

## Article

# Chicken Swarm-Based Feature Subset Selection with Optimal Machine Learning Enabled Data Mining Approach

Monia Hamdi <sup>1</sup>, Inès Hilali-Jaghdam <sup>2</sup>, Manal M. Khayyat <sup>3</sup>, Bushra M. E. Elnaim <sup>4</sup>, Sayed Abdel-Khalek <sup>5,6</sup> and Romany F. Mansour <sup>7,\*</sup>

<sup>1</sup> Department of Information Technology, College of Computer and Information Sciences, Princess Nourah Bint Abdulrahman University, P.O. Box 84428, Riyadh 11671, Saudi Arabia; mshamdi@pnu.edu.sa

<sup>2</sup> Department of Computer Sciences and Information Technology, Applied College, Princess Nourah Bint Abdulrahman University, P.O. Box 84428, Riyadh 11671, Saudi Arabia; imalihilali@pnu.edu.sa

<sup>3</sup> Department of Information Systems, College of Computers and Information Systems, Umm Al-Qura University, P.O. Box 7607, Makkah 24382, Saudi Arabia; mmkhayat@uqu.edu.sa

<sup>4</sup> Department of Computer Science, College of Science and Humanities in Al-Sulail, Prince Sattam Bin Abdulaziz University, P.O. Box 173, Al-Kharj 16278, Saudi Arabia; b.elamin@psau.edu.sa

<sup>5</sup> Department of Mathematics, Faculty of Science, Sohag University, Sohag 82524, Egypt; sayedquantum@yahoo.co.uk or sabotalb@tu.edu.sa

<sup>6</sup> Department of Mathematics, College of Science, Taif University, P.O. Box 11099, Taif 21944, Saudi Arabia

<sup>7</sup> Department of Mathematics, Faculty of Science, New Valley University, El-Kharga 72511, Egypt

\* Correspondence: romanyf@sci.nvu.edu.eg



**Citation:** Hamdi, M.; Hilali-Jaghdam, I.; Khayyat, M.M.; Elnaim, B.M.E.; Abdel-Khalek, S.; Mansour, R.F. Chicken Swarm-Based Feature Subset Selection with Optimal Machine Learning Enabled Data Mining Approach. *Appl. Sci.* **2022**, *12*, 6787. <https://doi.org/10.3390/app12136787>

Academic Editors: Stefanos Ougiaroglou and Dionisis Margaritis

Received: 9 April 2022

Accepted: 29 June 2022

Published: 4 July 2022

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

**Abstract:** Data mining (DM) involves the process of identifying patterns, correlation, and anomalies existing in massive datasets. The applicability of DM includes several areas such as education, healthcare, business, and finance. Educational Data Mining (EDM) is an interdisciplinary domain which focuses on the applicability of DM, machine learning (ML), and statistical approaches for pattern recognition in massive quantities of educational data. This type of data suffers from the curse of dimensionality problems. Thus, feature selection (FS) approaches become essential. This study designs a Feature Subset Selection with an optimal machine learning model for Educational Data Mining (FSSML-EDM). The proposed method involves three major processes. At the initial stage, the presented FSSML-EDM model uses the Chicken Swarm Optimization-based Feature Selection (CSO-FS) technique for electing feature subsets. Next, an extreme learning machine (ELM) classifier is employed for the classification of educational data. Finally, the Artificial Hummingbird (AHB) algorithm is utilized for adjusting the parameters involved in the ELM model. The performance study revealed that FSSML-EDM model achieves better results compared with other models under several dimensions.

**Keywords:** feature subset selection; data mining; educational data mining; artificial intelligence; machine learning; metaheuristics

## 1. Introduction

Data mining (DM) is the procedure of understanding data through cleaning raw data, discovering patterns, producing models, and testing the models. It comprises of several fields such as statistics, machine learning (ML), and database systems. Education DM (EDM) is an emergent field with an arising strategy to investigate the various types of data which are obtained from an education background [1]. It is an interdisciplinary area which inspects data mining (DM), man-fabricated consciousness, and measurable demonstrating with the data produced using an academic organization [2]. EDM uses a calculation method for taking care of elucidating academic data considering a definitive point of examining academic enquiries. To make a nation stand out among different nations across the globe, education frameworks should encounter an essential advancement by re-planning their design. The concealed data and examples from various data sources can be extricated

by adjusting the strategies for DM. For summing up the outcomes of students with their qualifications, they investigate the abuse of DM in the academic fields. Crude data can be altogether moved through DM models. The data achieved from an education organization go through examination of various DM strategies [3]. The strategy identifies the conditions wherein students can strive to have a positive impact [4].

Student performance prediction (SPP) has different definitions according to troublesome perspectives; however, the measured assessment assumes a significant part in current education establishments. SPP seems effective in aiding all partners in the education interaction. For students, SPP can assist them with picking reasonable courses or activities and make their arrangements for academic durations [5]. For educators, SPP can assist with changing learning material and presenting programs compatible with the students' capacity, and help identify struggling students. For education chiefs, SPP can assist with checking the education program and enhancing the course framework. Generally, partners in the education advancement have intentions to further develop the education outcome. Moreover, the data-driven SPP study provides a goal of reference for the education framework. Weka, a compelling DM technique was utilized to produce the outcome [6].

The increment of educational information from distinct sources has resulted in desperation for the EDM research [7]. This can help to further objectives and characterize specific goals of education and highlight subset determinations by disposing of the component that is repetitive/is not important. The set of components chosen should follow the Occam Razor rule to provide the best outcome in light of the goal [8]. The data size to be dealt with has expanded in the past five years; hence, the choice is turning into a necessity before any sort of arrangement happens. It is not quite the same as the element extraction strategy because the determination method does not change the first portrayal of the data [9]. The least complex method includes choice, where how much quality in an examination is diminished by choosing just the main view of the circumstances such as a more elevated level of exercises [10].

Since the high dimensionality raises computational costs, it is essential to define a way to reduce the number of considered features. Feature selection (FS) allows reducing a high dimensionality problem and selecting a suitable number of features. This study designed a Feature Subset Selection with optimal machine learning for the Educational Data Mining (FSSML-EDM) model. The proposed FSSML-EDM model involves the Chicken Swarm Optimization-based Feature Selection (CSO-FS) technique for electing feature subsets. Next, the extreme learning machine (ELM) classifier was employed for the classification of educational data. Finally, the Artificial Hummingbird (AHB) algorithm was utilized for adjusting the parameters involved in the ELM model. The performance study revealed the effectual outcomes of the FSSML-EDM model over the compared models under several dimensions. Our contributions are summarized as follows:

- We propose a model comprising data preprocessing, CSO-FS, ELM classification, and AHB parameter;
- We designed a new CSO-FS technique to reduce the curse of the dimensionality problem and enhanced the classification performance;
- We employed the ELM classification model with the AHB-based parameter optimization technique for the EDM process;
- We validated the performance of the FSSML-EDM model using the benchmark dataset from the UCI repository.

## 2. Literature Review

Injadat et al. [11] explored and analyzed two distinct datasets at two distinct phases of course delivery (20% and 50%) utilizing several graphical, statistical, and quantitative approaches. The feature analysis offers understanding as to the nature of distinct features regarded and utilizes in the selection of ML techniques and their parameters. Moreover, this work presents a systematic model dependent upon the Gini index and *p*-value for selecting an appropriate ensemble learner in a group of six potential ML techniques. Ashraf et al. [12]

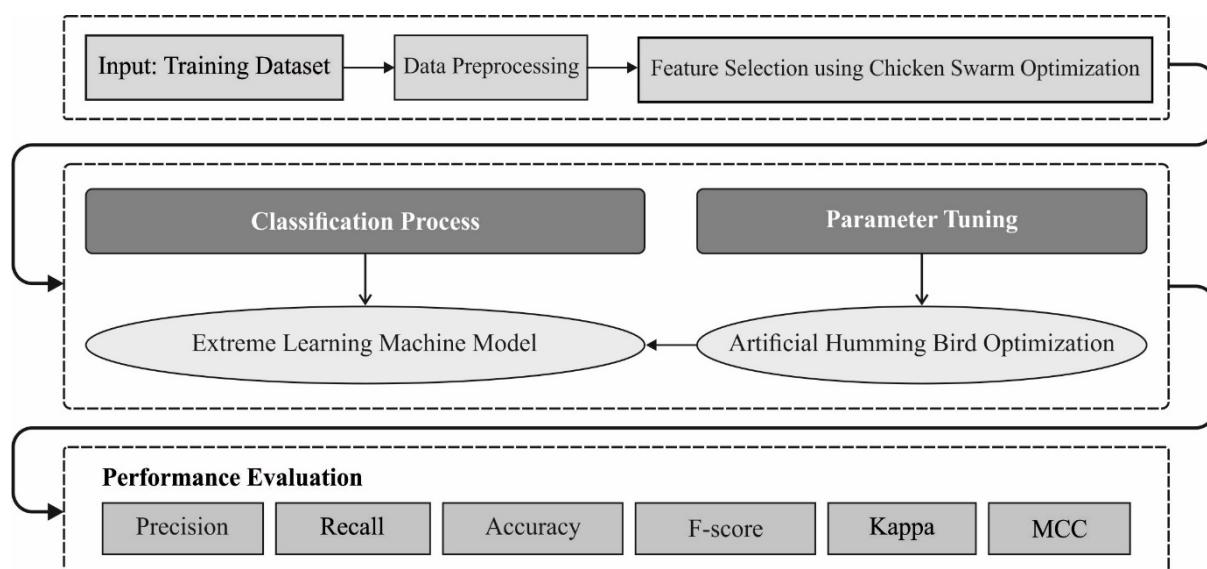
progressed an accurate prediction pedagogical method, considering the pronounced nature and novelty of presented approach in Educational Data Mining. The base classifications containing RT, j48, kNN, and naïve Bayes (NB) were estimated on a 10-fold cross-validation model. In addition, the filter procedure as over-sampling (SMOTE) and under-sampling (Spread subsampling) were exploited to examine some important alterations in outcomes amongst meta and base classifications.

Dabhade et al. [13] forecasted student academic performance in a technical institution in India. The data pre-processed and factor analysis were executed on the attained dataset for removing the anomaly from the data, decreasing the dimensionality of the data, and attaining the most correlated feature. Nahar et al. [14] generated two datasets concentrating on two distinct angles. In the primary dataset classification and forecast, the type of students (bad, medium, and good) on a particular course was dependent upon its prerequisite course efficiency. This can be executed during the artificial intelligence (AI) course. The secondary dataset also classified and forecasted the last grade (A, B, C) of an arbitrary subject; our data can be established in such a way that the data are only concentrated on the efficiency of the midterm exam.

Despite all the studies performed on FS process, to the best of our knowledge, only few works have carried out an FS-based classification model for EDM. Earlier works have used ML models for EDM without contributing much significance to the FS process. At the same time, the parameters involved in the ML models (i.e., ELM) considerably affect the overall classification performance. Since the trial-and-error method for parameter tuning is a tedious and erroneous process, metaheuristic algorithms can be applied. Therefore, metaheuristic optimization algorithms can be designed to optimally tune the parameters related to the ML models to improve the overall classification performance.

### 3. The Proposed Model

In this study, a new FSSML-EDM technique was developed for mining educational data. The proposed FSSML-EDM model involves data preprocessing at the initial stage to transform the input data into a compatible format. Then, the preprocessed data are passed into the CSO-FS technique for electing feature subsets. Next, the ELM classifier can be employed for the effective identification and classification of educational data. Finally, the AHB algorithm is utilized for effectively adjusting the parameters involved in the ELM model. The outcome of the ELM model is the classification output. Figure 1 depicts the block diagram of FSSML-EDM technique.



**Figure 1.** Block diagram of FSSML-EDM technique.

### 3.1. Process Involved in CSO-FS Technique

At the initial stage, the presented FSSML-EDM model incorporates the design of the CSO-FS technique for electing feature subsets. The CSO algorithm is chosen over other optimization algorithms due to its simplicity and high parallelism. The CSO simulates the chicken movement and the performance of the chicken swarm; the CSO is explained as follows: CSO has several groups, and all the groups have a dominant rooster, some hens, and chicks [15]. The rooster, hen, and chick from the group are found dependent upon their fitness value. The rooster (group head) is the chicken which is an optimum fitness value. However, the chick is the chicken which has the worse fitness value. The majority of chickens are hens and it can be selected arbitrarily to stay in that group. The dominance connection and mother-child connection from the group remain unaltered and upgrade during ( $G$ ) time steps. The movement of chickens are expressed under the equation which utilizes to the rooster place upgrade provided by Equation (1):

$$X_{ij}^{r+1} = X_{ij}^t \times \left(1 + \text{randn}(0, \sigma^2)\right) \quad (1)$$

whereas:

$$\sigma^2 = \begin{cases} 1 & \text{if } f_i \leq f_k \\ \exp\left(\frac{f_k - f_i}{|f_i| + \varepsilon}\right) & \text{Otherwise} \end{cases}$$

In which  $k \in [1, N_r]$ ,  $k \neq i$ , and  $N_r$  refers the amount of chosen roosters.  $X_{ij}$  signifies the place of rooster number  $i$  in  $j$ th dimensional under  $t$  and  $t + 1$  iteration,  $\text{randn}(0, \sigma^2)$  is utilized for generating Gaussian arbitrary numbers with mean 0 and variance  $\sigma^2$ ;  $\varepsilon$  refers to the constant with minimum value; and  $f_i$  is the fitness value to the equivalent rooster  $i$ . The equation which utilizes the hen place upgrade is provided by Equations (2)–(4):

$$X_{ij}^{t+1} = X_{ij}^t + S_1 \text{randn}\left(X_{r1,j}^t - X_{ij}^t\right) + S_2 \text{randn}\left(X_{r2,j}^t - X_{ij}^t\right) \quad (2)$$

In which:

$$S_1 = \exp\left(\frac{f_i - f_{r1}}{|f_i| + \varepsilon}\right) \quad (3)$$

and:

$$S_2 = \exp(f_{r2} - f_i) \quad (4)$$

where  $r1, r \in [1, \dots, N]$ ,  $r1 \neq rr$  refers to the index of the rooster, but  $r2$  implies the chicken in the swarm which is a rooster or hen and a uniform arbitrary number is created by  $\text{randn}$ . Finally, the equation that utilizes the chick place upgrade is provided by Equation (5):

$$X_{ij}^{t+1} = X_{ij}^t + FL\left(X_{m,j}^t - X_{ij}^t\right), \quad FL \in [0, 2] \quad (5)$$

where  $X_{m,j}^t$  signifies the place of  $i$ th chick mother.

### 3.2. ELM Based Classification

At this stage, the ELM classifier can be employed for the effective identification and classification of educational data. The ELM model has  $n$  input layers,  $l$  hidden layers, and  $m$  output layers. Initially, considering the training instance  $\{X, Y\} = \{x_i, y_i\} (i = 1, 2, \dots, Q)$ , and it is comprised of the input feature  $X = [x_{i1} x_{i2} \dots x_{iQ}]$  and matrix  $Y = [y_{j1} y_{j2} \dots y_{jQ}]$  with training instance, where the matrix  $X$  and  $Y$  are expressed by [16]:

$$X = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1Q} \\ x_{21} & x_{22} & \cdots & x_{2Q} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{nQ} \end{bmatrix}, \quad (6)$$

$$Y = \begin{bmatrix} y_{11} & y_{12} & \cdots & y_{mQ} \\ y_{21} & y_{22} & \cdots & y_{mQ} \\ \vdots & \vdots & \ddots & \vdots \\ y_{m1} & y_{m2} & \cdots & y_{mQ} \end{bmatrix},$$

where  $n$  and  $m$  parameters denote the dimension of input and output matrix. Next, ELM set the weights amongst input and hidden layers randomly:

$$w = \begin{bmatrix} w_{11} & w_{12} & \cdots & w_{1n} \\ w_{21} & w_{22} & \cdots & w_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ w_{l1} & w_{l2} & \cdots & w_{ln} \end{bmatrix}, \quad (7)$$

where  $w_{ij}$  denotes the weight from  $i$ th hidden and  $j$ th input layers. Figure 2 showcases the framework of ELM. Then, ELM considers the weight from output and hidden layers that can be shown below:

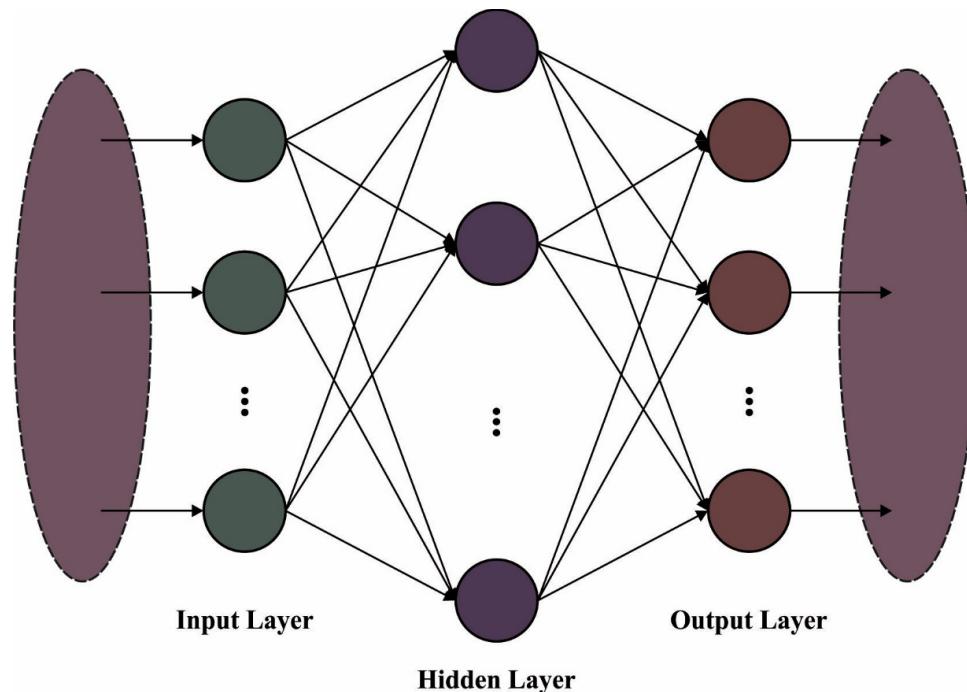
$$\beta = \begin{bmatrix} \beta_{11} & \beta_{12} & \cdots & \beta_{1m} \\ \beta_{21} & \beta_{22} & \cdots & \beta_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ \beta_{l1} & \beta_{l2} & \cdots & \beta_{lm} \end{bmatrix}, \quad (8)$$

where  $\beta_{jk}$  indicates the weight from  $j$ th hidden and  $i$ th output layers. Next, ELM set a bias of hidden layers randomly:

$$B = [b_1 b_2 \cdots b_n]^T. \quad (9)$$

After that, ELM chooses the network activation function. According to, output matrix  $T$  is characterized by:

$$T = [t_1, t_2, \dots, t_Q]_{m \times Q}. \quad (10)$$



**Figure 2.** Framework of ELM.

The column vector of output matrix  $T$  is shown in the following:

$$t_j = \begin{bmatrix} t_{1j} \\ t_{2j} \\ \vdots \\ t_{mj} \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^l \beta_{i1}g(w_i x_j + b_i) \\ \sum_{i=1}^l \beta_{i2}g(w_i x_j + b_i) \\ \vdots \\ \sum_{i=1}^l \beta_{im}g(w_i x_j + b_i) \end{bmatrix} (j = 1, 2, 3, \dots, Q). \quad (11)$$

Moreover, considering Equations (10) and (11), it is modelled in the following:

$$H\beta = T', \quad (12)$$

where  $T'$  indicates transpose of  $T$  and  $H$  represents simulation outcomes of hidden neurons. To achieve better solution with less error, the least square model is employed for determining the weight matrix measure of  $\beta$  [17,18].

$$\beta = H^+ T'. \quad (13)$$

To enhance the normalization ability of the system and provide stable outcomes, the regularization parameter of  $\beta$  is used. The amount of hidden layers is minimal in contrast with the amount of training samples,  $\beta$  is shown in the following [19]:

$$\beta = \left( \frac{I}{\lambda} + H^T H \right)^{-1} H^T T'. \quad (14)$$

When the amount of hidden layers is maximal than the amount of training samples,  $\beta$  is denoted as follows [20]:

$$\beta = H^T \left( \frac{I}{\lambda} + H H^T \right)^{-1} T'. \quad (15)$$

### 3.3. AHB Based Parameter Optimization

Finally, the AHB algorithm is utilized for effectively adjusting the parameters involved in the ELM model with the goal of attaining maximum classification performance [21]. The AHB algorithm is an optimization approach stimulated from the foraging and flight of hummingbirds. The three major models are provided as follows: in a guided foraging model, three flight behaviors are utilized in foraging (axial, diagonal, and omnidirectional flight). It can be defined as follows:

$$v_i(t+1) = x_{i,ta}(t) + h \cdot b \cdot (x_i(t) - X_{i,ta}(t))h \sim N(0, 1) \quad (16)$$

where  $x_{i,ta}(t)$  characterizes the location of the targeted food source,  $h$  signifies the guiding factor, and  $x_i(t)$  represents the location of  $i$ th food source at time  $t$ . The location updating of the  $i$ th food source is provided by:

$$x_{Ai}(t) = \begin{cases} x_i(t) & f(x_i(t)) \leq f(v_i(t+1)) \\ v_i(t+1) & f(x_i(t)) > f(v_i(t+1)) \end{cases} \quad (17)$$

where  $f(x_i(t))$  and  $f(v_i(t+1))$  denote the value of function fitness for  $x_i(t)$  and  $v_i(t+1)$ . the local search of hummingbirds in the territorial foraging strategy is provided in the following:

$$v_i(t+1) = x_{i(t)} + g \cdot b \cdot (\chi_{i(t)})g \sim N(0, 1) \quad (18)$$

where  $g$  represents the territorial factor. The arithmetical formula for the migration foraging of hummingbirds is provided by:

$$x_{wor}(t+1) = lb + r \cdot (ub - lb) \quad (19)$$

where  $x_{wor}$  indicates the source of food with worst population rate of nectar refilling,  $r$  represents a random factor, and  $ub$  and  $lb$  denote the upper and lower limits, respectively.

#### 4. Experimental Validation

The proposed FSSML-EDM model was simulated using a benchmark dataset from UCI repository, which comprises of 649 samples with 32 features and 2 class labels as illustrated in Table 1. The parameter settings are provided as follows: learning rate, 0.01; dropout, 0.5; batch size, 5; and number of epochs, 50. For experimental validation, the dataset is split into 70% training (TR) data and 30% testing (TS) data.

**Table 1.** Dataset details.

Details	Values
Number of samples	649
Feature count	32
Class count	2
Source	<a href="https://archive.ics.uci.edu/ml/datasets/student+performance">https://archive.ics.uci.edu/ml/datasets/student+performance</a> (accessed on 30 January 2022)

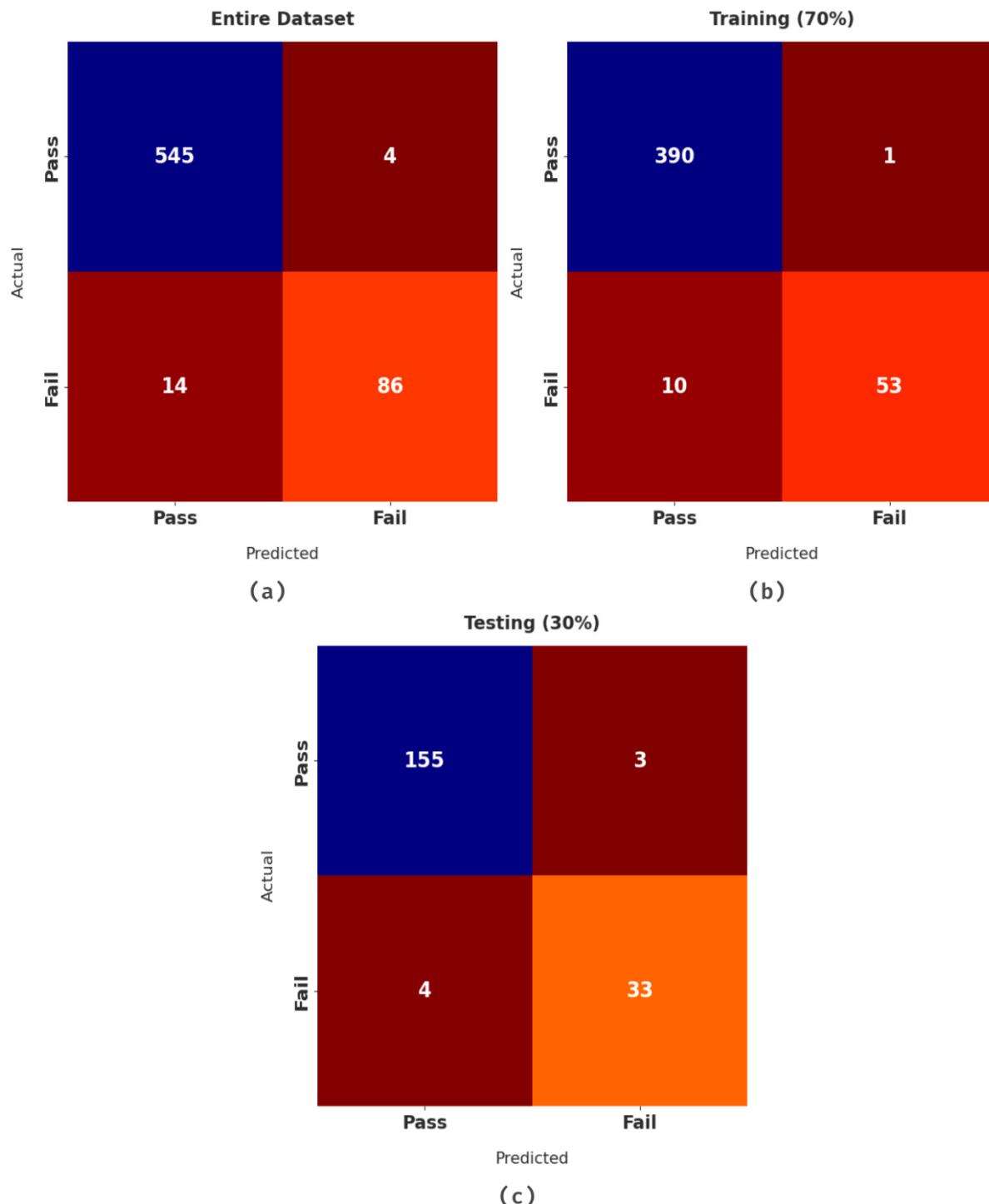
Figure 3 highlights a set of confusion matrices produced by the FSSML-EDM model on the test data. The figure indicates that the FSSML-EDM model resulted in effectual outcomes. On entire dataset, the FSSML-EDM model identified 545 samples into pass and 86 samples into fail. In addition, on 70% of the training dataset, the FSSML-EDM model identified 390 samples into pass and 53 samples into fail. Moreover, on 30% of testing dataset, the FSSML-EDM model identified 155 samples into pass and 33 samples into fail.

Table 2 offers a comprehensive EDM outcome of the FSSML-EDM model on test dataset. The experimental values indicated that the FSSML-EDM model accomplished maximum outcomes on all datasets. Figure 4 provides brief classification results of the FSSML-EDM model on entire dataset. It can be inferred from the figure that the FSSML-EDM model classified pass instances for  $accu_y$ ,  $prec_n$ ,  $reca_l$ ,  $F_{score}$ ,  $MCC$ , and  $\kappa$  of 97.23%, 97.50%, 99.27%, 98.38%, and 89.08% respectively. Moreover, the figure shows that the FSSML-EDM model classified fail instances for  $accu_y$ ,  $prec_n$ ,  $reca_l$ ,  $F_{score}$ ,  $MCC$ , and  $\kappa$  of 97.23%, 95.56%, 86%, 90.53%, and 89.08%, respectively.

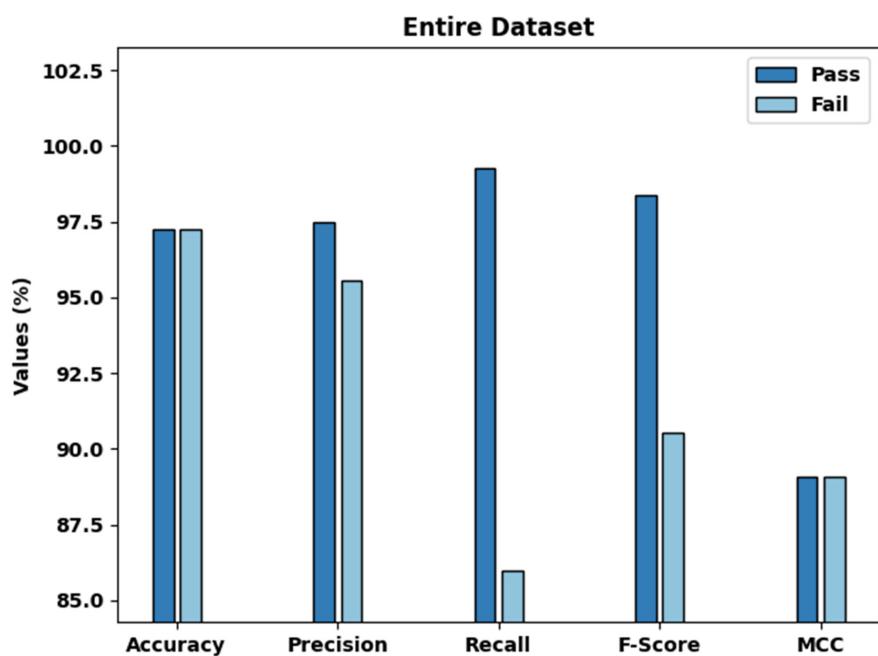
**Table 2.** Result analysis of FSSML-EDM technique with distinct measures and datasets.

Class Labels	Accuracy	Precision	Recall	F-Score	MCC	Kappa Score
Entire Dataset						
Pass	97.23	97.50	99.27	98.38	89.08	-
Fail	97.23	95.56	86.00	90.53	89.08	-
Average	97.23	96.53	92.64	94.45	89.08	88.91
Training Set (70%)						
Pass	97.58	97.50	99.74	98.61	89.57	-
Fail	97.58	98.15	84.13	90.60	89.57	-
Average	97.58	97.82	91.94	94.60	89.57	89.22
Testing Set (30%)						
Pass	96.41	97.48	98.1	97.79	88.22	-
Fail	96.41	91.67	89.19	90.41	88.22	-
Average	96.41	94.58	93.65	94.1	88.22	88.2

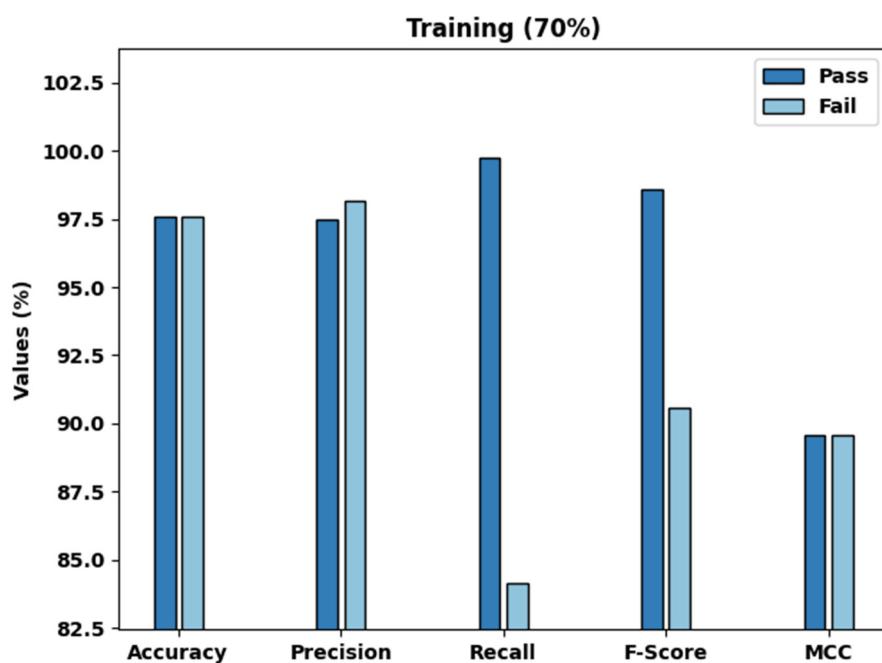
Figure 5 provides detailed classification results of the FSSML-EDM model on 70% of the training dataset. The figure reveals that the FSSML-EDM model classified pass instances for  $accu_y$ ,  $prec_n$ ,  $reca_l$ ,  $F_{score}$ ,  $MCC$ , and  $\kappa$  of 97.58%, 97.50%, 99.74%, 98.61%, and 89.57%, respectively. In addition, the figure shows that the FSSML-EDM model classified fail instances for  $accu_y$ ,  $prec_n$ ,  $reca_l$ ,  $F_{score}$ ,  $MCC$ , and  $\kappa$  of 97.58%, 98.15%, 84.13%, 90.60%, and 89.57% respectively.



**Figure 3.** Confusion matrix of FSSML-EDM technique on test data. (a) Entire dataset, (b) 70% of training dataset, and (c) 30% of testing dataset.



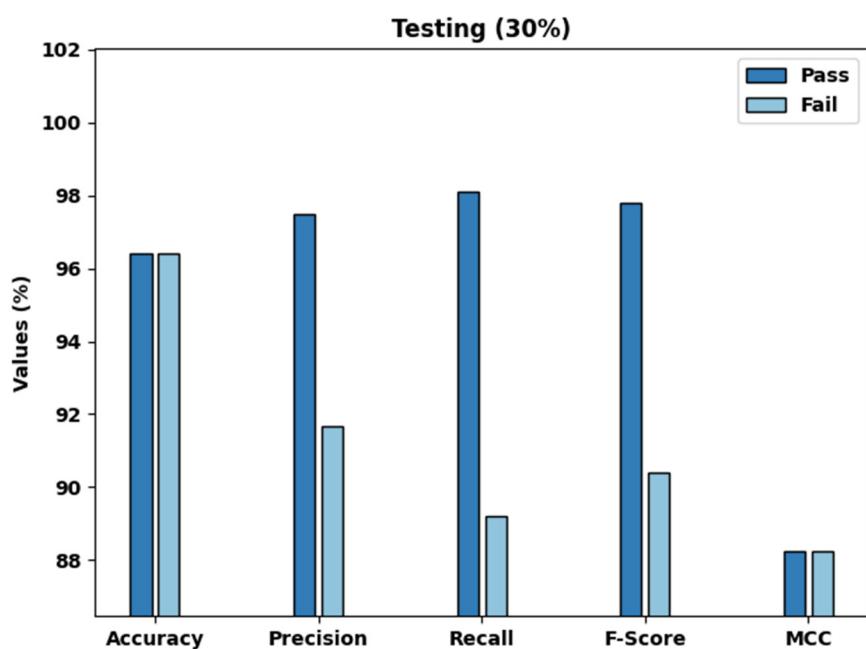
**Figure 4.** Result analysis of FSSML-EDM technique on entire dataset.



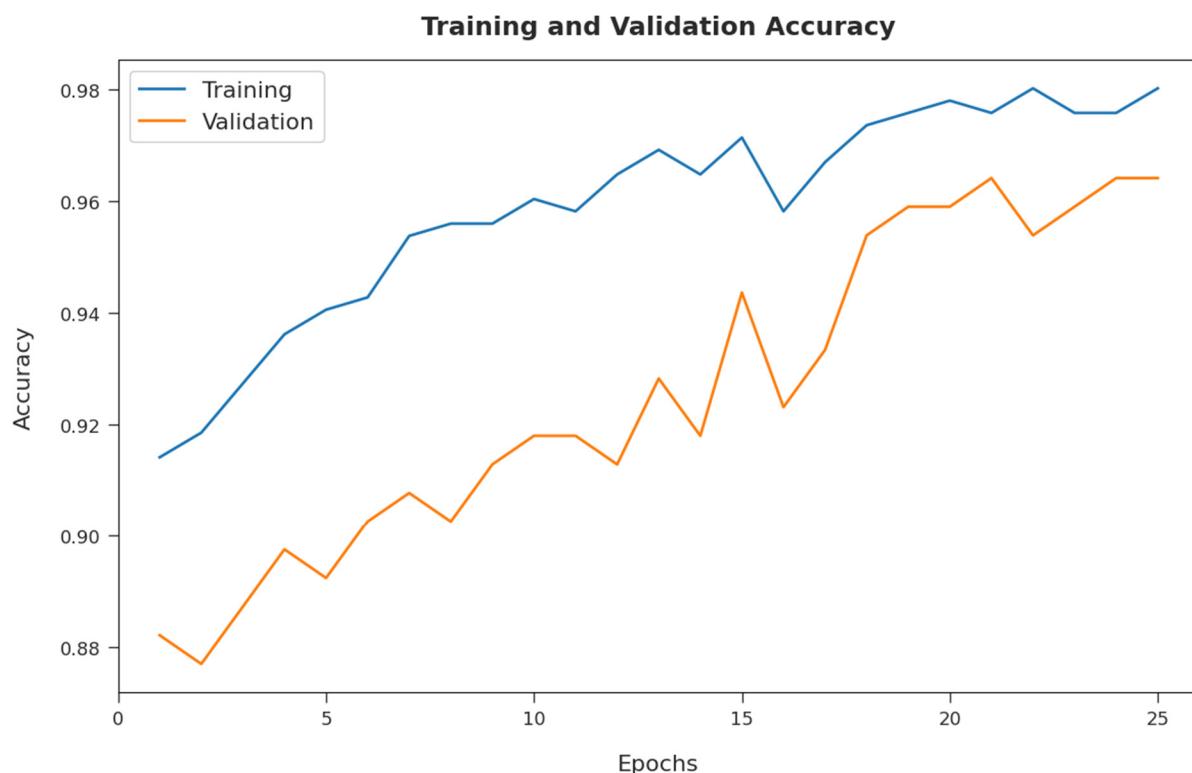
**Figure 5.** Result analysis of FSSML-EDM technique on 70% of training dataset.

Figure 6 offers brief classification results of the FSSML-EDM approach on 30% of testing dataset. The figure exposes that the FSSML-EDM algorithm classified pass instances for  $accu_y$ ,  $prec_n$ ,  $reca_l$ ,  $F_{score}$ ,  $MCC$ , and  $\kappa$  of 96.41%, 97.48%, 98.1%, 97.798%, and 88.22%, respectively. Moreover, the figure shows that the FSSML-EDM approach classified fail instances for  $accu_y$ ,  $prec_n$ ,  $reca_l$ ,  $F_{score}$ ,  $MCC$ , and  $\kappa$  of 96.41%, 91.67%, 89.19%, 90.41%, and 88.22%, respectively.

Figure 7 illustrates the training and validation accuracy inspection of the FSSML-EDM technique on the applied dataset. The figure of the FSSML-EDM approach offers maximum training/validation accuracy on the classification process.



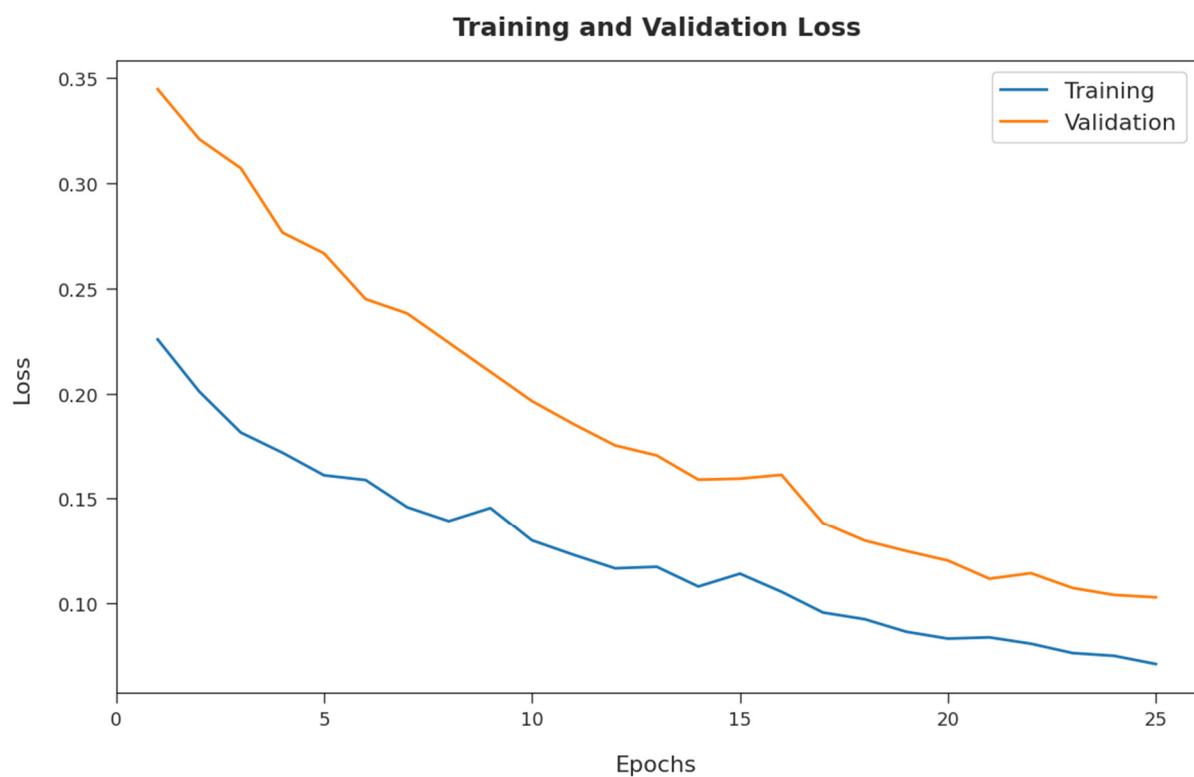
**Figure 6.** Result analysis of FSSML-EDM technique on 30% of testing dataset.



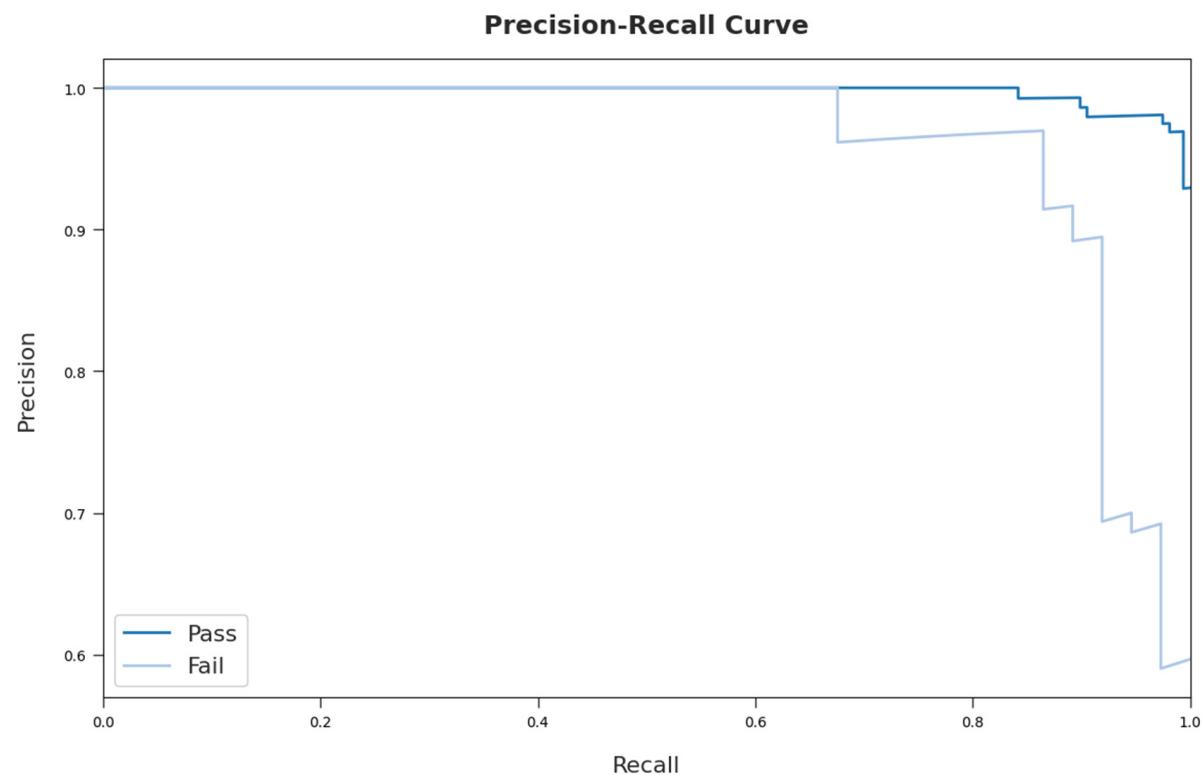
**Figure 7.** Accuracy graph analysis of FSSML-EDM technique.

Next, Figure 8 reveals the training and validation loss inspection of the FSSML-EDM approach on the applied dataset. The figure shows that the FSSML-EDM algorithm offers reduced training/accuracy loss on the classification process of the test data.

A brief precision-recall examination of the FSSML-EDM model on the test dataset is portrayed in Figure 9. By observing the figure, it is noticed that the DLBTDC-MRI model accomplished maximum precision-recall performance under all classes.

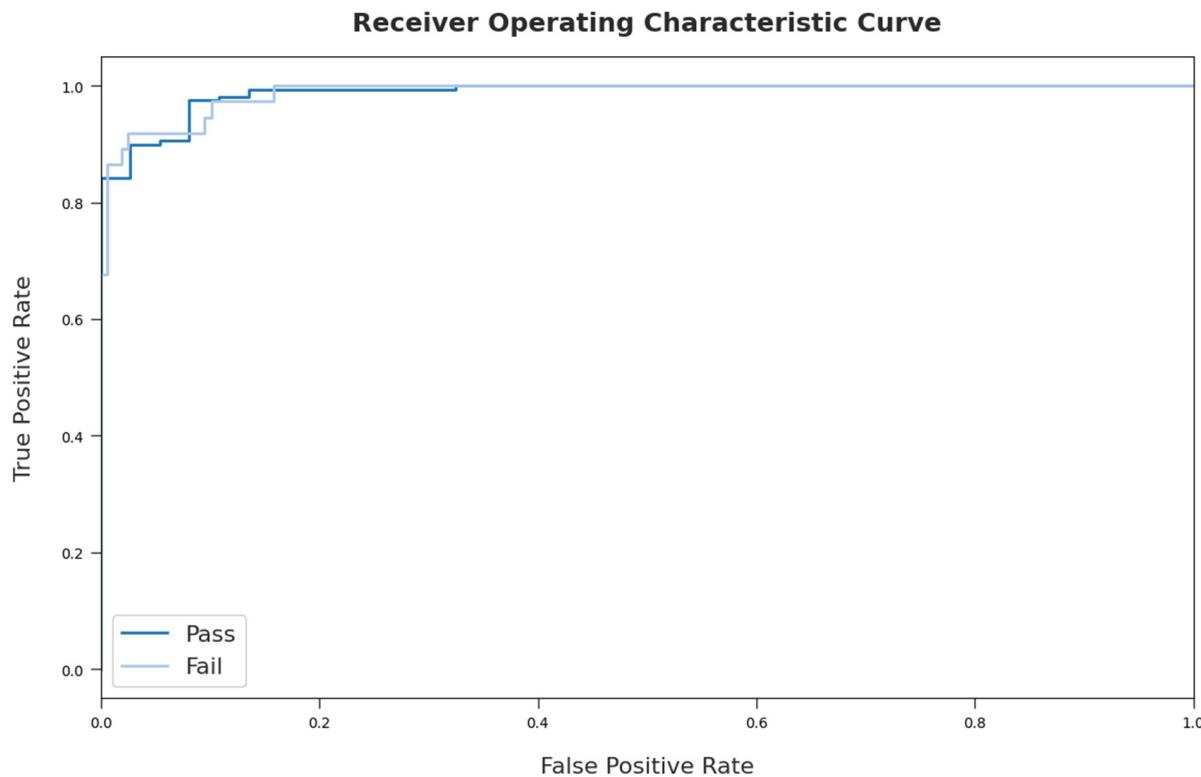


**Figure 8.** Loss graph analysis of FSSML-EDM technique.



**Figure 9.** Precision-recall curve analysis of FSSML-EDM technique.

A detailed ROC investigation of the FSSML-EDM approach on the distinct datasets is portrayed in Figure 10. The results indicate that the FSSML-EDM technique exhibited its ability in categorizing two different classes such as pass and fail on the test datasets.



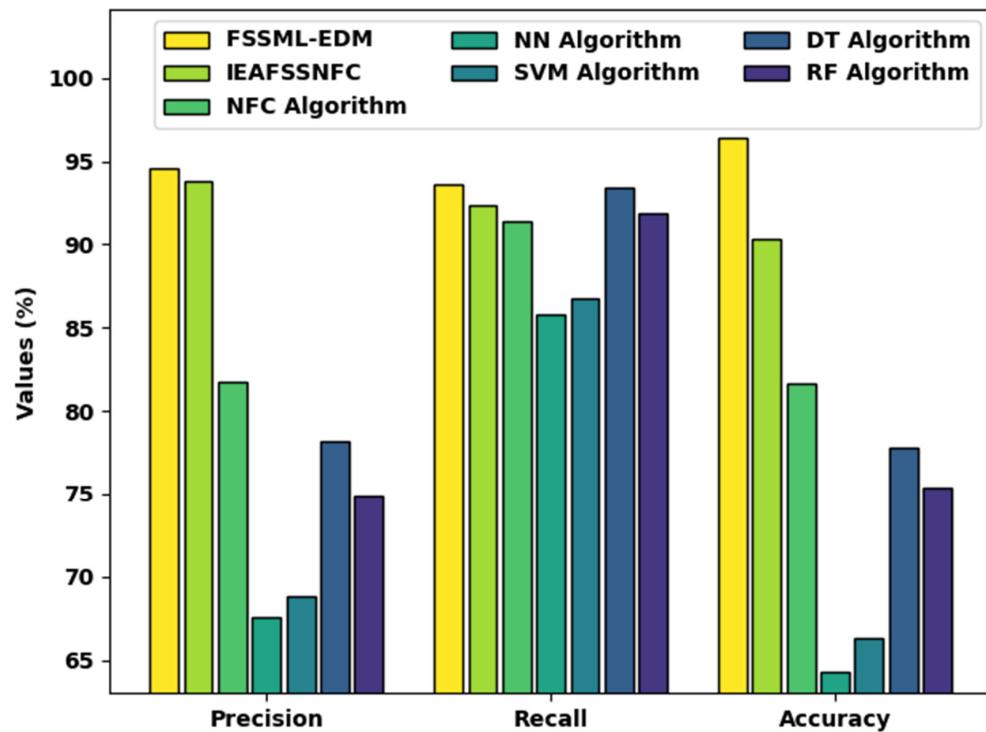
**Figure 10.** ROC curve analysis of FSSML-EDM technique.

Table 3 reveals an extensive comparative study of the FSSML-EDM model with existing models such as improved evolutionary algorithm-based feature subsets election with neuro-fuzzy classification (IEAFSSNFC) [22], neuro-fuzzy classification (NFC) [21], neural network (NN), support vector machines (SVM), decision tree (DT), and random forest (RF). Figure 11 inspects the comparative  $prec_n$ ,  $reca_l$ , and  $accu_y$  investigation of the FSSML-EDM model with recent methods. The figure reveals that the NN and SVM models showed poor performance with lower values of  $prec_n$ ,  $reca_l$ , and  $accu_y$ . The NFC, DT, and RF models have showed slightly improved values of  $prec_n$ ,  $reca_l$ , and  $accu_y$ . Moreover, the IEAFSSNFC model resulted in reasonable  $prec_n$ ,  $reca_l$ , and  $accu_y$  of 93.81%, 92.39%, and 90.33%, respectively. Furthermore, the FSSML-EDM model accomplished effectual outcomes with maximum  $prec_n$ ,  $reca_l$ , and  $accu_y$  of 94.58%, 93.65%, and 96.41%, respectively.

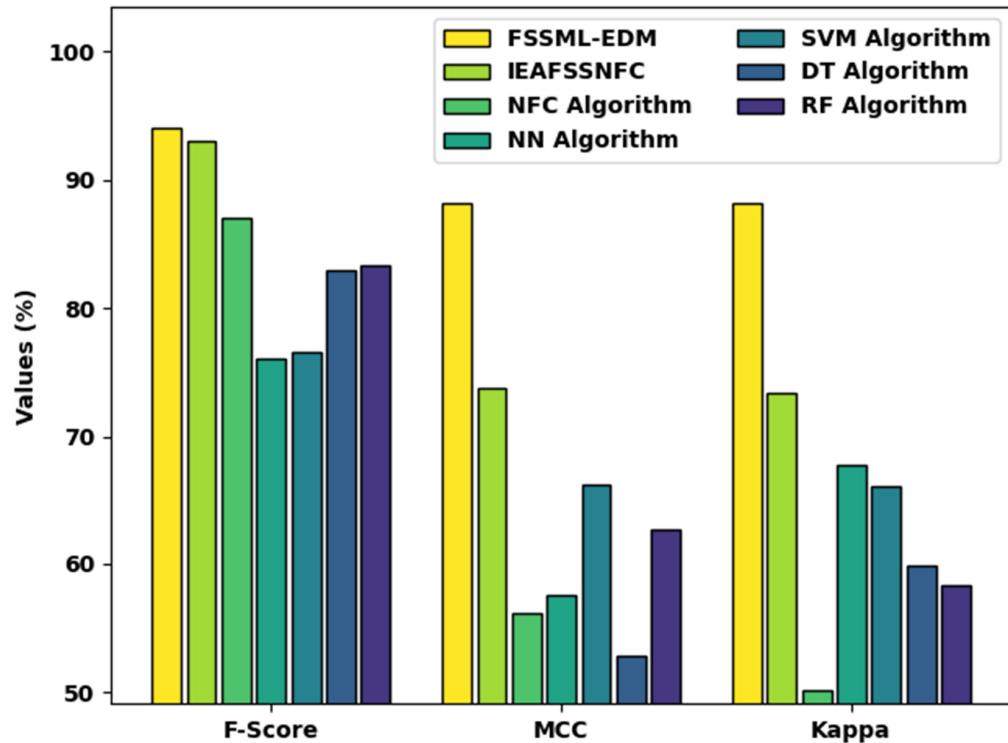
**Table 3.** Comparative analysis of FSSML-EDM approach with existing methods [21].

Methods	Precision	Recall	Accuracy	F-Score	MCC	Kappa
FSSML-EDM	94.58	93.65	96.41	94.10	88.22	88.20
IEAFSSNFC	93.81	92.39	90.33	93.01	73.78	73.37
NFC Algorithm	81.76	91.41	81.66	86.99	56.21	50.10
NN Algorithm	67.58	85.76	64.30	76.03	57.52	67.82
SVM Algorithm	68.83	86.77	66.36	76.63	66.27	66.17
DT Algorithm	78.22	93.40	77.81	83.02	52.87	59.80
RF Algorithm	74.95	91.90	75.40	83.39	62.68	58.34

Figure 12 inspects the comparative  $F_{score}$ ,  $MCC$ , and  $kappa$  analysis of the FSSML-EDM method with existing algorithms. The figure reveals that the NN and SVM methods showed poor performance with lower values of  $F_{score}$ ,  $MCC$ , and  $kappa$ . Similarly, the NFC, DT, and RF approaches showed slightly improved values of  $F_{score}$ ,  $MCC$ , and  $kappa$ .



**Figure 11.**  $Prec_n$ ,  $recal_l$ , and  $acc_y$  analysis of FSSML-EDM technique.



**Figure 12.**  $F_{score}$ ,  $MCC$ , and  $kappa$  analysis of FSSML-EDM technique.

Moreover, the IEAFSSNFC model resulted in reasonable  $F_{score}$ ,  $MCC$ , and  $kappa$  of 93.01%, 73.78%, and 73.37%, respectively. Additionally, the FSSML-EDM methodology accomplished effectual outcomes with maximum  $F_{score}$ ,  $MCC$ , and  $kappa$  of 94.10%, 82.22%, and 88.20%, respectively. Therefore, the FSSML-EDM model has the capability of assessing student performance in real time.

## 5. Conclusions

In this study, a new FSSML-EDM technique was developed for mining educational data. The proposed FSSML-EDM model involves three major processes. At the initial stage, the presented FSSML-EDM model incorporates the design of CSO-FS technique for electing feature subsets. Next, ELM classifier can be employed for the effective identification and classification of educational data. Finally, the AHB algorithm is utilized for effectively adjusting the parameters involved in the ELM model. The performance study revealed the effectual outcomes of the FSSML-EDM model over the compared models under several dimensions. Therefore, the FSSML-EDM model can be used as an effectual tool for EDM. In the future, feature reduction and outlier removal models can be employed to improve performance. In addition, the proposed model is presently tested on small-scale dataset, which needs to be explored. As a part of the future scope, the performance of the proposed model will be evaluated on a large-scale real-time dataset.

**Author Contributions:** Conceptualization, M.H.; Data curation, I.H.-J.; Formal analysis, I.H.-J.; Investigation, M.M.K.; Methodology, M.H. and R.F.M.; Project administration, M.M.K.; Resources, B.M.E.E.; Software, B.M.E.E.; Supervision, S.A.-K.; Validation, S.A.-K.; Visualization, S.A.-K.; Writing—original draft, M.H.; Writing—review & editing, R.F.M. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by Princess Nourah bint Abdulrahman University Researchers Supporting Project number (PNURSP2022R125), Princess Nourah bint Abdulrahman University, Riyadh, Saudi Arabia; Also, the authors would like to thank the Deanship of Scientific Research at Umm Al-Qura University for supporting this work by Grant Code: (22UQU4400271DSR07).

**Data Availability Statement:** Data sharing not applicable to this article as no datasets were generated during the current study.

**Conflicts of Interest:** The authors declare that they have no conflict of interest. The manuscript was written through contributions of all authors. All authors have given approval to the final version of the manuscript.

## References

- Prakash, B.A.; Ramakrishnan, N. Leveraging Propagation for Data Mining: Models, Algorithms and Applications. In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Francisco, CA, USA, 13–17 August 2016; pp. 2133–2134.
- Jalota, C.; Agrawal, R. Analysis of Educational Data Mining Using Classification. In Proceedings of the 2019 International Conference on Machine Learning, Big Data, Cloud and Parallel Computing (COMITCon), Faridabad, India, 14–16 February 2019; pp. 243–247. [[CrossRef](#)]
- Kenthapadi, K.; Mironov, I.; Thakurta, A.G. Privacy-preserving data mining in industry. In Proceedings of the Twelfth ACM International Conference on Web Search and Data Mining, WSDM’19: The Twelfth ACM International Conference on Web Search and Data Mining, Melbourne, VIC, Australia, 11–15 February 2019; pp. 840–841.
- Yan, D.; Qin, S.; Bhattacharya, D.; Chen, J.; Zaki, M.J. 20th International Workshop on Data Mining in Bioinformatics (BIOKDD 2021). In Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery & Data Mining, Singapore, 14–18 August 2021; pp. 4175–4176.
- Aldowah, H.; Al-Samarraie, H.; Fauzy, W.M. Educational data mining and learning analytics for 21st century higher education: A review and synthesis. *Telemat. Inform.* **2019**, *37*, 13–49. [[CrossRef](#)]
- Fernandes, E.; Holanda, M.; Victorino, M.; Borges, V.R.P.; Carvalho, R.; Van Erven, G. Educational data mining: Predictive analysis of academic performance of public school students in the capital of Brazil. *J. Bus. Res.* **2019**, *94*, 335–343. [[CrossRef](#)]
- De Andrade, T.L.; Rigo, S.J.; Barbosa, J.L.V. Active Methodology, Educational Data Mining and Learning Analytics: A Systematic Mapping Study. *Inform. Educ.* **2021**, *20*, 171–204. [[CrossRef](#)]

8. Sáiz-Manzanares, M.; Rodríguez-Díez, J.; Díez-Pastor, J.; Rodríguez-Arribas, S.; Marticorena-Sánchez, R.; Ji, Y. Monitoring of Student Learning in Learning Management Systems: An Application of Educational Data Mining Techniques. *Appl. Sci.* **2021**, *11*, 2677. [[CrossRef](#)]
9. Anand, N.; Sehgal, R.; Anand, S.; Kaushik, A. Feature selection on educational data using Boruta algorithm. *Int. J. Comput. Intell. Stud.* **2021**, *10*, 27. [[CrossRef](#)]
10. Shrestha, S.; Pokharel, M. Educational data mining in moodle data. *Int. J. Inform. Commun. Technol. (IJ-ICT)* **2021**, *10*, 9. [[CrossRef](#)]
11. Injadat, M.; Moubayed, A.; Nassif, A.B.; Shami, A. Systematic ensemble model selection approach for educational data mining. *Knowl.-Based Syst.* **2020**, *200*, 105992. [[CrossRef](#)]
12. Ashraf, M.; Zaman, M.; Ahmed, M. An Intelligent Prediction System for Educational Data Mining Based on Ensemble and Filtering approaches. *Procedia Comput. Sci.* **2020**, *167*, 1471–1483. [[CrossRef](#)]
13. Dabhadhe, P.; Agarwal, R.; Alameen, K.; Fathima, A.; Sridharan, R.; Gopakumar, G. Educational data mining for predicting students' academic performance using machine learning algorithms. *Mater. Today Proc.* **2021**, *47*, 5260–5267. [[CrossRef](#)]
14. Nahar, K.; Shova, B.I.; Ria, T.; Rashid, H.B.; Islam, A. Mining educational data to predict students performance. *Educ. Inf. Technol.* **2021**, *26*, 6051–6067. [[CrossRef](#)]
15. Deb, S.; Gao, X.-Z.; Tammi, K.; Kalita, K.; Mahanta, P. Recent Studies on Chicken Swarm Optimization algorithm: A review (2014–2018). *Artif. Intell. Rev.* **2019**, *53*, 1737–1765. [[CrossRef](#)]
16. Zhang, L.; Yang, H.; Jiang, Z. Imbalanced biomedical data classification using self-adaptive multilayer ELM combined with dynamic GAN. *Biomed. Eng. Online* **2018**, *17*, 181. [[CrossRef](#)] [[PubMed](#)]
17. Roozbeh, M.; Arashi, M.; Hamzah, N.A. Generalized Cross-Validation for Simultaneous Optimization of Tuning. *Iran. J. Sci. Technol. Trans. A Sci.* **2020**, *44*, 473–485. [[CrossRef](#)]
18. Amini, M.; Roozbeh, M. Optimal partial ridge estimation in restricted semiparametric regression models. *J. Multivar. Anal.* **2015**, *136*, 26–40. [[CrossRef](#)]
19. Roozbeh, M. Optimal QR-based estimation in partially linear regression models with correlated errors using GCV criterion. *Comput. Stat. Data Anal.* **2018**, *117*, 45–61. [[CrossRef](#)]
20. Roozbeh, M.; Babaie-Kafaki, S.; Aminifard, Z. Improved high-dimensional regression models with matrix approximations applied to the comparative case studies with support vector machines. *Optim. Methods Softw.* **2022**, 1–18. [[CrossRef](#)]
21. Zhang, Z.; Huang, C.; Huang, H.; Tang, S.; Dong, K. An optimization method: Hummingbirds optimization algorithm. *J. Syst. Eng. Electron.* **2018**, *29*, 386–404. [[CrossRef](#)]
22. Duhayyim, M.A.; Marzouk, R.; Al-Wesabi, F.N.; Alrajhi, M.; Hamza, M.A.; Zamani, A.S. An Improved Evolutionary Algorithm for Data Mining and Knowledge Discovery. *CMC-Comput. Mater. Contin.* **2022**, *71*, 1233–1247.