation:

$$Precision = \frac{TP}{TP+FP}. \tag{22}$$

Recall: Recall is the ratio of correctly identified positive occurrences to the overall number of positive occurrences present in the dataset. It is computed using the following equation:

$$Recall = \frac{TP}{TP+FN}. \tag{23}$$

F-measure: The F-measure is a mathematical average that combines both recall and precision using a harmonic mean. It is computed using the following equation:

$$F - measure = \frac{2*(Precision*Recall)}{(Precision+Recall)} \quad . \tag{24}$$

AUC: The AUC, also known as the area under the ROC curve, evaluates how effectively a model is able to discriminate between positive and negative classes. To calculate AUC, we plot the true positive rate against the false positive rate and determine the area under the resulting curve. The range of AUC values extends from 0 to 1, with a score of 1 signifying flawless discrimination and 0 signifying no discrimination.

4. Results and discussion

The Python 3.6.5 language was utilized to develop the proposed method for detecting credit card fraud. This approach was implemented on a PC equipped with an i7-3740QM processor, GeForce 1050Ti graphics card with 4 GB memory, 8 GB RAM, a 250 GB SSD and a 1 TB hard disk drive.

The proposed approach employed specific parameter values, including a dropout rate of 0.2, a learning rate of 0.001, a batch size of 32, 300 epochs and the ReLU activation function. These values were carefully selected to optimize the performance of our approach. In contrast, for the other approaches used for comparison, we utilized the default parameter settings available in Python. By employing consistent default settings, we ensure a fair and unbiased evaluation of our proposed approach against the alternative methods. In this part, the suggested HASC-DLCCFD method's experimental findings are examined based on a widely used credit card fraud dataset [32]. The dataset consisted of 24 attributes and 1,852,394 credit card transactions. It was generated by Brandon Harris' Sparkov Data Generation tool, simulating transactions between January 1, 2019, and December 31, 2020. The dataset comprises 1,842,743 genuine transactions and 9,651 fraudulent transactions, involving 1000 clients and 800 merchants. The included features encompass transaction details such as index, credit card number, date and time, merchant information (name and category), transaction amount, cardholder's details (name, gender, address, latitude and longitude), job, date of birth, transaction number and merchant's latitude and longitude, along with the target class. Among the 24 features, the HASC technique specifically selected nine.

The dataset used in this study exhibits an imbalance issue as the number of fraudulent transactions is considerably smaller than that of genuine transactions. To address this skewed distribution problem, the SMOTE approach was employed, resulting in a balanced dataset. Table 1 presents the number of fraudulent and genuine credit card transactions following the application of the SMOTE approach, aiming to rectify the skewed distribution issue.

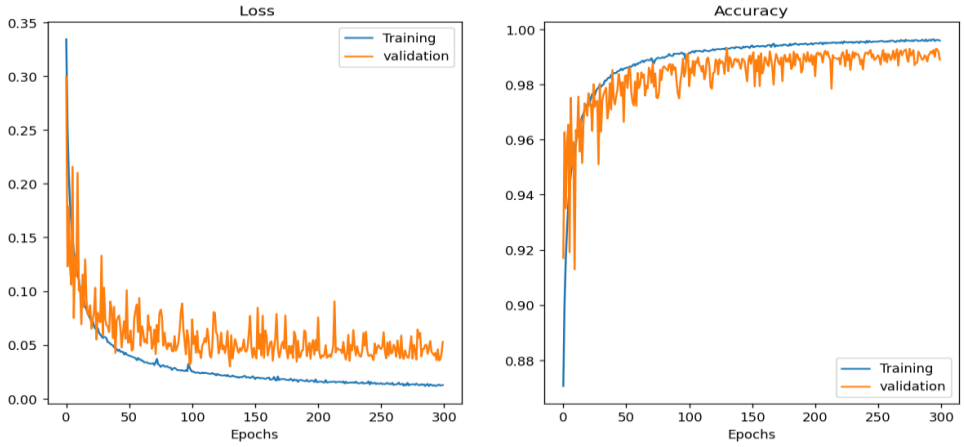
Table 1. Details of the balanced credit card fraud dataset.

Class	No. of samples
Fraud	70000
Normal	70000
Total	140000

During the initial phase of the experiments, the well-balanced dataset was divided into two sets: 80% for training and 20% for testing. The accuracy and loss curves of the HASC-DLCCFD method are depicted in Figure 4. Figure 4(a) indicates that the suggested method HASC-DLCCFD attained high accuracy scores as the number of epochs increased. Figure 4(b) illustrates the loss curve of the HASC-

DLCCFD method. The loss function employed in HASC-DLCCFD serves as a means of gauging the distinction between the anticipated result and the outcome generated by the HASC-DLCCFD. It quantifies the extent to which an estimated value deviates from its true value [33]. The loss function calculates the logarithm of the output index based on the provided ground truth. Consequently, the loss is computed only once per instance, eliminating the need for summation and thereby enhancing speed [33]. The sparse categorical cross entropy loss is determined using the following formula:

$$J(w) = -\log(\hat{y}_y) \tag{25}$$



(a). Accuracy curve (b). Loss curve

Figure 4. (a) Accuracy curve and (b) loss curve of the HASC-DLCCFD.

To explore the performance of the suggested credit card fraud identification approach, HASC-DLCCFD is compared with several classification methods, deep multi perceptron neural network (DNN), DT, KNN, Naïve Bayes (NB) and AdaBoost models.

The findings in Table 2 and Figure 5 highlight the remarkable performance of the suggested credit card fraud identification approach. The achieved accuracy, precision, recall, f-measure and AUC scores of 99.5%, 99.45%, 99.43%, 99.31% and 98.54%, respectively, demonstrate the effectiveness of integrating Hawks and sine and cosine algorithms for feature selection and combining CNN and LSTM for addressing the credit card fraud detection problem.

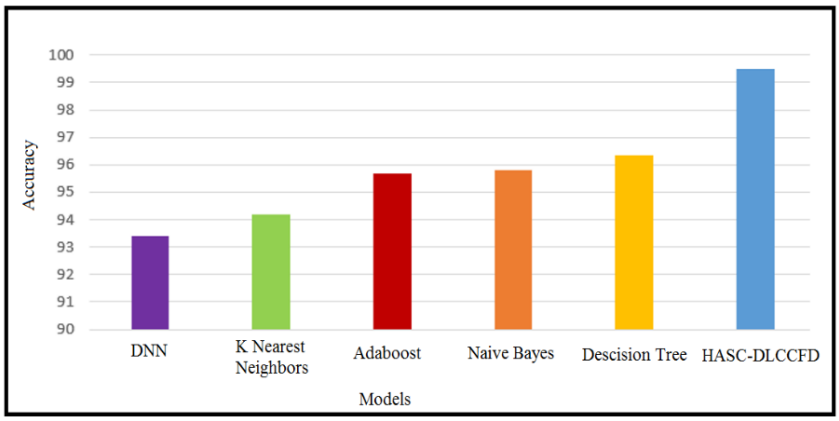


Figure 5. Comparison between the HASC-DLCCFD approach and other systems.

Table 2. Comparative analysis results of the suggested HASC-DLCCFD method and alternative systems.

Model	Accuracy	Precision	Recall	F1-measure	AUC
K Nearest Neighbors Model	94.20	97.81	93.43	96.51	95.80
Naïve Bayes Model	95.80	99.32	95.50	97.93	87.08
Decision Tree Model	96.34	98.10	97.42	98.21	88.41
AdaBoost Model	95.70	99.30	95.71	97.87	93.92
DNN Model	93.42	99.10	93.28	96.51	96.05
HASC-DLCCFD	99.50	99.45	99.43	99.31	98.54

It is worth mentioning that the results presented in Table 2 were obtained using the proposed approach with 9 selected features through the HASC method. However, it is important to note that the other approaches in Table 2 used the complete set of 24 features. By presenting the performance of the proposed approach in comparison to these other approaches, the effectiveness of the feature selection process carried out by HASC is demonstrated. The high accuracy score of 99.5% indicates that the proposed approach successfully classified 99.5% of credit card transactions accurately. This suggests that the model can effectively distinguish between fraudulent and legitimate transactions, minimizing the misclassification of fraudulent activities as legitimate ones. Precision, which measures the proportion of correctly predicted fraudulent transactions among all predicted frauds, achieved an impressive value of 99.4%. This implies that the proposed approach has a low false positive rate, thereby minimizing the instances where legitimate transactions are falsely flagged as fraudulent. Furthermore, the recall score of 99.3% demonstrates the model's capability to identify and correctly classify a significant majority of actual fraudulent transactions. The high recall score indicates a low false negative rate, meaning that the proposed approach minimizes the chances of missing fraudulent transactions. The F-measure, which combines precision and recall into a single metric, achieved a commendable value of 99.3%. This indicates that the proposed approach strikes a well-balanced performance between precision and recall, effectively handling both false positives and false negatives. Additionally, the AUC score of 98.5% signifies the strong discriminative power of the proposed approach [34]. It implies that the model can effectively rank and differentiate between genuine and fraudulent transactions, making it a reliable tool for credit card fraud identification.

The findings illustrated in Figure 6 illustrate the performance of the proposed approach in comparison to other existing methods, as measured by the AUC. A closer examination of the curves reveals that the suggested method consistently performs better than the alternative methods, as evidenced by achieving the highest curve positioned at the graph's upper-left corner. The positioning of the proposed approach's curve in the top left corner signifies its superiority in terms of both sensitivity and specificity. This favorable outcome indicates that the proposed approach has successfully struck a balance between correctly identifying positive instances (true positives) and accurately classifying negative instances (true negatives). By achieving the highest AUC, our approach demonstrates its ability to discern between positive and negative instances more effectively than the other methods tested in this research.

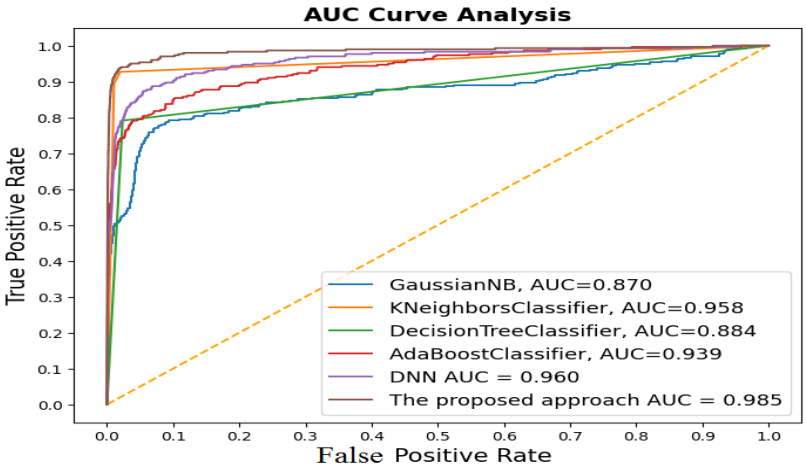


Figure 6. AUC curves of the suggested HASC-DLCCFD method and other approaches.

The confusion matrix resulting from the suggested HASC-DLCCFD approach technique is explained in Figure 7. The figure emphasizes the effectiveness of the suggested HASC-DLCCFD approach in terms of identifying 280 fraudulent credit card transactions and 29391 genuine transactions.

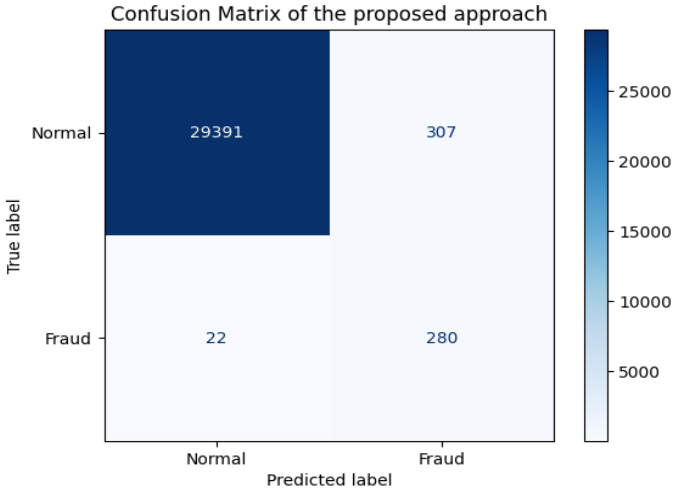


Figure 7. The confusion matrix of the suggested HASC-DLCCFD approach for credit card fraud identification.

As shown in Table 3, the experiments in this research also demonstrated that the proposed approach achieved superior accuracy in comparison to other publications in the literature. The proposed HASC-DLCCFD method surpasses conventional machine learning techniques like [8, 19, 20, 21, 22,35,36], as well as deep learning models such as [23, 24], attaining the highest accuracy of 99.5% in detecting credit card fraud. The proposed research in this paper introduces a novel deep learning-based hybrid Harris Hawks with Sine Cosine method for credit card fraud detection system (HASC-DLCCFD). It combines Harris hawks optimization (HHO) with the sine cosine algorithm (SCA) for feature selection and integrates a convolutional neural network with long short-term memory (CNN-LSTM) to distinguish between fraudulent credit card transactions and normal transactions. In [25], the authors enhanced credit card fraud detection by developing a model that

utilizes users' historical transactions to accurately identify fraudulent patterns. The model incorporates improvements to long short-term memory and introduces a time-aware gate to capture behavioral changes resulting from consecutive transactions. Additionally, a current-historical attention module establishes connections between current and past transactional behaviors, enabling the model to capture behavioral periodicity. On the other hand, in [26], the authors proposed a different model for credit card fraud detection, focusing on extracting transactional behaviors and learning new representations. By incorporating time-aware gates and a time-aware attention module, the model captures both long- and short-term habits of users, extracting valuable behavioral information from historical transactions. Furthermore, an interaction module enhances the comprehensiveness of the learned representations.

Table 3. Comparing the proposed HASC-DLCCFD approach with selected state-of-the-art studies.

Author	**Accuracy**	Precision	Recall	F1 measure,	AUC
Taha and Malebary [8]	98.40%	97.34%	40.59	56.95%	90.94%
Sudha et al. [19]	98.00%	94.00%	86.00%	90.00%	-
Wang and Zhao [20]	94.00%	-	-	-	98.24%
Afriyie et al. [21]	96.00%	90.00%	97.00%	17.00%	98.90%
Wng et al. [22]	96.40%.	-	-	-	-
Kewei et al. [23]	95.80%	-	-	-	91.00%
Zhang et al. [24]	98.25 %	-	-	-	-
Xie et al. [25]	-	88.00%	90.80%	88.10%	94.40
Xie et al. [26]	-	88.30%	93.80%	90.50%	95.20%
Jiang et al. [35]	-	97.95%	75.53%	85.29%	95.15%
Baabdullah et al. [36]	99.00%	94.00%	80.00%	-	90.00%
H. Ahmad et al. [37]	96.60%	96.40%	-	93.00%	-
Mniai et al. [38]	97.00%	-	-	-	94.00%
HASC-DLCCFD	99.50%	99.45	99.43	99.31	98.54

5. Conclusions

Detecting fraudulent credit card transactions is crucial for maximizing the effectiveness of credit cards in diverse e-commerce applications and digital payment solutions. In this research, the author has established a novel HASC-DLCCFD method for identifying fraud in the credit card transactions. The HASC-DLCCFD approach introduces a novel approach to feature selection by merging the HHO and SCA techniques. To detect fraud, the CNN-LSTM model is employed, and the Adam optimizer is utilized for hyperparameter tuning in the final stage. The effectiveness of the HASC-DLCCFD system is verified using a publicly accessible credit card fraud dataset, demonstrating its superior performance compared to other algorithms. In future work, enhancing the efficiency of the HASC-DLCCFD approach can be achieved by utilizing ensemble fusion-based deep learning techniques. Moreover, addressing the imbalance problem with overlap in credit card fraud detection will also be considered in the future work.

Use of AI tools declaration

The authors declare they have not used artificial intelligence (AI) tools in the creation of this article.

Acknowledgments

This research work was funded by Institutional Fund Projects under grant no. (IFPIP: 1524-830-1443). The author gratefully acknowledges technical and financial support provided by the Ministry of Education and King Abdulaziz University, DSR, Jeddah, Saudi Arabia.

Conflict of interest

The author declares no conflict of interest

References

1. R. Van Belle, B. Baesens, J. De Weerdt, CATCHM: A novel network-based credit card fraud detection method using node representation learning, *Decis. Support Syst.,* **164** (2023), 113866. https://doi.org/10.1016/j.dss.2022.113866
2. G. Zhang, Z. Li, J. Huang, J. Wu, C. Zhou, J. Yang, et al., efraudcom: An e-commerce fraud detection system via competitive graph neural networks, *ACM T. Inform. Syst.,* **40** (2022), 1–29. https://doi.org/10.1145/3474379
3. N. Prabhakaran, R. Nedunchelian, Oppositional Cat Swarm optimization-based feature selection approach for credit card fraud detection, *Comput. Intell. Neurosci.*, **2023** (2023), Article ID 2693022. https://doi.org/10.1155/2023/2693022
4. H. Fanai, H. Abbasimehr, A novel combined approach based on deep Autoencoder and deep classifiers for credit card fraud detection, *Expert Syst. Appl.*, **217** (2023), 119562. https://doi.org/10.1016/j.eswa.2023.119562
5. H. Tingfei, C. Guangquan, H. Kuihua, Using variational auto encoding in credit card fraud detection, IEEE Access, **8** (2020), 149841–149853. https://doi.org/10.1109/ACCESS.2020.3015600
6. I. D. Mienye, Y. Sun, A deep learning ensemble with data resampling for credit card fraud detection, *IEEE Access,* **11** (2023), 30628–30638. https://doi.org/10.1109/ACCESS.2023.3262020
7. E. F. Malik, K. W. Khaw, B. Belaton, W. P. Wong, X. Chew, Credit card fraud detection using a new hybrid machine learning architecture, *Mathematics*, **10** (2022), 1480. https://doi.org/10.3390/math10091480
8. A. A.Taha, S. J. Malebary, An intelligent approach to credit card fraud detection using an optimized light gradient boosting machine, *IEEE Access*, **8** (2020), 25579–25587. https://doi.org/10.1109/ACCESS.2020.2971354
9. E. Btoush, X. Zhou, R. Gururaian, K. C. Chan, X. Tao, A survey on credit card fraud detection techniques in banking industry for cyber security, In: *Proceedings of the 2021 8th International Conference on Behavioral and Social Computing (BESC)*, Doha, Qatar, 2021. https://doi.org/10.1109/BESC53957.2021.9635559
10. A. Cherif, A. Badhib, H. Ammar, S. Alshehri, M. Kalkatawi, A. Imine, Credit card fraud detection in the era of disruptive technologies: A systematic review, *J. King Saud Univ. Comput. Inf. Sci.,* **35** (2022), 145–174. https://doi.org/10.1016/j.jksuci.2022.11.008
11. I. D. Mienye, Y. Sun, Performance analysis of cost-sensitive learning methods with application to imbalanced medical data, *Inform. Med. Unlocked,* **25** (2021), 100690. https://doi.org/10.1016/j.imu.2021.100690
12. E. Strelcenia, S. Prakoonwit, Improving classification performance in credit card fraud detection by using new data augmentation, *AI*, 4 (2023), 172–198. https://doi.org/10.3390/ai4010008

13. M. Afif, R. Ayachi, Y. Said, M. Atri, An evaluation of EfficientDet for object detection used for indoor robots assistance navigation, *J. Real Time Image Process*., **19** (2022), 651–661. https://doi.org/10.1007/s11554-022-01212-4
14. R. Ayachi, M. Afif, Y. Said, A. Ben Abdelali, Drivers fatigue detection using EfficientDet in advanced driver assistance systems, In: *Proceedings of the 18th International Multi-Conference on Systems, Signals & Devices, Monastir, Tunisia,* 2021. https://doi.org/10.1109/SSD52085.2021.9429294
15. N. Ayoobi, D. Sharifrazi, R. Alizadehsani, A. Shoeibi, J. M. Gorriz, H. Moosaei, et al., Time series forecasting of new cases and new deaths rate for COVID-19 using deep learning methods, *Results Phys.*, **27** (2021), 104495. https://doi.org/10.1016/j.rinp.2021.104495
16. R. Bin Sulaiman, V. Schetinin, P. Sant, Review of machine learning approach on credit card fraud detection, *Human-Centric Intel. Syst.*, **2** (2022), 55–68. https://doi.org/10.1007/s44230-022-00004-0
17. M. Alamri, M. Ykhlef, Survey of credit card anomaly and fraud detection using sampling techniques, *Electronics*, **11** (2022), 4003. https://doi.org/10.3390/electronics11234003
18. E. Strelcenia, S. Prakoonwit, A survey on GAN techniques for data augmentation to address the imbalanced data issues in credit card fraud detection, *Mach. Learn. Knowl. Extr,* **5** (2023), 304–329. https://doi.org/10.3390/make5010019
19. C. Sudha, D. Akila, Credit Card Fraud Detection System based on Operational Transaction features using SVM and Random Forest Classifiers, In: *Proceedings of 2021 2nd International Conference on Computation, Automation and Knowledge Management (ICCAKM)*, Dubai, UAE, 19-21 January 2021. https://doi.org/10.1109/ICCAKM50778.2021.9357709
20. T. Wang, Y. Zhao, Credit Card Fraud Detection using Logistic Regression, In: Proceedings of 2022 International Conference on Big Data, Information and Computer Network (BDICN), Sanya, China, 20 Jan 2022. 301–305. https://doi.org/10.1109/BDICN55575.2022.00064
21. J. K. Afriyie, K. Tawiah, W. A. Pels, S. Addai-Henne, H. A. Dwamena, E. O. Owiredu, et al., Supervised machine learning algorithm for detecting and predicting fraud in credit card transactions, *Decis. Anal.,* **6** (2023), 100163. https://doi.org/10.1016/j.dajour.2023.100163
22. C. Wang, Y. Wang, Z. Ye, L. Yan, W. Cai, S. Pan, Credit Card Fraud Detection Based on Whale Algorithm Optimized BP Neural Network, In: *Proceedings of 2018 13th International Conference on Computer Science Education (ICCSE)*, Colombo, Sri Lanka, 8-11August, 2018. https://doi.org/10.1109/ICCSE.2018.8468855
23. X. Kewei, B. Peng, Y. Jiang, T. Lu, A Hybrid Deep Learning Model For Online Fraud Detection, In: *Proceedings of 2021 IEEE International Conference on Consumer Electronics and Computer Engineering (ICCECE),* Guangzhou, China, 15–17 January 2021. https://doi.org/10.1109/ICCECE51280.2021.9342110
24. X. Zhang, Y. Han, W. Xu, Q. Wang, Hoba: A novel feature engineering methodology for credit card fraud detection with a deep learning architecture., *Inf. Sci.,* **557** (2021), 302–316. https://doi.org/10.1016/j.ins.2019.05.023
25. Y. Xie, G. Liu, C. Yan, C. Jiang, M. Zhou, M. Li, Learning transactional behavioral representations for credit card fraud detection, *IEEE T. Neural Net. Lear.,* 2022, 1–14. https://doi.org/10.1109/TNNLS.2022.3208967
26. Y. Xie, G. Liu, C. Yan, C. Jiang, M. Zhou, Time-aware attention-based gated network for credit card fraud detection by extracting transactional behaviors, *IEEE Trans. Comput. Soc.,* **10** (2022), 1004–1016. https://doi.org/10.1109/TCSS.2022.3158318

27. Z. Li, M. Huang, G. Liu, C. Jiang, A hybrid method with dynamic weighted entropy for handling the problem of class imbalance with overlap in credit card fraud detection, *Expert Syst. Appl.,* **175** (2021), 114750. https://doi.org/10.1016/j.eswa.2021.114750
28. H. Q. Abdulrab, F. A. Hussin, I. Ismail, M. Assaad, A. Awang, H. Shutari, et al., Hybrid Harris Hawks with Sine Cosine for Optimal Node Placement and Congestion Reduction in an Industrial Wireless Mesh Network, *IEEE Access*, **11** (2023), 2500–2523. https://doi.org/10.1109/ACCESS.2023.3234109
29. Y. Liu, C. Yang, K. Huang, W. Liu, A multi-factor selection and fusion method through the CNN-LSTM network for dynamic price forecasting, *Mathematics*, **11** (2023), 1132. https://doi.org/10.3390/math11051132
30. R. Gao, J. Xu, Y. Chen, K. Cho, Heterogeneous feature fusion module based on CNN and transformer for multiview stereo reconstruction, *Mathematics*, **11** (2023), 112. https://doi.org/10.3390/math11010112
31. W. Lu, J. Li, J. Wang, L. Qin, A CNN-BiLSTM-AM method for stock price prediction, *Neural Comput. Appl.*, **33** (2020), 4741–4753. https://doi.org/10.1007/s00521-020-05532-z
32. Credit card fraud dataset. Available from: https://www.kaggle.com/datasets/kartik2112/fraud-detection. (Accessed on 3 March 2023).
33. I. Goodfellow, Y. Bengio, A. Courville, Deep learning, *Genet. Program. Evol. M.*, **19** (2018), 305–307. https://doi.org/10.1007/s10710-017-9314-z
34. P. B. Le, Z. T. Nguyen, ROC curves, loss functions, and distorted probabilities in binary classification, *Mathematics*, **10** (2022), 1410. https://doi.org/10.3390/math10091410
35. S. Jiang, R. Dong, J. Wang, M. Xia, Credit card fraud detection based on unsupervised attentional anomaly detection network, *Systems*, **1** (2023), 1–14. https://doi.org/10.3390/systems11060305
36. T. Baabdullah, D. B. Rawat, C. Liu, A. Alzahrani, An Ensemble-Based Machine Learning for Predicting Fraud of Credit Card Transactions, In: *Proceedings of Intelligent Computing: In Proceedings of 2022 Computing Conference*, London, United Kingdom, July 14–15, 2022. https://doi.org/10.1007/978-3-031-10464-0_14
37. H. Ahmad, B. Kasasbeh, B. Aldabaybah, E. Rawashdeh, Class balancing framework for credit card fraud detection based on clustering and similarity-based selection (SBS), Int. J. Inf., **15** (2023), 325–333. https://doi.org/10.1007/s41870-022-00987-w
38. A. Mniai, K. Jebari, Credit Card Fraud Detection by Improved SVDD, In: *Proceedings of 2022 World Congress on Engineering*, WCE 2022, London, U.K., 6–8 July, 2022. ISBN: 978-988-14049-3-0