IEEE Access
Multidisciplinary Rapid Review Open Access Journal

Received December 16, 2021, accepted January 10, 2022, date of publication January 13, 2022, date of current version January 19, 2022.

Digital Object Identifier 10.1109/ACCESS.2022.3143035

An Efficient Data Mining Technique for Assessing Satisfaction Level With Online Learning for Higher Education Students During the COVID-19

HANAN E. ABDELKADER[1], AHMED G. GAD[2], (Graduate Student Member, IEEE), AMR A. ABOHANY[2], AND SHAYMAA E. SOROUR[3]

[1]Faculty of Specific Education, Mansoura University, Mansoura 35516, Egypt
[2]Faculty of Computers and Information, Kafrelsheikh University, Kafrelsheikh 33516, Egypt
[3]Faculty of Specific Education, Kafrelsheikh University, Kafrelsheikh 33516, Egypt

Corresponding authors: Ahmed G. Gad (ahmed.gad@fci.kfs.edu.eg) and Shaymaa E. Sorour (shaymaasorour@gmail.com)

ABSTRACT All the educational organizations mainly aim at elevating the academic performance of students for improving the overall quality of education. In this direction, Educational Data Mining (EDM) is a rapidly trending research area that utilizes the essence of Data Mining (DM) concepts to help academic institutions figure out useful information on the Student Satisfaction Level (SSL) with the Online Learning process (OL) during COVID-19 lock-down. Different practices have been tried with EDM to predict students' behaviors to reach the best educational settings. Therefore, Feature Selection (FS) is typically employed to find the most relevant subset of features with minimum cardinality. As the predictive accuracy of a satisfaction model is significantly influenced by the FS process, the effectiveness of the SSL model is elaborately studied in this paper in connection with FS techniques. In this connection, a dataset was first collected online via a questionnaire of student reviews on OL courses. Using this datatset, the performance of wrapper FS techniques in DM and classification algorithms was analyzed in terms of fitness values. Ultimately, the goodness of subsets with different cardinalities is evaluated in terms of prediction accuracy and number of selected features by measuring the quality of 11 wrapper-based FS algorithms and the k-Nearest Neighbor (k-NN) and Support Vector Machine (SVM) as base-line classifiers. Based on the experiments, the optimal dimensionality of the feature subset was revealed, as well as the best method. The findings of the present study evidently support the well-known conjunction of the existence of minimum number of features and an increase in predictive accuracy. It is remarkable the relevancy of FS for high-accuracy SSL prediction, as the relevant set of features can effectively assist in deriving constructive educational strategies. Our study contributes a feature size reduction of up to 80% along with up to 100% classification accuracy on the adopted real-time dataset.

INDEX TERMS Classification, COVID-19, educational data mining (EDM), feature selection (FS), machine learning (ML), online learning (OL), student satisfaction level (SSL).

I. INTRODUCTION

The COVID-19 epidemic emergence has sparked enormous public-health worries. Many countries have therefore decided lock-downs to decrease social contact and limit transmission due to these emergency circumstances [1], [2]. Higher Education Institutions (HEIs) have been highly affected by the COVID-19. Based on the impacts of this epidemic and the need for alternative solutions, many unconventional educational solutions have been proposed to preserve the educational process continuity. One of the solutions was Online Learning (OL), known as learning in a synchronous or asynchronous situation by utilizing various equipment, such as mobile devices, computers, etc. with an Internet connection. Students can learn and communicate with teachers and other peers from any place by applying these platforms [3]. OL has increased in the recent decade as it provides better flexibility in time and location, speed of study,

The associate editor coordinating the review of this manuscript and approving it for publication was Ali Shariq Imran.

more accessibility, more active access to a wide range and a better quantity of knowledge, as well as the low monetary costs [4].

OL platforms and courses were the most notable aspect of the changeover. However, we continue to face various barriers and hurdles. Although robust digital infrastructure and platforms are required for online course delivery and data gathering interactivity, low Internet connectivity disrupts student learning worldwide. Both students and teachers need new technology to seamlessly interact with active and self-directed learning. Therefore, a reliable assessment method was a must to ensure the overall quality of online education in terms of learning outcomes. In the epidemic age, quality is measured by the achievement of learning objectives and the development of the emotional and social aspects [5], [6]. This consequently requires a tool to assess the learning process as a whole and the roles and interactions of teachers, learners, and teaching resources in post-digital learning environments specifically. The ability of universities and colleges to meet learning objectives and increase Student Satisfaction Level (SSL) is critical because these factors indirectly demonstrate the efficiency of OL systems at those institutions [7]. SSL can be described as a component of the relative degree of experiences and perceived performance related to educational services during the study time. In a nutshell, SSL is determined by how students rate their educational experiences, services, and facilities [8]. According to [9], SSL can only be achieved when there is no gap between what is expected and what is introduced by the service provider.

In this regard, it is necessary to point out that Educational Data Mining (EDM) could be very beneficial for the educational research process [10], [11]. The collected data should thus be structured and thoroughly evaluated to offer interpretable results. Choosing the suitable approach for the analysis is also critical to the success of EDM methods. Feature Selection (FS) is counted one of the most efficient, relevant data analytics tools. In applied models, high-dimensional data might lead to some negative consequences. These include increasing training time with advanced features and model processing [12]. FS performs a vital role in Machine Learning (ML) and Data Mining (DM), especially with high-dimensional datasets that include noisy, redundant, and irrelevant features. FS aims to select a subset of variables from the inputs that can more accurately characterize the data while minimizing the impact of noise and irrelevant features while still providing high predictive results. The feature subset selection is an essential issue in knowledge discovery and acceleration of DM techniques and performance improvement [13], [14]. FS has also been a successful and efficient data preparation strategy for preprocessing the high-dimensional data in many DM and ML problems. FS goals may span creating more precise models, increasing data-mining speed, and delivering comprehensible data [15]

It is worth noting that FS is utilized for many reasons: first, to simplify data for straightforward interpretation; second, to take the shortest time to select the most significant features; third, to overcome the curse of dimensionality; and finally, to develop generalization via hyper-allocation reduction. Once these four factors are met, the deceptive, ineffective, and redundant features are significantly reduced, resulting in a well-productive collection of features [16]. As a crucial stage in DM, FS is one of the first steps in tackling numerous challenges, like data classification [17], image-processing [18], data clustering [19], [20], bioinformatics data analysis [21], and fault diagnosis [22], [23].

The evaluation criteria and the search approach are the two primary merits that characterize FS methods. In terms of evaluation criteria, there are two kinds of FS methods: filter techniques and wrapper techniques. When assessing the constructed feature subsets, a learning algorithm (such as classifiers) is used to distinguish these two methodologies. Filter approaches concentrate on correlations between conditional features and classes. Relief [24], and hybrid Relief [25] are two popular FS filters. Wrapper approaches are evaluated using a learning algorithms, such as k-Nearest Neighbors (k-NN) [26] and Support Vector Machine (SVM) [27], which produces more precise findings than filter methods. In addition, the filter methods may need more computational resources than the wrapper ones. Moreover, the performance of wrapper method greatly depends on the learning algorithm [28]. Because the FS is categorized as an NP-hard checking issue, finding the ideal (minimum) feature subset when working with high-dimensional data is not an easy task [29]. As a result, due to their individualities (e.g., population-based search, solution diversity, exploration-exploitation capabilities), meta-heuristic approaches have been often more appropriate for handling this complex problem than exact methods [30], [31], and they were therefore opted to tackle the FS problem in this study. Various global search strategies have been developed to address the FS problem. Swarm Intelligence (SI) is a bit class of meta-heuristics mimicked by natural occurrences. There are numerous instances of SI algorithms used to solve FS problems, including Artificial Bees Colony (ABC) [6] Grey Wolf Optimizer (GWO) [13], Harmony Search Algorithm [32] and Salp Swarm Algorithm [33]. New SI algorithms recently developed, the Antlion Optimizer (ALO) is used as a wrapper method for the FS method [34]. In a word, meta-heuristic optimization algorithms are continuously gaining popularity in engineering applications due to their efficiency. They are based on simple and easy to execute concepts; do not need gradient information; avoid local optima; and apply to diverse challenges from several areas [35].

Thus, we can conclude that current environment requires extensive and innovative research and thinking about every area of education during epidemics, as all HEIs mainly aim to improve the education quality and enhance the overall performance. EDM is a significant study subject that assists HEIs in meeting their objectives. Therefore, this study aims to measure SSL with OL services using FS techniques with a real-time dataset collected from heterogeneous groups via an educational platform. The authors collected this dataset through a questionnaire prepared solely to define how students are affected by measuring their SSL with OL during the COVID-19.

In this paper, an ML model was developed to get the highest accuracy results throughout experimentation. Two ML Classifiers were adopted, k-NN and SVM, to build the proposed model. Historically, k-NN and SVM are the most widely employed classification systems [36]–[38]. In addition, this study used 11 meta-heuristic algorithms to identify the critical features in the used dataset, enabling to build a new model to predict SSL more correctly and reliably. These include a group of classical, state-of-the-art, and recent algorithms such as ABC [39], Particle Swarm Optimization (PSO) [40], Bat Algorithm (BA) [41], GWO [13] Whale Optimization Algorithm (WOA) [35] Grasshopper Optimization Algorithm (GOA) [42] SailFish Optimizer (SFO) [43], Harris Hawks Optimization (HHO) [44], Bird Swarm Algorithm (BSA) [45], Atom Search Optimization [46], and Henry Gas Solubility Optimization (HGSO) [47]

In empirical sense, the findings of this paper can be used to improve OL services in the future so that universities can build more inclusive, sustainable, and equitable education once the epidemic has gone away. Thus, this paper contributes to future development and increases universities and college's capability to achieve the goals of OL in terms of high SSL, to enhance the effectiveness of the OL system. Following are the key contributions of this work:

1. Proposed a new real-time dataset to determine and extract relevant information regarding SSL with OL to foresee whether the students are satisfactory or not during the COVID-19 outbreak.
2. Experimenting 11 meta-heuristic algorithms for FS to select the most critical features from the dataset at hands.
3. Testing two comparable ML classification algorithms (k-NN and SVM) through experimental evaluations in terms of the number of selected features, accuracy, fitness, and computational time, both for the whole features and those elected by each of the 11 algorithms.
4. The best model with the highest accuracy was then finally realized to estimate students' satisfaction likelihood of OL.

The remainder of this paper is organized as follows. Section II introduces the literature review. The proposed techniques are presented in Section III. In Section IV, findings and their analysis of experimental results are addressed. Finally, Section VI concludes the paper with implications stemmed from results.

II. LITERATURE REVIEW

This section reviews the most recent literature based on the effect of the COVID-19 in education community: Section II-A gives an overview of the COVID-19 and OL in HEIs, SSL with OL is introduced in Section II-B, and Section II-C discusses the FS process and EDM and its application to SSL estimation.

A. ONLINE LEARNING IN HIGHER EDUCATION DURING THE COVID-19

As a result of the COVID-19 epidemic, many countries discontinued in-person classes and shifted to various OL settings. Numerous studies have been conducted in the education field to check the impact of the COVID-19 on the educational process [48]–[52]. Underlined the necessity for further investigation to report the epidemic's impact on the educational process. This emphasized the need for HEIs to develop robust methods to meet students' learning requirements beyond the traditional classroom. In this regard, Qazi *et al.* [53] investigated that remote learning adoption comes from preparedness and situational perception gained through trust in OL information sources, such as media information and interpersonal OL communication programs related to perception. Clark *et al.* [54] concluded that children who received online education during the COVID-19 academically performed better than those who did not receive physical educational assistance from their HEI. Maqableh and Alia [55] assessed the impact of OL on undergraduate students during the COVID-19 epidemic and examined the OL's positive and negative points from student perspectives compared to traditional learning. The findings showed that students faced several challenges when transitioning to OL during the COVID-19 epidemic: technological, time management, and education. In addition, they proposed recommendations to improve student OL experiences. For example, for students and instructors to benefit from OL, educational institutions must provide global training. During the epidemic, university academic decisions must also consider obstacles that may face students' communication.

In the time of COVID-19 pandemic, each university must provide campus health and safety measures for all staff and students. Hussein *et al.* [56] conducted experiments in the United Arab Emirates. The research was conducted to investigate undergraduate student attitudes toward OL. Participants were asked to fill in semi-guided essays. The responses were analyzed, and according to the research findings, OL has both positives and negatives. The benefits include safety, appropriateness, time, and increased participation. The negatives (issues) are Internet connectivity, reduced focus, and technical aspects. Kapasia *et al.* [11] conducted experiments from West Bengal colleges and Indian universities. They gathered data from 232 undergraduate and postgraduate students to examine the impact of the COVID-19 crisis. They paid a special attention to the OL status and problems. The results concluded that students encountered some issues while participating in OL due to the COVID-19 crisis. They suggested that colleges and universities can adopt a strategy for education continuity so as to deliver continuous high-quality OL, in order to alleviate the overall education system quality.

B. ASSESSMENT OF PIVOTAL ELEMENTS AFFECTING STUDENT SATISFACTION WITH THE ONLINE LEARNING (OL)

Student satisfaction is an essential factor to determine whether online learners, courses, and programs work or not [57]. However, it is a complex process to design and execute an effective educational environment involving many satisfaction factors, including instructor support, student interaction and collaboration, and student autonomy [58]. Previous works demonstrated that there is a steady relationship between student participation and the satisfaction with OL. Students' most popular OL activities include online tests and quizzes, searching, and online group discussions. As a result, it is critical to improve student activities with OL [59]. Giray [60] analyzed the contentment of undergraduate students in Turkish universities during the COVID-19 epidemic. The data was examined quantitatively and qualitatively and the questionnaire garnered 290 usable responses. The findings suggested that interactivity is an essential component of satisfaction and persistence for online learners.

One of the OL quality measures is to continue with the improvement process [61]. Gopal *et al.* [62] investigated the elements affecting SSL and OL performance in the COVID-19 era. The data was collected from 544 respondents via an online survey. The proposed hypotheses were analyzed using structural equation modelling. According to the findings, instructor quality, course design, and student expectations positively improve SSL. Nevertheless, there is a lack of research on the main factors influencing the SSL and the online classroom performance, especially during the COVID-19 epidemic [63].

The COVID-19 epidemic has forced instructors to incorporate new methods into their learning style to maintain a good educational quality amid the epidemic's limitations, giving a special emphasis to SSL. From the literary treatment in this section and the precedent one, it is apparent that researches have duly highlighted the significance to investigate the factors that may influence SSL in higher and undergraduate institutions [64], [65]. However, none of the current research studies has examined the impact of course content and design, instructor quality, evaluation system, and E-Tests on the overall SSL with OL during the COVID-19 epidemic. The current study consequently strives to bridge this gap in an intelligent way based on the EDM paradigm.

C. FEATURE SELECTION (FS) AND EDUCATIONAL DATA MINING (EDM)

Academic performance assessment is a crucial procedure to ensure that students completely achieved their studies on time, perfectly as well. Many applications of ML algorithms have been proposed to predict student academic performance and satisfaction. The prediction is made by analyzing historical, educational datasets of student opinions via questionnaires. The resulting dataset comprises many attributes, which in turn increases the model's complexity and may reduce its performance because not all attributes may be relevant [66]. FS is one of the prerequisites adopted on mining student's questionnaire to assess the SSL. It mainly aims to reduce the high computational costs required by heavy mining tasks by discarding any noisy, redundant, or irrelevant features that may degrade classification accuracy [67].

Farissi and Dahlan [68] proposed a classification method based on the Genetic Algorithm (GA) to predict student academic performance. The data used in the experiment is from Kaggle.[1] The proposed method achieved impressive prediction results. Punlumjeak and Rachburee [69] compared four algorithms for selecting the best feature subsets. These algorithms include SVM, information gain, GAs, and minimum redundancy and maximum correlation. They used four supervised classifiers such as Decision Trees (DT), Naïve Bayes (NB), favorite neighbors, and Artificial Neural Networks (ANN). Ajibade *et al.* [70] provided a comparative analysis of various FS algorithms (e.g.,Discriminant Analysis (DISC) Discriminant Analysis (DISC), Sequential Forward Selection (SFS), Sequential Backward Selection (SBS), and Discriminant Analysis (DISC). NB, DT and k-NN were used to further improve the prediction accuracy of the classifier. These algorithms include related FS, relay method, Kullback-Leibler divergence, backward sequence selection, forward sequence selection, and the Differential Evolution algorithm (DE).

The studies [71]–[73] focused on FS as a method to increase the prediction model performance. They compared and analyzed the outcomes of a set of feature selectors combined with numerous classifiers. The experimental results suggested that algorithms executed better on lower dimensionality datasets. Zaffar *et al.* [74] examined the filter FS performance and classification algorithms on two independent student datasets. They applied Bayesian Network (BN), NB, NB Updateable (NBU), MLP, Simple Logistic (SL), SMO, DT, OneR J rip, Decsion Stump (DS), J48, Random Forest (RF), RandomTree (RT), and REPtree (RepT). They proposed 6 FS algorithm: CfsSubsetEval, ChiSquaredAttributeEval, FilteredAttributeEval, GainRatioAttributeEval, Principal Components, and ReliefAttributeEval. The findings indicated there was a 10% difference in prediction accuracies for datasets with various quantitative characteristics. Nguyen *et al.* [75] presented a detailed assessment of current studies using SI to acquire FS in classification, emphasizing representation and search procedures. They also recommend that researchers could explore different approaches, offer practitioners suggestions on selecting appropriate methods for usage in actual situations, and outline probable restrictions and concerns for future work. Nalić *et al.* [76] suggested a hybrid EDM model (Generalized Linear Model (GLM) + DT that integrates several FS (e.g., Classifier Feature Evaluation (ClassFE), Correlation Feature Evaluator (CorrelationFE), Gain Ratio Feature

[1]https://www.kaggle.com

Evaluator (GainRFE), Information Gain Feature Evaluator (InfoGainFE), Relief Feature Evaluator (RefilefFE)), and learning classification methods to help in decision making in the context of dimensionality reduction.

Chitra and Rashmi [77] aimed to predict academic student performance using two filter selection approaches, correlation FS (CFS) and wrapper-based FS (WFS), to demonstrate the significance of feature subset selection. For the underlined classification problem, SMO and J48 had the highest accuracy measures with the correlation FS algorithms, while naïve Bayes had the highest accuracy measures with the wrapper subset FS algorithms for predicting high, medium, and low grades for the students. The results will assist the researcher in determining the best combination of filter FS algorithms and associated classifiers. According to the above-discussed literature review, a comparison of the related work is shown in Table 1, highlighting the merits of present study compared to those ones.

Because SSL depends on the discernment and experience of students, SSL is challenging in HEIs. Therefore, SSL with OL should be adequately measured to guide the continuous improvement in quality. As a result, the HEIs must figure out aspects that may lead to beyond SSL during the COVID-19. Furthermore, FS has been demonstrated to be an excellent and efficient DM method for various ML problems. Some of the previous studies indicated that there are many studies to measure SSL. However, studies on SSL with OL during the COVID-19 period are still scarce, especially using FS techniques to analyze the SSL. This paper aims to build a more straightforward and precise model, enhance EDM performance, as well as, provide clean, interpretable results obtained from FS by assessing SSL with OL during the COVID-19.

III. PROPOSED MODEL

In this section, the proposed framework is presented in detail. The main objective is to build a prediction model that exploits ML techniques to achieve the highest accuracy results on SSL. Hence, a set of efficient meta-heuristic algorithms are applied to select the most relevant features from a given dataset. In addition, two classification techniques, k-NN and SVM, are employed to allow the system to recursively evaluate the relevance of the selected featured in terms of prediction accuracy and number of selected features. The framework of the SSL prediction system is depicted in Fig. 1. It is divided into four main stages: i) data preprocessing, ii) the FS process, iii) ML classification, and iv) evaluating the produced ML models.

A. DATA PREPROCESSING

Data preprocessing is one of the most critical processes that efficiently encode data for the ML algorithm, which must be trained and evaluated effectively. Preprocessing techniques like missing value elimination and Min-Max scalar are useful in the classification task. The Min-Max scalar algorithm shifts the data so that all feature values fall between 1 and 5. In very early stage, these preprocessing techniques are applied.

B. META-HEURISTIC ALGORITHMS FOR FS

The merit of applying wrapper-based meta-heuristic FS algorithms is that they can identify the critical features in corpus data. In the current study, we have used 11 different meta-heuristic optimization algorithms to indicate which essential features must be in to anticipate SSL successfully and which features, if removed, will enhance, or even preserve the system's prediction capabilities the highest. Each of those algorithms is briefly described in the following.

1) ARTIFICIAL BEE COLONY (ABC)

ABC algorithm is one of the most promising optimization algorithms under SI based optimization algorithms. It was presented by Karaboga [39], inspired by honey bees' intelligent behavior. Each cycle of the search in the ABC algorithm consists of three main steps: i) sending the hired bees to the food source and measuring their amount of nectar after sharing the information of the employed bees by the spectators, and ii) determining the amount of nectar in the food, based on which the final food source is selected, and iii) and then sending the scout bee to potential food sources.

2) PARTICLE SWARM OPTIMIZATION (PSO)

PSO is a simple, robust optimization algorithm modelled after animal social behavior such as fish and birds. It has been successfully used in vast domains of scientific and engineering applications, including ML, image processing, DM, robotics, and many others. For instance, Eberhart and Kennedy first introduced PSO in 1995 [40]. They proposed a model that describes the social behavior of animals, such as flocks of birds and fish. Therefore, PSO has been utilized in a variety of industries to solve optimization problems.

3) BAT ALGORITHM (BA)

BA is a relatively recent meta-heuristic optimization technique based on bat echolocation, developed by Yang in 2010 [41]. It was inspired by microbat echolocation, which has varied pulse rates of emission and loudness [78]. In addition, it is based on SI and inspired by bat observations. For example, bats often hunt at night by emitting short, loud sound impulses and listening for the echo reflecting an obstacle or prey. In addition, a bat's particular hearing apparatus can detect the size and location of an object.

4) GREY WOLF OPTIMIZATION (GWO)

In 2014, Zorarpacı and Özel [79] designed the GWO algorithm, one of the currently most popular SI-based algorithms. Grey wolves inspired the GWO algorithm in nature, which seek the most efficient way to pounce a prey. GWO algorithm follows the pack hierarchy to organize many roles in the wolf pack. Based on the wolf role that aids in the hunting process, GWO separates pack members into

TABLE 1. Comparison of the related works against present study.

Author	Classifier(s) employeed	FS algorithm(s) applied	Best accuracy	Selected features (% ratio)
Farissi *et al.* [68]	DT, ANN, RF, Voting, Bagging, and Boosting	GA	81.18% for RF	6 (38%)
Punlumjeak and Rachburee [69]	NB, DT, k-NN, SVM, and ANN	GA	91.12% for k-NN	10 (–)
Ajibade *et al.* [70]	NB, DT, KNN, and DISC	SFS, SBS, and DE	83.09% for DISC	6 (38%)
Zaffar *et al.* [74]	BN, NB, NBU, MLP, SL, SMO, DT, OneR J rip, DS, J48, RF, RT, and RepT.	CfsSubsetEval, ChiSquared, GainRatio, and Relief	76.39% for Chi-MLP	24 (67%)
Jalota and Agrawal [77]	SVM and ANN-AdaBoost	WFS and CFS	91% for NB + WFS	9 (56%)
Nalić *et al.* [76]	GLM, DT, SVM, and NB	ClassFE, CorrelationFE, GainRFE, InfoGainFE, and RefilefFE	87.69% for PCA + GLM + DT	26 (81%)
Present study	k-NN and SVM	11 wrapper-based FS algorithms	100% for k-NN and SVM	4 (18.8%)

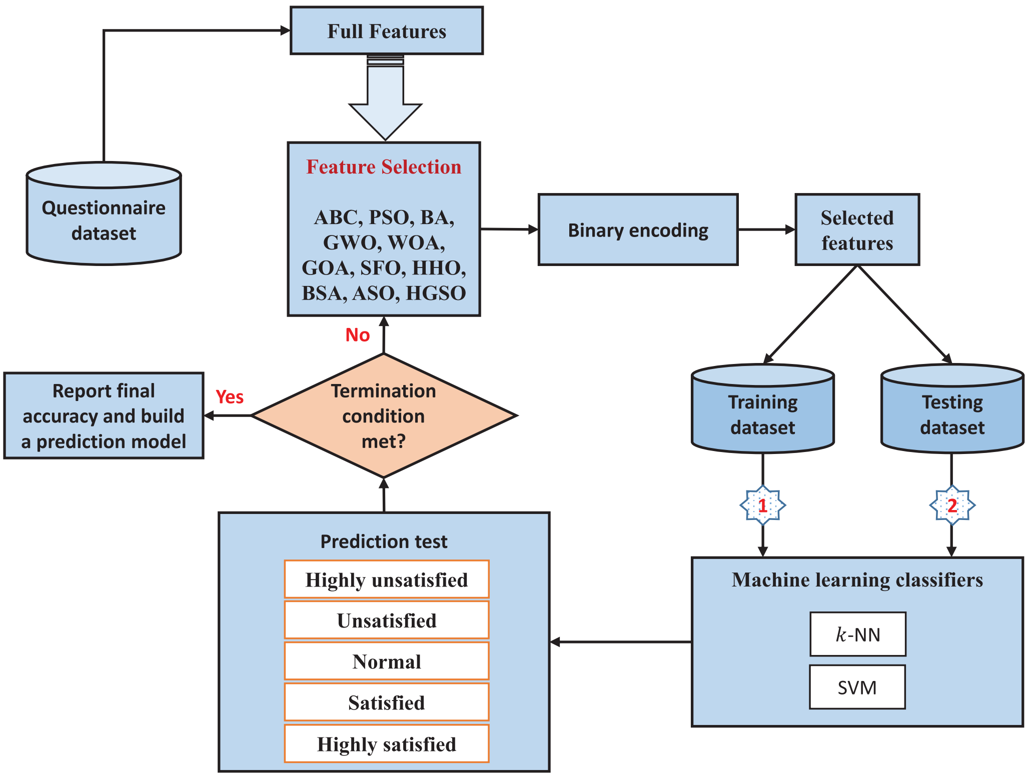


FIGURE 1. The framework of SSL assessment model.

four divisions. These four groups are alpha, beta, delta, and omega, each represents a potential optimal hunting solution yet identified.

5) WHALE OPTIMIZATION ALGORITHM (WOA)

WOA was proposed by Mirjalili and Lewis [35] for solving numerical optimization problems. It has three operators that simulate humpback whale foraging behavior: searching for prey, circling prey, and using bubble nets. WOA employs a population of search agents to find the global best solution to a given optimization problem. The rules that improve the candidate solutions in each optimization step distinguish WOA from other algorithms. WOA replicates the hunting behavior of humpback whales by identifying and attacking prey via a technique known as bubble-net feeding.

6) GRASSHOPPER OPTIMIZATION ALGORITHM (GOA)

GOA was proposed by Saremi *et al.* [42]. It is an SI system inspired by grasshoppers' natural foraging and swarming activity. Furthermore, this algorithm's unique adaptive mechanism smoothly balances exploration and exploitation. Because of these qualities, the GOA algorithm may overcome the challenges of a multi-objective search space and outperform other methods. Furthermore, the computational complexity is lower than that of several existing optimization approaches.

7) SAILFISH OPTIMIZER (SFO)

SFO is a population-based meta-heuristic algorithm inspired by the alternative attack strategy of sailfish hunting sardines [43]. This hunting approach provides hunters with an advantage by allowing them to save energy. It considers two populations: sailfish and sardines. Candidate solutions are sailfishes. The algorithm makes every possible effort to ensure that the movement of search agents (sailfish and sardine) is stochastic. Thus, sailfish are assumed to be dispersed in the search space, whereas sardine placements aid in identifying the best solution.

8) HARRIS HAWKS OPTIMIZATION (HHO)

Heidari *et al.* [44] developed the HHO. HHO is an SI-based optimization approach. Its basic concept is to simulate the activity and reaction of Hawk's team collaboration for hunting and coping with prey escaping. Hawks pursuing actions represent agents in the search space, while a prey represents the best position. Thus, HHO helps to resolve a variety of real-world optimization problems. Furthermore, the HHO may be utilized to tackle unknown forms of search space and solve issues involving discrete and continuous areas, give higher solution quality, and extract optimal parameters with high accuracy [80].

9) BIRD SWARM ALGORITHM (BSA)

Meng *et al.* [45] introduced BSA, a new meta-heuristic approach for continuous optimization problems. BSA is based on SI, which is derived from the social behaviors and interactions of bird swarms. It imitates the foraging, attentiveness, and flying behaviors of birds. As a result, SI can be effectively retrieved from bird swarms to solve different optimization problems.

10) ATOM SEARCH OPTIMIZATION (ASO)

ASO is introduced as a molecular dynamics-inspired optimization approach. In ASO, the position of each atom within the search space represents a solution whose fitness is measured by its mass heaviness [46]. Based on their distance, all atoms in the population will attract or repel to each other, causing the lighter atoms to gravitate toward the heavier ones. Furthermore, as heavier atoms move slower, they seek better solutions in local spaces more aggressively. Conversely, lighter atoms move faster, boosting them to search more widely for new locations across the search space.

11) HENRY GAS SOLUBILITY OPTIMIZATION (HGSO)

Hashim *et al.* [47] proposed the HGSO algorithm in 2019. HGSO is a meta-heuristic algorithm inspired by Henry's law to imitate gas particles' behavior [81]. The HGSO simulates gas huddling behavior to balance exploitation and exploration in the search space and avoid local optima.

C. MACHINE LEARNING TECHNIQUES

The ML classifiers utilized in this study are described in this section.

1) k-NEAREST NEIGHBORS (k-NN)

The k-NN algorithm [26] is a simple, supervised ML algorithm to solve classification and regression problems. Compared to other complex supervised ML algorithms, it has the advantage of simple implementation, so it is widely used [82], [83]. In pattern recognition paradigm, k-NN is applied in many fields, including image recognition [84], finance [85], healthcare [86], forestry [87], etc. The k-NN operates by assigning unlabeled observations to the class of the most near labelled examples. In addition, characteristics of observations are collected for both training and test datasets [88].

The k-NN algorithm can be thought of as a packaging method in which training examples generate classification rules. Once k-NN learns from the training process, the unknown patterns in the test set are approximated according to their proximity degree to the patterns in the training set. The unlabeled samples can be further classified according to the maximum probability of the category. As the choice of k in k-NN is critical, in the empirical experiment of this study, the selected feature subset is verified using the k-NN classifier ($k = 5$) [30], [89], [90] with the Euclidean distance metric.

2) SUPPORT VECTOR MACHINE (SVM)

SVM [91] is a supervised ML algorithm capable of solving classification and regression problems. Its primary application has been to solve classification problems. It is highly preferred by the data mining community as it, with less computational resources, can produce significant classification accuracy. The goal of linear SVM is to find the best hyperplane capable of categorizing the dataset into two categories. It divides the dataset into two classes, 0 and 1, located on opposing sides of the hyperplane. SVM is popular in the data science community because it can classify with high accuracy while using fewer computing resources. This process is accomplished by mapping the primary data from the original input space using the non-linear function into a higher dimensional space. Linear separation of the data can occur by finding a hyperplane with the maximal margin in this higher dimensional space > 0 for discovering the boundaries between the input classes. However, this technique opposes two significant key challenges, appropriate primary function selection and parameter adjustment [92]. Choosing the optimum decision plane is primarily regarded as an optimization process that aids a kernel function in determining the ideal space, wherein categories are frequently separated linearly through one non-linear transformation. Therefore, the current study set the polynomial kernel at an adequate value, 2.

D. TRANSFER FUNCTIONS (TFs)

A Transfer Function (TF) [93] determines the probability of changing the continuous values at elements of a position vector x to 0s or 1s. Thus, particles are forced to move in a binary space via TFs. In the current research, the proposed model adopted the sigmoid function as a TF whose a mathematical expression is shown in (1).

$$TF(x) = \frac{1}{1 + e^{-x}} \tag{1}$$

IV. EXPERIMENTAL METHODOLOGY

This section presents the experimental results. Specifically, Section IV-A describes the collected dataset that is used to validate the efficacy of the proposed model. Determination of a consensus decision is presented in Section IV-B. Parameter settings are prescribed in Section IV-C. Section IV-D presents performance measures adopted to evaluate the final results. Sections IV-E and IV-F discusses comparisons based on the k-NN and SVM classifier, respectively. Finally, Section IV-H is dedicated to further verify the robustness of the proposed 11 techniques based on a set of benchmark datasets from the UCI repository. So overall, several tests and experiments have been carried out determining the efficiency, advantage, and limitations of the proposed methods.

A. DATASET DESCRIPTION

This section describes the real-time dataset mainly used in this study for building a student satisfaction prediction model for the OL services during the COVID-19 epidemic. The dataset builds on an online questionnaire to collect students' reviews on OL courses from multiple academic institutions for the academic year 2020-2021. Fig. 2 is a screenshot of the online questionnaire form. The data was collected from 7 educational programs: Art Education, Home Economics, Educational Technology (General Division), Educational Technology (Computer Teacher), Educational Media, Educational Media (Theater), and Music Education at two different faculties from Kafrelsheikh University[2] and Mansoura University[3] in Egypt.

The dataset is described in details in Table 2 in terms of the per-feature frequency of the five satisfaction levels (5 as highly satisfied to 1 as highly unsatisfied). It contains 20 features and 18691 records. The dataset builds on heterogeneous data sources gathered from the OL services using educational platforms. The purpose of those specific features in the proposed dataset is to identify whether a student is satisfied or not with the OL in terms of several educational quality metrics/features (questions with ordinal responses). Moreover, each record in the proposed dataset is labelled with a flag class indicating a final satisfaction observation from high '5' to low '1'.

This dataset is mainly used to practice the ML techniques adopted herein for predicting SSL. This dataset contains 20 variables as features and one dependent variable as the SSL's class label. As shown in Table 3, the class label has five potential values, 5 representing highly satisfied, and 1 representing highly unsatisfied. Fig. 3 visualizes the balanced representation of different SSLs based on the dataset features.

[2] https://www.kfs.edu.eg
[3] https://www.mans.edu.eg

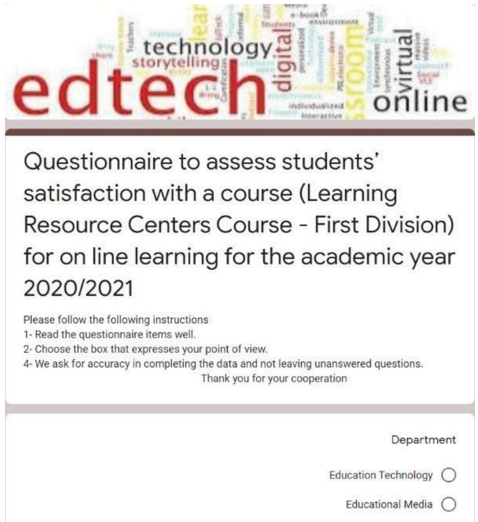


FIGURE 2. The online form to quest students' reviews on OL courses.

B. DETERMINATION OF CONSENSUS DECISION

We adopt a methodology to compute a decision of consensus (final satisfaction level for each instance/student) in our dataset that is based on survey responses of Likert-type questionnaires. The approach is based on a geometric framework applied to proxy economic uncertainty [94] to determine the likelihood of disagreement among election outcomes [95].

Let us assume a Likert-type questionnaire with N reply options, where S_i denotes the responses count of i in each X_i, $i = 1, 2, 3, 4, 5$, a natural representation of the vector X_i containing all the information from the surveyed units as a point on a simplex [96]. R_i denotes the aggregate responses in each category S_i, T_i is the sum of s_i, a natural representation of the vector X_i containing all the information from the surveyed units as a point on a simplex [96]. The interior of this simplex encompasses all possible combinations of reply options, which correspond to the barycentric coordinates of each point in time. Each of the N vertexes correspond to a point of maximum consensus.

We propose measuring the level of agreement as the ratio between the distance of the point to the barycentre and that from the barycentre to the nearest vertex. Hence, the measure of consensus can be estimated as:

$$W_i = \sum_{i=1}^{n} \frac{R_i}{T_i}, \tag{2}$$

where W_i is the overall weight of the selected record response categories. This measure incorporates the share of neutral

TABLE 2. Dataset features and their description.

Feature	Frequency of SSLs				
	5	4	3	2	1
University, College, Scientific Department, and Course Name	Categorical/nominal features				
The faculty member is always committed to the contents of the course.	10641	6299	1155	245	350
The faculty member is always obligated to upload the lecture weekly on the educational platform according to the announced academic schedule dates.	10676	3780	2415	910	909
The faculty member encourages students to ask questions and express their points of view.	8261	4935	3430	1085	979
The faculty member deals with the topics of the course in-depth.	9556	5390	3359	140	245
I feel that the faculty member is always well prepared for the lecture.	5740	7840	1995	1365	1750
The lectures are presented in an attractive style.	7000	4970	4620	1225	875
The amount of information provided in the lectures is sufficient.	4660	7105	3570	1260	420
The lectures included practical cases.	6230	8575	2205	630	1050
The teaching method in this course encourages to participate actively during the lectures.	9030	7560	945	1085	70
Students are assigned to prepare assignments.	8505	7350	2205	350	280
The scientific material is uploaded with a clear explanation presentation.	7070	7840	1890	1365	525
The lecturer presents practical videos that are useful in the educational process.	9310	5145	2660	560	1015
The lecturer holds meetings with the students through zoom/google meeting/classroom/MS team.	6825	6020	2555	1505	1785
I really enjoy the distance learning experience.	9608	3990	1540	1750	665
Students were trained on how to solve exams online by preparing an experimental quiz.	8108	4410	2380	770	1190
An experimental quiz has been prepared.	8140	3710	2345	1400	1365
The quiz has been prepared in degrees.	6335	5670	4620	455	1610
The lecturer helped the students to solve their problems during exams.	4620	6895	5635	1225	315
Exams are subjective.	5950	4550	5320	1995	875
Exam time is appropriate.	7035	7700	3115	280	560
Exams cover the contents of the course.	9608	3990	1540	1750	665
Decision label	**HS**	**S**	**N**	**U**	**HU**

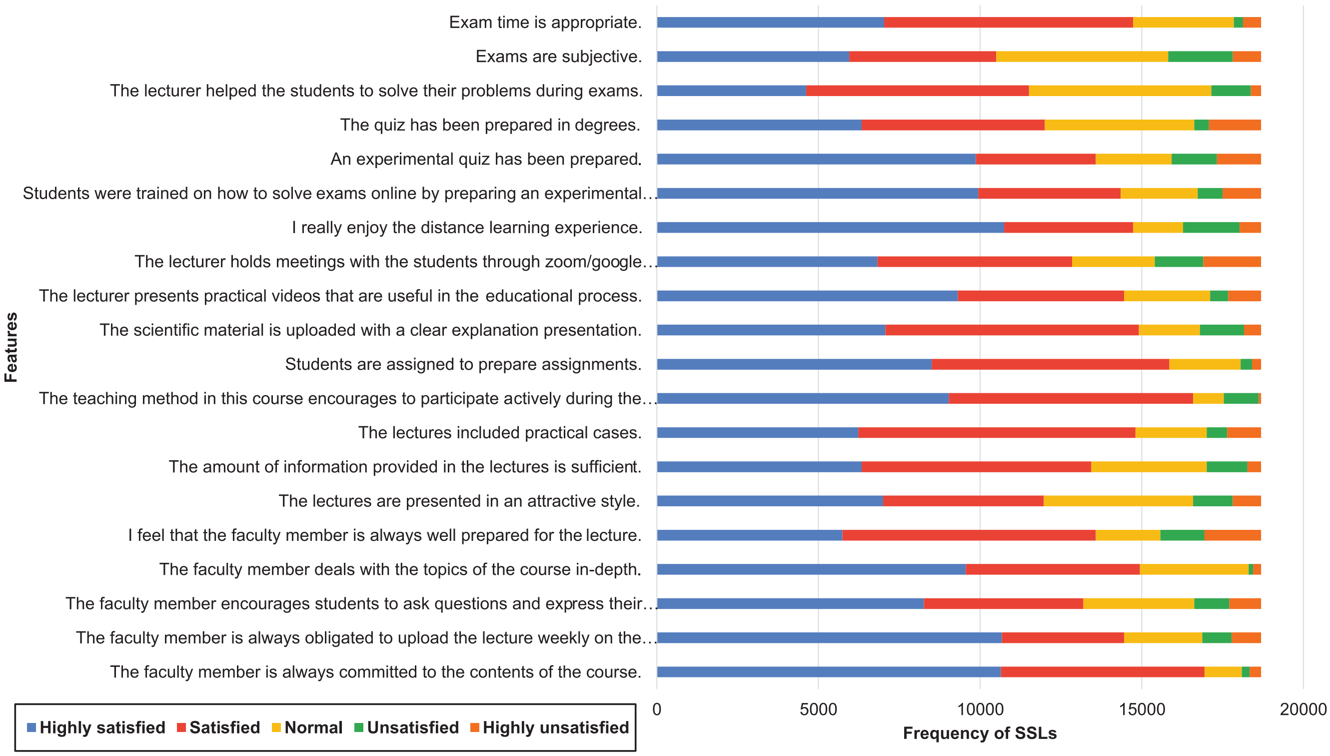


FIGURE 3. The frequency of every SSL per each feature.

TABLE 3. Dataset output label key.

Value	Label
1	Highly Unsatisfied (HU)
2	Unsatisfied (U)
3	Normal (N)
4	Satisfied (S)
5	Highly Satisfied (HS)

responses and allows to capture the trajectories of the different state. D_i is the overall decision of the selected record response categories, given that:

$$\begin{aligned}&\text{If } 5 \geq W_i > 4 \text{ then } D_i = HS,\\&\text{If } 4 \geq W_i > 3 \text{ then } D_i = S,\\&\text{If } 3 \geq W_i > 2 \text{ then } D_i = N,\\&\text{If } 2 \geq W_i > 1 \text{ then } D_i = U,\\&\text{If } 1 \geq W_i > 0 \text{ then } D_i = HU.\end{aligned}$$

C. PARAMETER SETTINGS

This paper suggests 11 meta-heuristic algorithms, including ABC, PSO, BA, GWO, WOA, GOA, SFO, HHO, BSA, ASO, and HGSO, along with two ML classifiers k-NN and SVM. Regarding the ML classifiers, k-NN employs the Euclidean distance metric $k = 5$, whereas SVM uses a polynomial kernel with degree $d = 2$. Furthermore, because meta-heuristics have a stochastic nature, 30 independent runs were conducted for each algorithm. Table 4 lists the general settings for all algorithms and the parameter values of each algorithm. Python was used to execute all experiments in this paper.

D. PERFORMANCE MEASURES

To validate the performance of the proposed models, they must be judged by standard measures to ensure that the experimental outcomes are statistically significant and meaningful. So, some essential performance measures were used as follows:

- **Mean accuracy** ($Mean_{Acc}$): It measures the rate at which data is classified correctly. The average classification accuracy is estimated by running each algorithm 30 times ($N = 30$) as expressed in (3).

$$Mean_{Acc} = \frac{1}{N}\frac{1}{M}\sum_{k=1}^{M}\sum_{r=1}^{N} match\,(PL_r, AL_r)\,, \tag{3}$$

  where M is the sample size in the dataset, PL_r and AL_r respectively represent the output label of the predicted class and the reference class label for sample r, while $match\,(PL_r, AL_r)$ denotes a discriminant comparison function. If $PL_r = AL_r$, then the function value is 1, otherwise 0.
- **Mean fitness** ($Mean_{Fit}$)**:** It measures the mean fitness gained by running the algorithm 30 times, highlighting the strong link between minimizing the number of selected features by excluding irrelevant ones and reducing the classification error rate. The less the value, the more reliable the obtained solution, which is shown in (4).

$$Mean_{Fit} = \frac{1}{N}\sum_{i=1}^{N} f_*^i, \tag{4}$$

  where f_*^i denotes the optimum fitness value obtained so far in the i-th run.
- **Mean number of selected features** ($Mean_{Feat}$): It shows the ratio of the number of the selected features to the total number of features, which is calculated by (5).

$$Mean_{Feat} = \frac{1}{N}\sum_{i=1}^{N}\frac{size(g_*^i)}{D}, \tag{5}$$

  where $size(g_*^i)$ is the number of selected features in the i-th run, and D is the total features' number in the original dataset.
- **Mean computational time** (T_z): It shows the execution time in seconds for each algorithm validated over 30 different runs, which can be computed by (6).

$$T = \frac{1}{N}\sum_{i=1}^{N} RunTime_i, \tag{6}$$

TABLE 4. Parameter setup for all algorithms.

Algorithm	Parameter
Common settings	Dimensionality D = number of features in the original datasets
	Runs' number $= 30$
	Maximum iterations' number $T = 100$
	Population size $N = 10$
PSO	Inertia weight ($\omega_{\min} = 0.4, \omega_{\max} = 0.9$)
	Acceleration coefficients ($c_1 = c_2 = 1.2$)
ABC	Number of employed bees = 16
	Number of scout bees = 3
	Number of onlooker bees = 4
GWO	a is linearly decreased from 2 to 0
BA	Loudness $A = 0.8$
	Lower and upper pulse frequencies $= \{0, 10\}$
	Pulse emission rate $r = 0.95$
GOA	$C_{\min} = 0.00004$ and $C_{\max} = 1$
WOA	a is linearly decreased from 2 to 0
	$b = 1.0$
	$p = 0.5$
HHO	Energy of rabbit $E \in [-1, 1]$
SFO	Ratio between sailfish and sardines $pp = 0.1$
	$\varepsilon = 0.0001$
	$A = 1$
ASO	Multiplier weight $\beta = 0.2$
	Depth weight $\alpha = 50$
BSA	Flight frequency $ff = 10$
	Acceleration coefficients ($c_1 = c_2 = 1.5$)
	Probability of foraging for food $p = 0.8$
	Followed coefficient $fl = 0.5$
	Effect on birds' vigilance behaviors ($a_1 = a_2 = 1.0$)
HGSO	Number of clusters $= 2$
	$l_1 = 5E-03$, $l_2 = 1E+02$, and $l_3 = 1E-02$
	$\alpha = \beta = 0.1$ and $K = 1.0$

TABLE 5. Comparisons of different algorithms based on k-NN in terms of average accuracy.

Algorithm	Best	Mean	Worst	Std
ABC	**1**	**1**	**1**	**0**
PSO	**1**	**1**	**1**	**0**
BA	**1**	**1**	**1**	**0**
GWO	**1**	**1**	**1**	**0**
WOA	**1**	**1**	**1**	**0**
GOA	**1**	**1**	**1**	**0**
SFO	**1**	**1**	**1**	**0**
HHO	**1**	**1**	**1**	**0**
BSA	**1**	**1**	**1**	**0**
ASO	**1**	**1**	**1**	**0**
HGSO	**1**	**1**	**1**	**0**

where N is the runs' number, and $RunTime_i$ is the computational time in seconds at run i.

E. COMPARISONS BASED ON THE k-NN CLASSIFIER USING THE PROPOSED REAL-TIME DATASET

This section discusses the performance analysis based on the k-NN classifier according to the average fitness value, the average classification accuracy, the average number of selected features, and the average computational time. Table 5 presents the average classification accuracy results for k-NN with the different comparison algorithms. From this table, it is clear that all involved algorithms achieved 100% classification accuracy on the proposed real-time dataset. This undoubtedly implies the robustness of meta-heuristic algorithms compared to other exact or deterministic peers by achieving the highest possible classification accuracy rate.

As presented in Table 6, the SFO algorithm has the best ability of exploration and exploitation, given the minimal features' number selected by the wrapper-based SFO algorithm. In the best conditions, the SFO algorithm determined only four features (the most critical):

- "*The lectures are presented in an attractive style*".
- "*The teaching method in this course encourages me to participate actively during the classes*".
- "*The quiz has been prepared in degrees*".
- "*Students trained on how to solve exams online by designing an experimental quiz*".

It should be also pointed out the following. The feature "*Students trained on how to solve exams online by designing an experimental quiz*" was selected by 6 algorithms, "*The lectures are presented in an attractive style*" was opted in by four algorithms, and "*The quiz has been prepared in degrees" and "The teaching method in this course encourages me to participate actively during the classes*" was selected by two algorithms. Thus, as those features give the best results over most of the involved algorithms, they can be considered the most informative features and should be paid a special attention by decision makers in educational institutions during the COVID-19.

Table 7 presents the average values of fitness for all algorithms based on the k-NN classifier. As shown in the table, the SFO algorithm achieved the minimum mean classification error compared to other algorithms.

TABLE 6. Comparisons of different algorithms based on k-NN in terms of the average number of selected features.

Algorithm	Best	Mean	Worst	Std	Selection rate
ABC	5	6.0	7	0.7746	0.240
PSO	5	5.1	6	0.3000	0.204
BA	5	6.4	9	1.3565	0.256
GWO	5	5.3	6	0.4583	0.212
WOA	5	5.6	6	0.4899	0.224
GOA	6	6.7	7	0.6403	0.268
SFO	**4**	**4.7**	**5**	**0.4583**	**0.188**
HHO	6	6.8	8	0.6000	0.272
BSA	6	6.6	8	0.6633	0.264
ASO	6	6.3	7	0.4583	0.252
HGSO	7	7.8	9	0.7483	0.312

TABLE 7. Comparisons of different algorithms based on k-NN in terms of average fitness.

Algorithm	Best	Mean	Worst	Std
ABC	0.0020	0.0024	0.0028	0.0003
PSO	0.0020	0.0020	0.0024	**0.0001**
BA	0.0020	0.0026	0.0039	0.0006
GWO	0.0020	0.0021	0.0024	0.0002
WOA	0.0020	0.0022	0.0024	0.0002
GOA	0.0024	0.0028	0.0028	**0.0001**
SFO	**0.0016**	**0.0019**	**0.0020**	0.0002
HHO	0.0024	0.0027	0.0032	0.0002
BSA	0.0024	0.0026	0.0032	0.0003
ASO	0.0024	0.0025	0.0028	0.0002
HGSO	0.0028	0.0031	0.0036	0.0003

Table 8 presents the average values of computational time for different algorithms with the k-NN classifier. The HHO algorithm achieved the minimum mean computational time compared with other methods. In this study, although SFO did not work the best in terms of computational time, a higher priority is given to the classification accuracy and the number of selected features over computational time and this can be considered rational.

F. COMPARISONS BASED ON THE SVM CLASSIFIER

This section is to analyze the performance of the SVM classifier according to the averages fitness value, classification accuracy, number of selected features, and computational time. Table 9 shows the average accuracy outcomes for SVM with 11 FS methods. From Table 9, it is demonstrated that all methods obtained 100% classification accuracy on the adopted dataset.

As shown in Table 10, the SFO method has the best exploration ability than others based on the average number of features selected, which was supported by selected minimum features numbers. According to best conditions, the SFO method selected only five features:

- "*Exam time is appropriate*".
- "*The lectures are presented in an attractive style*".

TABLE 8. Comparisons of different algorithms based on k-NN in terms of computational time (in seconds).

Algorithm	Best	Mean	Worst	Std
ABC	1794286	1863934	1835980	57518
PSO	467189	480804	557390	35966
BA	562529	501013	557390	37061
GWO	554109	527983	541440	30893
WOA	609426	611430	672602	43390
GOA	594179	627841	636280	38286
SFO	4343391	4151502	2806539	519926
HHO	**358539**	**356440**	**358132**	**16384**
BSA	426368	419659	431039	20837
ASO	519569	721933	978056	186267
HGSO	1016440	934474	819713	142263

TABLE 9. Comparisons of different algorithms based on SVM in terms of average accuracy.

Algorithm	Best	Mean	Worst	Std
ABC	**1**	**1**	**1**	**0**
PSO	**1**	**1**	**1**	**0**
BA	**1**	**1**	**1**	**0**
GWO	**1**	**1**	**1**	**0**
WOA	**1**	**1**	**1**	**0**
GOA	**1**	**1**	**1**	**0**
SFO	**1**	**1**	**1**	**0**
HHO	**1**	**1**	**1**	**0**
BSA	**1**	**1**	**1**	**0**
ASO	**1**	**1**	**1**	**0**
HGSO	**1**	**1**	**1**	**0**

TABLE 10. Comparisons of different algorithms using SVM in terms of the average number of selected features.

Algorithm	Best	Mean	Worst	Std	Selection rate
ABC	6	6.6	7	0.4899	0.264
PSO	6	6.4	7	0.4899	0.256
BA	7	7.6	9	0.6633	0.304
GWO	6	6.2	7	0.4000	0.248
WOA	6	6.6	7	0.4899	0.264
GOA	7	7.4	9	0.8000	0.296
SFO	**5**	**5.3**	**5.5**	**0.4583**	**0.212**
HHO	6	7.5	8	0.6708	0.300
BSA	7	7.7	9	0.7810	0.308
ASO	7	7.6	9	0.6633	0.304
HGSO	8	8.6	10	0.6633	0.344

- "*The teaching method in this course encourages me to participate actively during the classes*".
- "*The quiz has been prepared in degrees*".
- "*Students trained on how to solve exams online by designing an experimental quiz*".

Table 11 shows the average fitness values for different methods utilizing the SVM classifier. As shown in Table 11, the SFO algorithm had the minimum fitness value compared to others. ABC, PSO, GWO, WOA, and HHO algorithms come second.

Table 12 shows the average computational time values based on the 11 FS methods with the SVM classifier. First, the ASO obtained the minimum average computational time, then the BA algorithm ranked second.

TABLE 11. Comparisons of different algorithms based on SVM in terms of average fitness.

Algorithm	Best	Mean	Worst	Std
ABC	0.0024	0.0026	0.0028	**0.0002**
PSO	0.0024	0.0026	0.0028	**0.0002**
BA	0.0028	0.0030	0.0036	0.0003
GWO	0.0024	0.0026	0.0028	0.0002
WOA	0.0024	0.0025	0.0028	0.0002
GOA	0.0028	0.0030	0.0036	0.0003
SFO	**0.0016**	**0.0019**	**0.0020**	**0.0002**
HHO	0.0024	0.0030	0.0032	0.0003
BSA	0.0028	0.0031	0.0036	0.0003
ASO	0.0028	0.003	0.0036	0.0003
HGSO	0.0032	0.0035	0.0040	0.0003

TABLE 12. Comparisons of 11 FS algorithms based on SVM in terms of computational time (in seconds).

Algorithm	Best	Mean	Worst	Std
ABC	22981739	17867785	17144705	2726585
PSO	109201474	62008794	40628865	53698852
BA	4669413	11232803	7547947	10707005
GWO	5249360	6688068	6566769	1524636
WOA	5747104	25510864	5819433	30735841
GOA	9159423	7358292	10106963	1873218
SFO	209201474	72008794	60628865	63698852
HHO	5855739	7773606	4742538	2891184
BSA	5492264	6692333	6514122	1354725
ASO	**4152734**	**6050998**	**3038738**	2104077
HGSO	6818171	7147302	7374711	**352211**

G. COMPARISONS BASED ON ALL AND SELECTED FEATURES USING k-NN AND SVM

Fig. 4 depicts the average overall accuracy for k-NN and SVM using all and selected features. The average overall accuracy for k-NN achieved 98%, while the SVM reached 92%. By applying meta-heuristic algorithms for FS, the SFO with k-NN and SVM achieved an overall 100% accuracy for the proposed dataset over counterparts.

H. BENCHMARK DATASETS RESULTS

In this section, the quality of the proposed 11 techniques is further validated using 6 multi-scale datasets from the UCI ML data repository [97] in many areas (e.g., biology, game, and physics). Table 13 shows the number of features, number of instances, and domain for each dataset. In the following tables, "Std" denotes the standard deviation metric, while "W|T|L" stands for the sum of wins, ties, ans losses fro each algorithm over all datastes.

1) COMPARISONS BASED ON THE k-NN CLASSIFIER

All comparisons based on the 6 UCI datasets are tabulated in Tables 14, 15, and 16 based on the k-NN classifier according to the average fitness value, average classification accuracy,

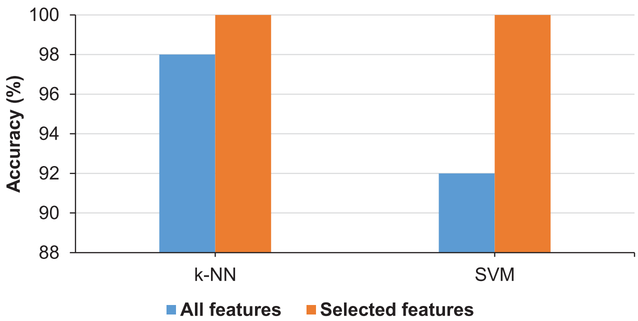


FIGURE 4. The average accuracy of *k*-NN and SVM for all and selected features.

TABLE 13. Number of features and instances for each dataset.

#	Dataset	No. of features	No. of instances	Domain
1	Exactly	13	1000	Biology
2	Exactly2	13	1000	Biology
3	KrvskpEW	36	3196	Game
4	M-of-n	13	1000	Biology
5	Tic-tac-toe	9	958	Game
6	WaveformEW	40	5000	Physics

and average number of selected features, respectively. As can be seen from Table 14, it can be easily noted that ABC outperforms other algorithms.

Table 15 shows the average accuracy results of k-NN obtained by each algorithm, it is noted that ABC scored the highest accuracy results 97%

In addition, Table 16 shows the average number of selected features based on k-NN for the different methods. As shown in Table 16, GWO, WOA, SFO, BSA, ASO, and HGSO won in one dataset by selecting the fewest number of features.

2) COMPARISONS BASED ON THE SVM CLASSIFIER

The comparisons between the 6 UCI datasets are recorded in Tables 17,18, and 19 based on the SVM classifier according to the average fitness value, average classification accuracy, and average number of selected features, respectively. According to the results of fitness value, as given in Table 17, GWO has the smallest fitness value among the tested datasets.

Table 18 shows the average accuracy results of SVM for the different methods. From Table 18, it is noted that ABC and SFO ranked first by excelling in one dataset and a tie over other 3 datasets. GWO ranked second by winning over one dataset and a tie over two datasets. Finally, GOA, HHO, and BSA ranked third by achieving a tie over three datasets in terms of average classification accuracy.

In addition, Table 19 shows the average number of selected features based on SVM with the different FS methods. As shown in Table 19, it is noted that ASO ranked first by winning over 2 datasets. BA, GWO, and WOA ranked second by winning over one dataset. Finally, SFO and HHO tied over one dataset in terms of the average number of selected features, ranked so third.

V. DISCUSSION

Due to the COVID-19 epidemic, instructors have been enforced to incorporate new methods into their courses in order to emphasise SSL despite the epidemic's limitations. According to the discussions in Section II, students encountered some difficulties while participating in OL as a result of the COVID-19 crisis. Researchers proposed that colleges and universities should develop an education continuity strategy to ensure continuous high-quality OL. Many applications of ML algorithms have been proposed to predict student academic performance. When mining a student's SSL, FS is one of those recommended prerequisites. It aims to reduce the high computational costs required by heavy mining tasks by removing any noisy, redundant, or irrelevant features that may degrade classification accuracy.

Nonetheless, no current research has looked at the impact of FS techniques on SSL and awareness of: course content and design, instructor quality, evaluation system, and how many and which features affected the most on SSL with OL during the COVID-19 epidemic. This paper developed two more straightforward and precise models based on k-NN anf SVM with 11 meta-heuristic algorithms in order to improve EDM performance. We used 4 features (rather than the original 20 features) which should really affect SSL and so important enough to predict SSL with OL with 100% accuracy.

Summarizing, this study advocated proposing new educational real-time dataset, as well as integrating two ML strategies, k-NN and SVM, in 11 meta-heuristics for the wrapper-based FS job. The comparative analysis performed in the previous section, using the adopted real-time and some other benchmark datasets, indicate that the proposed methodologies have proven effective. k-NN with SVM efficiently achieved an overall accuracy of 100% on the real-dataset (as shown in Tables 5 and 9) with a feature size reduction of up to 80% (as shown in Tables 6 and 10), and a relatively good results with the benchmark datasets (as shown in Tables 14–19). Apart from better accuracy and low feature size, k-NN and SVM models also exhibited faster convergence behavior that can be easily found out based on fitness values and time consumed on most of the datasets (the real-time one and benchmarks). This can be noted by inspecting Tables 7, 8, 11, and 7 for the real time dataset, and Tables 14 and 17 for the benchmark ones. Moreover, computational cost SVM-based models were notably higher than k-NN, given weights' updating time taken by SVM using a learning step whereas k-NN simply classifies based on computing of distance. Lastly, the proposed models are designed in such a straightforward way that it will be simple to implement any potential to improve the methods.

Apart from effectivenesses demonstrated above, the proposed models also maintain certain limitations:

- Some of the suggested meta-heuristics for FS may comprise several control parameters, which may hurt their applicability.

TABLE 14. Comparisons of different algorithms based on k-NN in terms of average fitness.

Benchmark	Metric	ABC	PSO	BA	GWO	WOA	GOA	SFO	HHO	BSA	ASO	HGSO
Exactly	$Mean_{Fit}$	**0.0132**	0.1232	0.1546	0.0260	0.0472	0.0390	0.0147	0.0302	0.0439	0.2710	0.1476
	Std_{Fit}	**0.0097**	0.0764	0.0915	0.0354	0.0518	0.0342	0.0178	0.0374	0.0480	0.0535	0.0796
Exactly2	$Mean_{Fit}$	0.2294	0.2332	0.2418	0.2324	0.2331	0.2324	**0.2274**	0.2369	0.2328	0.2450	0.2438
	Std_{Fit}	0.0023	0.0049	0.0054	0.0041	0.0049	0.0037	**0.0000**	0.0053	0.0038	0.0063	0.0063
KrVsKpEW	$Mean_{Fit}$	**0.0254**	0.0407	0.0397	0.0265	0.0309	0.0307	0.0279	0.0265	0.0316	0.0611	0.0364
	Std_{Fit}	**0.0022**	0.0085	0.0095	0.0044	0.0053	0.0050	0.0028	0.0057	0.0040	0.0185	0.0066
M-of-n	$Mean_{Fit}$	0.0070	0.0484	0.0491	**0.0049**	0.0135	0.0124	0.0054	0.0080	0.0097	0.1338	0.0515
	Std_{Fit}	0.0048	0.0334	0.0508	**0.0010**	0.0140	0.0127	0.0021	0.0066	0.0082	0.0408	0.0242
Tic-tac-toe	$Mean_{Fit}$	**0.1544**	0.1576	0.1611	0.1547	0.1547	0.1551	0.1544	0.1610	0.1554	0.1908	0.1613
	Std_{Fit}	**0.0000**	0.0050	0.0099	0.0018	0.0018	0.0026	**0.0000**	0.0112	0.0031	0.0166	0.0103
WaveformEW	$Mean_{Fit}$	0.1624	0.1762	0.1794	0.1636	0.1671	0.1672	**0.1618**	0.1622	0.1653	0.1932	0.1787
	Std_{Fit}	0.0040	0.0053	0.0067	0.0057	0.0059	0.0048	**0.0028**	0.0066	0.0067	0.0077	0.0045
Ranking	W|T|L	**3|0|3**	0|0|6	0|0|6	1|0|5	0|0|6	0|0|6	2|0|4	0|0|6	0|0|6	0|0|6	0|0|6

TABLE 15. Comparisons of different algorithms based on k-NN in terms of average accuracy.

Benchmark	Metric	ABC	PSO	BA	GWO	WOA	GOA	SFO	HHO	BSA	ASO	HGSO
Exactly	$Mean_{Acc}$	**0.9917**	0.8818	0.8502	0.9788	0.9577	0.9660	0.9902	0.9747	0.9610	0.7333	0.8575
	Std_{Acc}	**0.0094**	0.0764	0.0919	0.0353	0.0518	0.0341	0.0176	0.0373	0.0480	0.0541	0.0795
Exactly2	$Mean_{Acc}$	0.7737	0.7695	0.7607	0.7703	0.7698	0.7707	**0.7750**	0.7658	0.7695	0.7572	0.7590
	Std_{Acc}	0.0022	0.0049	0.0057	0.0045	0.0051	0.0038	**0.0000**	0.0058	0.0039	0.0073	0.0066
KrVsKpEW	$Mean_{Acc}$	**0.9808**	0.9649	0.9660	0.9794	0.9751	0.9753	0.9781	0.9794	0.9742	0.9448	0.9701
	Std_{Acc}	**0.0022**	0.0088	0.0101	0.0046	0.0055	0.0051	0.0029	0.0057	0.0042	0.0191	0.0067
M-of-n	$Mean_{Acc}$	0.9978	0.9577	0.9563	**0.9998**	0.9917	0.9928	0.9995	0.9970	0.9955	0.8717	0.9545
	Std_{Acc}	0.0046	0.0330	0.0508	**0.0009**	0.0136	0.0124	0.0020	0.0064	0.0079	0.0419	0.0239
Tic-tac-toe	$Mean_{Acc}$	**0.8542**	0.8493	0.8453	0.8536	0.8536	0.8531	0.8542	0.8460	0.8526	0.8135	0.8453
	Std_{Acc}	**0.0000**	0.0075	0.0119	0.0028	0.0028	0.0039	0.0000	0.0128	0.0047	0.0167	0.0121
WaveformEW	$Mean_{Acc}$	0.8424	0.8277	0.8249	0.8405	0.8369	0.8370	**0.8426**	0.8422	0.8390	0.8112	0.8258
	Std_{Acc}	0.0040	0.0054	0.0071	0.0057	0.0061	0.0048	**0.0029**	0.0067	0.0067	0.0079	0.0046
Ranking	W|T|L	**3|0|3**	0|0|6	0|0|6	1|0|5	0|0|6	0|0|6	2|0|4	0|0|6	0|0|6	0|0|6	0|0|6

TABLE 16. Comparisons of different algorithms based on k-NN in terms of average number of selected features.

Benchmark	Metric	ABC	PSO	BA	GWO	WOA	GOA	SFO	HHO	BSA	ASO	HGSO
Exactly	$Mean_{Feat}$	006.47	008.10	008.20	006.57	006.90	007.00	**006.43**	006.67	006.93	009.13	008.50
	Std_{Feat}	**000.50**	001.04	001.28	000.67	000.79	000.63	000.56	000.70	000.73	001.82	001.18
Exactly2	$Mean_{Feat}$	006.97	006.47	006.37	006.53	006.77	007.00	006.00	006.53	006.00	**005.97**	006.83
	Std_{Feat}	001.02	001.02	001.74	001.09	001.23	001.39	**000.00**	001.61	001.15	002.23	001.85
KrVsKpEW	$Mean_{Feat}$	023.13	021.77	021.87	021.77	022.47	022.60	022.40	021.90	**021.73**	023.23	024.33
	Std_{Feat}	002.35	002.65	003.52	002.50	002.28	002.65	**001.89**	002.74	002.45	004.34	003.01
M-of-n	$Mean_{Feat}$	006.37	008.50	007.67	**006.20**	006.83	006.90	006.37	006.50	006.77	008.73	008.40
	Std_{Feat}	000.55	001.48	001.27	**000.40**	000.86	000.75	000.48	000.56	000.72	002.02	000.99
Tic-tac-toe	$Mean_{Feat}$	009.00	007.53	007.13	008.83	008.83	008.70	009.00	007.70	008.53	**005.60**	**007.33**
	Std_{Feat}	**000.00**	002.25	002.16	000.90	000.90	001.13	000.00	001.93	001.41	001.28	002.12
WaveformEW	$Mean_{Feat}$	025.30	022.60	024.20	022.83	**022.53**	023.40	023.70	023.93	023.57	025.10	025.03
	Std_{Feat}	003.84	002.68	003.95	003.01	003.04	002.63	**002.44**	002.74	002.60	004.46	002.54
Ranking	W|T|L	0|0|6	0|0|6	0|0|6	**1|0|5**	**1|0|5**	0|0|6	**1|0|5**	0|0|6	**1|0|5**	**1|0|5**	**1|0|5**

- Additionally, the subset of selected features may change at each time of execution, given the stochasticity nature of optimization techniques, which may confuse the user which feature sets to realize.

TABLE 17. Comparisons of different algorithms based on SVM in terms of average fitness.

Benchmark	Metric	ABC	PSO	BA	GWO	WOA	GOA	SFO	HHO	BSA	ASO	HGSO
Exactly	$Mean_{Fit}$	0.2612	0.3006	0.3248	0.2835	0.2969	0.2892	**0.2589**	0.3044	0.3012	0.3321	0.3160
	Std_{Fit}	0.0137	0.0311	0.0135	0.0301	0.0305	0.0318	0.0084	0.0310	0.0289	**0.0040**	0.0236
Exactly2	$Mean_{Fit}$	0.2486	0.2492	0.2494	**0.2483**	0.2487	0.2485	**0.2483**	**0.2483**	0.2487	0.2499	0.2497
	Std_{Fit}	0.0004	0.0005	0.0006	0.0001	0.0005	0.0004	**0.0000**	**0.0000**	0.0004	0.0006	0.0006
KrVsKpEW	$Mean_{Fit}$	0.0266	0.0395	0.0402	**0.0256**	0.0315	0.0324	0.0284	0.0278	0.0304	0.0539	0.0362
	Std_{Fit}	0.0031	0.0083	0.0083	0.0040	0.0048	0.0037	**0.0023**	0.0050	0.0031	0.0166	0.0037
M-of-n	$Mean_{Fit}$	0.0053	0.0266	0.0419	**0.0050**	0.0102	0.0056	0.0051	0.0052	0.0056	0.1149	0.0116
	Std_{Fit}	**0.0004**	0.0511	0.0642	0.0005	0.0251	**0.0004**	**0.0004**	0.0006	0.0006	0.0701	0.0269
Tic-tac-toe	$Mean_{Fit}$	**0.1017**	0.1206	0.1197	**0.1017**	0.1051	0.1018	**0.1017**	0.1018	**0.1017**	0.1602	0.1129
	Std_{Fit}	**0.0000**	0.0191	0.0251	**0.0000**	0.0090	0.0004	**0.0000**	0.0004	0.0002	0.0320	0.0120
WaveformEW	$Mean_{Fit}$	**0.1315**	0.1468	0.1435	0.1328	0.1357	0.1380	0.1333	0.1320	0.1352	0.1536	0.1412
	Std_{Fit}	0.0029	0.0051	0.0076	0.0033	0.0045	0.0050	**0.0028**	0.0044	0.0046	0.0070	0.0048
Ranking	W|T|L	1|1|4	0|0|6	0|0|6	**2|2|2**	0|0|6	0|0|6	1|2|3	0|1|5	0|1|5	0|0|6	0|0|6

TABLE 18. Comparisons of different algorithms based on SVM in terms of average accuracy.

Benchmark	Metric	ABC	PSO	BA	GWO	WOA	GOA	SFO	HHO	BSA	ASO	HGSO
Exactly	$Mean_{Acc}$	0.7415	0.7022	0.6777	0.7187	0.7060	0.7135	**0.7435**	0.6973	0.7013	0.6683	0.6870
	Std_{Acc}	0.0132	0.0311	0.0139	0.0305	0.0304	0.0318	0.0081	0.0316	0.0292	**0.0051**	0.0234
Exactly2	$Mean_{Acc}$	**0.7500**	**0.7500**	**0.7500**	**0.7500**	**0.7500**	**0.7500**	**0.7500**	**0.7500**	**0.7500**	**0.7500**	**0.7500**
	Std_{Acc}	**0.0000**	**0.0000**	**0.0000**	**0.0000**	**0.0000**	**0.0000**	**0.0000**	**0.0000**	**0.0000**	**0.0000**	**0.0000**
KrVsKpEW	$Mean_{Acc}$	0.9799	0.9661	0.9657	**0.9807**	0.9742	0.9738	0.9773	0.9781	0.9757	0.9518	0.9700
	Std_{Acc}	0.0033	0.0087	0.0089	0.0043	0.0052	0.0038	**0.0025**	0.0053	0.0031	0.0173	0.0038
M-of-n	$Mean_{Acc}$	**1.0000**	0.9795	0.9642	**1.0000**	0.9953	**1.0000**	**1.0000**	**1.0000**	**1.0000**	0.8905	0.9950
	Std_{Acc}	**0.0000**	0.0523	0.0650	0.0000	0.0251	**0.0000**	**0.0000**	**0.0000**	**0.0000**	0.0718	0.0269
Tic-tac-toe	$Mean_{Acc}$	**0.9062**	0.8868	0.8875	0.9062	0.9028	**0.9062**	**0.9062**	**0.9062**	**0.9062**	0.8457	0.8948
	Std_{Acc}	**0.0000**	0.0199	0.0261	0.0000	0.0094	**0.0000**	**0.0000**	**0.0000**	**0.0000**	0.0332	0.0128
WaveformEW	$Mean_{Acc}$	**0.8738**	0.8579	0.8611	0.8722	0.8694	0.8669	0.8718	0.8728	0.8696	0.8512	0.8639
	Std_{Acc}	0.0028	0.0052	0.0079	0.0037	0.0045	0.0049	**0.0028**	0.0045	0.0047	0.0075	0.0051
Ranking	W|T|L	1|3|2	0|1|5	0|1|5	1|2|3	0|1|5	0|3|3	**1|3|2**	0|3|3	0|3|3	0|1|5	0|1|5

TABLE 19. Comparisons of different algorithms based on SVM in terms of average number of selected features.

Benchmark	Metric	ABC	PSO	BA	GWO	WOA	GOA	SFO	HHO	BSA	ASO	HGSO
Exactly	$Mean_{Feat}$	006.87	007.53	007.40	006.53	007.60	007.27	006.43	006.23	007.17	**004.83**	008.03
	Std_{Feat}	000.85	001.67	002.32	001.54	001.58	001.59	**000.67**	002.69	001.49	002.65	001.62
Exactly2	$Mean_{Feat}$	001.43	002.27	002.47	001.03	001.50	001.30	**001.00**	**001.00**	001.53	003.13	002.80
	Std_{Feat}	000.50	000.63	000.76	000.18	000.62	000.46	**000.00**	**000.00**	000.50	000.81	000.79
KrVsKpEW	$Mean_{Feat}$	024.03	021.47	022.73	023.47	**021.40**	023.20	021.57	021.87	022.70	022.13	023.37
	Std_{Feat}	**002.59**	003.69	003.55	003.61	003.53	003.13	002.87	003.15	002.79	002.77	002.74
M-of-n	$Mean_{Feat}$	006.83	008.13	008.40	**006.53**	007.23	007.27	006.57	006.73	007.30	008.43	008.67
	Std_{Feat}	000.58	001.41	001.62	000.62	000.88	000.57	**000.50**	000.73	000.82	001.76	000.98
Tic-tac-toe	$Mean_{Feat}$	008.00	007.67	007.53	008.00	008.00	008.13	008.00	008.13	008.03	**006.63**	007.87
	Std_{Feat}	**000.00**	000.83	000.92	**000.00**	000.45	000.34	**000.00**	000.34	000.18	001.14	000.85
WaveformEW	$Mean_{Feat}$	026.30	024.37	**024.10**	024.80	025.53	024.93	025.57	024.20	024.50	025.27	025.53
	Std_{Feat}	002.35	002.76	003.17	002.98	**002.23**	002.62	002.72	002.29	002.55	003.47	003.01
Ranking	W|T|L	0|0|6	0|0|6	1|0|5	1|0|5	1|0|5	0|0|6	0|1|5	0|1|5	0|0|6	**2|0|4**	0|0|6

- Apart from the 11 meta-heuristic methods adopted herein, this study used k-NN as classification algorithm in a wrapper-based FS strategy due to its ease of implementation. However, its performance is often degraded for being a slow learner and thus vulnerable to noisy data. On the other hand, SVM is complex in nature and so revealed overall higher performance in terms of accuracy and selection ratio. It is noteworthy that switching to other classifiers may exacerbate the running time.

VI. CONCLUSION

In this paper, an SSL prediction model was proposed to develop the educational process during COVID 19 and solve issues impeding OL progress. Our model consists of four components: data preprocessing, FS, ML classifiers, and evaluating ML models. The dataset was collected via a questionnaire particularly designed to specifiy how students are affected by OL. The current study utilized some standard SSL evaluation criteria, including faculty member obligation (online), teaching and lectures (online), assessment systems, and E-Tests. The best set of features were selected using 11 wrapper-based FS algorithms. In addition, two ML classifiers, k-NN and SVM, were applied to all features and the selected ones to sense the difference.

The findings demonstrated that overall accuracy based on the selected features had been improved by 2% and 8% for k-NN and SVM, respectively, compared to using all features; the overall mean accuracy for k-NN and SVM achieved 100% with FS algorithms. The SFO algorithm with k-NN and SVM performs the best in terms of exploration and exploitation abilities (fitness). It only determined four features. We conclude that four features (instead of the 20 features) affected SSL and are sufficient to predict SSL with OL with a 100% prediction accuracy. The minimal, yet crucial, selected features are: "*The lectures are presented in an attractive style*", "*The teaching method in this course encourages me to participate actively during the classes*", "*The quiz has been prepared in degrees*", and "*Students trained on how to solve exams online by designing an experimental quiz*". This could help HEIs to predict SSL at an early stage and present the diagnosis and therapy to avoid hitches in the educational process and achieve the most significant possible outcomes during acute crises like the COVID-19. In the future, as Random Forest (RF) model can perfectly fit the input-output relationship with unlimited high complexity, it can be tried with the 11 FS methods on the proposed real-time dataset.

REFERENCES

[1] G. Bonaccorsi, F. Pierri, M. Cinelli, A. Flori, A. Galeazzi, F. Porcelli, A. L. Schmidt, C. M. Valensise, A. Scala, W. Quattrociocchi, and F. Pammolli, "Economic and social consequences of human mobility restrictions under COVID-19," *Proc. Nat. Acad. Sci. USA*, vol. 117, no. 27, pp. 15530–15535, Jul. 2020.

[2] H. Fang, L. Wang, and Y. Yang, "Human mobility restrictions and the spread of the novel coronavirus (2019-nCoV) in China," *J. Public Econ.*, vol. 191, Nov. 2020, Art. no. 104272.

[3] V. Singh and A. Thurman, "How many ways can we define online learning? A systematic literature review of definitions of online learning (1988–2018)," *Amer. J. Distance Educ.*, vol. 33, no. 4, pp. 289–306, Oct. 2019.

[4] C. Khurana, "Exploring the role of multimedia in enhancing social presence in an asynchronous online course," Ph.D. dissertation, Rutgers Univ.-Graduate School, New Brunswick, NJ, USA, 2016.

[5] M. A. Peters, H. Wang, M. O. Ogunniran, Y. Huang, B. Green, J. O. Chunga, E. A. Quainoo, Z. Ren, S. Hollings, C. Mou, S. W. Khomera, M. Zhang, S. Zhou, A. Laimeche, W. Zheng, R. Xu, L. Jackson, and S. Hayes, "China's internationalized higher education during COVID-19: Collective student autoethnography," *Postdigital Sci. Educ.*, vol. 2, no. 3, pp. 968–988, Oct. 2020.

[6] A. S. Abdulamir and R. R. Hafidh, "The possible immunological pathways for the variable immunopathogenesis of COVID-19 infections among healthy adults, elderly and children," *Electron. J. Gen. Med.*, vol. 17, no. 4, p. em202, Mar. 2020.

[7] A. Rasouli, Z. Rahbania, and M. Attaran, "Students' readiness for e-learning application in higher education," *Malaysian Online J. Educ. Technol.*, vol. 4, no. 3, pp. 51–64, 2016.

[8] S. Raime, M. F. Shamsudin, R. A. Hashim, and N. A. Rahman, "Students' self-motivation and online learning students' satisfaction among unitar college students," *Asian J. Res. Educ. Social Sci.*, vol. 2, no. 3, pp. 62–71, 2020.

[9] S. Wilkins and M. S. Balakrishnan, "Assessing Student satisfaction in transnational higher education," *Int. J. Educ. Manage.*, vol. 27, no. 2, pp. 143–156, Feb. 2013.

[10] A. Dutt, M. A. Ismail, and T. Herawan, "A systematic review on educational data mining," *IEEE Access*, vol. 5, pp. 15991–16005, 2017.

[11] N. Kapasia, P. Paul, A. Roy, J. Saha, A. Zaveri, R. Mallick, B. Barman, P. Das, and P. Chouhan, "Impact of lockdown on learning status of undergraduate and postgraduate students during COVID-19 pandemic in West Bengal, India," *Children Youth Services Rev.*, vol. 116, Sep. 2020, Art. no. 105194.

[12] M. Canayaz, "MH-COVIDNet: Diagnosis of COVID-19 using deep neural networks and meta-heuristic-based feature selection on X-ray images," *Biomed. Signal Process. Control*, vol. 64, Feb. 2021, Art. no. 102257.

[13] Q. Al-Tashi, H. Rais, and S. Jadid, "Feature selection method based on grey wolf optimization for coronary artery disease classification," in *Proc. Int. Conf. Reliable Inf. Commun. Technol.* Springer, 2018, pp. 257–266.

[14] A. I. Hammouri, M. Mafarja, M. A. Al-Betar, M. A. Awadallah, and I. Abu-Doush, "An improved dragonfly algorithm for feature selection," *Knowl.-Based Syst.*, vol. 203, Sep. 2020, Art. no. 106131.

[15] Q. Al-Tashi, H. M. Rais, S. J. Abdulkadir, S. Mirjalili, and H. Alhussian, "A review of grey wolf optimizer-based feature selection methods for classification," in *Evolutionary Machine Learning Techniques*, 2020, pp. 273–286.

[16] J. Tang, S. Alelyani, and H. Liu, "Feature selection for classification: A review," in *Data Classification: Algorithms and Applications*, 2014, p. 37.

[17] R. Alabdallah, A. Jaradat, I. Doush, and Y. Jaradat, "A binary classifier based on firefly algorithm," *Jordanian J. Comput. Inf. Technol.*, vol. 3, no. 3, p. 172, 2017.

[18] B. Chen, L. Chen, and Y. Chen, "Efficient ant colony optimization for image feature selection," *Signal Process.*, vol. 93, no. 6, pp. 1566–1576, 2013.

[19] A. I. Hammouri and S. Abdullah, "Biogeography-based optimisation for data clustering," in *New Trends in Software Methodologies, Tools and Techniques*. Amsterdam, The Netherlands: IOS Press, 2014, pp. 951–963.

[20] H.-J. Yu and D.-S. Huang, "Normalized feature vectors: A novel alignment-free sequence comparison method based on the numbers of adjacent amino acids," *IEEE/ACM Trans. Comput. Biol. Bioinf.*, vol. 10, no. 2, pp. 457–467, Mar. 2013.

[21] Y.-F. Gao, B.-Q. Li, Y.-D. Cai, K.-Y. Feng, Z.-D. Li, and Y. Jiang, "Prediction of active sites of enzymes by maximum relevance minimum redundancy (mRMR) feature selection," *Mol. BioSyst.*, vol. 9, no. 1, pp. 61–69, 2013.

[22] R. S. Broetto and F. M. Varejao, "Heterogeneous feature models and feature selection applied to detection of street lighting lamps types and wattages," in *Proc. 42nd Annu. Conf. IEEE Ind. Electron. Soc. (IECON)*, Oct. 2016, pp. 933–938.

[23] K. Zhang, Y. Li, P. Scarf, and A. Ball, "Feature selection for high-dimensional machinery fault diagnosis data using multiple models and radial basis function networks," *Neurocomputing*, vol. 74, no. 17, pp. 2941–2952, Oct. 2011.

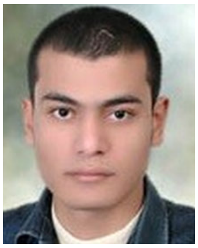
[24] T. T. Le, R. J. Urbanowicz, J. H. Moore, and B. A. McKinney, "Statistical inference relief (STIR) feature selection," *Bioinformatics*, vol. 35, no. 8, pp. 1358–1365, Apr. 2019.

[25] Z. Huang, C. Yang, X. Zhou, and T. Huang, "A hybrid feature selection method based on binary state transition algorithm and ReliefF," *IEEE J. Biomed. Health Informat.*, vol. 23, no. 5, pp. 1888–1898, Sep. 2019.

[26] T. Cover and P. Hart, "Nearest neighbor pattern classification," *IEEE Trans. Inf. Theory*, vol. IT-13, no. 1, pp. 21–27, Jan. 1967.
[27] K. K. Kandaswamy, G. Pugalenthi, M. K. Hazrati, K.-U. Kalies, and T. Martinetz, "BLProt: Prediction of bioluminescent proteins based on support vector machine and ReliefF feature selection," *BMC Bioinf.*, vol. 12, no. 1, pp. 1–7, Dec. 2011.
[28] E. Pashaei and N. Aydin, "Binary black hole algorithm for feature selection and classification on biological data," *Appl. Soft Comput.*, vol. 56, pp. 94–106, Jul. 2017.
[29] H. Faris, A.-Z. Ala'M, A. A. Heidari, I. Aljarah, M. Mafarja, M. A. Hassonah, and H. Fujita, "An intelligent system for spam detection and identification of the most relevant features based on evolutionary random weight networks," *Inf. Fusion*, vol. 48, pp. 67–83, Aug. 2019.
[30] M. Mafarja, I. Aljarah, A. A. Heidari, H. Faris, P. Fournier-Viger, X. Li, and S. Mirjalili, "Binary dragonfly optimization for feature selection using time-varying transfer functions," *Knowl.-Based Syst.*, vol. 161, pp. 185–204, Dec. 2018.
[31] I. Zelinka, "A survey on evolutionary algorithms dynamics and its complexity—Mutual relations, past, present and future," *Swarm Evol. Comput.*, vol. 25, pp. 2–14, Dec. 2015.
[32] I. A. Doush, M. A. Al-Betar, M. A. Awadallah, A. I. Hammouri, R. M. Al-Khatib, S. ElMustafa, and H. ALkhraisat, "Harmony search algorithm for patient admission scheduling problem," *J. Intell. Syst.*, vol. 29, no. 1, pp. 540–553, Apr. 2018.
[33] N. Neggaz, A. A. Ewees, M. A. Elaziz, and M. Mafarja, "Boosting Salp swarm algorithm by sine cosine algorithm and disrupt operator for feature selection," *Expert Syst. Appl.*, vol. 145, May 2020, Art. no. 113103.
[34] S. Mirjalili, "The ant lion optimizer," *Adv. Eng. Softw.*, vol. 83, pp. 80–98, May 2015.
[35] S. Mirjalili and A. Lewis, "The whale optimization algorithm," *Adv. Eng. Softw.*, vol. 95, pp. 51–67, Feb. 2016.
[36] S. Maldonado, R. Weber, and F. Famili, "Feature selection for high-dimensional class-imbalanced data sets using support vector machines," *Inf. Sci.*, vol. 286, pp. 228–246, Dec. 2014.
[37] B. Xue, M. Zhang, W. N. Browne, and X. Yao, "A survey on evolutionary computation approaches to feature selection," *IEEE Trans. Evol. Comput.*, vol. 20, no. 4, pp. 606–626, Aug. 2016.
[38] Z. Manbari, F. A. Tab, and C. Salavati, "Hybrid fast unsupervised feature selection for high-dimensional data," *Expert Syst. Appl.*, vol. 124, pp. 97–118, Jun. 2019.
[39] D. Karaboga, "An idea based on honey bee swarm for numerical optimization," Tech. Rep., 2005.
[40] R. Eberhart and J. Kennedy, "A new optimizer using particle swarm theory," in *Proc. 6th Int. Symp. Micro Mach. Human Sci.*, 1995, pp. 39–43.
[41] X.-S. Yang, "A new metaheuristic bat-inspired algorithm," in *Nature Inspired Cooperative Strategies for Optimization (NICSO)*. Springer, 2010, pp. 65–74.
[42] S. Saremi, S. Mirjalili, and A. Lewis, "Grasshopper optimisation algorithm: Theory and application," *Adv. Eng. Softw.*, vol. 105, pp. 30–47, Mar. 2017.
[43] S. Shadravan, H. R. Naji, and V. K. Bardsiri, "The sailfish optimizer: A novel nature-inspired metaheuristic algorithm for solving constrained engineering optimization problems," *Eng. Appl. Artif. Intell.*, vol. 80, pp. 20–34, Apr. 2019.
[44] A. A. Heidari, S. Mirjalili, H. Faris, I. Aljarah, M. Mafarja, and H. Chen, "Harris hawks optimization: Algorithm and applications," *Future Gener. Comput. Syst.*, vol. 97, pp. 849–872, Aug. 2019.
[45] X.-B. Meng, X. Z. Gao, L. Lu, Y. Liu, and H. Zhang, "A new bio-inspired optimisation algorithm: Bird swarm algorithm," *J. Exp. Theor. Artif. Intell.*, vol. 28, no. 6, pp. 673–687, Jun. 2015.
[46] W. Zhao, L. Wang, and Z. Zhang, "Atom search optimization and its application to solve a hydrogeologic parameter estimation problem," *Knowl.-Based Syst.*, vol. 163, pp. 283–304, Jan. 2019.
[47] F. A. Hashim, E. H. Houssein, M. S. Mabrouk, W. Al-Atabany, and S. Mirjalili, "Henry gas solubility optimization: A novel physics-based algorithm," *Future Gener. Comput. Syst.*, vol. 101, pp. 646–667, Dec. 2019.
[48] M. Usak, A. R. Masalimova, E. I. Cherdymova, and A. R. Shaidullina, "New playmaker in science education: Covid-19," *J. Baltic Sci. Educ.*, vol. 19, no. 2, p. 180, 2020.
[49] W. Bao, "COVID-19 and online teaching in higher education: A case study of Peking university," *Human Behav. Emerg. Technol.*, vol. 2, no. 2, pp. 113–115, Apr. 2020.
[50] E. J. Sintema, "Effect of COVID-19 on the performance of grade 12 students: Implications for STEM education," *Eurasia J. Math., Sci. Technol. Educ.*, vol. 16, no. 7, Apr. 2020, Art.no. em1851.
[51] C. M. Toquero, "Challenges and opportunities for higher education amid the COVID-19 pandemic: The Philippine context," *Pedagog. Res.*, vol. 5, no. 4, Apr. 2020, Art. no. em0063.
[52] M. L. Holshue *et al.*, "First case of 2019 novel coronavirus in the united states," *New England J. Med.*, pp. 929–936, Jan. 2020.
[53] A. Qazi, J. Qazi, K. Naseer, M. Zeeshan, S. Qazi, O. Abayomi-Alli, I. S. Ahmad, M. Darwich, B. A. Talpur, G. Hardaker, U. Naseem, S. Yang, and K. Haruna, "Adaption of distance learning to continue the academic year amid COVID-19 lockdown," *Children Youth Services Rev.*, vol. 126, Jul. 2021, Art. no. 106038.
[54] A. E. Clark, H. Nong, H. Zhu, and R. Zhu, "Compensating for academic loss: Online learning and student performance during the COVID-19 pandemic," *China Econ. Rev.*, vol. 68, Aug. 2021, Art. no. 101629.
[55] M. Maqableh and M. Alia, "Evaluation online learning of undergraduate students under lockdown amidst COVID-19 pandemic: The online learning experience and students' satisfaction," *Children Youth Services Rev.*, vol. 128, Sep. 2021, Art. no. 106160.
[56] E. Hussein, S. Daoud, H. Alrabaiah, and R. Badawi, "Exploring undergraduate students' attitudes towards emergency online learning during COVID-19: A case from the UAE," *Children Youth Services Rev.*, vol. 119, Dec. 2020, Art. no. 105699.
[57] D. U. Bolliger, "Key factors for determining student satisfaction in online courses," *Int. J. E-Learning*, vol. 3, no. 1, pp. 61–67, 2004.
[58] J. Macnish, S. Trinidad, D. Fisher, and J. Aldridge, "The online learning environment of a technology-rich secondary college," in *Proc. Annu. Meeting Amer. Educ. Res. Assoc.*, Chicago, IL, USA, Apr. 2003.
[59] Z. Zamakhsari and A. Ridzuan, "An investigation on students participation and satisfaction towards online learning," in *Proc. IEEE Conf. E-Learn., E-Manage. E-Services (IC3e)*, Aug. 2015, pp. 143–147.
[60] G. Giray, "An assessment of student satisfaction with e-learning: An empirical study with computer and software engineering undergraduate students in Turkey under pandemic conditions," *Educ. Inf. Technol.*, vol. 26, pp. 6651–6673, Mar. 2021.
[61] M. N. A. Rahman, S. N. A. S. Zamri, and L. K. Eu, "A meta-analysis study of satisfaction and continuance intention to use educational technology," *Int. J. Academic Res. Bus. Social Sci.*, vol. 7, no. 4, pp. 1059–1072, May 2017.
[62] R. Gopal, V. Singh, and A. Aggarwal, "Impact of online classes on the satisfaction and performance of students during the pandemic period of COVID 19," *Educ. Inf. Technol.*, vol. 26, pp. 6923–6947, Apr. 2021.
[63] Y. B. Rajabalee and M. I. Santally, "Learner satisfaction, engagement and performances in an online module: Implications for institutional e-learning policy," *Educ. Inf. Technol.*, vol. 26, no. 3, pp. 2623–2656, May 2021.
[64] M. Maqableh, M. Jaradat, and A. Azzam, "Exploring the determinants of students' academic performance at university level: The mediating role of internet usage continuance intention," *Educ. Inf. Technol.*, vol. 26, no. 4, pp. 4003–4025, Jul. 2021.
[65] A. A. Yunusa and I. N. Umar, "A scoping review of critical predictive factors (CPFs) of satisfaction and perceived learning outcomes in e-learning environments," *Educ. Inf. Technol.*, vol. 26, no. 1, pp. 1223–1270, Jan. 2021.
[66] C. Sammut and G. I. Webb, *Encyclopedia of Machine Learning*. Springer, 2011.
[67] L. Yin, Y. Ge, K. Xiao, X. Wang, and X. Quan, "Feature selection for high-dimensional imbalanced data," *Neurocomputing*, vol. 105, no. 3, pp. 3–11, 2013.
[68] A. Farissi and H. M. Dahlan, "Genetic algorithm based feature selection with ensemble methods for student academic performance prediction," *J. Phys., Conf. Ser.*, vol. 1500, no. 1, Apr. 2020, Art. no. 012110.
[69] W. Punlumjeak and N. Rachburee, "A comparative study of feature selection techniques for classify student performance," in *Proc. 7th Int. Conf. Inf. Technol. Electr. Eng. (ICITEE)*, Oct. 2015, pp. 425–429.
[70] S.-S.-M. Ajibade, N. B. Ahmad, and S. M. Shamsuddin, "An heuristic feature selection algorithm to evaluate academic performance of students," in *Proc. IEEE 10th Control Syst. Graduate Res. Colloq. (ICSGRC)*, Aug. 2019, pp. 110–114.
[71] Z. Davoodabadi and A. Moeini, "Building customers' credit scoring models with combination of feature selection and decision tree algorithms," *Adv. Comput. Sci., Int. J.*, vol. 4, pp. 97–103, Mar. 2015.

[72] J. Nalic and A. Svraka, "Importance of data pre-processing in credit scoring models based on data mining approaches," in *Proc. 41st Int. Conv. Inf. Commun. Technol., Electron. Microelectron. (MIPRO)*, May 2018, pp. 1046–1051.
[73] W. S. Walusala, R. Rimiru, and C. Otieno, "A hybrid machine learning approach for credit scoring using PCA and logistic regression," *Int. J. Comput.*, vol. 27, no. 1, pp. 84–102, 2017.
[74] M. Zaffar, M. Ahmed, K. S. Savita, and S. Sajjad, "A study of feature selection algorithms for predicting students academic performance," *Int. J. Adv. Comput. Sci. Appl.*, vol. 9, no. 5, pp. 541–549, 2018.
[75] B. H. Nguyen, B. Xue, and M. Zhang, "A survey on swarm intelligence approaches to feature selection in data mining," *Swarm Evol. Comput.*, vol. 54, May 2020, Art. no. 100663.
[76] J. Nalić, G. Martinović, and D. Žagar, "New hybrid data mining model for credit scoring based on feature selection algorithm and ensemble classifiers," *Adv. Eng. Informat.*, vol. 45, Aug. 2020, Art. no. 101130.
[77] C. Jalota and R. Agrawal, "Feature selection algorithms and student academic performance: A study," in *Proc. Int. Conf. Innov. Comput. Commun.* Springer, 2021, pp. 317–328.
[78] T. L. Best, "Bats: Biology and behaviour," *J. Mammal.*, vol. 78, no. 3, p. 986, 1997.
[79] E. Zorarpacı and S. A. Özel, "A hybrid approach of differential evolution and artificial bee colony for feature selection," *Expert Syst. Appl.*, vol. 62, pp. 91–103, Nov. 2016.
[80] H. M. Alabool, D. Alarabiat, L. Abualigah, and A. A. Heidari, "Harris hawks optimization: A comprehensive review of recent variants and applications," *Neural Comput. Appl.*, vol. 33, pp. 8939–8980, Feb. 2021.
[81] J. Staudinger and P. V. Roberts, "A critical review of Henry's law constants for environmental applications," *Crit. Rev. Environ. Sci. Technol.*, vol. 26, no. 3, pp. 205–297, Aug. 1996.
[82] N. S. Altman, "An introduction to kernel and nearest-neighbor nonparametric regression," *Amer. Statist.*, vol. 46, no. 3, pp. 175–185, Jan. 1992.
[83] Y. Wu, K. Ianakiev, and V. Govindaraju, "Improved *k*-nearest neighbor classification," *Pattern Recognit.*, vol. 35, no. 10, pp. 2311–2318, Oct. 2002.
[84] C. Cariou, S. Le Moan, and K. Chehdi, "Improving K-nearest neighbor approaches for density-based pixel clustering in hyperspectral remote sensing images," *Remote Sens.*, vol. 12, no. 22, p. 3745, Nov. 2020.
[85] Y. Chen and Y. Hao, "A feature weighted support vector machine and K-nearest neighbor algorithm for stock market indices prediction," *Expert Syst. Appl.*, vol. 80, pp. 340–355, Sep. 2017.
[86] M. Li, H. Xu, X. Liu, and S. Lu, "Emotion recognition from multichannel EEG signals using K-nearest neighbor classification," *Technol. Health Care*, vol. 26, no. S1, pp. 509–519, Jul. 2018.
[87] G. Chirici, M. Mura, D. McInerney, N. Py, E. O. Tomppo, L. T. Waser, D. Travaglini, and R. E. McRoberts, "A meta-analysis and review of the literature on the k-nearest neighbors technique for forestry applications that use remotely sensed data," *Remote Sens. Environ.*, vol. 176, pp. 282–294, Apr. 2016.
[88] B. Zhang and S. N. Srihari, "Fast k-nearest neighbor classification using cluster-based trees," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 26, no. 4, pp. 525–528, Apr. 2004.
[89] E. Emary, H. M. Zawbaa, and A. E. Hassanien, "Binary ant lion approaches for feature selection," *Neurocomputing*, vol. 213, pp. 54–65, Nov. 2016.
[90] T. Thaher, A. A. Heidari, M. Mafarja, J. S. Dong, and S. Mirjalili, "Binary Harris hawks optimizer for high-dimensional, low sample size feature selection," in *Evolutionary Machine Learning Techniques*. Springer, 2020, pp. 251–272.
[91] V. Vapnik, *The Nature of Statistical Learning Theory*. Springer, 2013.
[92] A. Tharwat, A. E. Hassanien, and B. E. Elnaghi, "A BA-based algorithm for parameter optimization of support vector machine," *Pattern Recognit. Lett.*, vol. 93, pp. 13–22, Jul. 2017.
[93] S. Mirjalili and A. Lewis, "S-shaped versus V-shaped transfer functions for binary particle swarm optimization," *Swarm Evol. Comput.*, vol. 9, pp. 1–14, Apr. 2013.
[94] O. Claveria, E. Monte, and S. Torra, "Economic uncertainty: A geometric indicator of discrepancy among experts' expectations," *Social Indicators Res.*, vol. 143, no. 1, pp. 95–114, May 2019.
[95] D. G. Saari, "Complexity and the geometry of voting," *Math. Comput. Model.*, vol. 48, nos. 9–10, pp. 1335–1356, Nov. 2008.
[96] H. Coxeter, "Affinely regular polygons," in *Abhandlungen aus dem Mathematischen Seminar der Universität Hamburg*. Springer, 1969, pp. 38–58.
[97] A. Asuncion and D. Newman, "UCI machine learning repository," Tech. Rep., 2007.

HANAN E. ABDELKADER received the Ph.D. degree from the Computer Teacher Preparation Department, Faculty of Specific Education, Mansoura University, in 2013. Since October 2017, she has been a Certified Instructor at the Professional Academy for Teachers, Ministry of Education, Egypt. She worked as a Humanities Sector Coordinator at the Office of Vocational Guidance for Rehabilitation and Continuous Training headed by the Vice President for Education and Student Affairs, from January 2017 to January 2018. She is currently an Assistant Professor of computer science and education with the Department of Computer Teacher Preparation, Faculty of Specific Education, Mansoura University, Egypt. Her research interests include computer science, artificial intelligence, and web mining.



AHMED G. GAD (Graduate Student Member, IEEE) received the B.Sc. degree (Hons.) from the Faculty of Computers and Information, Mansoura University, Egypt, in 2013. Since 2017, he has been a full-time Teaching Assistant of information technology with the Faculty of Computers and Information, Kafrelsheikh University, Kafrelsheikh, Egypt. His current research interests include meta-heuristics, optimization, machine learning, data mining, cloud computing, scheduling, and blockchain.

AMR A. ABOHANY received the B.Sc. degree from the Faculty of Computers and Information, Zagazig University, Egypt, in 2007, and the M.Sc. and Ph.D. degrees from the Faculty of Computers and Information, Helwan University, Egypt, in 2014 and 2018, respectively. He has more than 11 scientific research articles published in prestigious international journals in the topics of information systems. His current research interests include optimization, machine learning, and the Internet of Things.

SHAYMAA E. SOROUR received the Ph.D. degree in computer science and education from the Department of Advanced Information Technology, Faculty of Information Science and Electrical Engineering, Kyushu University, Japan, in 2016. She is currently an Assistant Professor of computer science and education with the Department of Educational Technology (Computer Teacher), Faculty of Specific Education, Kafrelsheikh University, Egypt. She is also the Director of the Quality Assurance Unit, Faculty of Specific Education, Kafrelsheikh University. Her research interests include computer science, artificial intelligence, and machine learning algorithms. She received the best paper awarded in the 5th IIAI International Congress on Advanced Applied Informatics, in July 2016.