Computer Systems Science & Engineering
DOI: 10.32604/csse.2023.028107
Tech Science Press
Article

Optimal Machine Learning Enabled Performance Monitoring for Learning Management Systems

Ashit Kumar Dutta[1,*], Mazen Mushabab Alqahtani[2], Yasser Albagory[3], Abdul Rahaman Wahab Sait[4] and Majed Alsanea[5]

[1]Department of Computer Science and Information Systems, College of Applied Sciences, AlMaarefa University, Ad Diriyah, 13713, Riyadh, Kingdom of Saudi Arabia
[2]Physical Therapy Department, Majmaah University, Majmaah, 11952, Kingdom of Saudi Arabia
[3]Department of Computer Engineering, College of Computers and Information Technology, Taif University, Taif, 21944, Kingdom of Saudi Arabia
[4]Department of Archives and Communication, King Faisal University, Al Ahsa, 31982, Hofuf, Kingdom of Saudi Arabia
[5]Department of Computing, Arabeast Colleges, Riyadh, 11583, Kingdom of Saudi Arabia
*Corresponding Author: Ashit Kumar Dutta. Email: adotta@mcst.edu.sa
Received: 02 February 2022; Accepted: 30 March 2022

Abstract: Learning Management System (LMS) is an application software that is used in automation, delivery, administration, tracking, and reporting of courses and programs in educational sector. The LMS which exploits machine learning (ML) has the ability of accessing user data and exploit it for improving the learning experience. The recently developed artificial intelligence (AI) and ML models helps to accomplish effective performance monitoring for LMS. Among the different processes involved in ML based LMS, feature selection and classification processes find beneficial. In this motivation, this study introduces Glowworm-based Feature Selection with Machine Learning Enabled Performance Monitoring (GSO-MFWELM) technique for LMS. The key objective of the proposed GSO-MFWELM technique is to effectually monitor the performance in LMS. The proposed GSO-MFWELM technique involves GSO-based feature selection technique to select the optimal features. Besides, Weighted Extreme Learning Machine (WELM) model is applied for classification process whereas the parameters involved in WELM model are optimally fine-tuned with the help of Mayfly Optimization (MFO) algorithm. The design of GSO and MFO techniques result in reduced computation complexity and improved classification performance. The presented GSO-MFWELM technique was validated for its performance against benchmark dataset and the results were inspected under several aspects. The simulation results established the supremacy of GSO-MFWELM technique over recent approaches with the maximum classification accuracy of 0.9589.

Keywords: Learning management system; data mining; performance monitoring; machine learning; feature selection

1 Introduction

Nowadays, education techniques have been increasingly adopted in higher education institutions since teaching is no longer limited to Face-to-Face (F2F) physically-interactive sessions [1]. For university courses, the integration of F2F teaching and e-learning increases the flexibility, choices, and accessibility for communication [2]. This achievement in instructional productivity is made possible by Learning Management System (LMS) that is commonly employed as the environment to support e-learning and hybrid online F2F courses. This advanced LMS has the potential to accomplish conventional instructional activities through online such as course material management, dissemination of information, evaluation and collection of students' feedback. After the outbreak of COVID-19, universities have increasingly adopted LMS to support the course [3]. Education Data Mining (EDM) is an emerging field that is concerned with emergent techniques to evaluate distinct types of information acquired from educational setting. This technique is also used to develop the settings where both faculty and student share knowledge [4]. EDM is a multi-disciplinary research domain that investigates statistical modeling, Data Mining (DM), and Artificial Intelligence (AI) with information generated from education institution [5]. Fig. 1 shows the objectives of an LMS.

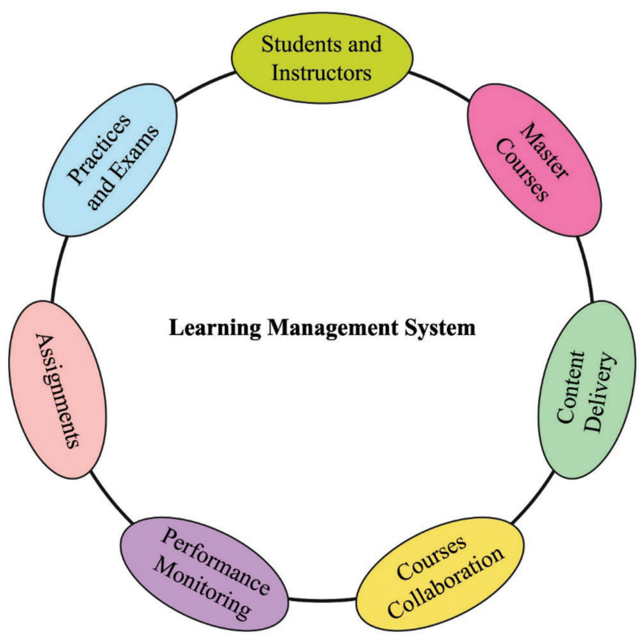

Figure 1: Goals in learning management system

EDM uses computational method to handle education data in order to investigate the educational queries as its final goal [6]. In order to make a country stand exclusive among other countries worldwide, the education system needs to have knowledge as a principal progression which can be achieved through improving the learning pedagogy from time to time. The hidden pattern in the data, collected from various information sources, is extracted by adopting DM method. In order to summarize the performance of students with their credentials, the researchers checked how DM can be exploited in educational field. Each educational institute generates huge volumes of information every year. New information is transferred considerably when using DM method. The information accomplished from educational institutes undergo inspection with the help of distinct DM models [7]. The technique detects the environment, where a student gets better inspiration to lead a useful life [8]. Weka, an efficient DM model was proposed earlier and utilized to generate substantial outcomes. The drastic growth of educational information [9] from heterogeneous sources results in an urgent need for EDM study. This could help in achieving the objectives so as to determine some educational purposes. In Machine Learning (ML) with increasing dimensions of the data, there is an increasing number of information required to provide a reliable analysis. Feature subset selection works by removing irrelevant or redundant features. The subset of features selected should follow Occam Razor principle so that it can offer important outcome. In some instances, NP remains a challenging issue and is resolved by metaheuristic approach [10].

The aim of the study conducted by Ahmmed et al. [11] is to carry out student visa processing via ML. The ML technique can be implemented at the place where the visas get rejected or approved for higher studies abroad. In this study, the researchers predicted the visa information for higher studies on the basis of student's background information. Next, the information is processed (through transformation, cleaning, standardization, feature selection, and integration). Then, various classifier techniques were used such as k-nearest neighbor (KNN), C4.5 (j48), random forest (RF), naïve bayes (NB), neural network (NN), and support vector machine (SVM), to classify the model. Hung et al. [12] employed EDM to explore the learning behavior of students in blended learning courses through the data collected from them. The experiment information was gathered from first-year students enrolled in Python programming courses at a university located in north Taiwan. During a semester, high-risk learners might be forecasted precisely by the data produced from blended education platform. In literature [13], the researchers presented a method to predict the student's dropout with NB Classifier method in R language. The study also investigated the reasons for students drop out at an earlier stage. The model forecasted whether a student may drop out or not in future. This study cited many factors that affect a student to drop out from the course.

Sarra et al. [14] estimated the helpfulness of a certain class method i.e., Bayesian Profile Regression, for identification of students who are likely to drop out from their courses. Students' resilience, performance, and motivation were considered and the method allowed to draw a student's profile with high risks of educational failure. In literature [15], EDM method with KNN and Multilayer Perceptron (MLP) approaches are used to predict the performance of the learner. The classifier outcomes were satisfactory while the kNN classification attained the optimal outcomes. The experiment outcome shows that the performance of the learner can be evaluated. Further, the researchers observed a relationship between learning performance and video sequence viewing behavior. Ramaswami et al. [16] established a generic prediction method that can identify at-risk students over a wide range of courses. The experiment was implemented by a variety of approaches and when generic method was used, it produced an efficient outcome. It was found to be an outstanding candidate to provide solutions in this field, given the fact that it can flawlessly manage missing and categorical information.

In this background, the current study introduces a Glowworm based Feature Selection with Machine Learning Enabled Performance Monitoring (GSO-MFWELM) technique for LMS. The major intention of the proposed GSO-MFWELM technique is to effectively monitor the performance in LMS. GSO-MFWELM technique primarily designs a GSO-based Feature Selection (FS) technique to select the optimal features. Besides, a Weighted Extreme Learning Machine (WELM) model is applied for classification process whereas the parameters involved in WELM model are optimally fine-tuned with the help of Mayfly Optimization (MFO) algorithm. GSO-MFWELM technique was validated for its performance against benchmark dataset and the results were inspected under several aspects.

2 The Proposed Model

In current study, a new GSO-MFWELM technique has been developed to monitor the performance of LMS. The proposed GSO-MFWELM technique encompasses three major processes namely, feature subset selection, WELM-based classification, and MFO-based parameter tuning. The weight values of WELM model can be optimally elected by MFO algorithm with classification error rate as the objective function. Fig. 2 illustrates the working process of GSO-MFWELM technique.

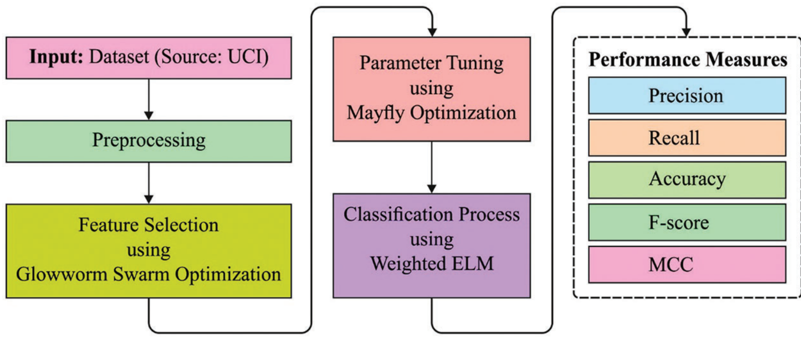

Figure 2: Overall process of GSO-MFWELM technique

2.1 Steps Involved in GSO-FS Technique

In this stage, the learning data is fed into GSO-FS technique to elect an optimal subset of features. GSO [17] is a smart optimization technique that functions according to the phenomenon in which the light, emitted by glowworm, is utilized as a signal to attract glowworms. This approach includes a collection of glowworms that are arbitrarily distributed. All the glowworms are considered to be potential solutions, characterized by their location. The glowworm, with high luminosity, exhibits high brightness and so it attracts glowworms with low-brightness. In this manner, global optimization is accomplished. The elementary steps are as follows.

Step 1. Initialize the elementary parameter of GSO. This parameter includes fluorescein update rate γ, population size g, fluorescein volatilization factor ρ, set of glowworms $N_i(t)$ in the decision domain, update rate β of the dynamic decision domain, perception radius r_s, move step s and threshold n_t for the number of glowworms in the neighborhood.

Step 2. The fitness value of glowworm i during t^{th} iteration is transformed into fluorescein value as given herewith.

$$l_i(t) = (1 - \rho)l_i(t - 1) + \gamma J(X(t)) \tag{1}$$

where as ρ represents the fluorescein decay constant that belongs to (0, 1) and γ indicates the fluorescein enhancement constant.

Step 3. All the glowworms select individuals with high brightness than the dynamic decision radius $r_d^i(t)$ to form the neighbor set $N_i(t)$.

Step 4. Evaluate the probability $p_{ij}(t)$ of glowworm $X_i(t)$ that moves towards the glowworm $X_j(t)$ in a dynamic decision radius as follows.

$$p_{ij}(t) = \frac{l_j(t) - l_i(t)}{\sum_{k \in N_i(t)} l_k(t) - l_i(t)} \tag{2}$$

Step 5. Upgrade the location of glowworm $X(t)$ as follows

$$X_i(t + 1) = X_i(t) + s \times \left[\frac{X_j(t) - X_i(t)}{||X_j(t) - X_i(t)||}\right] \tag{3}$$

Step 6. Upgrade the dynamic decision radius of glowworm $X(t)$ as follows.

$$r_d^i(t+1) = \{\min r_s,\ \{\max 0,\ \beta \times (n_t - |N_i(t)|)\}\} \tag{4}$$

The arithmetical formula of FS is presented. In general, the classification (that is., supervised learning) of data sets sized $N_S \times N_F$ is performed whereas N_F denotes the amount of features and N_S represents the amount of samples. The primary goal of FS is to choose a subset of features S from the overall amount of features (N_F) in which the size of S is lesser, when compared to N_F. It is attained by minimizing the subsequent objective function.

$$Fit = \lambda \times \gamma_S + (1-\lambda) \times \left(\frac{|S|}{N_F}\right) \tag{5}$$

Here γ_S represents the classification error that use S whereas $|S|$ indicates the number of selected features. λ is utilized for balancing between $\left(\frac{|S|}{N_P}\right)$ and γ_S.

2.2 WELM Based Classification

During classification process, WELM model is applied for the classification of learning data. WELM is an extended version of Extreme Learning Machine (ELM) [18]. With mapped data set $\chi_i',\ y_j \in \Re^p \times \Re^c (i = 1,\ 2,\ \ldots,\ n)$, the outcome of generalization in Single Layer Feed Forward Networks (SLFN) with q hidden node and activation function $h(x')$ are properly written as follows

$$o_i = \sum_{k=1}^{q} \beta_k h_k(\chi_i') = \sum_{k=1}^{q} \beta_k h(w_k,\ b_k,\ \mathrm{x}_\mathrm{i}'), \tag{6}$$

where $i = 1,\ 2,\ \ldots,\ n$, $w_k = [w_{k1},\ w_{k2},\ \ldots,\ w_{kp}]^T$ refers to input weight vector that links the input node and k^{th} hidden node, b_k refers to the bias of k^{th} hidden node, $\beta_k = [\beta_{k1},\ \beta_{k2},\ \ldots,\ \beta_{kc}]^T$ signifies the resultant weight vector that links the resultant node and k^{th} hidden node, and 0_i refers to the predictable output of i^{th} sample. The activation function generally utilized from ELM contains multiquadric, sigmoid, hard limit, and Gaussian RBF functions. Eq. (6) is equivalently modified as follows.

$$H\beta = O, \tag{7}$$

where H defines the hidden state resultant matrix of SLFNs and is determined as follows.

$$H = H(w_1,\ \ldots,\ w_q,\ b_1,\ \ldots,\ b_q,\ x_1',\ \ldots,\ x_n') = \begin{bmatrix} h(x_1') \\ \vdots \\ h(x_n') \end{bmatrix} \tag{8}$$

$$= \begin{bmatrix} h(w_1,\ b_1,\ x_1') & \cdots & h(w_q,\ b_q,\ x_1') \\ \vdots & \ddots & \vdots \\ h(w_1,\ b_1,\ x_1') & \cdots & h(w_q,\ b_q,\ x_1') \end{bmatrix}_{n \times q}$$

where the i^{th} row of H refers to the resultant vector of hidden state, in terms of input sample x_i', and the k^{th} column of H refers to the resultant vector of k^{th} hidden node in terms of input instances $x_1',\ x_2',\ \ldots,\ x_n'$. β implies weight matrix that links the hidden as well as output layers which are determined as follows

$$\beta = \begin{bmatrix} \beta_1^T \\ \vdots \\ \beta_q^T \end{bmatrix}_{q\times c} \tag{9}$$

O stands for the predictable label matrix while all the rows signify the resultant vector of one instance. O is determined as given herewith.

$$O = \begin{bmatrix} o_1^T \\ \vdots \\ T_{O_n} \end{bmatrix} = \begin{bmatrix} o_{11} & \cdots & 0_{1c} \\ \vdots & \ddots & \vdots \\ o_{n1} & \cdots & 0_{nc} \end{bmatrix} \tag{10}$$

While the target of trained SLFN is to minimize the outcome error, for instance, similar to the input instances with zero error as below.

$$\sum_{i=1}^{n} \|0_i - y_i\| = \|O - Y\| = 0 \tag{11}$$

where $y = \begin{bmatrix} y_1^T \\ \vdots \\ Ty_n \end{bmatrix} = \begin{bmatrix} y_{11} & \cdots & y_{1c} \\ \vdots & \ddots & \vdots \\ y_{n1} & \cdots & y_{nc} \end{bmatrix}$ represents the target resultant matrix.

$$H\beta = Y \tag{12}$$

In ELM, the weight w_k of input connection and bias b_k of hidden state nodes are arbitrarily and individually selected. Once this parameter is allocated, Eq. (12) is changed to suit the linear system and the resultant weight matrix β is logically defined as finding the least-square solutions of linear system.

$$\min_{\beta} \|H\beta - Y\| \tag{13}$$

An optimum solution of Eq. (13) as:

$$\hat{\beta} = H^{\dagger} Y = (H^T H) \tag{14}$$

where $H\dagger$ demonstrates Moore-Penrose generalization inverse of the hidden state resultant matrix, H. Here, $\hat{\beta}$ is made sure to be a less trained error and optimum generalized capability is obtained herewith. Further, it also avoids plunging the local optimally since $\hat{\beta}$ is unique. At last, it takes the classifier function of ELM as follows.

$$f(x\prime) = h(x\prime)\hat{\beta} = h(x\prime)H^{\dagger y} \tag{15}$$

If ELM classifier is constructed, it can be determined a $n \times n$ diagonal matrix W, in which the diagonal element W_{ii} refers to the weight of the trained instance χ_i'. In particular, when χ_i' appears to a majority class, the weight W_{ii} is comparatively lesser than the instance that appears to minority class. Based on the KKT statement, Eq. (14) is modified as follows.

$$\hat{\beta} = H^{\dagger} Y = (H^T W H)^{-1} H^T W T \tag{16}$$

Afterward, Eq. (15) is developed as follows.

$$f(x') = h(x')\hat{\beta} = h(x')(H^T W H)^{-1} H^T W T \tag{17}$$

Mostly, there are two structures exist to assign the weights to instances of two classes as follows.

$$W1 = W_{ii} = \begin{cases} 1/n_P & \textit{if } x_i' \in \text{minority}_{\text{class}} \\ 1/n_N & \textit{if } x_i' \in \text{majority}_{\text{class}} \end{cases} \tag{18}$$

or

$$W2 = W_{ii} = \begin{cases} 0.618/n_P & \textit{if } x_i' \in \text{minority}_{\text{class}} \\ 1/n_N & \textit{if } x_i' \in \text{majority}_{\text{class}} \end{cases}, \tag{19}$$

where Wl and $W2$ refere to two weighting processes, n_P and n_N signify the amount of samples of minority as well as majority classes correspondingly.

2.3 MFO Based Parameter Optimization

Finally, MFO algorithm is employed to fine tune the parameters of WELM technique, thereby enhancing the classifier results. The recently-designed MFO is an alteration of the familiar Particle Swarm Optimization (PSO) [19]. Since it is a combination of strength found in different optimization methods, it is assumed as a hybrid model. MFO algorithm is inspired by the social behavior of mayflies. Mayflies form as adults while the fittest one survives. Initially, two groups of population are created such as male and female populations. The candidate is characterized as a d-dimension vector $x = (x_1, \ldots, x_d)$. Then, the fitness of the candidate is estimated by the fitness function, $f(x)$. The velocity $v = (v_1, \ldots, v_d)$ represents the change in candidate position. All the candidates modify their trajectory based on their optimal position (pbest) and the optimal position of each mayfly (gbest) [20].

The congregation of male mayfly reflects the experience of all the males in defining its position, regarding the *neighbors^I* location. In order to determine x_i^t as the existing position of candidate solution i at time t, the location can be modified by adding a velocity v_i^{t+1}

$$x_i^{t+1} = x_i^t + v_i^{t+1} \tag{20}$$

With $x_i^0 U(x_{\min}, x_{\max})$.

Assume the lower velocity of the male population while the velocity is estimated as follows

$$v_{ij}^{t+1} = v_{ij}^t + a_1 e^{-\beta r_p^2}(pbest_{ij} - x_{ij}^t) + a_2 e^{-\beta r_g^2}(gbest_i - x_{ij}^t) \tag{21}$$

Here, v_{ij}^t represents the velocity of mayfly i, x_{ij}^t denotes the location of mayfly i, a_1 and a_2 are determined as positive constants that represent the attraction. $pbest_i$ denotes the optimal position achieved by the candidate solution i, and $pbest_{ij}$ at the subsequent phase while $t + 1$ is defined as follows.

$$pbest_i = \begin{cases} x_i^{t+1}, & \textit{if } f(x_i^{t+1}) < f(pbest_i) \\ \textit{same as before}, & \textit{otherwise} \end{cases} \tag{22}$$

Whereas $f{:}\mathbb{R}^n \Rightarrow \mathbb{R}$ denotes the function to be minimalized, $gbest$ indicates the global optimal position reached for the problem ever, at time t. The coefficient in Eq. (21) limits a *population^{I}s* visibility. r_p characterizes the distance between x_i and $pbest_i$. In the meantime, r_g describes the distance from x_i to gbest. r_p and r_g are defined as follows.

$$\|x_i - X_i\| = \sqrt{\sum_{j=1}^{n} (x_{ij} - X_{ij})^2} \tag{23}$$

Whereas x_{ij} represents the j^{th} element of i^{th} candidate. X_i is correlated to pbest.

The optimal fit candidate keeps implanting up and down movement by varied velocity. The velocity is defined as follows.

$$v_{ij}^{t+1} = v_{ij}^{t} + d * r \tag{24}$$

Here d represents a coefficient correlated to up and down movements and r indicates a random value between − 1 and 1.

Female mayfly does not gather, but tend to move towards the male mayflies. Here, y_i^t is defined as the existing position of female mayfly i at time t.

$$y_i^{t+1} = y_i^t + v_i^{t+1} \tag{25}$$

with $y_i^0 \ U \ (x_{\min}, \ x_{\max})$.

The velocity of the female mayfly is defined as follows.

$$v_{ij}^{t+1} = \begin{cases} v_{ij}^t + a_2 e^{-\beta r_{mf}^2}(x_{ij}^t - y_{ij}^t), & iff(y_i) > f(x_i) \\ v_{ij}^t + fl * r, & if\ f(y_i) \le f(x_i) \end{cases} \tag{26}$$

Here, v_{ij}^t indicates the velocity of i^{th} female at time t, y_{ij}^t shows the position of i^{th} female candidate solution at time t, a_2 represent a positive constant, β denotes a fixed coefficient, r_{mf} symbolizes the distance among the male candidates' solutions and the female one is calculated by Eq. (23), fl denotes a coefficient that corresponds to the female which remains unattracted, and r implies an arbitrary value between − 1 and 1.

Mating can be denoted by the operator i.e., crossover operator. A couple of male and female mayfly parents is selected. Next, the crossover operator creates two offspring as follows.

$$offspringl = L * male + (1 - L) * female$$

$$offspring2 = L * female + (1 - L) * male \tag{27}$$

Here, L represents an arbitrary value. At first, the velocity of offspring is equal to zero.

3 Experimental Validation

The proposed model was validated for its performance against benchmark dataset from UCI repository (available at https://archive.ics.uci.edu/ml/datasets/student+performance). The dataset includes 649 samples with 33 attributes under two classes. Class 0 (pass) includes 549 samples, whereas class 1 (fail) includes 100 instances.

Fig. 3 shows the correlation matrix generated for the test dataset. Tab. 1 shows the feature selection results of Chaotic Whale Optimization Algorithm (CWOA)-FS and GSO-FS techniques. The results show the outcomes of GSO-FS technique since it selected 15 features, whereas the presented CWOA-FS technique selected 20 features out of 32 and established its supremacy.

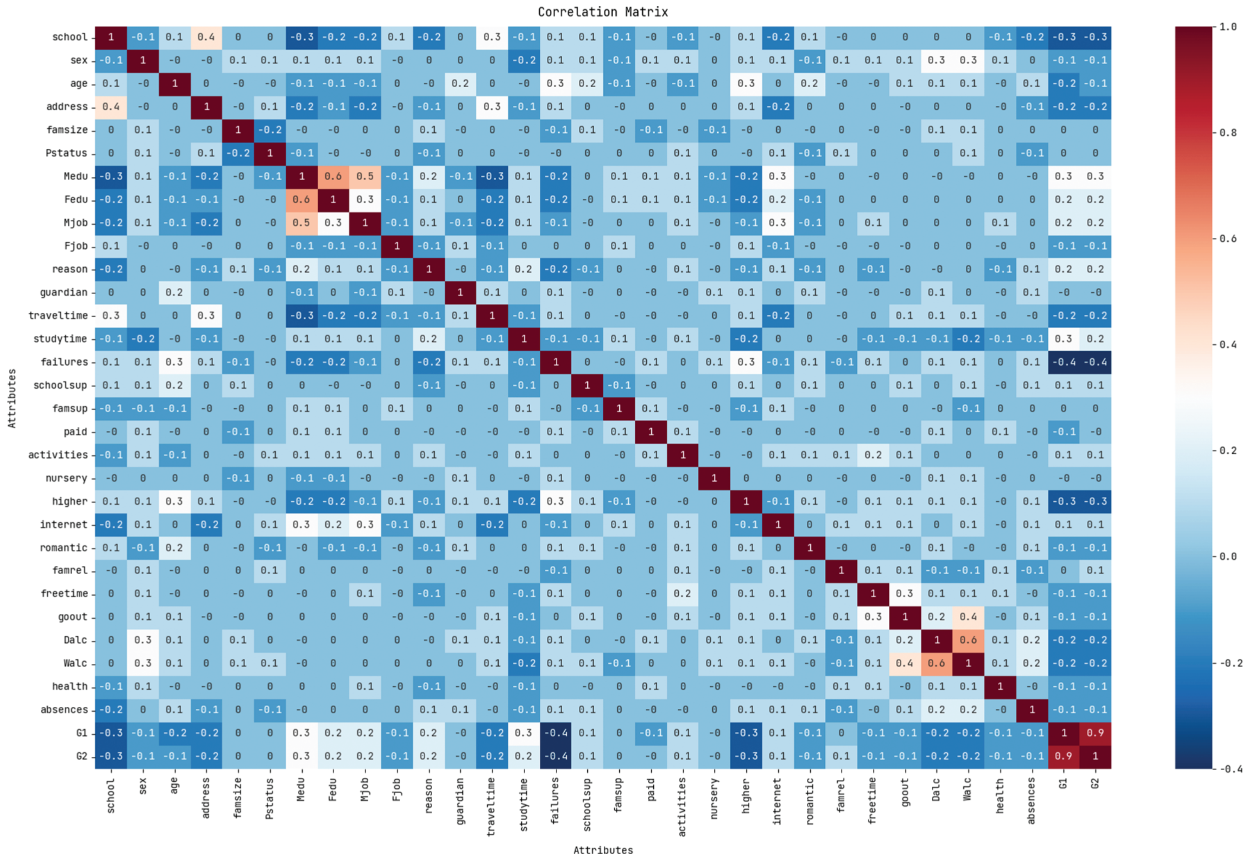

Figure 3: Correlation matrix of the proposed method

Table 1: Feature selection results of the proposed method

Methods	Total features	No. of features
CWOA-FS	32	20
GSO-FS	32	15

Fig. 4 demonstrates the confusion matrices generated by the proposed GSO-MFWELM technique under various Hidden Units (HUs). With HU-1, GSO-MFWELM technique categorized 529 instances under class 0 and 94 instances under class 1. Eventually, with HU-3, the proposed GSO-MFWELM approach categorized 531 instances under class 0 and 93 instances under class 1. Meanwhile, with HU-6, the presented GSO-MFWELM system categorized 528 instances under class 0 and 93 instances under class 1. The values in the confusion matrix are tabulated in Tab. 2 in terms of True Positive (TP), True Negative (TN), False Positive (FP), and False Negative (FN).

Tab. 3 offers the classification results of the analysis accomplished by GSO-MFWELM technique under distinct Hidden Units (HUs).

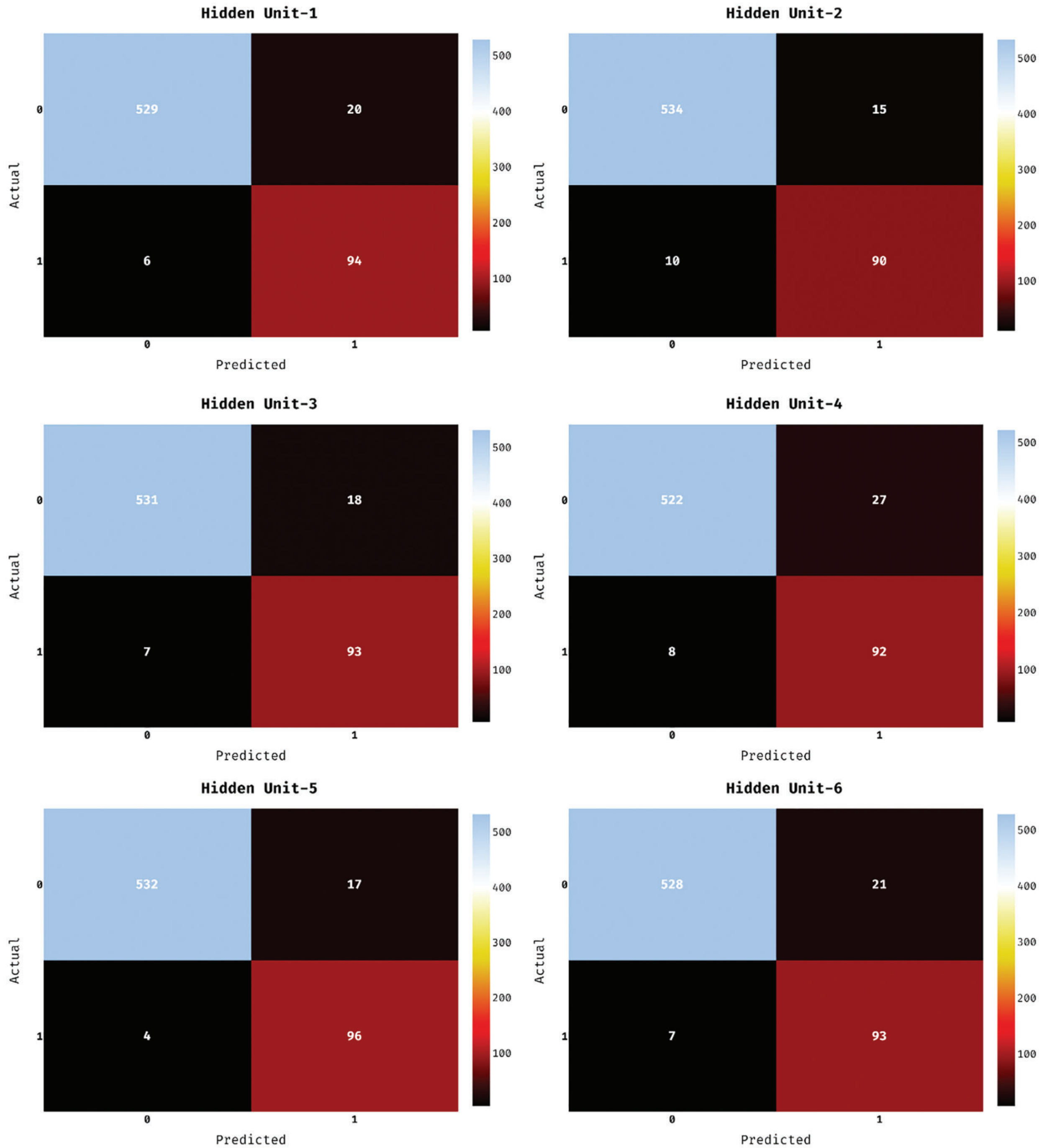

Figure 4: Confusion matrix of GSO-MFWELM technique under various HUs

Table 2: Confusion matrix

No. of hidden units	TP	FN	FP	TN
Hidden-unit-1	529	20	6	94
Hidden-unit-2	534	15	10	90
Hidden-unit-3	531	18	7	93
Hidden-unit-4	522	27	8	92
Hidden-unit-5	532	17	4	96
Hidden-unit-6	528	21	7	93

Table 3: Results of the analysis of GSO-MFWELM technique under various HUs

No. of hidden units	Precision	Recall	Accuracy	F-score	MCC
Hidden unit-1	0.9888	0.9636	0.9599	0.9760	0.8573
Hidden unit-2	0.9816	0.9727	0.9615	0.9771	0.8556
Hidden unit-3	0.9870	0.9672	0.9615	0.9770	0.8603
Hidden unit-4	0.9849	0.9508	0.9461	0.9676	0.8125
Hidden unit-5	0.9925	0.9690	0.9676	0.9806	0.8845
Hidden unit-6	0.9869	0.9617	0.9569	0.9742	0.8461
Average	0.9870	0.9642	0.9589	0.9754	0.8527

Fig. 5 shows the results of $prec_n$, $reca_l$, and $accu_y$ analysis attained by GSO-MFWELM technique under different HUs. The experimental values indicate that the proposed GSO-MFWELM technique obtained effectual classification performance. For instance, with HU-1, GSO-MFWELM technique attained $prec_n$, $reca_l$, and $accu_y$ values such as 0.9888, 0.9636, and 0.9599 respectively. In addition to this, with HU-6, GSO-MFWELM approach reached $prec_n$, $reca_l$, and $accu_y$ values namely, 0.9869, 0.9617, and 0.9569.

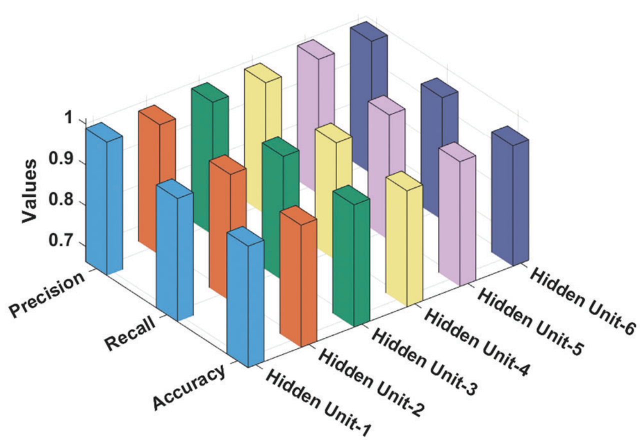

Figure 5: Results of the analysis of GSO-MFWELM technique under various HUs

Fig. 6 offers the F_{score} and Mathew Correlation Coefficient (MCC) analysis results accomplished by GSO-MFWELM technique under various HUs. The experimental values point out the effectual classification performance of GSO-MFWELM technique. For instance, with HU-1, the proposed GSO-MFWELM technique obtained F_{score} and MCC values such as 0.9760 and 0.8573 correspondingly. Besides, with HU-6, the proposed GSO-MFWELM technique achieved F_{score} and MCC values namely, 0.9742 and 0.8461.

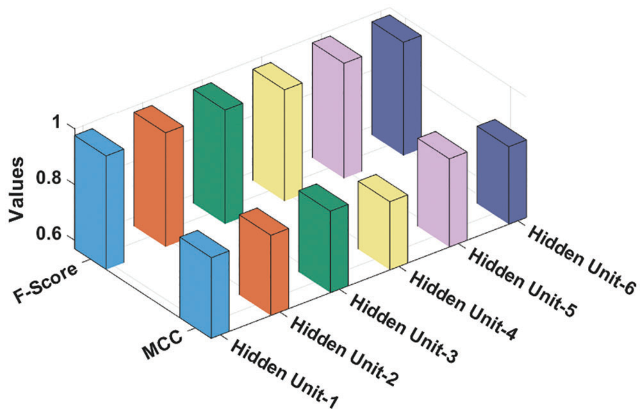

Figure 6: F-score and MCC analysis results of GSO-MFWELM technique under various HUs

Fig. 7 portrays the average analysis results achieved by the proposed GSO-MFWELM technique under distinct HUs. The figure infers that the proposed GSO-MFWELM technique achieved average $prec_n$ of 0.9870, $reca_l$ of 0.9642, $accu_y$ of 0.9589, F_{score} of 0.9754, and MCC of 0.8527.

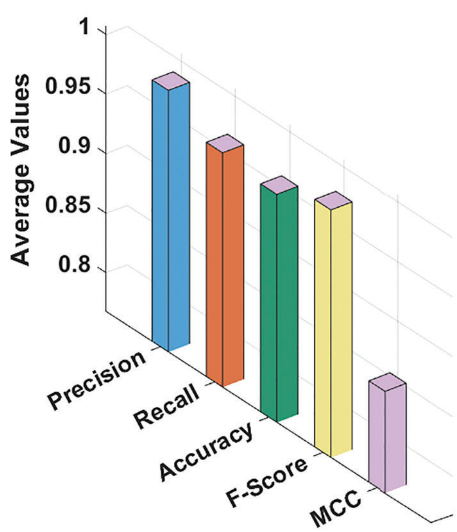

Figure 7: Average analysis results of GSO-MFWELM technique with distinct measures

Fig. 8 demonstrates the analysis results of GSO-MFWELM system against recent approaches in terms of $prec_n$ and $reca_l$. The results illustrate that NN and SVM methods achieved minimal $prec_n$ and $reca_l$ values. Besides, Decision Tree (DT) and Random Forest (RF) techniques obtained certainly increased values in terms of $prec_n$ and $reca_l$. Furthermore, mproved Evolutionary Algorithm based Feature Subset Selection with Neuro Fuzzy Classifier (IEAFSS-NFC) and Neural fuzzy classifier (FC) techniques obtained reasonable $prec_n$ and $reca_l$ values. At last, the proposed GSO-MFWELM methodology surpassed all other techniques and achieved $prec_n$ and $reca_l$ values such as 0.9870 and 0.9642.

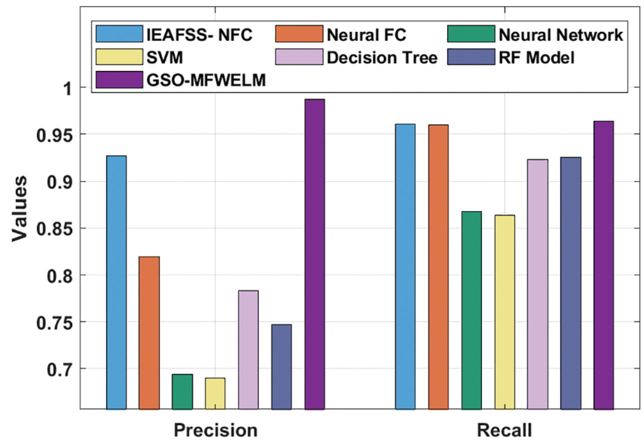

Figure 8: $Prec_n$ and $Reca_l$ analysis results of GSO-MFWELM technique with recent approaches

Fig. 9 showcases the results of the analysis of GSO-MFWELM technique against recent methods in terms of $accu_y$. The results show that both NN and SVM models achieved the least $accu_y$ values. At the same time, DT and RF models achieved certainly increased values of $accu_y$. Moreover, IEAFSS-NFC and Neural FC techniques obtained reasonable $accu_y$ values. However, the proposed GSO-MFWELM technique surpassed all other methods and achieved an $accu_y$ of 0.9589.

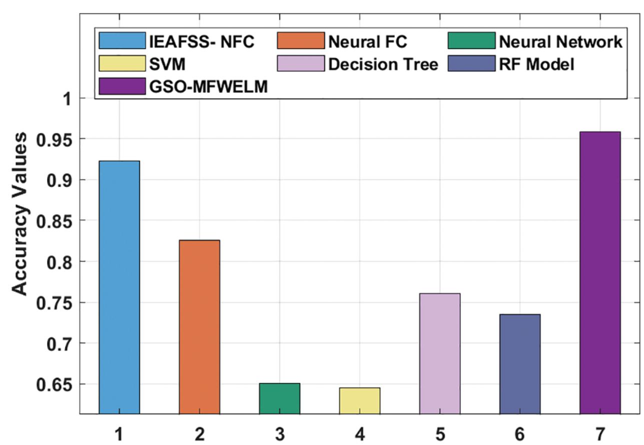

Figure 9: Acc_y analysis results of GSO-MFWELM technique with recent approaches

Tab. 4 highlights the comparison results of GSO-MFWELM technique against existing models [21].

Table 4: Comparative analysis results of GSO-MFWELM technique against recent methods

Methods	Precision	Recall	Accuracy	F-Score	MCC
IEAFSS- NFC	0.9271	0.9607	0.9229	0.9531	0.7461
Neural FC	0.8196	0.9598	0.8260	0.8884	0.5436

(Continued)

Table 4 (continued)

Methods	Precision	Recall	Accuracy	F-Score	MCC
Neural Network	0.6939	0.8678	0.6510	0.7712	0.0880
SVM	0.6903	0.8633	0.6450	0.7672	0.0697
Decision Tree	0.7832	0.9227	0.7610	0.8472	0.3396
RF Model	0.7468	0.9255	0.7350	0.8266	0.3233
GSO-MFWELM	0.9870	0.9642	0.9589	0.9754	0.8527

Fig. 10 depicts the results achieved by GSO-MFWELM algorithm in analysis against recent methodologies with respect to F_{score} and MCC. The results infer that NN and SVM models obtained the least F_{score} and MCC values. Simultaneously, DT and RF approaches reached certainly increased F_{score} and MCC values. In addition, IEAFSS-NFC and Neural FC methods achieved reasonable F_{score} and MCC values. Finally, the proposed GSO-MFWELM algorithm surpassed all other methods with the highest F_{score} and MCC values such as 0.9754 and 0.8527. Therefore, it has been established from the above discussed results that the proposed model has the ability to attain maximum results over existing techniques.

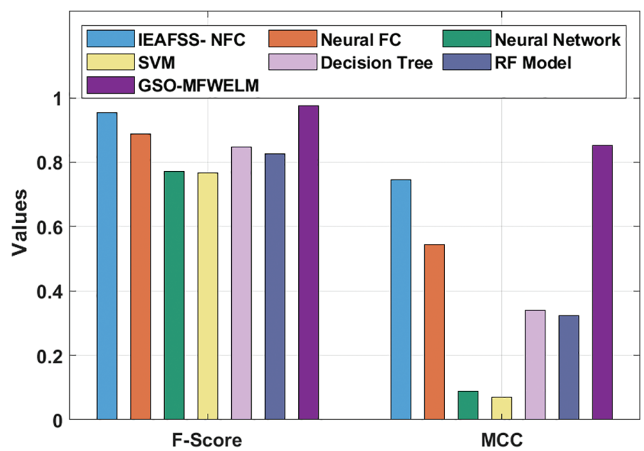

Figure 10: F_{score} and MCC analysis results of GSO-MFWELM technique with recent approaches

4 Conclusion

In current study, a new GSO-MFWELM technique has been developed to monitor LMS performance. The proposed GSO-MFWELM technique encompasses three major processes namely, feature subset selection, WELM based-classification, and MFO-based parameter tuning. The weight values of the WELM model can be optimally elected by MFO algorithm with classification error rate as an objective function. The proposed GSO-MFWELM technique was validated for performance using benchmark dataset and the results were inspected under several aspects. The simulation results infer the supremacy of the proposed GSO-MFWELM technique over recent approaches under different measures. In future, clustering techniques can be integrated into GSO-MFWELM technique to achieve enhanced performance.

Acknowledgement: The authors deeply acknowledge the Researchers supporting program (TUMA-Project-2021-27) Almaarefa University, Riyadh, Saudi Arabia for supporting steps of this work.

Funding Statement: This research was supported by the Researchers Supporting Program (TUMA-Project-2021-27) Almaarefa University, Riyadh, Saudi Arabia. Taif University Researchers Supporting Project number (TURSP-2020/161), Taif University, Taif, Saudi Arabia.

Conflicts of Interest: The authors declare that they have no conflicts of interest to report regarding the present study.

References

[1] A. H. Blanco, B. H. Flores, D. Tomás and B. Navarro-Colorado, "A systematic review of deep learning approaches to educational data mining," *Complexity*, vol. 2019, pp. 1–22, 2019.

[2] P. Guleria and M. Sood, "A convergence of mining and machine learning: The new angle for educational data mining," in *Cognitive Computing Systems*, 1st ed., vol. 1, Apple Academic Press, pp. 97–118, 2021.

[3] W. Sun, G. Z. Dai, X. R. Zhang, X. Z. He, X. Chen, "TBE-Net: A three-branch embedding network with part-aware ability and feature complementary learning for vehicle re-identification," *IEEE Transactions on Intelligent Transportation Systems*, pp. 1–13, 2021. https://doi.org/10.1109/TITS.2021.3130403.

[4] W. Sun, L. Dai, X. R. Zhang, P. S. Chang, X. Z. He, "RSOD: Real-time small object detection algorithm in UAV-based traffic monitoring," *Applied Intelligence*, pp. 1–16, 2021. https://doi.org/10.1007/s10489-021-02893-3.

[5] R. Ahuja, A. Jha, R. Maurya and R. Srivastava, "Analysis of educational data mining," in *Harmony Search and Nature Inspired Optimization Algorithms, Advances in Intelligent Systems and Computing Book Series (AISC)*, vol. 741, Springer, Singapore, pp. 897–907, 2018.

[6] E. Fernandes, M. Holanda, M. Victorino, V. Borges, R. Carvalho *et al.,* "Educational data mining: Predictive analysis of academic performance of public school students in the capital of Brazil," *Journal of Business Research*, vol. 94, pp. 335–343, 2019.

[7] R. S. Baker, "Challenges for the future of educational data mining: The baker learning analytics prizes," *Journal of Educational Data Mining*, vol. 11, no. 1, pp. 1–17, 2019.

[8] B. Bakhshinategh, O. R. Zaiane, S. ElAtia and D. Ipperciel, "Educational data mining applications and tasks: A survey of the last 10 years," *Education and Information Technologies*, vol. 23, no. 1, pp. 537–553, 2018.

[9] C. A. Palacios, J. A. R. Suárez, L. A. Bearzotti, V. Leiva and C. Marchant, "Knowledge discovery for higher education student retention based on data mining: Machine learning algorithms and case study in Chile," *Entropy*, vol. 23, no. 4, pp. 1–23, 2021.

[10] S. Sweta, "Educational data mining Techniques with Modern Approach," in *Modern Approach to Educational Data Mining and its Applications, SpringerBriefs in Applied Sciences and Technology*, Springer, Singapore, pp. 25–38, 2021.

[11] A. Ahmmed, T. Sultan, H. I. Shad, J. Mia and S. Mazumder, "An automated visa prediction technique for higher studies using machine learning in the context of bangladesh," in *Data Engineering for Smart Systems, Lecture Notes in Networks and Systems (LNNS)*, Springer, Singapore, vol. 238, pp. 557–567, 2022.

[12] H. C. Hung, I. F. Liu, C. T. Liang, and Y. S. Su, "Applying educational data mining to explore students' learning patterns in the flipped learning approach for coding education," *Symmetry*, vol. 12, no. 2, pp. 1–14, 2020.

[13] V. Hegde and P. P. Prageeth, "Higher education student dropout prediction and analysis through educational data mining," in *2018 2nd Int. Conf. on Inventive Systems and Control (ICISC)*, Coimbatore, India, pp. 694–699, 2018.

[14] A. Sarra, L. Fontanella and S. Di Zio, "Identifying students at risk of academic failure within the educational data mining framework," *Social Indicators Research*, vol. 146, no. 1–2, pp. 41–60, 2019.

[15] H. El Aouifi, M. El Hajji, Y. Es-Saady and H. Douzi, "Predicting learner's performance through video sequences viewing behavior analysis using educational data-mining," *Education and Information Technologies*, vol. 26, no. 5, pp. 5799–5814, 2021.

[16] G. Ramaswami, T. Susnjak and A. Mathrani, "On developing generic models for predicting student outcomes in educational data mining," *Big Data and Cognitive Computing*, vol. 6, no. 1, pp. 1–16, 2022.

[17] K. N. Krishnanand and D. Ghose, "Glowworm swarm optimization for simultaneous capture of multiple local optima of multimodal functions," *Swarm Intelligence*, vol. 3, no. 2, pp. 87–124, 2009.

[18] S. Zheng, J. Gai, H. Yu, H. Zou and S. Gao, "Software defect prediction based on fuzzy weighted extreme learning machine with relative density information," *Scientific Programming*, vol. 2020, pp. 1–18, 2020.

[19] K. Zervoudakis and S. Tsafarakis, "A mayfly optimization algorithm," *Computers & Industrial Engineering*, vol. 145, pp. 1–79, 2020.

[20] M. A. M. Shaheen, H. M. Hasanien, M. S. El Moursi and A. A. El-Fergany, "Precise modeling of PEM fuel cell using improved chaotic MayFly optimization algorithm," *International Journal of Energy Research*, vol. 45, no. 13, pp. 18754–18769, 2021.

[21] M. A. Duhayyim, R. Marzouk, F. N. Al-Wesabi, M. Alrajhi, M. A. Hamza *et al.,* "An improved evolutionary algorithm for data mining and knowledge discovery," *Computers, Materials & Continua*, vol. 71, no. 1, pp. 1233–1247, 2022.