Mathematical Statistician and Engineering Applications
ISSN: 2094-0343
2326-9865

Hunger Search Algorithm with Optimal Deep Learning Driven Credit Card Fraud Detection and Classification Model

G. K. Arun[1], P. Rajesh[2]

[1]Department of Computer and Information Science, Annamalai University, Annamalai Nagar – 608 002, Tamil Nadu, India.

[2]PG Department of Computer Science, Government Arts College, Chidambaram – 608 102 (Deputed from Department of Computer and Information Science, Annamalai University, Annamalai Nagar), Tamil Nadu, India

[1]arunnura2370@gmail.com, [2]rajeshdatamining@gmail.com

Article Info
Page Number: 387-406
Publication Issue:
Vol. 71 No. 4 (2022)

Article History
Article Received: 25 March 2022
Revised: 30 April 2022
Accepted: 15 June 2022
Publication: 19 August 2022

Abstract

With the development of e-commerce websites, persons and financial companies depend on online services for executing its transaction which has performed an exponential improvement from credit card fraud (CCF). The fraudulent credit card transaction gives rise to a loss of massive count of money. The proposal of effectual fraud detection method was essential for reducing the losses suffered by the customer and financial company. The research is complete on several methods for preventing and detecting CCFs. Feature of CCFs role a vital play if machine learning (ML) was utilized to credit card fraud detection (CCFD), and it can be selected accurately. This article focuses on the design of hunger search algorithm with optimal deep learning is driven credit card fraud detection and classification (HSAODL-CCFC) model. The major intention of the HSAODL-CCFC model is to properly distinguish credit card transactions from legitimate or fraud. To accomplish this, the presented HSAODL-CCFC model applies data pre-processing at the initial stage. Next, a new binary version of HSA based feature selection (BHSA-FS) approach was executed for electing feature subsets. Finally, root means square propagation (RMSProp) with deep belief network (DBN) technique was utilized for the identification and classification of the CCFs. The simulation analysis of the HSAODL-CCFC model is tested using 2 benchmark datasets. The comparative analysis reported the promising performance of the HSAODL-CCFC model over recent approaches.

Keywords: Credit card transactions; Fraud detection; Machine learning; Metaheuristics; Deep learning; Feature selection

1. Introduction

In past decades, there was an exponential development of the Internet [1]. This has activated the growth and rise in the usage of services like tap and pay systems, e-commerce, online bill payment systems, and much more. As a result, intruders are also raised actions for attacking transactions which can be done with the help of credit cards [2]. There comes various modalities utilized for protecting credit card transactions which include credit card data tokenization and encryption. Credit card transactions were highly common at present but then it has certain sets of issues [3]. There exist many issues faced at the time of fraud detection [4]. Thus, the process implanted for the identification of a fraudulent transaction must be very quick and efficient. One more issue is there coming a massive number of similar kinds of transactions taking place simultaneously. As a result, it becomes tough to observe all transactions separately [5]. Therefore, an effective Fraud Detection System has to be enforced for differentiating fraud and genuine transactions. Though these methodologies were efficient in many scenarios, they cannot completely protect credit card transaction from fraud [6].

Machine Learning (ML) is a sub-domain of Artificial Intelligence (AI) which permits computers for learning in earlier information and improve its prediction capabilities without being explicitly programmed to do so [7]. In this article the application of ML methodologies for credit card fraud detection (CCFD). The one main problem in implementing ML techniques is the CCFD issue is many publications were impractical to reproduce [8]. The reason behind that is credit card transactions were more confidential. Thus, the datasets which were utilized for developing ML techniques for CCFD consist of anonymized attributes [9]. Moreover, CCFD becomes a challenging one owing to the constantly changing patterns and nature of fraudulent transactions. In addition to this, prevailing ML methodologies for CCFD endure a low detection accuracy and cannot solve the highly skewed characteristics of CCF datasets [10]. Thus, it becomes necessary for developing ML techniques which could execute optimally and could detect CCF with more accuracy.

The authors in [11] aimed to protect credit card transactions; so people might utilize e-banking easily and safely. To detect the CCF there are different methods that depend on Naive Bayesian, Deep learning, Logistic Regression, Neural Network, Artificial Immune systems, Support Vector Machine (SVM), Data Mining, K Nearest Neighbor, Fuzzy logic based Systems, Genetic Algorithm Decision Tree, and so on. The authors in [12] developed and designed a fraud detection model to Stream Transaction Data, with the objective of analyzing the historical transaction details of the customer and extracting the behavioural pattern. Where cardholder is clustered into distinct classes according to their transaction amount. Next, utilizing sliding window strategy, for aggregating the transaction made by the cardholder from distinct classes such that the behavioural pattern of the group is extracted correspondingly. In [13], ML algorithm is utilized for detecting CCF. Firstly, standard model is initially utilized. Next, hybrid method uses majority voting and AdaBoost methods are employed. To estimate the model efficiency, an open source credit card dataset was utilized. Next, a real-time credit card dataset from a financial corporation is investigated. Furthermore, noise is included in the data sample for additionally assessing the strength of the algorithm.

The authors in [14] proposed multiple algorithms of ML namely ANN, SVM, and KNN in forecasting the incidence of the fraud. Furthermore, it is conducted a differentiation of attained supervised ML and DL approaches for differentiating between fraud and non-fraud transactions. Azhan and Meraj [15] presented a study that further analyzed the fraudulent activity related to credit cards. While all of them could not be handled concurrently, the study discussed how ML and NN methods are utilized for determining the potential fraudster with reference to the details of previous fraudsters and earlier mistakes. ML techniques namely RF Regression, Multi-nomial NB, LR, NN, and SVM are also utilized.

This article focuses on the design of hunger search algorithm with optimal deep learning is driven credit card fraud detection and classification (HSAODL-CCFC) model. The presented HSAODL-CCFC model applies data pre-processing at the initial stage. Next, a new binary version of HSA based feature selection (BHSA-FS) approach was executed for electing feature subsets. Finally, root means square propagation (RMSProp) with deep belief network (DBN) technqiue was utilized for the identification and classification of the CCFs. The simulation analysis of the HSAODL-CCFC model is tested using two benchmark datasets and examines the results under distinct dimensions.

2. The Proposed Model

In this article, an effective HSAODL-CCFC model has been introduced to properly distinguish credit card transactions from legitimate or fraud. The presented HSAODL-CCFC model primarily applied data pre-processing at the initial stage. Then, the BHSA-FS technique is applied to elect feature subsets. Next, the RMSProp with DBN technique was utilized for the identification and classification of the CCFs. Fig. 1 depicts the overall procedure of HSAODL-CCFC approach

2.1. Data Pre-processing

Initially, pre-processing of input credit information executes from 2 stages like format as well as data conversions. In format conversion, input data in .csv format is converted to. arff format. Next, during the data conversion stage, the arithmetical number is altered to same categorical value, using 0's and 1's value has been converted to worse and optimum credits. If the data has been pre-processed, it can be broadcasted to FS approach for chosen count of features.

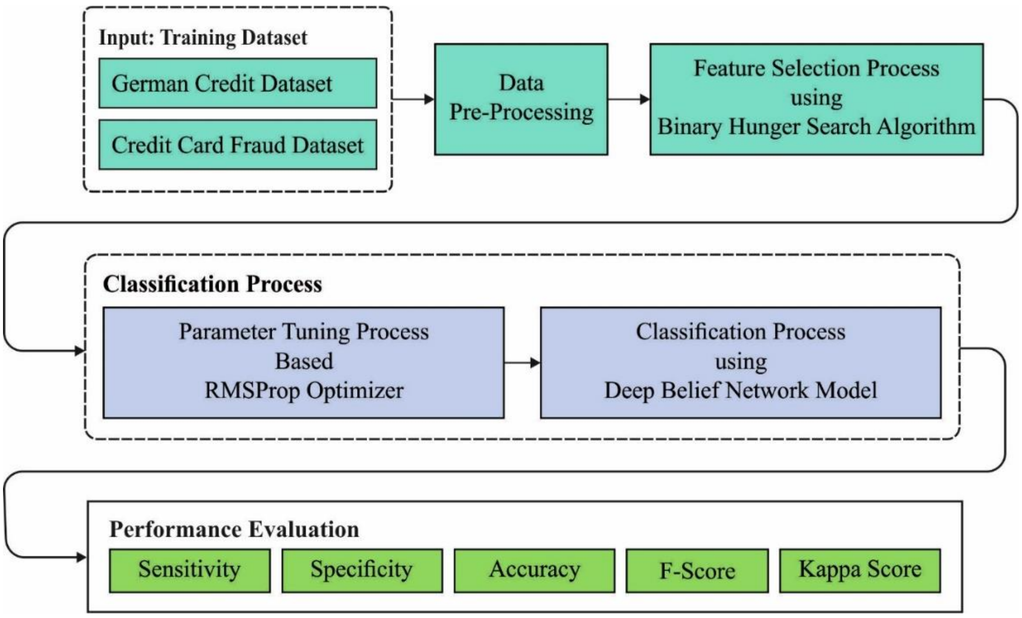

Fig. 1. Overall Process of HSAODL-CCFC approach

2.2. Process involved in BHSA-FS Technique

At this stage, the pre-processed data was passed into the BHSA-FS model to elect feature subsets. The HSA is an optimization technique to model animal behavior proposed by hunger Yang et al. [16]. Hunger's capability to become the primary homeostatic reason for decision, behavior, and action in the animal existence describes HGS. HGS mathematical modelling initiates with a population of N solutions, X, and proceeds to the objective function value for solution. The subsequent equation is utilized for accomplishing the modernization stage:

$$X = \begin{cases} X(t) \times (1 + rand), r_1 < l \\ W_1 \times X_b + R \times W_2 \times |X_b - X(t)|, r_1 > l, r_2 > E \\ W_1 \times X_b - R \times W_2 \times |X_b - X(t)|, r_1 > l, r_2 < E \end{cases} \tag{1}$$

r_1 and r_2 denotes arbitrary numbers, and the parameter rand produces number from a standard distribution, and R indicates a parameter that value is defined as the range $[-a, a]$, and based on the iteration count the below equation is formulated:

$$R = 2 \times s \times rand - s, s = 2 \times \left(1 - \frac{t}{T}\right) \tag{2}$$

While, the variable E, in Eq. (1), indicates the control variable viz., determined by the following equation:

$$E = sech(\quad |Fit_i - Fit_b) \tag{3}$$

Fit_b denotes the optimal value of the objective function and Sech corresponding to the hyperbolic function

$$(x) = \frac{2}{e^x - e^{-x}}.$$

Further, W_1 and W_2 represents the hunger weight provided as follows.

$$W_1 = \begin{cases} H_i \times \frac{N}{SH} \times r_4, r_3 < l \\ 1, r_3 > l \end{cases} \tag{4}$$

$$W_2 = 2\left(1 - e^{(-|H_i - SH|)}\right) \times r_5 \tag{5}$$

r_3, r_4, and r_5 represents arbitrary value that ranges from 0 to 1, and the parameter SH corresponding to the solution of the hunger feeling summation provided in the following:

$$SH = \sum_i H_i\, 0 \tag{6}$$

Moreover, the parameter H_i corresponding the solution to hunger H_i shown as follows:

$$H_i = \begin{cases} 0, Pit_i = Fit_b \\ H_j + H_n, otherwise \end{cases} \tag{7}$$

The optimal value for an objective is supplied by Fit and the existing solution X_i has an objective provided by Fit, and the novel hunger is provided as the parameter H_n:

$$H_n = \begin{cases} LH \times (1+r)\,, TH < LH \\ TH, otherwise \end{cases} \tag{8}$$

$$TH = 2\frac{Fit_i - Fit_b}{Fit_w - Fit_b} \times r_6 \times (UB - LB) \tag{9}$$

Fit_w provides a low value to the objective function, and $r_6 \in [0,1]$ indicates an arbitrary parameter that represents hunger has harmful or positive effects depending on various aspects. Because HGS was firstly developed for continuous global optimization, it is essential to transform the structural model to manage feature selection problems, in which solution is constrained to binary value [0, 1]. Thus, a binary hunger game search (bHGS) is presented for addressing the present study. The building block in designing a binary variant is added a transfer function, which transforms the initial value of agent location into the range of [0, 1]. In another word, the output of transfer function provides likelihood value for classifying [0, 1] into {0, 1}. In general, two well-known classes of transfer function (that is., V-and S-shaped) are adapted. Especially, the S-shaped sigmoidal function receives better efficiency and is denoted by:

$$S\left(x_i^d(t)\right) = \frac{1}{1 + e^{-x_i^d(t)}} \tag{10}$$

From the expression, $x_i^d(t)$ denotes the continuous-value location of i^{th} agent in d^{th} dimensional vector. The output of Eq. (10) is considered by the threshold (likelihood values), the threshold values change as the slope of transfer function increases. Next, the continuous value is transformed to binary '0' or '1' as follows.

$$\chi_i^d(t) = \begin{cases} 0 & if\ rand \le S\left(x_i^d(t)\right) \\ 1 & if\ rand > S\left(\chi_i^d(t)\right) \end{cases} \tag{11}$$

Here, *rand* indicates a value produced within [0, 1].

Algorithm 1: Pseudocode of HGS Algorithm
Begin with the iterations number T, solutions number N
Position initiation of solution X.
while $t \le T$ do
Determine objective value for solution X_i.
Determine optimal solution X_b, Fit_b, Fit_W
Increase H_i
Repeat and upgrade W_1 and W_2
for $doi = 1{:}\,N$
Update R
Update E

Update X_i $t = t + 1$ Display X_b.

2.3. Optimal DBN based Classification

In this study, the RMSProp with DBN approach was utilized for the identification and classification of CCFs. The DBN is type of DNN using hidden layer and a huge amount of hidden units. The typical DBN corresponds to the stacked RBM model with output units [17]. Each RBM contains visible layer ν and hidden state h, connected with un-directed weight. The variable set of RMB is $= (w, b, a)$, where w_{ij} indicates the weights amongst ν_i and h_j. b_i and a_j denotes the bias of layer.

$$E(\theta) = -\sum_i b_i \nu_i - \sum_i a_j h_j - \sum_i \sum_j w_{ij} \nu_i h_j, \tag{12}$$

and joint probability distribution of ν and h are given below

$$p(\nu, h|\theta) = \frac{\exp\left(-E(\nu, h|\theta)\right)}{\sum_{\nu,h} \exp\left(-E(\nu, h|\theta)\right)}. \tag{13}$$

As well, the marginal likelihood distribution of ν is shown below

$$p(\nu|\theta) = \frac{\sum_{\nu,h} \exp\left(-E(\nu, h|\theta)\right)}{\sum_{\nu,h} \exp\left(-E(\nu, h|\theta)\right)}. \tag{14}$$

For acquiring the best θ for individual dataset vector ν, the gradient of log probability approximation is evaluated as follows:

$$\frac{\partial \log p(\nu|\theta)}{\partial w_{ij}} = \langle \nu_i h_j \rangle_{data} - \langle \nu_i h_j \rangle_{model},$$

$$\frac{\partial \log p(\nu|\theta)}{\partial a_j} = \langle h_j \rangle_{data} - \langle h_j \rangle_{mode1}, \tag{15}$$

$$\frac{\partial \log p(\nu|\theta)}{\partial b_i} = \langle \nu_i \rangle_{data} - \langle \nu_i \rangle_{model}.$$

Now, $\langle \cdot \rangle$ indicates the probability in the distribution represented by subscript. As there is no connection amongst the units from the same layer, $\langle \cdot \rangle_{data}$ attain the conditional probability distribution:

$$pp(\nu_i|h, \theta) = \frac{1}{1 + \exp\left(-\sum_j w_{ij} h_j - b_i\right)}, \tag{16}$$

$$(h_j|\nu, \theta) = \frac{1}{1 + \exp\left(-\sum_j w_{ij} \nu_j - a_j\right)}.$$

For $\langle\cdot\rangle_{model}$ the contrastive divergence (CD) learning method is utilized for minimizing the divergence of two Kullback-Leibler divergences (KL). The weight in DBN layer is trained using unlabelled dataset with the abovementioned greedy and faster unsupervised methodology. For modelling a prediction, a supervised layer needs to add above DBN for altering the learned feature by labelled dataset with an up-down fine-tuning approach. Now, the FC layers are selected for implementing as the topmost layers and exploiting the sigmoid function. Fig. 2 showcases the framework of DBN.

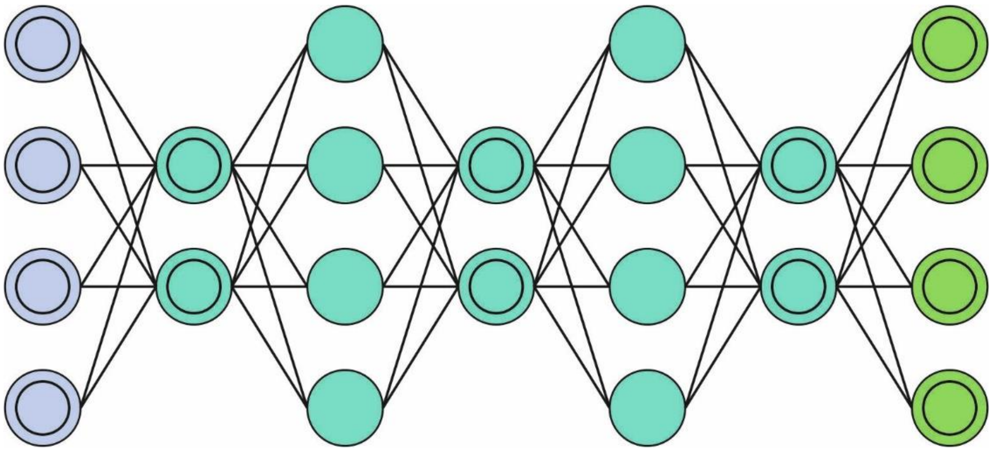

Fig. 2. Framework of DBN model

To tune the hyperparameters related to the DBN model, the RMSProp optimizer is used. An adaptive learning approach which purpose for enhancing AdaGrad is RMSprop [18]. For AdaGrad's cumulative sum of squared gradients, the 'exponential moving average' was utilized.

$$w_{t+1} = w_t - \frac{\alpha_t}{(v_t + \varepsilon)^{\frac{1}{2}}} * \left[\frac{\delta L}{\delta w_t}\right] \tag{17}$$

whereas

$$v_t = \beta v_{t-1} + (1-\beta) * [\frac{\delta L}{\delta w_t}]^2 \tag{18}$$

W_t refers the weighted at time t

W_{t+1} denotes the weighted at time $t+1$

α_t implies the rate of learning at time t

∂L represents the derivative of Loss Function

∂W_t stands for the derivative of weights at time t

V_t defines the sum of square of past gradients.

β demonstrates the moving average parameter (const, 0.9)

ε signifies the smaller positive constant (10^{-8})

3. Experimental Validation

In this section, the experimental validation of the HSAODL-CCFC approach is tested using two datasets namely German Credit [19] and Credit Fraud Detection [20] dataset. Table 1 depicts the details dataset description.

Table 1 Dataset Description

Descriptions	**German Credit Dataset**	**Credit Fraud Detection Dataset**
Source	UCI	Kaggle
# of instances	1000	284807
# of attributes	20	30
# of class	2	2
Classes: Good/Bad	700/300	284315/492

Table 2 gives a brief result analysis of the BHSA-FS model on the feature selection process. On the test German Credit dataset, the BHSA-FS model has chosen a total of 6 features with the best cost of 0.106. In addition, on the test Credit Fraud Detection dataset, the BHSA-FS approach has chosen a total of 10 features with best cost of 0.539.

Table 2 Proposed FS based Selected Features and its Best Cost

Methods	**Dataset**	**Selected Features**	**Best Cost**
BHSA-FS	German Credit	1,5,8,10,14,16	0.106
	Credit Fraud Detection	1,3,6,8,10,13,15,17,22,28	0.539

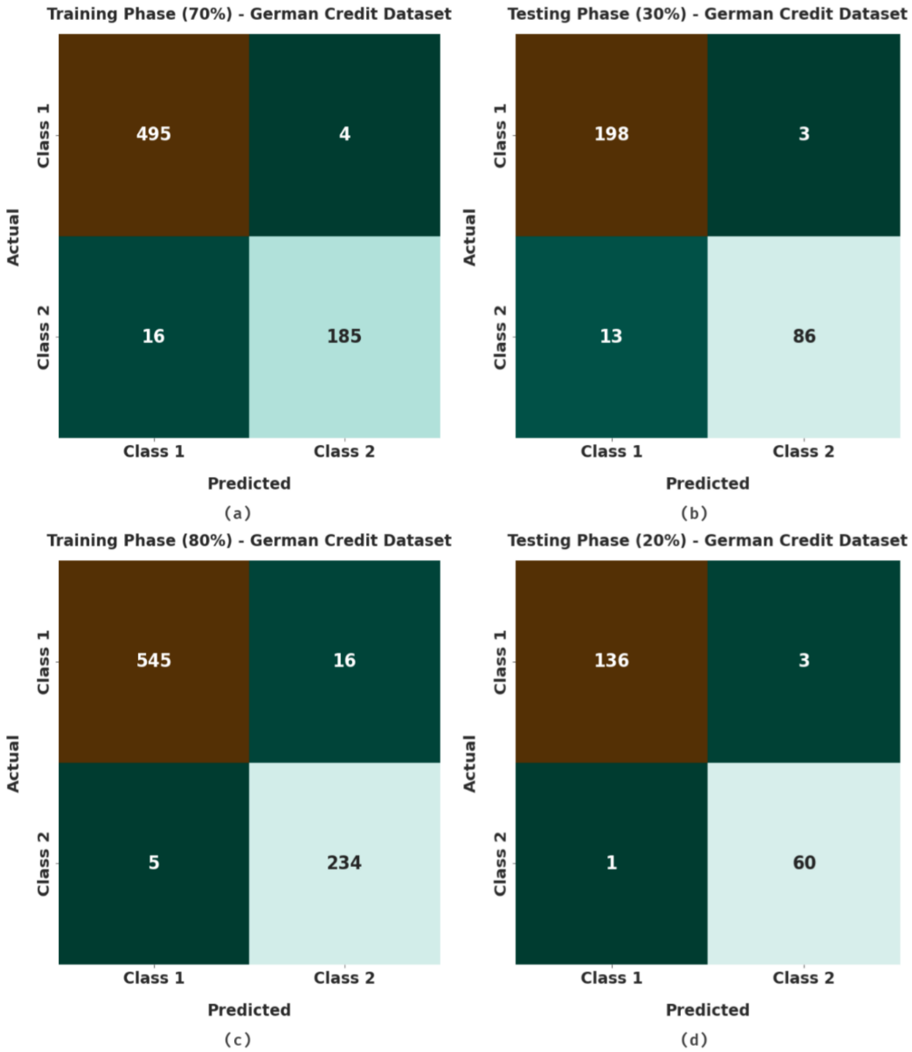

Fig. 3. Confusion matrices of HSAODL-CCFC technique on German Credit dataset (a) 70% of TR data, (b) 30% of TS data, (c) 80% of TR data, and (d) 20% of TS data

Fig. 3 offers the confusion matrices produced by the HSAODL-CCFC model on the executed German Credit dataset. With 70% of training (TR) data, the HSAODL-CCFC model has identified 495 samples under class 1 and 185 samples under class 2. Also, with 30% of testing (TS) data, the HSAODL-CCFC approach has identified 198 samples under class 1 and 86 samples under class 2. In addition, with 80% of TR data, the HSAODL-CCFC systen has identified 545 samples under class 1 and 234 samples under class 2.

Table 3 and Fig. 4 demonstrate a detailed outcome analysis of the HSAODL-CCFC model on German credit dataset. The results indicated that the HSAODL-CCFC approach has reached effectual outcomes in all aspects. For instance, on 70% of TR data, the HSAODL-CCFC model has offered average $accu_y$ of 97.14%, $sens_y$ of 95.62%, $spec_y$ of 95.62%, F_{score} of 96.45%, and kappa of 92.89%. At the same time, on 80% of TR data, the HSAODL-CCFC methodology has obtainable average $accu_y$ of 97.38%, $sens_y$ of 97.53%, $spec_y$ of 97.53%, F_{score} of 96.91%, and kappa of 93.82%. In line with, 20% of TS data, the HSAODL-CCFC system has accessible average $accu_y$ of 98%, $sens_y$ of 98.10%, $spec_y$ of 98.10%, F_{score} of 97.66%, and kappa of 95.33%.

Table 3 Result analysis of HSAODL-CCFC approach on German credit dataset

Labels	Accuracy	Sensitivity	Specificity	F-Score	Kappa Score
Training Phase (70%)					
Class 1	97.14	99.2	92.04	98.02	-
Class 2	97.14	92.04	99.20	94.87	-
Average	**97.14**	**95.62**	**95.62**	**96.45**	**92.89**
Testing Phase (30%)					
Class 1	94.67	98.51	86.87	96.12	-
Class 2	94.67	86.87	98.51	91.49	-
Average	**94.67**	**92.69**	**92.69**	**93.80**	**87.62**
Training Phase (80%)					
Class 1	97.38	97.15	97.91	98.11	-
Class 2	97.38	97.91	97.15	95.71	-
Average	**97.38**	**97.53**	**97.53**	**96.91**	**93.82**
Testing Phase (20%)					
Class 1	98.00	97.84	98.36	98.55	-
Class 2	98.00	98.36	97.84	96.77	-
Average	**98.00**	**98.10**	**98.10**	**97.66**	**95.33**

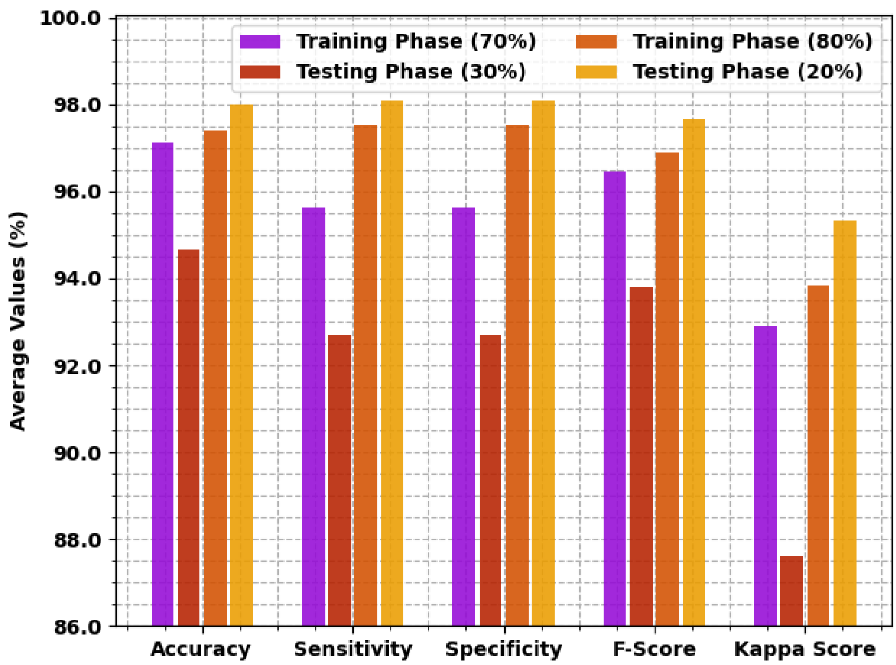

Fig. 4. Result analysis of HSAODL-CCFC approach on German credit dataset

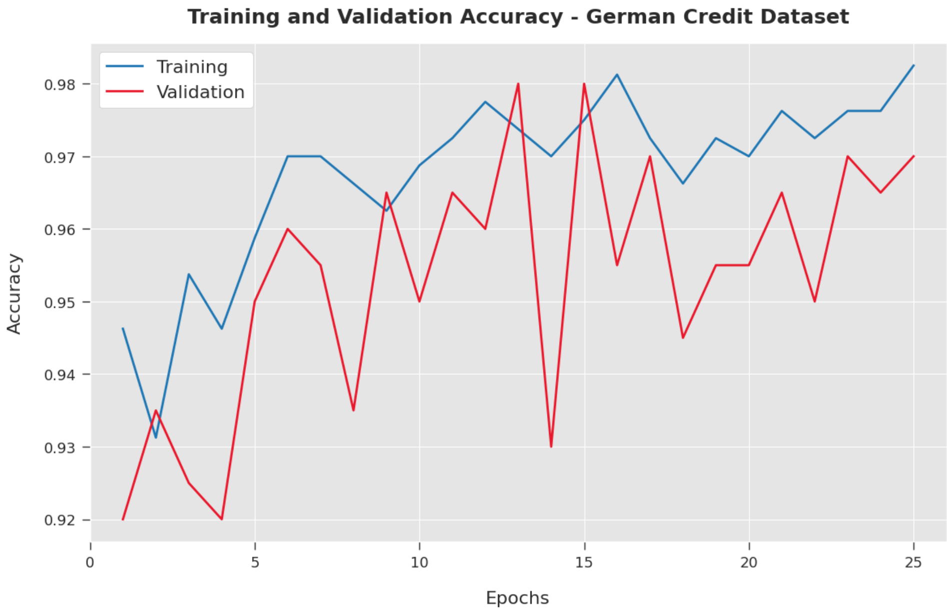

Fig. 5. TA and VA analysis of HSAODL-CCFC approach on German credit dataset

The training accuracy (TA) and validation accuracy (VA) attained by the HSAODL-CCFC approach on German credit dataset is demonstrated in Fig. 5. The experimental outcome implied that the HSAODL-CCFC methodology has gained maximal values of TA and VA. In specific, the VA seemed to be higher than TA.

The training loss (TL) and validation loss (VL) achieved by the HSAODL-CCFC approach on German credit dataset in Fig. 6. The experimental outcome inferred that the HSAODL-CCFC system has been able least values of TL and VL. In specific, the VL seemed to be lesser than TL.

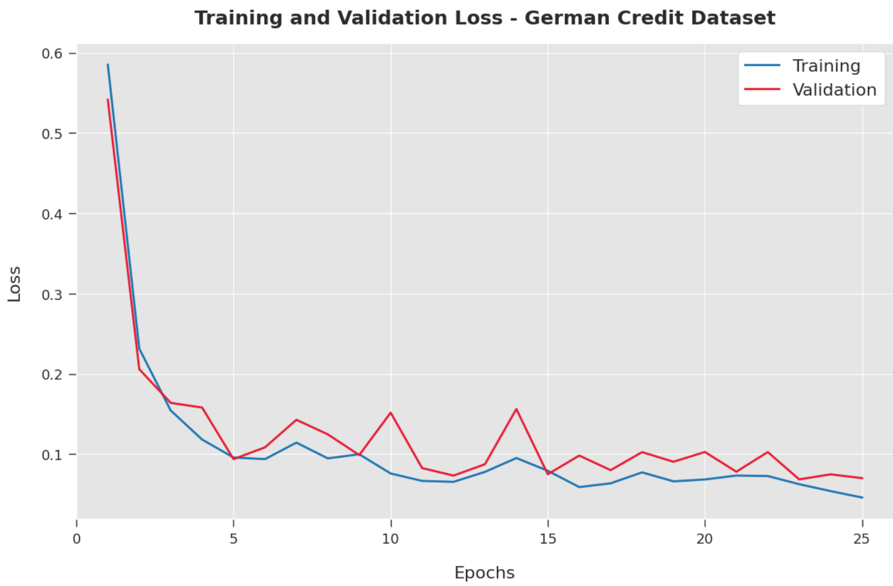

Fig. 6. TL and VL analysis of HSAODL-CCFC approach on German credit dataset

Fig. 7 provides a detailed accuracy analysis of the HSAODL-CCFC technique with recent systems on the test German Credit dataset. The figure indicated that the KNN-IGDFS, Decision Tree, MLP, RBFNetwork, Logistic Regression, KNN-GAW, ACO-DC, NB-GAW, NB-IGDFS, and SVM approches have obtained lower values of $accu_y$. Besides, SVM-GAW, GA-SVM, PSO-SVM, SVM-IGDFS, ABC-SVM, and NSGA II models have reported moderately closer values of $accu_y$. Though the SMOPSO, SSO-ANN, GRU, and BEPO-OGRU models have accomplished reasonably $accu_y$ values, the presented HSAODL-CCFC model has gained maximum $accu_y$ of 98%.

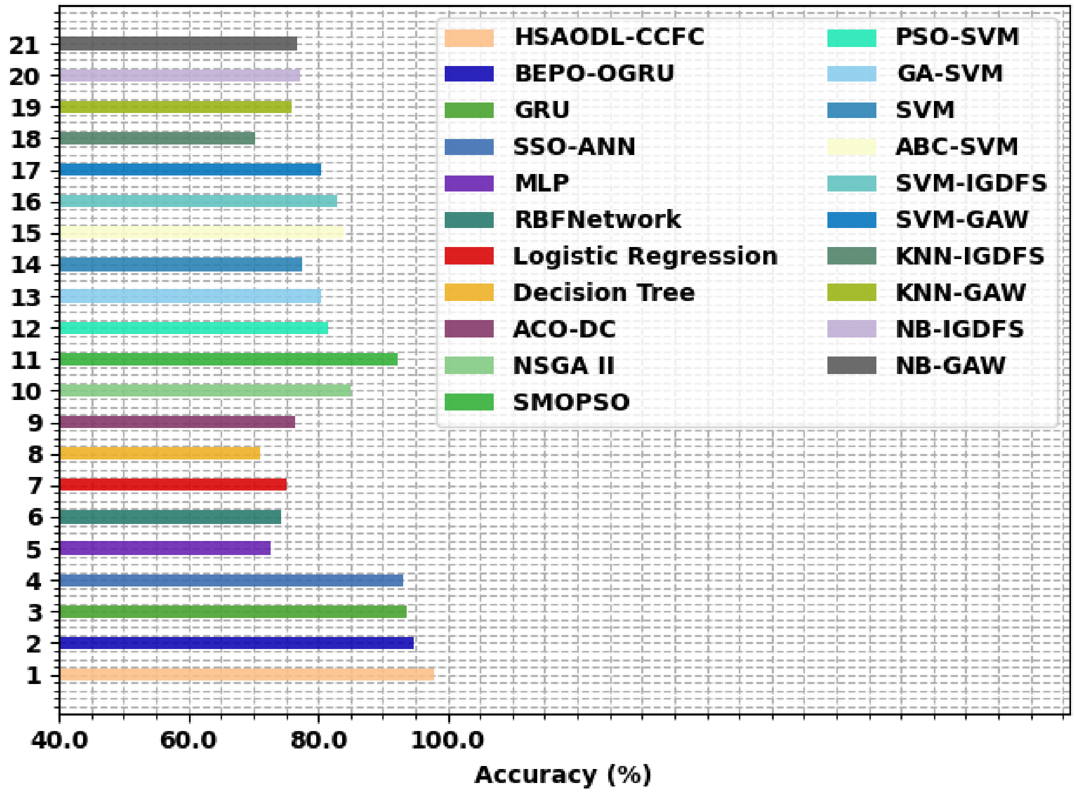

Fig. 7. $Accu_y$ analysis of HSAODL-CCFC approach on German Credit dataset

Fig. 8 offers a detailed $sens_y$ and $spec_y$ analysis of the HSAODL-CCFC system with existing methodologies on the test German Credit dataset. The figure revealed that the Decision Tree, ACO-DC, RBFNetwork, and Logistic Regression technqiues have reached lesser values of $sens_y$ and $spec_y$. At the same time, MLP, NSGA II, and SSO-ANN systems have reported moderately closer values of $sens_y$ and $spec_y$. Then, the SMOPSO, GRU, and BEPO-OGRU methodologies have accomplished reasonably $sens_y$ and $spec_y$ values, the projected HSAODL-CCFC model has reached maximal $sens_y$ and $spec_y$ of 98.10% and 98.10% correspondingly.

Fig. 9 demonstrates a detailed F_{score} and Kappa analysis of the HSAODL-CCFC system with recent models on the test German Credit dataset. The figure exposed that the Decision Tree, ACO-DC, RBFNetwork, and Logistic Regression models have reached minimal values of F_{score} and Kappa. Simultaneously, MLP and SSO-ANN models have reported moderately closer values of F_{score} and Kappa. Eventually, the GRU and BEPO-OGRU algorithms have accomplished reasonably F_{score} and Kappa values, the presented HSAODL-CCFC system has attained maximal F_{score} and Kappa of 97.66% and 95.33% correspondingly.

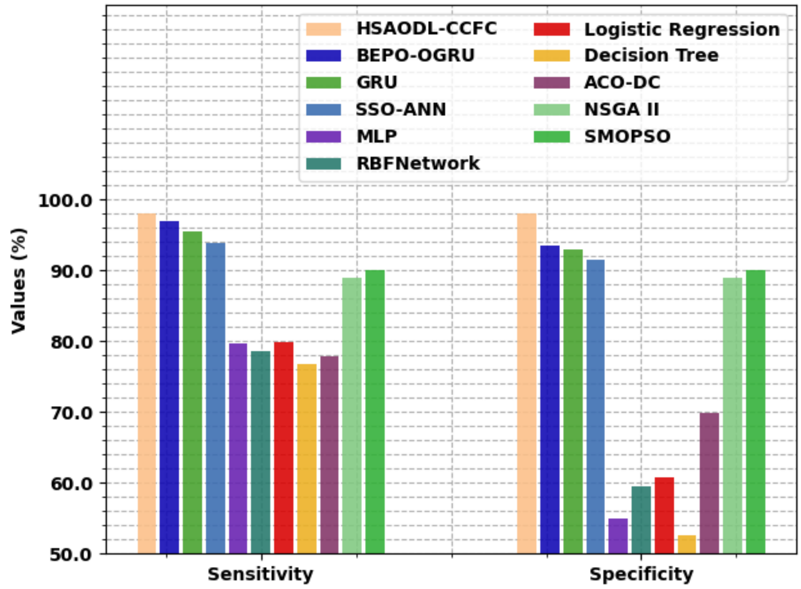

Fig. 8. $Sens_y$ and $Spec_y$ analysis of HSAODL-CCFC approach on German Credit dataset

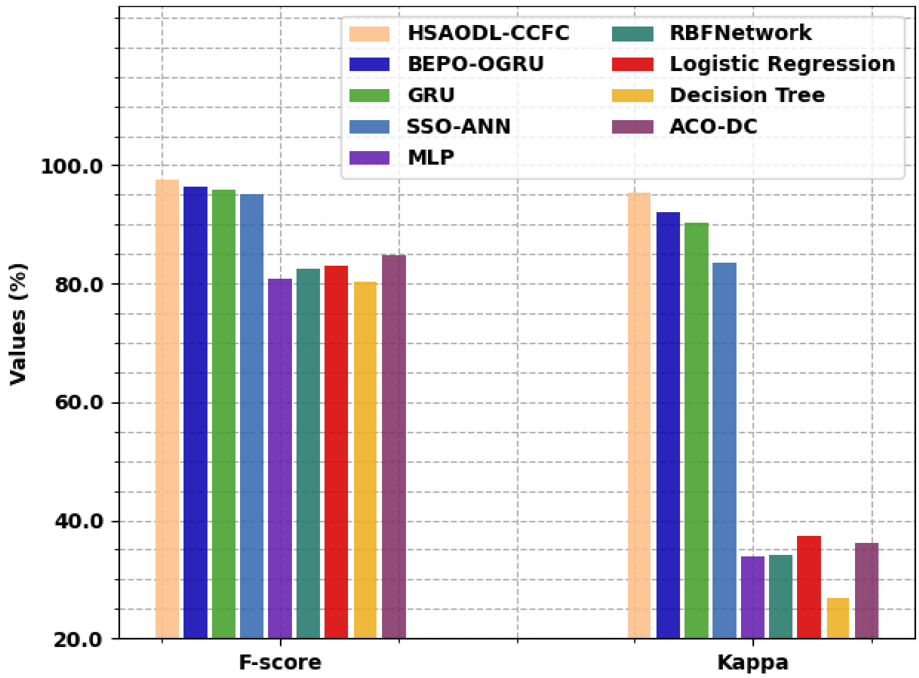

Fig. 9. F_{score} and kappa analysis of HSAODL-CCFC approach on German Credit dataset

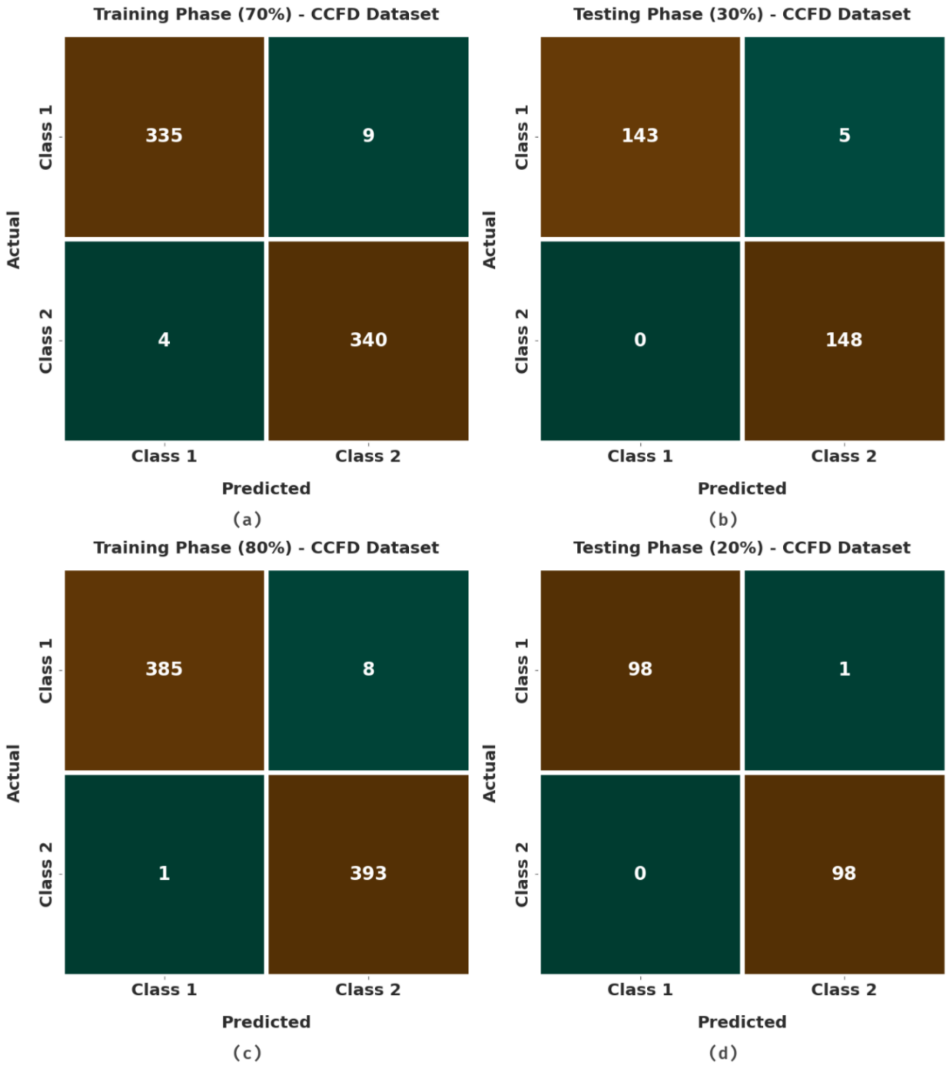

Fig. 10. Confusion matrices of HSAODL-CCFC technique on CCFD dataset (a) 70% of TR data, (b) 30% of TS data, (c) 80% of TR data, and (d) 20% of TS data

Fig. 10 gives the confusion matrices produced by the HSAODL-CCFC model on the applied CCFD dataset. With 70% of TR data, the HSAODL-CCFC approach has identified 335 samples under class 1 and 340 samples under class 2. Besides, with 30% of TS data, the HSAODL-CCFC model has identified 143 samples under class 1 and 148 samples under class 2. Additionally, with 80% of TR data, the HSAODL-CCFC system has identified 385 samples under class 1 and 393 samples under class 2.

Table 4 and Fig. 11 offer a detailed outcome analysis of the HSAODL-CCFC technique on CCFD dataset. The outcomes exposed that the HSAODL-CCFC system has reached effectual outcomes in all aspects. For instance, on 70% of TR data, the HSAODL-CCFC methodology has accessible average $accu_y$ of 98.11%, $sens_y$ of 98.11%, $spec_y$ of 98.11%, F_{score} of 98.11%, and kappa of 96.22%. Also, on 80% of TR data, the HSAODL-CCFC model has offered average $accu_y$ of 98.86%, $sens_y$ of 98.86%, $spec_y$ of 98.86%, F_{score} of 98.86%, and kappa of 97.71%. At last, on 20% of TS data, the HSAODL-CCFC approach has obtainable average $accu_y$ of 98%, $sens_y$ of 99.49%, $spec_y$ of 99.49%, F_{score} of 99.49%, and kappa of 98.98%.

Table 4 Result analysis of HSAODL-CCFC approach on CCFD dataset

Labels	Accuracy	Sensitivity	Specificity	F-Score	Kappa Score
Training Phase (70%)					
Class 1	98.11	97.38	98.84	98.1	-
Class 2	98.11	98.84	97.38	98.12	-
Average	**98.11**	**98.11**	**98.11**	**98.11**	**96.22**
Testing Phase (30%)					
Class 1	98.31	96.62	100.00	98.28	-
Class 2	98.31	100.00	96.62	98.34	-
Average	**98.31**	**98.31**	**98.31**	**98.31**	**96.62**
Training Phase (80%)					
Class 1	98.86	97.96	99.75	98.84	-
Class 2	98.86	99.75	97.96	98.87	-
Average	**98.86**	**98.86**	**98.86**	**98.86**	**97.71**
Testing Phase (20%)					
Class 1	99.49	98.99	100.00	99.49	-
Class 2	99.49	100.00	98.99	99.49	-
Average	**99.49**	**99.49**	**99.49**	**99.49**	**98.98**

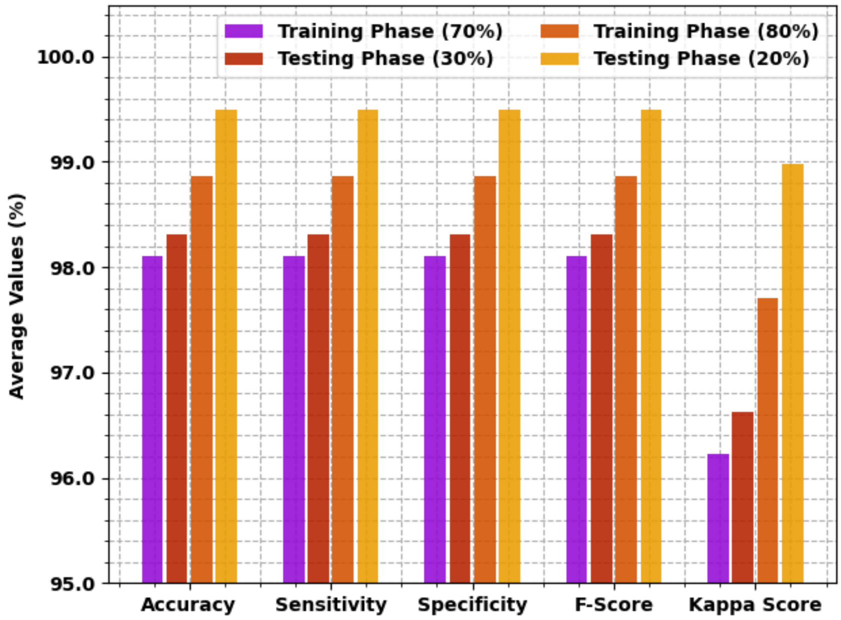

Fig. 11. Result analysis of HSAODL-CCFC approach on CCFD dataset

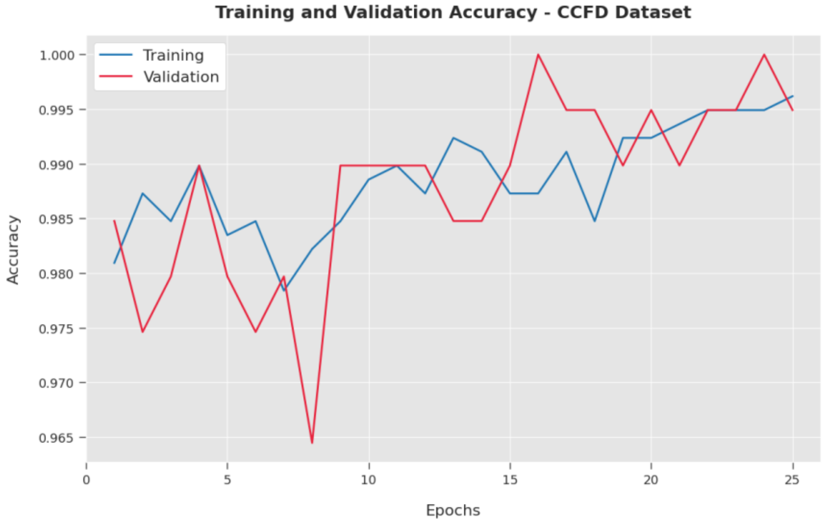

Fig. 12. TA and VA analysis of HSAODL-CCFC approach on CCFD dataset

The TA and VA attained by the HSAODL-CCFC approach on CCFD dataset are established in Fig. 12. The experimental outcome implied that the HSAODL-CCFC methodology has gained higher values of TA and VA. In specific, the VA seemed that superior to TA.

The TL and VL achieved by the HSAODL-CCFC approach on CCFD dataset in Fig. 13. The experimental outcome exposed that the HSAODL-CCFC system has been able least values of TL and VL. In specific, the VL seemed that lower than TL.

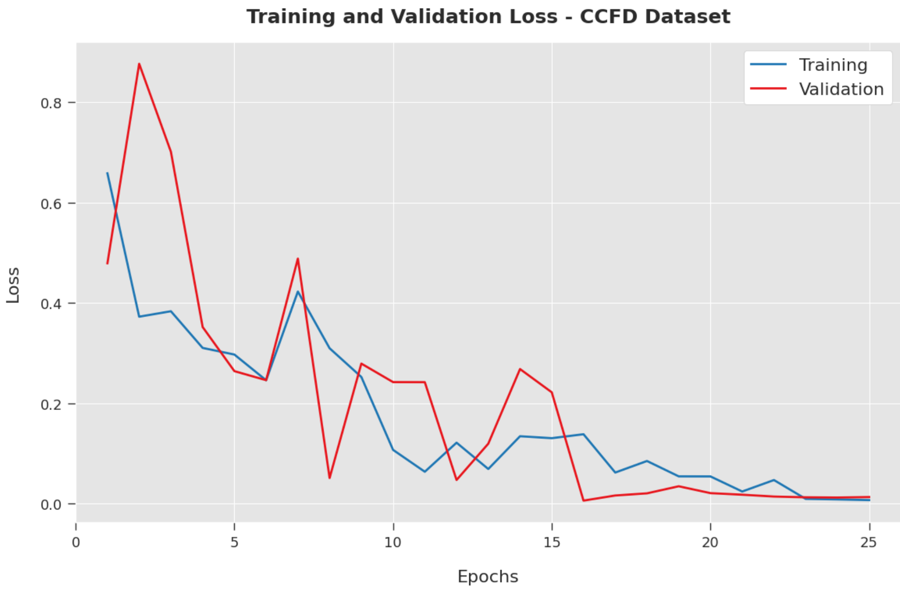

Fig. 13. TL and VL analysis of HSAODL-CCFC approach on CCFD dataset

Fig. 14 demonstrated a detailed accuracy analysis of the HSAODL-CCFC approach with recent systems on the test CCFD dataset [21-24]. The figure implied that the Decision Tree, Logistic Regression, RBFNetwork, and MLP approaches have obtained minimal values of $accu_y$. Concurrently, Naïve Bayes, Random forest, and SSO-ANN methodologies have reported moderately closer values of $accu_y$. Finally, the GRU and BEPO-OGRU techniques have accomplished reasonably $accu_y$ values, the presented HSAODL-CCFC approach has gained maximal $accu_y$ of 99.49%.

Fig. 15 offers a detailed $sens_y$ and $spec_y$ analysis of the HSAODL-CCFC method with recent methodologies on the test CCFD dataset. The figure indicated that the Naïve Bayes and Random Forest models have obtained minimal values of $sens_y$. Likewise, the Decision, Logistic Regression, RBFNetwork, and MLP techniques have reported moderately closer values of $sens_y$ and $spec_y$. Though the SSO-ANN, GRU, and BEPO-OGRU approaches have accomplished reasonably $sens_y$ and $spec_y$ values, the projected HSAODL-CCFC system has gained higher $sens_y$ and $spec_y$ of 99.49% and 99.49% correspondingly.

Fig. 16 illustrates a detailed F_{score} and Kappa analysis of the HSAODL-CCFC approach with recent models on the test CCFD dataset. The figure indicated that the Naïve Bayes and Random Forest models have attianed minimal values of F_{score}. Besides, Decision, Logistic Regression, RBFNetwork, and MLP algorithms have reported moderately closer values of F_{score} and Kappa. Though the SSO-ANN, GRU, and BEPO-OGRU systems have accomplished reasonably F_{score} and Kappa values, the presented HSAODL-CCFC approach has gained superior F_{score} and Kappa of 99.49% and 98.98% correspondingly.

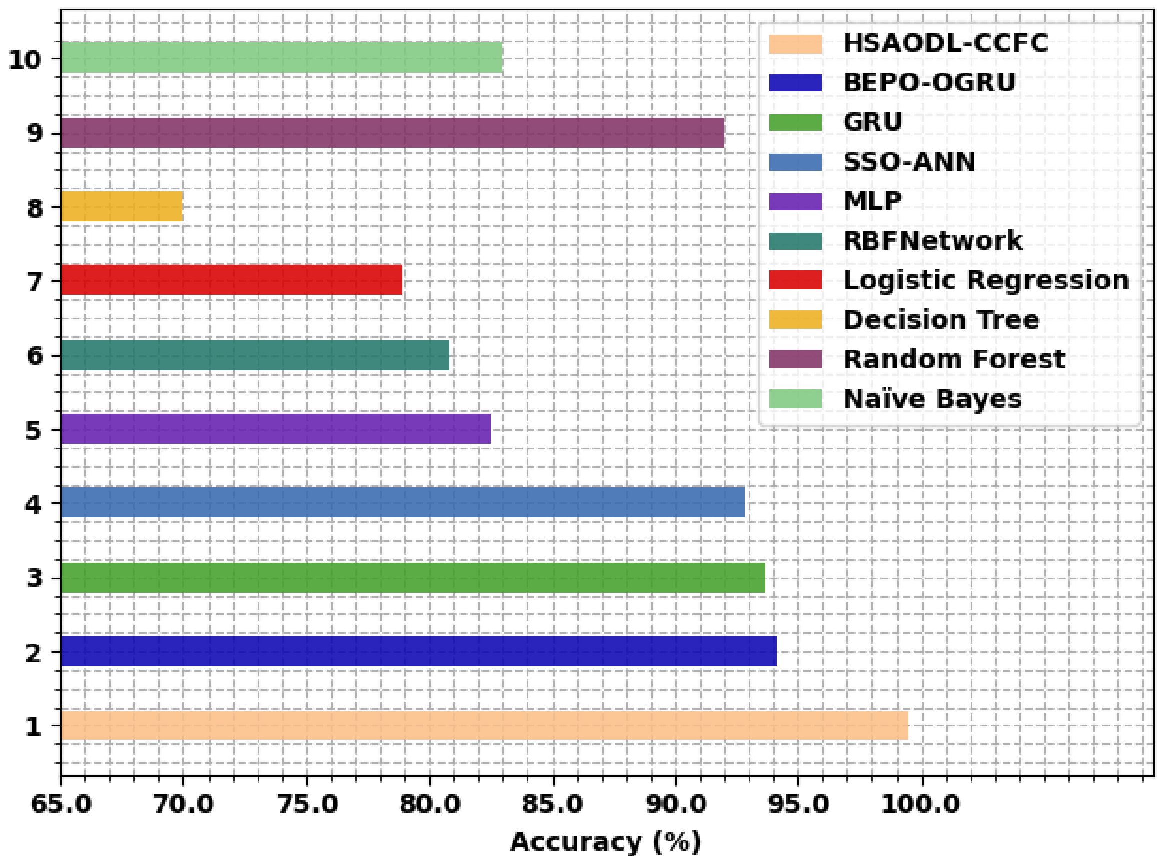

Fig. 14. $Accu_y$ analysis of HSAODL-CCFC approach under CCFD dataset

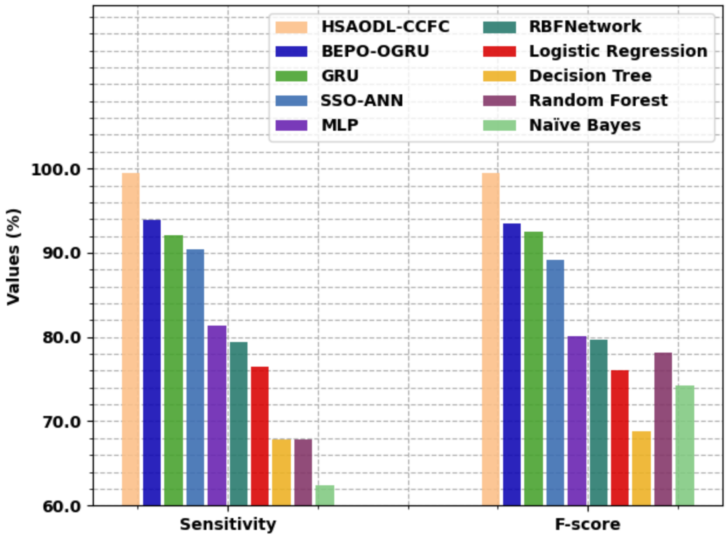

Fig. 15. $Sens_y$ and $Spec_y$ analysis of HSAODL-CCFC approach under CCFD dataset

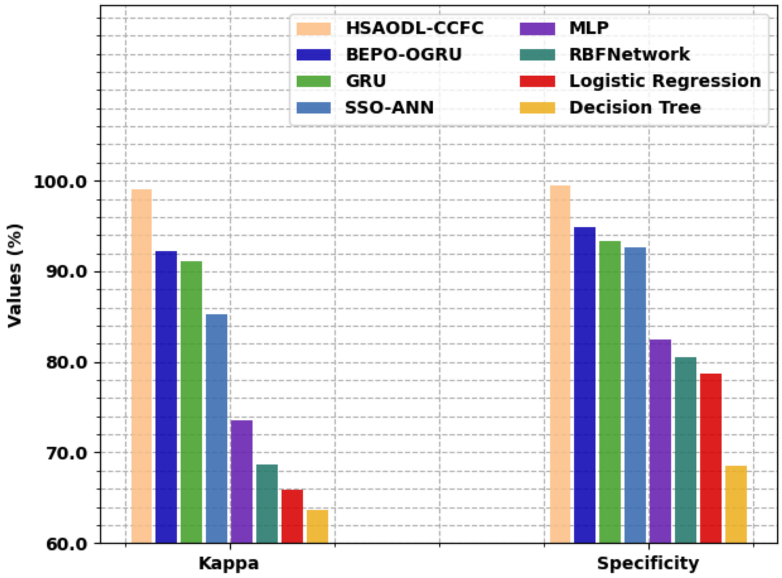

Fig. 16. F_{score} and kappa analysis of HSAODL-CCFC approach under CCFD dataset

From the comprehensive result analysis, it can be evident that the proposed approach has accomplished improved performance over the other existing models.

4. Conclusion

In this article, an effective HSAODL-CCFC method has been introduced to properly distinguish credit card transactions from legitimate or fraud. The presented HSAODL-CCFC model primarily applied data pre-processing at the initial stage. Then, the BHSA-FS technique is applied to elect feature subsets. Next, the RMSProp with DBN system was utilized for the identification and classification of CCFs. The simulation analysis of the HSAODL-CCFC model is tested using 2 benchmark datasets. The comparative analysis reported the promising efficiency of HSAODL-CCFC model on recent approaches. Thus, the presented HSAODL-CCFC model can be exploited as an effectual tool for CCFD. In future, an ensemble of DL based classification models can be designed to attain enhanced detection efficiency.

References

[1] Sailusha, R., Gnaneswar, V., Ramesh, R. and Rao, G.R., 2020, May. Credit card fraud detection using machine learning. In *2020 4th International Conference on Intelligent Computing and Control Systems (ICICCS)* (pp. 1264-1270). IEEE.

[2] Shirgave, S., Awati, C., More, R. and Patil, S., 2019. A Review On Credit Card Fraud Detection Using Machine Learning. *Int. J. Sci. Technol. Res*, *8*, pp.1217-1220.

[3] Awoyemi, J.O., Adetunmbi, A.O. and Oluwadare, S.A., 2017, October. Credit card fraud detection using machine learning techniques: A comparative analysis. In *2017 international conference on computing networking and informatics (ICCNI)* (pp. 1-9). IEEE.

[4] Thennakoon, A., Bhagyani, C., Premadasa, S., Mihiranga, S. and Kuruwitaarachchi, N., 2019, January. Real-time credit card fraud detection using machine learning. In *2019 9th International Conference on Cloud Computing, Data Science & Engineering (Confluence)* (pp. 488-493). IEEE.

[5] Maniraj, S.P., Saini, A., Ahmed, S. and Sarkar, S., 2019. Credit card fraud detection using machine learning and data science. *International Journal of Engineering Research and*, *8*(09).

[6] Shukur, H.A. and Kurnaz, S., 2019. Credit card fraud detection using machine learning methodology. *International Journal of Computer Science and Mobile Computing*, *8*(3), pp.257-260.

[7] Varmedja, D., Karanovic, M., Sladojevic, S., Arsenovic, M. and Anderla, A., 2019, March. Credit card fraud detection-machine learning methods. In *2019 18th International Symposium INFOTEH-JAHORINA (INFOTEH)* (pp. 1-5). IEEE.

[8] Lucas, Y. and Jurgovsky, J., 2020. Credit card fraud detection using machine learning: A survey. *arXiv preprint arXiv:2010.06479*.

[9] Adepoju, O., Wosowei, J. and Jaiman, H., 2019, October. Comparative evaluation of credit card fraud detection using machine learning techniques. In *2019 Global Conference for Advancement in Technology (GCAT)* (pp. 1-6). IEEE.

[10] Khatri, S., Arora, A. and Agrawal, A.P., 2020, January. Supervised machine learning algorithms for credit card fraud detection: a comparison. In *2020 10th International Conference on Cloud Computing, Data Science & Engineering (Confluence)* (pp. 680-683). IEEE.

[11] Popat, R.R. and Chaudhary, J., 2018, May. A survey on credit card fraud detection using machine learning. In *2018 2nd international conference on trends in electronics and informatics (ICOEI)* (pp. 1120-1125). IEEE

[12] Dornadula, V.N. and Geetha, S., 2019. Credit card fraud detection using machine learning algorithms. *Procedia computer science*, *165*, pp.631-641

[13] K. Randhawa, C. K. Loo, M. Seera, C. P. Lim and A. K. Nandi, "Credit Card Fraud Detection Using AdaBoost and Majority Voting," in *IEEE Access*, vol. 6, pp. 14277-14284, 2018, doi: 10.1109/ACCESS.2018.2806420

[14] Asha, R.B. and KR, S.K., 2021. Credit card fraud detection using artificial neural network. *Global Transitions Proceedings*, *2*(1), pp.35-41

[15] M. Azhan and S. Meraj, "Credit Card Fraud Detection using Machine Learning and Deep Learning Techniques," *2020 3rd International Conference on Intelligent Sustainable Systems (ICISS)*, 2020, pp. 514-518, doi: 10.1109/ICISS49785.2020.9316002

[16] Yang, Y., Wu, Y., Yuan, H., Khishe, M. and Mohammadi, M., 2022. Nodes clustering and multi-hop routing protocol optimization using hybrid chimp optimization and hunger games search algorithms for sustainable energy efficient underwater wireless sensor networks. *Sustainable Computing: Informatics and Systems*, *35*, p.100731.

[17] Chen, Z. and Li, W., 2017. Multisensor feature fusion for bearing fault diagnosis using sparse autoencoder and deep belief network. *IEEE Transactions on Instrumentation and Measurement*, *66*(7), pp.1693-1702.

[18] Li, D., Chen, C., Lv, Q., Gu, H., Lu, T., Shang, L., Gu, N. and Chu, S.M., 2018, April. AdaError: An adaptive learning rate method for matrix approximation-based collaborative filtering. In *Proceedings of the 2018 World Wide Web Conference* (pp. 741-751).

[19] https://www.kaggle.com/mlg-ulb/creditcardfraud

[20] https://archive.ics.uci.edu/ml/datasets/statlog+(german+credit+data)

[21] Arun, G.K. and Venkatachalapathy, K., 2020. Intelligent feature selection with social spider optimization based artificial neural network model for credit card fraud detection. *IIOABJ*, *11*(2), pp.85-91

[22] Rajesh, P., & Karthikeyan, M. (2019). Prediction of Agriculture Growth and Level of Concentration in Paddy—A Stochastic Data Mining Approach. In *Advances in Big Data and Cloud Computing* (pp. 127-139). Springer, Singapore

[23] Rajesh, P., Karthikeyan, M., & Arulpavai, R. (2019, December). Data mining approaches to predict the factors that affect the groundwater level using stochastic model. In *AIP Conference Proceedings* (Vol. 2177, No. 1, p. 020079). AIP Publishing LLC

[24] Rajesh, P. (2020). Developing Data Mining Based Stochastic Model for Investigate the Parameters for Increasing Wheat and Groundnut Productions. *Adalya Journal*, *9*, 319-327.