Received: 1 April 2022 | Revised: 16 August 2022 | Accepted: 12 September 2022

DOI: 10.1002/cae.22572

RESEARCH ARTICLE

WILEY

Predicting individual learning performance using machine-learning hybridized with the teaching-learning-based optimization

Mehrdad Arashpour[1] | **Emad M. Golafshani**[1] | **Rajendran Parthiban**[1] | **Julia Lamborn**[1] | **Alireza Kashani**[2] | **Heng Li**[3] | **Parisa Farzanehfar**[4]

[1]Faculty of Engineering, Monash University, Melbourne, Australia

[2]Faculty of Engineering, University of New South Wales, Sydney, Australia

[3]Department of Building and Real Estate, Hong Kong Polytechnic University, Hong Kong, Hong Kong

[4]Department of Medicine, Florey Institute of Neuroscience and Mental Health, Parkville, Australia

Correspondence

Mehrdad Arashpour, Department of Civil Engineering, Monash University, VIC 3800, Australia.

Email: Mehrdad.arashpour@monash.edu

Funding information

Monash University, Grant/Award Number: MDFI2021

Abstract

Reliable prediction of individual learning performance can facilitate timely support to students and improve the learning experience. In this study, two well-known machine-learning techniques, that is, support vector machine (SVM) and artificial neural network (ANN), are hybridized by teaching–learning-based optimizer (TLBO) to reliably predict the student exam performance (fail-pass classes and final exam scores). For the defined classification and regression problems, the TLBO algorithm carries out the feature selection process of both ANN and SVM techniques in which the optimal combination of the input variables is determined. Moreover, the ANN architecture is determined using the TLBO algorithm parallel to the feature selection process. Finally, four hybrid models containing anonymized information on both discrete and continuous variables were developed using a comprehensive data set for learning analytics. This study provides scientific utility by developing hybridized machine-learning models and TLBO, which can improve the predictions of student exam performance. In practice, individual performance prediction can help to advise students about their academic progress and to take appropriate actions such as dropping units in subsequent teaching periods. It can also help scholarship providers to monitor student progress and provision of support.

KEYWORDS

artificial neural networks (ANN), final exam scores, machine-learning methods, student engagement, support vector machines (SVM), teaching-learning-based optimizer (TLBO)

1 | INTRODUCTION

Early prediction of the individual learning performance facilitates the provision of necessary support [15]. This is particularly important in tertiary education settings where tailoring learning pathways can assist students in enhancing performance [37]. Provision of targeted assessment adjustments for those students at high risk of dropping out is another significant result in predicting student exam performance [38]. At the institution level, proper staffing and allocation of tutors can be targeted towards areas in which predicted performance needs short-term and long-term improvements.

Recent advancements in data mining and data availability from learning management systems (LMS) are beneficial for academic performance predictions [8, 21, 34]. Ethical considerations, however, should be the focal point of analyzing LMS data [17]. For the first step, the educational institution should ask students to provide consent to collecting and using data [26]. Additionally, detailed information about the data storage period and options to remove the data earlier should be provided to students [29]. Moreover, the use of data mining and learning analytics (e.g., for predicting exam performance) should not result in the creation of bias for teaching teams, and the single aim should be the provision of early support to students [1, 7].

Students' performance in final assessments often exhibits a strong correlation with engagement levels, demographics, and achieved results in ongoing assessments [10, 18, 28]. Previous studies have analyzed limited factors such as age/gender and investigated their impacts on academic performance [3, 30]. However, holistic modeling and analyzing comprehensive data sets have been done in sparse studies [13]. Holistic modeling can result in the accurate prediction of academic performance [14]. This will ensure the provision of timely support to students and improvements in learning experience/retention rates [4, 6].

There are two mainstreams in learning analytics research for modeling and predicting academic performance using statistical techniques and machine-learning (ML)-based models [12]. The commonly used statistical techniques include discriminant analysis [23], multiple regression [36], and stepwise regression [2, 20]. However, the assumption of data normality and variances being homoscedastic deteriorates the prediction performance of statistical techniques in handling real-world data [24]. ML techniques have been widely used in education research due to their excellent performance in handling large data sets. The most common algorithms for predicting academic performance include but are not limited to K-nearest neighbors [22], artificial neural network (ANN) [35], logistic regression [31], decision trees [32], support vector machine (SVM) [19], and Bayesian network [11]. Previous ML modeling results have shown that several factors, including demographics, engagement levels, and ongoing assessment scores, can affect future performance in final exams [33].

In the above-mentioned studies, the researchers used ML techniques to model the academic performance of students by considering a predefined number of effective input variables that were often identified based on previous studies or statistical-based methods. The selection of the effective input variables is a feature selection problem that can be defined as an optimization problem and solved using metaheuristic optimization algorithms. This study uses the teaching–learning-based optimizer (TLBO) [27] algorithm as a powerful metaheuristic optimization algorithm to carry out the feature selection task. This optimization algorithm considers different combinations of input variables in which the performance of each combination of input variables is examined using an ML method. For ML techniques, ANN and SVM techniques are used and hybridized by the TLBO to improve academic performance predictions. To maximize the generalizability of hybrid models, a comprehensive data set was utilized to represent a wide range of academic performances and input variables. The paper has the following structure. Section 2 focuses on the data used for the development of performance predictive models. Section 3 discusses the optimization and ML algorithms. Section 4 provides details of the proposed hybrid models. Section 5 focuses on the interpretation of results and discussions of implications. Conclusions and research limitations are provided in section 6.

2 | DATA DESCRIPTION

In this study, the Open University data set for learning analytics [16] has been selected and utilized. The main drivers for this selection are comprehensiveness of the data set and appropriate anonymization of students' records that satisfies ethical requirements. The data set contains anonymized information about discrete variables, including gender, age brackets, disability status, index of multiple deprivation (IMD), prior qualifications, and attempt counts in passing subjects. Table 1 provides data descriptions of the discrete variables and mapping against final exam results. As can be seen, the biggest age group is under 35 years (73.87%), followed by 35–55 years (24.94%) and older than 55 years (1.19%). A total of 329 students (9.53%) reported a level of disability. The majority of students live in regions with IMD scores between 20% and 80%. Within the whole cohort, 306 (8.86%) of enrolled students live in the most deprived regions and 762 (22.07%) live in the least deprived regions. The prior qualification for most students is at level A (46%), followed by lower A (49.19%). Around 17% of the population has a higher education (HE) qualification and a minority of less than 1% holds postgraduate qualifications.

The Open University data set for learning analytics also contains anonymized information about continuous variables, including the number of credit points, student engagement level (clickstream data within the learning management system), and academic performance

TABLE 1 Data descriptions of the discrete variables

Variables	Classes	Total Frequency	%	Exam result classes: Fail Frequency	%	Pass Frequency	%
Gender	Male	2145	62.14	752	35.06	1393	64.94
	Female	1307	37.86	439	33.59	868	66.41
Age (years)	≤35	2550	73.87	926	36.31	1624	63.69
	35 < Age ≤ 55	861	24.94	259	30.08	602	69.92
	>55	41	1.19	6	14.63	35	85.37
Disability	No	3123	90.47	1071	34.29	2052	65.71
	Yes	329	9.53	120	36.47	209	63.53
Index of multiple deprivation (IMD) (%)	IMD ≤ 20	306	8.86	135	44.12	171	55.88
	20 < IMD ≤ 40	817	23.67	301	36.84	516	63.16
	40 < IMD ≤ 60	806	23.35	303	37.59	503	62.41
	60 < IMD ≤ 80	761	22.05	234	30.75	527	69.25
	IMD > 80	762	22.07	218	28.61	544	71.39
Qualification	No formal	16	0.46	8	50.00	8	50.00
	Lower A	1121	32.47	527	47.01	594	52.99
	Level A	1698	49.19	505	29.74	1193	70.26
	Higher education	596	17.27	148	24.83	448	75.17
	Postgraduation	21	0.61	3	14.29	18	85.71
Previous attempts	0	2852	82.62	877	30.75	1975	69.25
	1	457	13.24	230	50.33	227	49.67
	>1	143	4.14	84	58.74	59	41.26

(quiz and final exam scores). Table 2 presents continuous variables and related descriptive statistics.

As Table 2 shows, students are on average enrolled in 80 credit points (a measure of study load). Clickstream data (engagement level) shows a mean value of 1223 clicks with a maximum of 15,931 on record. Average scores for in-term assessments are presented for six tests. For instance, the first quiz score has a mean value of 70.11 with a standard deviation of 23.02. The average score for the final exam is 63.20, with a standard deviation of 18.83.

3 | THE OPTIMIZATION AND ML ALGORITHMS

This section describes the optimization and ML algorithms used in this study. The TLBO algorithm is explained at first and then ANN and SVM techniques for modeling the final exam classes and scores of students are explained thoroughly.

3.1 | TLBO algorithm

TLBO algorithm, proposed by Rao et al. [27], is a metaheuristic optimization algorithm that mimics the learning process of trainees during the course of study by a teacher. Similar to other population-based optimization algorithms, the TLBO algorithm looks for the optimal solution of an optimization problem using a set of solutions and through an iterative process. In this algorithm, each student is a symbol of a possible solution and his/her knowledge represents the cost function.

Two stages are carried out in each iteration of the TLBO algorithm: (1) the teacher stage and (2) the learner stage. In the teacher stage, a good teacher attempts to reach the knowledge levels of students to his/her knowledge level for a subject. Because of the different learning capabilities of students, it is not possible to enhance the knowledge levels of all students to a certain level. However, a good teacher can increase the average knowledge level of a class. Suppose S_i^t is the ith student with the knowledge level L_i^t in the tth iteration. Because

TABLE 2 Data descriptions of the continuous variables

Statistical parameters	Variable: Credit points	Engagement level	1st assessment score	2nd assessment score	3rd assessment score	4th assessment score	5th assessment score	6th assessment score	Final exam score
Minimum	60	1	0	0	0	0	0	0	0
Maximum	280	15,931	100	100	100	100	100	100	100
Average	80.03	1223.33	70.11	68.42	67.21	58.33	49.47	41.18	63.20
Standard deviation	32.78	1255.80	23.02	28.35	31.08	33.28	33.48	34.55	18.83
Skewness	1.91	3.53	−1.68	−1.51	−1.28	−0.73	−0.40	−0.01	−0.18
Kurtosis	4.42	23.13	2.72	1.20	0.34	−0.86	−1.30	−1.57	−0.57

there is no information about the knowledge of the teacher in each iteration, the best student of the population (T^t) (the best solution found so far) is set as the teacher. The students are updated using the following equation:

$$S_{i,U}^t = S_i^t + r_i(T^t - T_F M^t), \tag{1}$$

where $S_{i,U}^t$ is the ith updated student and M^t is the mean values of the decision variables of students in the tth iteration. r_i is a random number between 0 and 1 and T_F is the teacher factor which can be set to 1 or 2 randomly. This equation attempts to move the mean knowledge of students towards the teacher's knowledge. If the cost value of $S_{i,U}^t$ is less than that of S_i^t, $S_{i,U}^t$ replaces the S_i^t in the $(t+1)$th iteration, otherwise S_i^t remains in the next population.

In the learner stage, students can improve their knowledge of each other through cooperation. For this, the ith student (S_i^t) randomly selects the jth student (S_j^t) of the population in the tth iteration and tries to improve his/her knowledge level based on the selected student's knowledge level. If the knowledge level of the ith student (L_i^t) is better than that of the jth student (L_j^t), he/she tries to update himself/herself with more attention to his/her position and vice versa, formulated as follows:

$$S_{i,U}^t = S_i^t - sign\left(L_i^t - L_j^t\right) r_i \left(S_i^t - S_j^t\right), \tag{2}$$

where $sign$ is the sign function and r_i is a random number between 0 and 1. If the updated student performs better than its previous status, it transfers to the next iteration; otherwise, its previous status is kept in the next population. The pseudocode of the TLBO algorithm is shown in Figure 1.

3.2 | ANN

To model the behavior of a system, ANN utilizes the concept of biological neurons of the human brain and their connections for doing activities. In multilayer ANN, the network is arranged in several successive layers, including input, hidden, and output layers, as demonstrated in Figure 2. The layers encompass several artificial neurons consisting of computational and noncomputational ones which are successively connected to each other by links called weights. The neurons of the input layer are the noncomputational neurons that represent the input variables of the system. As the computational neurons, the neurons of the hidden and output layers carry out the linear and nonlinear computations of the network. For this, each computational neuron sums the weighted input signals and activates this value using a transfer function.

The number of hidden layers and their related neurons determines the architecture of ANN. For a given architecture of an ANN model, the training process of an ANN model is to find the optimal weights of the ANN through an iterative process so that the prediction errors of the network are minimized. To do so, the feed-forward computation is carried out from the input to the output layers to calculate the network predictions for given input variables. Based on the prediction errors of the network, the weights of the model are modified using back-propagation optimization algorithms. For classification problems, one binary node is usually considered for each class of the system's output and the predicted outputs are rounded for continuously bounded transfer functions. In the case of regression problems, a neuron is assigned to each output variable of the network. This study uses sigmoid transfer functions for the hidden neurons of both classification and regression ANNs. Besides, softmax and linear transfer functions are served for the output layers of the classification and regression ANNs, respectively.

Begin
Set the students number (n_{class}) and maximum iteration number ($n_{Maxiteration}$)
Initialize students' population
Calculate knowledge level of students
For t =1:MI
  Calculate mean values of decision variables of students (M^t)
  % Teacher stage
  Set the best student of the population as the teacher (T^t)
  For i =1:N
    Calculate the ith updated student $S_{i,U}^t$ using Eq. (1)
    Check the boundaries of the decision variables of $S_{i,U}^t$
    Calculate the knowledge level of $S_{i,U}^t$ ($L_{i,U}^t$)
    If $L_{i,U}^t \leq L_i^t$
      Replace S_i^t by $S_{i,U}^t$
    Endif
  Endfor
  % Learner stage
  For i =1:N
  Select a random student of the population (S_j^t)
  Calculate the ith updated student $S_{i,U}^t$ using Eq. (2)
  Check the boundaries of the decision variables of $S_{i,U}^t$
  Calculate the knowledge level of $S_{i,U}^t$ ($L_{i,U}^t$)
    If $L_{i,U}^t \leq L_i^t$
      Replace S_i^t by $S_{i,U}^t$
    Endif
  Endfor
Endfor
Save the best student
Finish

FIGURE 1 Pseudocode of the teaching–learning-based optimizer algorithm

3.3 | SVM and support vector regression (SVR)

SVM is a powerful supervised ML algorithm developed to solve classification problems [9]. This algorithm aims to find the best hyperplane, so that maximizes the margin of data points of different classes, defined as follows:

$$w^T x - b = 0, \tag{3}$$

where w and b are the weight vector and bias of the hyperplane, respectively, and x is the input vector. Besides, the training data points that determine the maximum margin of SVM are called support vectors. In the linearly separable binary SVM classifier with two classes of −1 and +1, $\|w\|$ should be minimized subject to $y_i(w^T x_i - b) \geq 1$, in which x_i and y_i are the ith input vector and prediction,

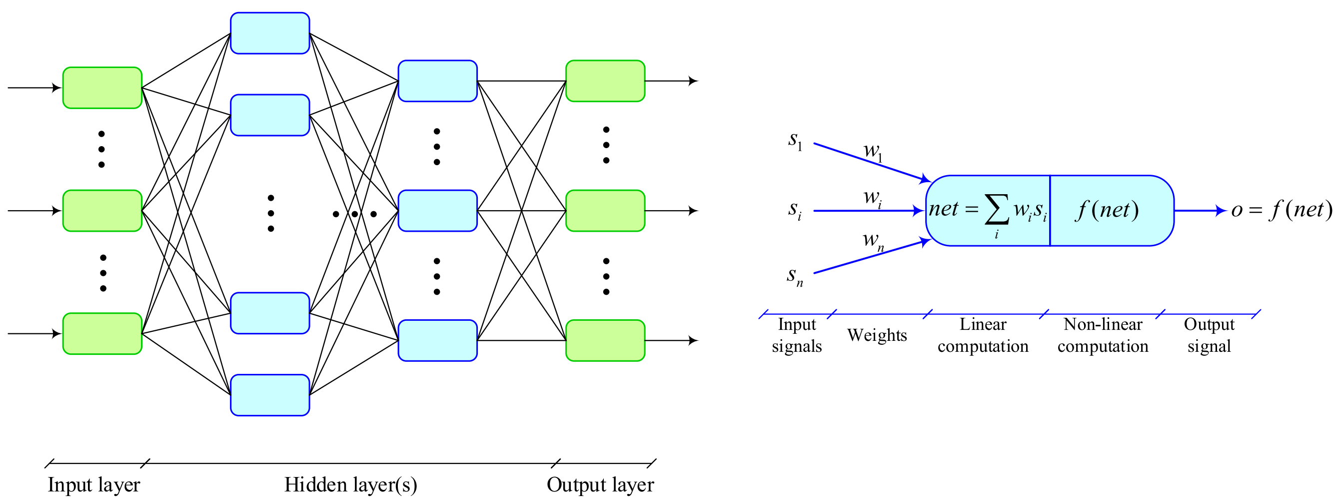


FIGURE 2 A schematic representation of an artificial neural network model and computations of an artificial neuron

respectively. For nonlinearly separable classification problems, the soft-margin SVM classifier is defined in which data points with incorrect class predictions are penalized by ξ_i. The primal optimization problem in this new condition can be rewritten as follows:

$$\min \quad \frac{1}{2}w^T w + C\sum_i \xi_i, \tag{4}$$

s.t.

$$y_i(w^T x_i - b) \geq 1 - \xi_i,$$

where C is the penalty coefficient and $\xi_i \geq 0$ is defined as $\max(0,1 - y_i(w^T x_i - b))$. To solve the defined optimization problem, the dual form of the primal problem is written using the Lagrange coefficients α_i, as follows:

$$\max \quad \sum_i \alpha_i - \frac{1}{2}\sum_i \sum_j \alpha_i \alpha_j y_i y_j \left(x_i^T x_j\right), \tag{5}$$

s.t.

$$\sum_i \alpha_i y_i = 0,$$

$$0 \leq \alpha_i \leq C.$$

To have a more efficient SVM model, the kernel trick can be used in which the data points are mapped into higher-dimension linearly separable space using kernel functions K. The revised dual objective function can be written as follows:

$$\max \quad -\frac{1}{2}\sum_i \sum_j \alpha_i \alpha_j y_i y_j K(x_i, x_j) + \sum_i \alpha_i. \tag{6}$$

To solve the above optimization problem, it is required to use a very high-computational and time-consuming quadratic programming (QP) algorithm. In this study, the sequential minimal optimization (SMO) algorithm served as the fast optimization technique in which the main QP is broken into several small possible QP problems, and the generated problems are solved analytically [5, 25]. The optimal Lagrange coefficients with values more than 0 and less than C specify the support vectors of the SVM model.

SVR is an extension of the SVM algorithm dealing with regression problems which aim to find the linear regression model of the high-dimensional input space, expressed as follows:

$$y = w^T x + b. \tag{7}$$

In SVR, an ε-insensitive loss function is defined, expressed as follows:

$$L_\varepsilon(t_i, y_i) = \begin{cases} 0, & |t_i - y_i| \leq \varepsilon \\ |t_i - y_i| - \varepsilon, & \text{Otherwise}, \end{cases} \tag{8}$$

where ε is a positive number. The data points with errors higher than ε are penalized and the mean penalized value of all data points determines the empirical risk of the SVR model. For each data point, two positive slack variables ξ_i^- and ξ_i^+ are defined so that $\xi_i^- \xi_i^+ = 0$ and show the violations of the ith data point from ε. The primal optimization problem defined for the SVR algorithm can be expressed as follows:

$$\min \quad \frac{1}{2}w^T w + C\sum_i \left(\xi_i^- + \xi_i^-\right), \tag{9}$$

s.t.

$$-\varepsilon - \xi_i^- \leq t_i - y_i \leq \varepsilon + \xi_i^+$$

$$\xi_i^-, \xi_i^+ \geq 0,$$

By defining the Lagrange coefficients α_i^- and α_i^+, the dual optimization problem is defined as follows:

$$\begin{aligned} \max \quad & -\frac{1}{2}\sum_i \sum_j \left(\alpha_i^+ - \alpha_i^-\right)\left(\alpha_j^+ - \alpha_j^-\right) x_i^T x_j \\ & + \sum_i \left(\alpha_i^+ - \alpha_i^-\right) t_i - \varepsilon \sum_i \left(\alpha_i^+ - \alpha_i^-\right), \end{aligned} \tag{10}$$

s.t.

$$\sum_i \left(\alpha_i^+ - \alpha_i^-\right) = 0,$$

$$0 \leq \alpha_i^-, \alpha_i^+ \leq C.$$

Serving the kernel trick, the revised objective function is rewritten as follows:

$$\begin{aligned} \max \quad & -\frac{1}{2}\sum_i \sum_j \left(\alpha_i^+ - \alpha_i^-\right)\left(\alpha_j^+ - \alpha_j^-\right) K\left(x_i, x_j\right) \\ & + \sum_i \left(\alpha_i^+ - \alpha_i^-\right) t_i - \varepsilon \sum_i \left(\alpha_i^+ - \alpha_i^-\right). \end{aligned} \tag{11}$$

The SMO algorithm can be used to solve the defined dual optimization problem and the optimal values of Lagrange coefficients are obtained. The data points with $0 < \alpha_i^-, \alpha_i^+ < C$ make the support vectors and the final SVR model is expressed as follows:

$$y = \sum_i \left(\alpha_i^+ - \alpha_i^-\right) K(x_i, x) + b. \quad (12)$$

3.4 | The proposed TLBO–ML algorithm

This study aims to serve the TLBO algorithm for developing the final exam classification and regression models using two ML techniques, including ANN and SVM. For classification and regression ANN models, two tasks are expected from the TLBO algorithm: (1) Feature selection and (2) ANN architecture determination. In the feature selection process, the optimized combinations of the input variables are achieved. For this, the TLBO algorithm selects a subset of input variables and changes them through a sophisticated process so that the prediction errors of the ANN, SVM, and SVR are minimized.

In addition, the optimized ANN architecture is determined by the TLBO algorithm parallel to the feature selection process. For SVM and SVR methods, the TLBO technique is only served to determine the optimal combination of the input variables through the feature selection process. Figure 3 depicts different steps of the proposed TLBO–ML algorithm and more descriptions are presented in the following.

- The control parameters regarding the ML and TLBO algorithms are set at first, including the maximum and minimum numbers of hidden layers and related neurons, optimization method, repetition number (n_{rep}) of ML methods' runs for each architecture, and ratios of training, validation, and testing subsets for ML algorithms, as well as the maximum iteration number and population size for the TLBO algorithm.
- The selected data set is split randomly into three subsets, including training, validation, and testing subsets. The training subset is used in the training phase of ML methods in which the effective input variables, architecture, and weights of ANN models are optimized. For the SVM and SVR techniques, the effective input variables and support vectors are determined in the training phase. The validation subset is served to inhibit the over-training of the ANN model and to select the best-developed SVM and SVR models among the repeated models for a certain set of input variables. To verify the developed ML models, the testing subset is employed to examine the accuracy of the developed models against unknown data.
- For the classification and regression ANN methods, each student of the TLBO algorithm contains three main sections, as illustrated in Figure 4. The first section of each student is related to the feature selection task. In this regard, a binary decision variable is assigned to each potential input variable. The values of zero and one of the decision variables show whether the corresponding input variable takes part in the ANN training process or not, respectively. The second and third sections of each student, represented as the binary and discrete decision variables, respectively, are attributed to the ANN architecture determination task. The numbers of decision variables in the second and third sections are the same and equal to the maximum number of hidden layers defined earlier. Each decision variable of the second section is related to one hidden layer and specifies if the hidden layer is active or not. For each hidden layer in the second section of a student, there is a unique decision variable in the third section representing the number of hidden neurons. For a system with the potential input variable number of n_i and the maximum hidden layer number of n_h, the number of decision variables (student size) is $n_i + 2n_h$. The representation of a student with the format of (1-1-0-1-0-0-1-1-0-8-6) for a system with n_i and n_h equal to, respectively, 7 and 2 indicates an ANN model with only one hidden layer with eight hidden neurons constructed by the first, second, fourth, and seventh potential input variables of the data set. It can give this opportunity to various architectures of ANN models to be modeled parallelly with certain input variables and the knowledge among the input and output variables is discovered. In the case of SVM and SVR algorithms, similar to the first section of a student in ANN, each student carries out the feature selection task, and its size is equal to the number of potential input variables.
- Because of the random nature of the ANN, SVM, and SVR techniques, a certain set of input variables is run several times, and the best model in the training and validation phases is saved. To evaluate the knowledge level of each student, the accuracy function and root mean squared error (RMSE) are used for the classification and regression models, defined later.
- In each iteration of the TLBO algorithm, new students are generated using the teacher and learner stages. The generated students are evaluated and then transferred into the next iteration if they perform better than their previous status. This process is followed whenever the termination conditions of the algorithm are satisfied. In this study, the maximum iteration number and the maximum successively unsuccessful iterations served as the termination conditions of the TLBO algorithm.

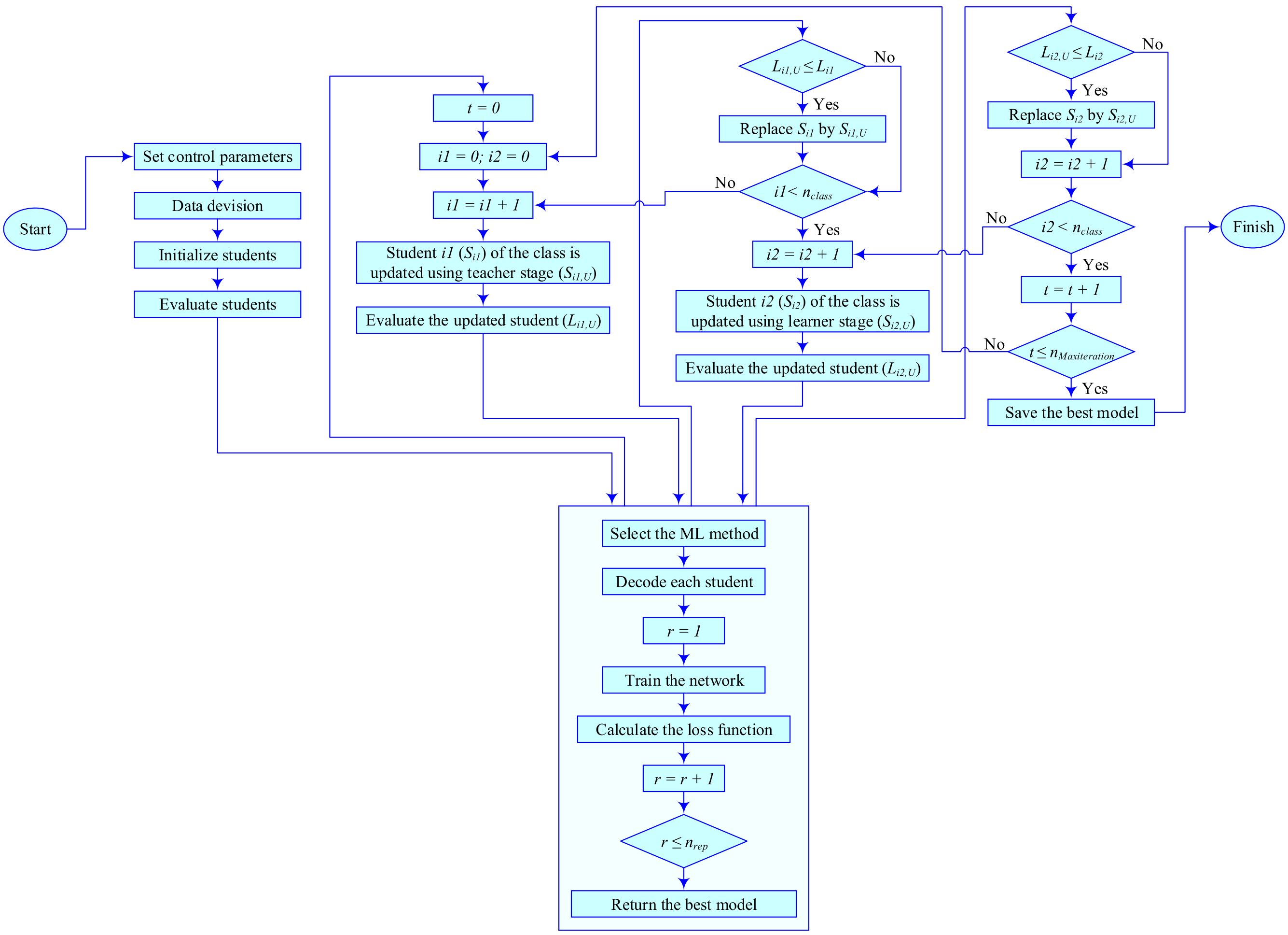


FIGURE 3 The flowchart of the proposed teaching–learning-based optimizer-based machine-learning algorithm

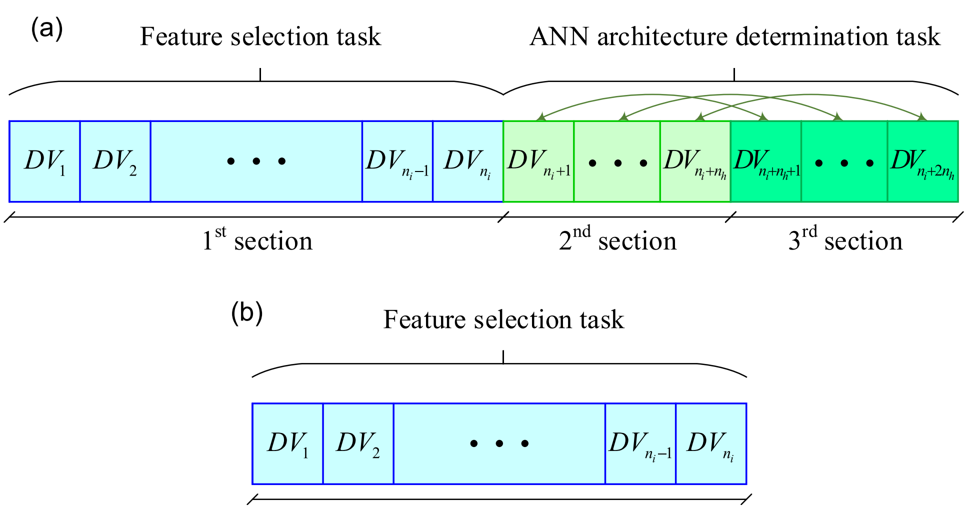


FIGURE 4 Decision variables (DVs) of a student in the teaching–learning-based optimizer algorithm defined for (a) artificial neural network (ANN) models and (b) support vector machine and support vector regression models

4 | THE PROPOSED TLBO–ML MODEL DEVELOPMENT

To develop the final exam result classification and regression models using the proposed algorithm, the control parameters of ML and TLBO algorithms should be set first. The control parameter values chosen in this study are given in Table 3. It is obvious that there are limited numbers of control parameters for the TLBO algorithm compared to other metaheuristic optimization algorithms showing its simplicity. Because of the randomness character of the TLBO, the algorithm was run 20 times, and the best ML models were selected as the final models. Moreover, the ANN architecture without any hidden layer was also considered in this study. For the SVM and SVR models, the student size is 14, while for the ANN models, the student size is 18, including 14 decision variables for the feature selection task and four decision variables for the ANN architecture determination task.

After coding and running the proposed ML model in the MATLAB environment, the final classification ANN model has two hidden layers with 23 and 16 neurons in the first and second hidden layers, respectively. Among 14 potential input variables, only student age does not take part in the optimal model. For the SVM model, eight input variables consisting of the number of credit points, all assessment scores (six variables), and engagement level are the most crucial variables affecting the pass and fail status of the final exam score of students. In the case of the final regression ANN model, the numbers of hidden neurons of the first and second hidden layers are 28 and 30, respectively. In this model, all 14 input variables, excluding IMD and the second assessment score, influence the final exam score of students. For the optimal SVR model, only IMD does not affect the final exam score of students.

5 | RESULTS AND DISCUSSION

The receiver operator characteristic (ROC) curve is used to demonstrate the detection capability of classification models by plotting the true-positive against the false-positive rates. The classification model close to the top-left corner of the ROC curve shows better performance than the classifier close to the 45° line. Figure 5 illustrates the ROC curves of the ANN and SVM models for the fail-pass prediction of the final exam of students for both training and testing phases. For the training phase, the area under the curve (AUC) of the developed

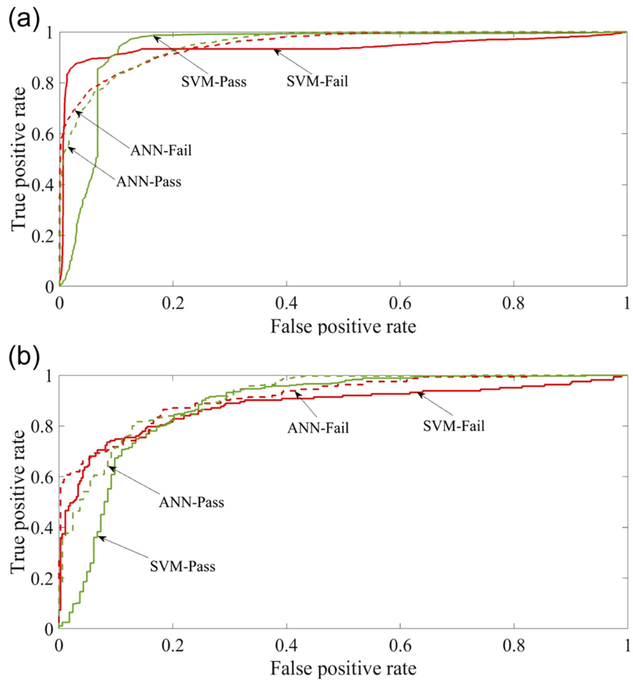


FIGURE 5 Receiver operator characteristics of the developed SVM and ANN models for (a) training and (b) testing phases. ANN, artificial neural network; SVM, support vector machine.

TABLE 3 Control parameters of the TLBO–ML algorithm adjusted in this study

TLBO parameters	Values	ML parameters	Values
Repetition number	20	Repetition number	20
Maximum iteration number	200	Training subset ratio	0.7
population size	50	Validation subset ratio	0.15
Maximum successively	20	Testing subset ratio	0.15
unsuccessful iteration		Range of hidden layer numbers of ANN	[0–2]
		Range of hidden node numbers of ANN	[1–30]
		Training algorithm of ANN	Levenberg- Marquardt

Abbreviations: ANN, artificial neural network; ML, machine-learning; TLBO, teaching–learning-based optimizer.

SVM and ANN models are 0.94 and 0.95, respectively. The AUC of the developed SVM and ANN models are 0.88 and 0.91, correspondingly, for the testing phase. As shown in this figure and by comparing AUC values, ANN and SVM models demonstrate comparable performances.

The confusion matrix is a graphical figure summarizing the prediction capabilities of a classifier. The confusion matrixes of the developed SVM and ANN models for classifying the students' performance are depicted in Figure 6. As shown, the prediction quality of the developed SVM model is better than the developed ANN model for both the training and testing phases.

Performance measures are used to evaluate the efficiency of the developed classification and regression ML models for correct final exam classification and score, respectively. In this study, Precision, Recall, Accuracy, F1-Score, Matthews correlation coefficient (MCC), and Fowlkes-Mallows index (FM) are served as the performance measures of the developed classification ML models. RMSE, scatter index (SI), mean bias error (MBE), mean absolute error (MAE), correlation coefficient (R), and Tstat are employed as the performance measures of the developed regression ML models, given in Table 4.

The performance measures of the developed classification and regression ML models for training, testing, and all data are given in Table 5. The best values of each data set are shown in bold font in this table. For the classification models, the SVM model outperforms the ANN model in most measures for all phases. Because of the competitive performances of the SVM and ANN models in the Precision and Recall measures, FM and

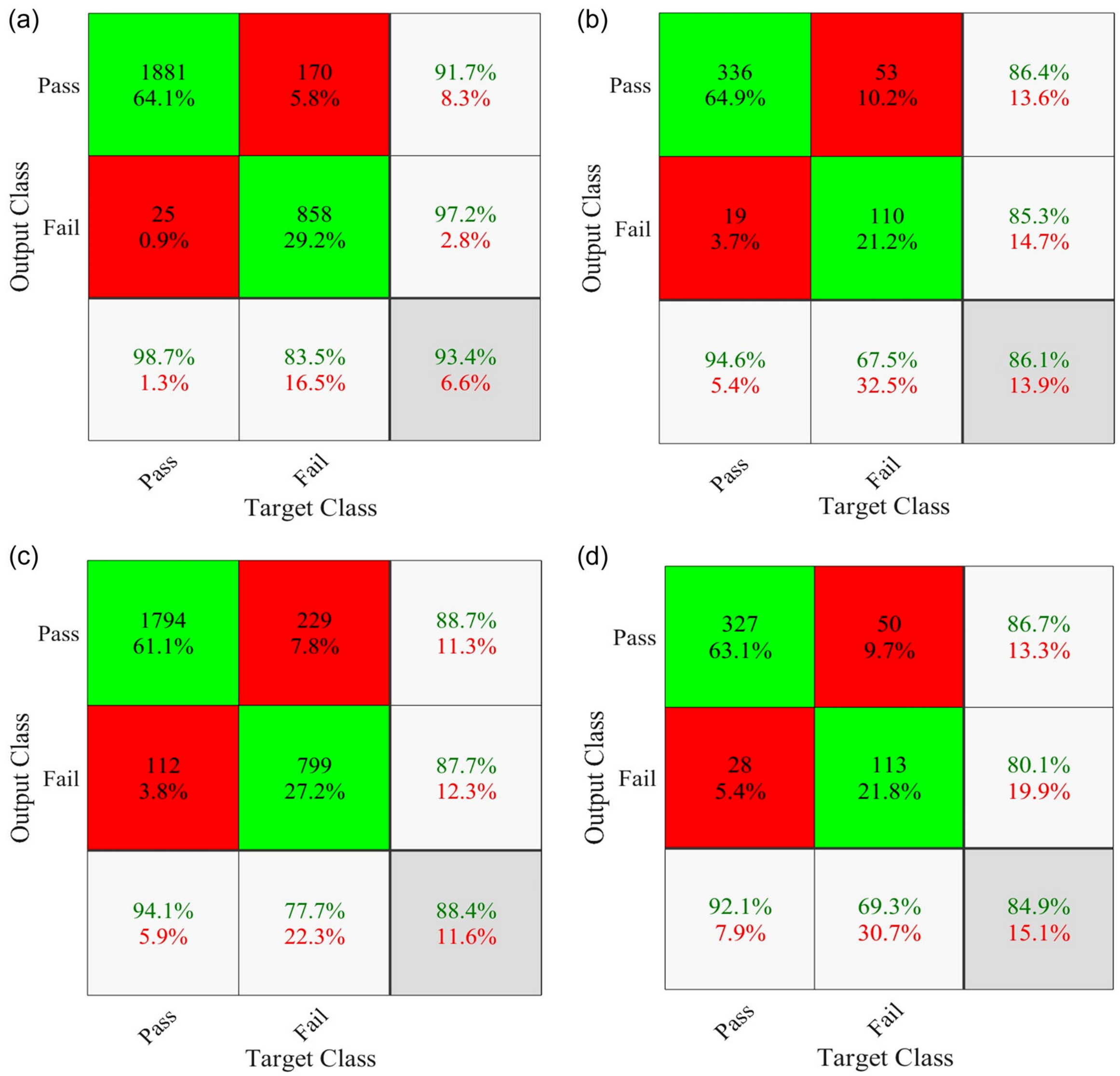


FIGURE 6 Confusion matrixes for (a) SVM-training phase, (b) SVM-testing phase, (c) ANN-training phase, and (d) ANN-testing phase. ANN, artificial neural network; SVM, support vector machine.

TABLE 4 Classification and regression performance measures used in this study

Types	Performance measures	Equations
Classification performance measures	Precision	$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}$
	Recall	$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}}$
	Accuracy	$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{FN} + \text{FP} + \text{TN}}$
	FM	$\text{FM} = \sqrt{\text{Precision} \times \text{Recall}}$
	F1-Score	$\text{F1} - \text{Score} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$
	MCC	$\text{MCC} = \frac{\text{TP} \times \text{TN} - \text{FP} \times \text{FN}}{\sqrt{(\text{TP}+\text{FP})(\text{TP}+\text{FN})(\text{TN}+\text{FP})(\text{TN}+\text{FN})}}$
Regression performance measures	RMSE	$\text{RMSE} = \frac{1}{\text{DN}}\sum_i (Y_i - T_i)^2$
	SI	$\text{SI} = \frac{\text{RMSE}}{\bar{T}}$
	MBE	$\text{MBE} = \frac{1}{\text{DN}}\sum_i (Y_i - T_i)$
	MAE	$\text{MAE} = \frac{1}{\text{DN}}\sum_i \lvert Y_i - T_i \rvert$
	R	$R = \frac{\sum_i (Y_i - \bar{Y})(T_i - \bar{T})}{\sqrt{\sum_i (Y_i - \bar{Y})^2 \sum_i (T_i - \bar{T})^2}}$
	Tstat	$T\text{stat} = \sqrt{\frac{(\text{DN} - 1)\text{MBE}^2}{(\text{RMSE}^2 - \text{MBE}^2)}}$

Abbreviations: DN, data number; FM, Fowlkes–Mallows index; FN, false negatives; FP, false positives; MAE, mean absolute error; MBE, mean bias error; MCC, Matthews correlation coefficient; R, correlation coefficient; RMSE, root mean squared error; SI, scatter index; $\bar{T}$, mean target scores; T_i, target scores; TN, true negatives; TP, true positives; $\bar{Y}$, mean predicted scores; Y_i, predicted scores.

F1-Score that define the combination of these measures are useful indicators. The performance of the SVM model is slightly better than the ANN model, comparing FMs and F1-Scores. The accuracy of the SVM model is almost 6%, 1%, and 5% better than the accuracy of the ANN model for the training, testing, and all data sets, respectively. Comparing the MCCs indicates that the correlation coefficient of the SVM model is significantly higher than the correlation coefficient of the ANN model in the training phase. However, there are negligible differences between the MCCs of these classifiers for the testing and all data sets. What distinguishes the performance of the SVM model compared to the ANN model is the number of input variables identified by the TLBO algorithm. The proposed SVM model with eight effective input variables led to a more reliable model compared to the ANN model with 13 identified influential input variables.

In the case of regression ML models, the performance of the SVR model is better than the regression ANN model in all data sets. In the training phase, the SVR model is about 6% and 15% better than the regression ANN model, respectively, comparing RMSE and MAE. For the testing phase, the SVR model performs almost 10% and 6% better than the ANN model, respectively, comparing RMSE and MAE. It means that the SVR model outperforms the ANN model in both the training and testing phases. The results of the regression models show that the developed SVR and ANN models have R values of more than 0.7 for the training and all data sets denoting the high correlation between the target and predicted scores. However, the correlation coefficients of developed models are lower in the testing phase compared to the other phases. MBE measure shows the mean prediction error of the developed models. The MBEs of the SVR model are significantly closer to zero than those of the ANN model. Besides, the developed ANN model underestimates the final exam score of students in all phases, while the SVR model overestimates and underestimates the final exam score for the testing and training phases, respectively. The SI values between 0.1 and 0.2 of the SVR and ANN models for the training and all data show that the regression ML models possess "good prediction capability," while their performances are "fair" in the testing phase. *Tstat* is a dimensionless measure with the combination of RMSE and MBE measures, and there is a distinguished distinction between the performances of SVR and ANN models.

To find the relationship between the input variables and fail–pass status and final exam score of students, an impact analysis was carried out based on the best

TABLE 5 Performance measures of the developed models for the training, testing, and all data

Classification models	Phases	Classification performance measures					
		Precision	Recall	Accuracy	FM	F1-Score	MCC
SVM model	Training data	91.71	98.69	93.35	0.95	95.07	0.85
	Testing data	94.65	86.38	86.10	0.90	90.32	0.67
	All data	92.14	96.60	92.27	0.94	94.32	0.82
ANN model	Training data	94.12	88.68	88.38	0.91	91.32	0.74
	Testing data	92.11	86.74	84.94	0.89	89.34	0.64
	All data	93.81	88.38	87.86	0.91	91.01	0.73
Regression models	**Phases**	**Regression performance measures**					
		RMSE	**SI**	**MBE**	**MAE**	***R***	***T* stat**
SVR model	Training data	11.52	0.18	0.24	8.16	0.79	0.99
	Testing data	15.80	0.25	−0.26	12.81	0.54	0.33
	All data	12.25	0.19	0.16	8.85	0.76	0.69
ANN model	Training data	12.26	0.19	−0.35	9.60	0.76	1.35
	Testing data	17.56	0.28	−0.90	13.62	0.48	1.02
	All data	13.19	0.21	−0.43	10.20	0.72	1.69

Abbreviations: ANN, artificial neural network; FM, Fowlkes–Mallows index; FN, false negatives; FP, false positives; MAE, mean absolute error; MBE, mean bias error; MCC, Matthews correlation coefficient; *R*, correlation coefficient; RMSE, root mean squared error; SI, scatter index; SVM, support vector machine; SVR, support vector regression.

classification and regression ML models, respectively. For the classification problem, the values of each identified input variable by the TLBO algorithm were perturbed between their allowable ranges while other variables were unchanged, and the predicted classes before and after perturbation were monitored. Next, the probability percentage change of each identified variable was calculated, as given in Figure 7. The increment of the number of enrolled credit points from 60 to 100 increases the probability percentage change by about 7%, followed by almost a 3% reduction for the enrolled credit points between 100 and 120. Next, the probability percentage change raises up to about 17% for the number of credit points of 170. For the first and second assessments, it does not be observed a significant improvement in the increment of scores from 0 to 100. Whatever goes to the last assessment, the effect of the assessment scores on the final fail-pass status of students increases. The students with scores of between almost 15 and 30 in the third last assessment are unlikely to pass the final exam. Besides, achieving high scores in the fifth and sixth assessments can alone enhance the passing chance of the final exam up to about 29%. Concerning the impact of engagement level, the students with a click number of less than 750 during the study period do not have a high chance of passing the subject. The impact of engagement level for the click numbers between about 750 and 1900 is considerable and it can reach the passing probability of about 29% for the click number of 5000.

For the regression problem, the mean predicted final exam scores after and before the perturbation of each identified variable were recorded, and the mean change percentages after the perturbation are calculated. In this study, the Wilcoxon test was carried out to find whether the difference between the estimated scores before and after perturbations are significant or not, with the associated p value of .05. Table 6 gives the impact analysis results for the input variables identified by the TLBO algorithm. The p value of less than .05 indicates a meaningful difference between the performances of the male and female students.

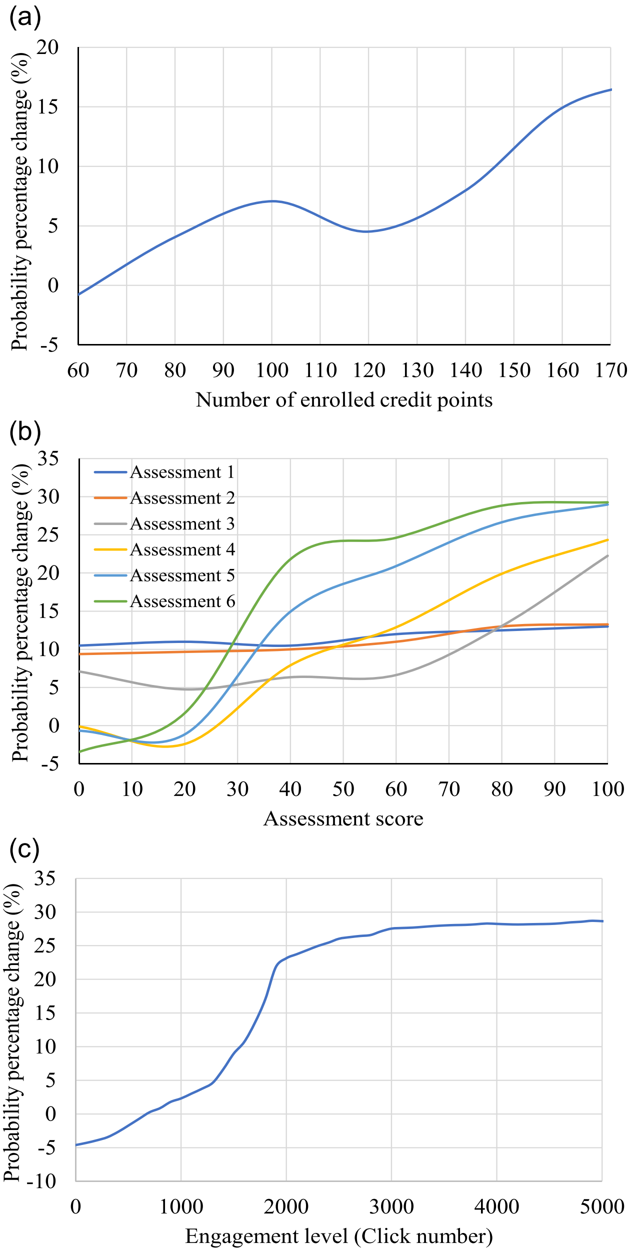


FIGURE 7 Impact analysis of the identified input variables influencing the final score exam in the classification problem (a) number of enrolled credit points, (b) assessment score and (c) Engagement level

Moreover, student age is observed to have an inverse relationship with academic performance. Moreover, there is not a significant difference between the performances of disabled and nondisabled students. The students with higher educational backgrounds demonstrate better performance than those with lower educational backgrounds. The number of previous attempts inversely affect the final exam score. The number of enrolled credit points does not have a meaningful effect on the final exam score. The engagement level of fewer than 100 clicks does not have a

TABLE 6 Sensitivity analysis of the identified input variables influencing the final exam score in the regression problem

Variables	Classes	Mean score	Mean change (%)	p Value (Wilcoxon test)
Gender	Female	60.89	-	-
	Male	65.33	+7.29	<.05
Age (years)	≤35	63.63	-	-
	35 < Age ≤ 55	62.67	−1.51	<.05
	>55	60.80	−4.45	<.05
Disability	No	63.20	-	-
	Yes	63.51	+0.05	>.05
Qualification	Lower A	61.21	-	-
	Level A	63.95	+4.48	<.05
	Higher education	63.24	+3.31	<.05
Previous attempts	0	63.21	-	-
	1	61.43	−2.83	<.05
	>1	59.43	−5.99	<.05
Credit points	60	63.29	-	-
	100	63.83	+0.85	>.05
	120	62.93	−0.57	>.05
	170	63.73	+0.69	>.05
Engagement level	10	61.44	-	-
	100	61.64	+0.31	>.05
	500	62.45	+1.63	<.05
	1000	63.29	+2.99	<.05
	2000	64.04	+4.21	<.05
Assessment 1	0	58.73	-	-
	50	57.89	−1.43	>.05
	100	65.90	+12.20	<.05
Assessment 2	0	59.45	-	-
	50	58.56	−1.50	>.05
	100	65.01	+9.36	<.05
Assessment 3	0	57.14	-	-
	50	58.24	+1.91	<.05
	100	65.32	+14.32	<.05
Assessment 4	0	55.44	-	-
	50	57.55	+3.80	<.05
	100	67.79	+22.28	<.05

(Continues)

TABLE 6 (Continued)

Variables	Classes	Mean score	Mean change (%)	p Value (Wilcoxon test)
Assessment 5	0	58.59	-	-
	50	60.35	+3.01	<.05
	100	67.80	+15.72	<.05
Assessment 6	0	59.17	-	-
	50	61.55	+4.02	<.05
	100	68.54	+15.83	<.05

significant impact on the final score. For more than 100 clicks, the effect of the engagement level is considerable. The scores of less than 50 for the first and second assessments do not significantly influence the final score. While achieving high scores in the first and second assessments can improve the final score by 12.20% and 9.36%, respectively. The final exam scores of students can improve by 1.91%, 3.80%, 3.01%, and 4.02% with the increment of the scores from 0 to 50 for the third, fourth, fifth, and sixth assessments, respectively. However, increasing the scores of the third, fourth, fifth, and sixth assessments from 50 to 100 can enhance the final exam scores by 14.32%, 22.28%, 15.72%, and 15.78%, respectively. It is obvious that the effects of the assessment scores between 50 and 100 are higher than 0 to 50, especially for the fourth assessment. Besides, the influence of the assessment scores on the final exam score is higher than other input variables, and the disability and the number of enrolled credit points do not have a meaningful impact on the final score.

6 | CONCLUSIONS

Important factors affecting academic performance can be traced in terms of low engagement (fewer than average clickstream records) and lower than average scores in ongoing assessments. Early analysis of these important factors results in the timely provision of support to the students regarding their academic progress and taking suitable actions such as underloading. It can also serve as a predictor to provide scholarship support. In this study, hybrid models were developed using the combination of TLBO and two ML methods, that is, ANN and SVM. Open University data set for learning analytics [16] was utilized, and the following conclusions were drawn:

- For the classification of student exam performance, the proposed SVM model with eight effective input variables leads to a more reliable model compared to the ANN model with 13 identified influential input variables.
- In the case of regression ML models, the SVR model outperforms the ANN model in both the training and testing phases. The results of the regression models show that the developed SVR and ANN models have R values of more than 0.7 for the training and all data sets denoting the high correlation between the target and predicted scores. The scatter index values between 0.1 and 0.2 of the SVR and ANN models for the training and all data show that the regression ML models possess "good prediction capability," while their performances are "fair" in the testing phase.
- Impact analysis shows that achieving high scores in the ongoing assessments closer to the final exam significantly increases the chance of ending up in the pass category for the subject. Concerning the impact of engagement level, the students with a click number of less than 750 during the study period do not have a high chance of passing the subject. The impact of engagement level for the click numbers between about 750 and 1900 is considerable, and it can reach the passing probability of about 29% for the click number of 5000.
- Overall, the influence of the continuous assessment on the final exam score is higher than other input variables.

As observed in this study, the development of hybrid optimization and ML algorithms provides the opportunity for early prediction of academic performance, which in turn facilitates the timely provision of support to students. In addition, the most impactful input variables are identified, with the developed classification, and regression models demonstrate acceptable performances based on monitored error metrics.

Some limitations in the current study should be reported. First, clickstream data to represent the level of study engagement was used. This is acceptable when the delivery of subjects is mainly online. Future research can investigate other avenues for capturing and analyzing students' engagement in different delivery modes. Second, in this study, hybrid models of TLBO optimization, and ML algorithms (ANN and SVM) were developed, and their prediction performances were evaluated. In the future, other types of hybrid models can be developed to achieve more reliable predictions of academic performance.

ACKNOWLEDGMENTS

Open access publishing facilitated by Monash University, as part of the Wiley - Monash University agreement via the Council of Australian University Librarians.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the corresponding author upon reasonable request.

ORCID

Mehrdad Arashpour http://orcid.org/0000-0003-4148-3160

REFERENCES

1. J. Angelva, T. Tepsa, and M. Mielikäinen, *Team teaching experiences in engineering education a project-based learning approach*, 45th Annu. Conf. Eur. Soc. Eng. Educ., SEFI 2017, J. Bernardino, J. Rocha, and J. C. Quadrado (eds.), European Society for Engineering Education (SEFI), Brussels, Belgium, 2017, pp. 1182–1189.
2. M. Arashpour, V. Kamat, A. Heidarpour, M. R. Hosseini, and P. Gill, *Computer vision for anatomical analysis of equipment in civil infrastructure projects: theorizing the development of regression-based deep neural networks*, Autom. Constr. **137** (2022), 104193.
3. M. Arashpour, J. Lamborn, and P. Farzanehfar, *Optimising collaborative learning and group work amongst tertiary students*. In: Ozevin D, ed. 10th Internat. Struct. Eng. Constr. Conf., ISEC, ISEC Press, Chicago, IL, USA, 2019.
4. M. Arashpour, J. Lamborn, and P. Farzanehfar, *Group dynamics in higher education: impacts of gender inclusiveness and selection interventions on collaborative learning*. In: Mostafa S, Rahnamayiezekavat P, eds. Claiming identity through redefined teaching in construction programs, IGI Global, Hershey, Pennsylvania, 2020.
5. M. Arashpour, T. Ngo, and H. Li, *Scene understanding in construction and buildings using image processing methods: a comprehensive review and a case study*, J. Build. Eng. **33** (2021), 101672.
6. J. Berg, A. Gottburgsen, and B. Kleimann, *Formalising organisational responsibility for refugees in German higher education: the case of first contact positions*, Stud. High. Educ. **47** (2022), no. 6, 1243–1255.
7. M. Brown and C. Klein, *Whose data? which rights? whose power? A policy discourse analysis of student privacy policy documents*, J. Higher Educ. **91** (2020), 1149–1178.
8. A. Cohen, *Analysis of student activity in web-supported courses as a tool for predicting dropout*, Educ. Technol. Res. Dev. **65** (2017), 1285–1304.
9. C. Cortes and V. Vapnik, *Support-vector networks*, Machine Learning **20** (1995), 273–297.
10. I. N. Z. Day, F. M. Van Blankenstein, P. M. Westenberg, and W. F. Admiraal, *University teachers' conceptions of their current and ideal intermediate assessment: an A+ is good, but speaking your mind is better*, Stud. High. Educ. **44** (2019), 2223–2234.
11. S. de Klerk, B. P. Veldkamp, and T. J. H. M. Eggen, *Psychometric analysis of the performance data of simulation-based assessment: a systematic review and a Bayesian network example*, Comput. Educ. **85** (2015), 23–34.
12. C. C. Gray and D. Perkins, *Utilizing early engagement and machine learning to predict student outcomes*, Comp. Educ. **131** (2019), 22–32.
13. C. Herodotou, B. Rienties, A. Boroowa, Z. Zdrahal, and M. Hlosta, *A large-scale implementation of predictive learning analytics in higher education: the teachers' role and perspective*, Educ. Technol. Res. Dev. **67** (2019), 1273–1306.
14. T. Icekson, O. Kaplan, and O. Slobodin, *Does optimism predict academic performance? Exploring the moderating roles of conscientiousness and gender*, Stud. High. Educ. **45** (2020), 635–647.
15. A. Jokhan, B. Sharma, and S. Singh, *Early warning system as a predictor for student performance in higher education blended courses*, Stud. High. Educ. **44** (2019), 1900–1911.
16. J. Kuzilek, M. Hlosta, and Z. Zdrahal, *Open university learning analytics dataset*, Scientific Data **4** (2017), 170171.
17. K. H. Kyritsi, V. Zorkadis, E. C. Stavropoulos, and V. S. Verykios, *The pursuit of patterns in educational data mining as a threat to student privacy*, J. Interact. Media Educ. **2019** (2019), 2019.
18. J. S. Lee, *The relationship between student engagement and academic performance: is it a myth or reality?* J. Educ. Res. **107** (2014), 177–185.
19. S. Li, S. P. Lajoie, J. Zheng, H. Wu, and H. Cheng, *Automated detection of cognitive engagement to inform the art of staying engaged in problem-solving*, Comput. Educ. **163** (2021), 104114.
20. J. Lim and J. C. Richardson, *Predictive effects of undergraduate students' perceptions of social, cognitive, and teaching presence on affective learning outcomes according to disciplines*, Comput. Educ. **161** (2021), 104063.
21. Q. Liu and S. Geertshuis, *Professional identity and the adoption of learning management systems*, Stud. High. Educ. **46** (2021), 624–637.
22. F. Marbouti, H. A. Diefes-Dux, and K. Madhavan, *Models for early prediction of at-risk students in a course using standards-based grading*, Comput. Educ. **103** (2016), 1–15.
23. P.-F. Pai, Y.-J. Lyu, and Y.-M. Wang, *Analyzing academic achievement of junior high school students by an improved rough set model*, Comput. Educ. **54** (2010), 889–900.
24. M. Paliwal and U. A. Kumar, *A study of academic performance of business school graduates using neural network and statistical techniques*, Expert Sys. Appl. **36** (2009), 7865–7872.
25. J. Platt. *Sequential minimal optimization: a fast algorithm for training support vector machines*. 1998.
26. P. Prinsloo, S. Slade, and M. Khalil, *Student data privacy in MOOCs: a sentiment analysis*, Distance Educ. **40** (2019), 395–413.
27. R. V. Rao, V. J. Savsani, and D. Vakharia, *Teaching–learning-based optimization: a novel method for constrained mechanical design optimization problems*, Comput.-Aided Des. **43** (2011), 303–315.
28. T. Rashid and H. M. Asghar, *Technology use, self-directed learning, student engagement and academic performance: examining the interrelations*, Comput. Human Behav. **63** (2016), 604–612.
29. J. R. Reidenberg and F. Schaub, *Achieving big data privacy in education*, Theory Res. Educ. **16** (2018), 263–279.
30. J. T. E. Richardson and A. Woodley, *Another look at the role of age, gender and subject as predictors of academic attainment in higher education*, Stud. High. Educ. **28** (2003), 475–493.
31. M. Riestra-González, M. D. P. Paule-Ruíz, and F. Ortin, *Massive LMS log data analysis for the early prediction of course-agnostic student performance*, Comput. Educ. **163** (2021), 104108.
32. S. Rizvi, B. Rienties, and S. A. Khoja, *The role of demographics in online learning; A decision tree based approach*, Comput. Educ. **137** (2019), 32–47.

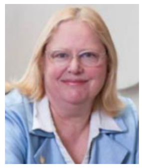
33. J.-E. Russell, A. Smith, and R. Larsen, *Elements of success: supporting at-risk student resilience through learning analytics*, Comput. Educ. **152** (2020), 103890.
34. D. A. Thomas and M. Nedeva, *Broad online learning EdTech and USA universities: symbiotic relationships in a post-MOOC world*, Stud. High. Educ. **43** (2018), 1730–1749.
35. N. Tomasevic, N. Gvozdenovic, and S. Vranes, *An overview and comparison of supervised data mining techniques for student exam performance prediction*, Comput. Educ. **143** (2020), 103676.
36. F. H. Wang, *Interpreting log data through the lens of learning design: second-order predictors and their relations with learning outcomes in flipped classrooms*, Comput. Educ. **168** (2021), 104209.
37. L. Wijnia, S. M. M. Loyens, E. Derous, N. S. Koendjie, and H. G. Schmidt, *Predicting educational success and attrition in problem-based learning: do first impressions count?* Stud. High. Educ. **39** (2014), 967–982.
38. K. L. Wilson, K. A. Murphy, A. G. Pearson, B. M. Wallace, V. G. S. Reher, and N. Buys, *Understanding the early transition needs of diverse commencing university students in a health faculty: informing effective intervention practices*, Stud. High. Educ. **41** (2016), 1023–1040.

AUTHOR BIOGRAPHIES

Mehrdad Arashpour is the Head of Construction Engineering Discipline at Monash University. He started his academic career from Grenoble University, France. Mehrdad then researched and lectured at the Royal Melbourne University of Technology before joining Monash University. His lab (ASCII) undertakes research on artificial intelligence, modern engineering education, robotics, prefabricated structures and vision technologies. He also conducts research on education projects that focus on inclusive learning and engineering communities of practice. Mehrdad has published several book chapters and more than 120 papers in top-tier journals and international conference proceedings. He received a teaching citation for Outstanding Contribution to Student Learning from Monash Engineering in 2019. He is an Australian Chartered Professional Engineer (CPEng), National Engineering Registrant and International Professional Engineer (IntPE).

Emadaldin M. Golafshani is a researcher in Civil Engineering at Monash University. As a multidisciplinary scholar, he has broad-based knowledge in civil engineering, construction and engineering management, transportation engineering, materials science, and especially artificial intelligence. His focus in artificial intelligence is the development of novel machine-learning and metaheuristic optimization algorithms.

Rajendran Parthiban is the Associate Dean (Education) of the Faculty of Engineering at Monash University. He received a teaching citation for Outstanding Contribution To Student Learning from the Australian Learning and Teaching Council in 2008. He is a senior member of Optica and the Institute of Electrical and Electronic Engineering (IEEE).



Julia Lamborn is the Director of Engineering Accreditation at Monash University. She received her BEng (Civil), Graduate Diploma, MEng and PhD in Environmental Engineering, all from the Swinburne University of Technology. Julia's educational research focuses on the relevance of mathematics to 1st-year students in engineering courses and critical elements for successful student projects.

Ali Kashani is a Senior Lecturer (Assistant Professor) in Sustainable Concrete and 3D printing and a Churchill Fellow in Construction 3D Printing with extensive experience in research, development, and commercialisation of advanced and sustainable construction materials. His main research areas are construction automation and sustainable construction materials for a circular economy.

Heng Li is a Chair Professor of Construction Informatics at the Hong Kong Polytechnic University. Heng started his academic career at Tongji University since 1987. Heng then researched and lectured at the University of Sydney, James Cook University and Monash University before joining Hong Kong Polytechnic University. He has conducted many funded research projects related to the innovative application and transfer of construction information technologies, and he has published two books, more than 300 journal papers and numerous conference papers.

Parisa Farzanehfar is a Medical Doctor currently practising at Northern Health Hospital (Melbourne). She also holds a PhD in Neuroscience and is an honorary research fellow at Florey Institute of Neuroscience, University of Melbourne. She is passionate about enhancing tertiary education and has published papers on communities of practice in medical settings and optimising collaborative learning and group work.

How to cite this article: M. Arashpour, E. M. Golafshani, R. Parthiban, J. Lamborn, A. Kashani, H. Li, and P. Farzanehfar, *Predicting individual learning performance using machine learning hybridized with the teaching-learning-based optimization*, Comput. Appl. Eng. Educ. 2023;31:83–99. https://doi.org/10.1002/cae.22572