TUNA-SWARM OPTIMIZATION ALGORITHM (TSOA) AND IMPROVED CONDITIONAL GENERATIVE ADVERSARIAL NETWORK-BASED ENSEMBLE DEEP LONG SHORT-TERM MEMORY (ICGAN-EDLSTM) FOR STUDENT PERFORMANCE PREDICTION DURING PANDAMIC

S. Sathiyapriya[1], Dr A. Kanagaraj[2], A. Shubha[3], N. Balakumar[4], P. Karthiga[5]

[1,3,4,5] Ph.D Research Scholar, Department of Computer Science, Nallamuthu Gounder Mahalingam College Pollachi, Tamil Nadu - 642001, India

[2] Assistant Professor, Department of Computer Science, Nallamuthu Gounder Mahalingam College Pollachi, Tamil Nadu - 642001, India.

s.sathipriya@gmail.com , a.kanagaraj@gmail.com

ABSTRACT - *The rapid growth of the COVID-19 epidemic put the education system in a problematic condition for examination analysis of students, particularly when the virtual engagement was challenging for learners. With the shift to mandatory virtual education during the outbreak, it has become especially more crucial to predict these students and provide educational programs to prevent dropping them ahead. This prediction is done by the different Data Mining (DM) and Artificial Intelligence (AI) algorithms in the previous research. However, only a limited study was available to estimate learners who fail in learning during the epidemic. Additionally, there is a need for students' performance data while participating in virtual learning during the outbreak. From this perspective, this article presents an automated student performance prediction framework, which implements partially available student learning records in the online education system. This framework encompasses four different phases: data gathering, attribute selection, prediction and performance analysis. At first, the student's database including student's sex, age, overall time spent in virtual classrooms, their emotions towards Covid-19, etc., is collected. After that, a new metaheuristic optimization algorithm called Tuna-Swarm Optimization Algorithm (TSOA) is introduced to choose the most relevant attributes and minimize the data dimension. Once all the relevant attributes are selected, an Improved Conditional Generative Adversarial Network-based Ensemble Deep Long Short-Term Memory (ICGAN-EDLSTM) classification system is developed to learn those attributes and predict the student's outcome during the Covid-19 outbreak. Finally, these algorithms are applied to the 3 different benchmark databases to evaluate their prediction efficiency compared to the other classical algorithms.*

KEYWORDS - Education data mining, Virtual learning, Tuna-swarm optimization, ICGAN, Long short-term memory network

I. INTRODUCTION

On March 11, 2020, the World Health Organization proclaimed the Covid-19 coronavirus a worldwide outbreak [1], forcing nations to make deliberate actions to deal with the virus. These actions influenced on the development of society [2]. In numerous nations, educational institutions were momentarily suspended. Information from the United Nations Educational, Scientific and Cultural Organization on April 2, 2020 [3] shows that over 1.5 billion pupils/learners (nearly 85% worldwide) were disrupted by classroom suspensions in 172 nations. So, there is a necessity to adopt valuable innovations in academia thus all learners could receive the greatest quality curriculum and one of the most significant such innovations is to provide video lectures (webinars). In today's world, online or virtual learning offers ample teaching programs not only for completing academic obligations but also for maintaining a relationship between academic personnel and learners [4].

With the uncertainty of economic problems and emerging developments in pace with the substantial modifications around the globe, many academic institutions have entered towards online learning education systems on a worldwide scale, which preferred to create and implement virtual curricula [5-6]. Any government's curriculum needs to be treated properly for it to develop. Since a variety of academic materials, all provinces have their curriculum and grading methods (smart learning, project-based training, virtual classes, webinars and so on). Nonetheless, when they are not properly examined, these methods will not succeed. So, each academic institution must have a well-defined grading system to certify its achievement [7]. When enrolling, all academic institutions generate a great volume of information on every pupil and if that information is not adequately processed, entire processes are discarded and no potential need for information exists [8-9].

In the current pandemic situation, educational institutions face substantial challenges in providing students with strong learning while simultaneously boosting their achievement. So, forecasting students who are at risk during the epidemic has become increasingly crucial for academic institutions, since learners may experience alone during virtual classes [10]. To avoid this inconvenience, it is critical to provide learners with positive and helpful comments in a moment. Learners at risk must be prioritized during web classes [11]. For more than a decade, academics have been focusing on implementing an Education Data Mining (EDM) system using DM and AI approaches to assess and estimate the performance of students, as well as, their main reason. Different DM and AI approaches were used to forecast diverse quality of education such as skills, achievements, attendance, dropouts and brilliance. Those approaches are immensely useful in the subject of virtual learning, particularly for assessing and forecasting the efficiency of online classes during pandemic situations [12-14].

Predicting learners' achievement at a prior phase of an examination is a highly valuable technique for premature interventions to improve their learning and to lower student dropouts after a semester. On the other hand, predicting students' academic success during the Covid-19 outbreak is a difficult task since several elements might influence student performance, like educational environment, which is prior educational success, social traits, financial conditions, personality tendencies and so on [15-16]. To the state of the art, only a limited study has been conducted that predicts learners who are ineffective intellectually during the epidemic.

From this viewpoint, a Fuzzy Distribution Elephant Herding Optimization (FDEHO)-based attribute selection was developed to choose the subgroup of behavioral and academic attributes for predicting the student's performance. The FDEHO was performed to choose every possible collection of attributes based on the objective value (i.e., accuracy). Once all attributes were chosen, an ICGAN-Based Ensemble Deep Support Vector Machine (ICGAN-EDSVM) classification system was introduced to estimate students' performance in school, university and family tutoring. This ICGAN-EDSVM system provides dual advantages for the setting of low model size in student's academic database, wherein the ICGAN augments the data size and the EDSVM classifier increases the accuracy of prediction using deep learning design. The EDSVM was utilized to train an SVM and the kernel initializations of support vectors were applied to train an SVM in the consecutive layers. In this EDSVM classifier, deep SVMs were an ensemble, where the product rule was utilized to determine the fused probability. In contrast, it needs additional attributes regarding students' performance in virtual classrooms during the outbreak.

To combat this challenge, a new prediction framework is designed in this article for predicting the student's performance in virtual classrooms during the Covid-19 pandemic. Initially, the student's database related to the online education system is collected, which comprises various attributes about the student's age, gender, time spent on social sites, interests and so on. Then, a new TSOA is introduced to choose the most relevant attributes and minimize the data dimension. After selecting the attributes, an ICGAN-EDLSTM classification system is developed to learn those attributes and predict the student's performance in e-learning during the Covid-19 epidemic efficiently. Thus, this newly presented framework can predict the student's learning quality during the Covid-19 pandemic more reliably. The following segments of this paper are prepared: Section II presents the comprehensive study of the existing students' performance prediction algorithms. Section III discusses the presented TSOA and ICGAN-EDLSTMalgorithms and their effectiveness is illustrated in Section IV. Section V summarizes the entire study and suggests possible improvements.

II. LITERATURE SURVEY

Chen et al. [17] designed a technique to estimate learning results from student behavior in online short programs. This technique was focused on behavior-based machine learning characteristics acquired by analyzing data gathered during the learning task and the learner's communication with the data, as well as, other online learning platforms. But, it needs more data and online learning platform characteristics to increase the prediction efficiency. Cazarez& Martin [18] suggested the neural network-based model to estimate student's performance in online education. But, it was not suitable for a huge amount of data since its training time was high. Wang et al. [19] presented a Decision Tree (DT) algorithm to create a training intervention framework of a dynamic training paradigm to estimate student behaviors and activity in online learning. But, it was time-consuming while increasing the number of data.

Francis and Babu [20] designed a novel prediction method to analyze students' performance in academia based on classification and clustering algorithms. First, the student data was pre-processed and various features were extracted. Then, these features were classified by the SVM, Naive Bayes (NB), DT and neural network classifiers to predict which of the features provide higher accuracy. After that, such features were fed to the K-means clustering and applied to a majority voting scheme to predict students' academic performance. But, it does not handle a huge variety of features from the student database. Li et al. [21] suggested the fuzzy C-means clustering algorithm to monitor the learner's achievement according to their examination results. However, it takes only the student's scores while the remaining characteristics of the students have been essential to enhance the prediction efficiency.

Alshabandar et al. [22] developed 2 predictive frameworks called student assessments grades and final student's performance to identify the factors that impact student's learning achievements in massive open online programs. In the initial framework, the student's past and present activities along with the past performance were considered to determine the learner's success. In the secondary framework, Multi-Layer Perceptron (MLP) was applied to estimate long-term learner achievement based on the behavioral, temporal and demographic attributes. But, it was not promising to integrate the temporal attributes with behavioral attributes since the dataset contains student temporal data for the entire duration.

A unified training [23] was developed by executing a pupil-centered learning scheme depending on the flicked learning and undersized personal virtual program. In this scheme, the training facilities were recreated to generate a unified training case; information, which represents student's training characteristics, was gathered. Such data were utilized to estimate their achievement using the multi-regression system. Also, learner's achievement was determined using the end classroom test and its probability in all semesters was examined for early interference. But, the correlation between interference period and predictive stability was needed to find using a huge number of tests.

Jiang & Wang [24] developed a preference cognitive analysis technique to formulate students' knowledge states. Initially, the direct-indirect technique was designed to obtain students' preferred learning materials. The relevant data was obtained from students' reading data, which defines their favorite for training resources to get such favorite training resources. Also, favorite training resources were obtained by examining the resemblance of students' training nature during the interpretation task. Besides, the student's favorite degree for knowledge ideas was determined according to the obtained favorite training resources and their talent was diagnosed for awareness ideas via the cognitive analysis technique. Moreover, such 2 factors were merged to formulate students' knowledge states and their scores on tests were estimated. But, it was complex to gather a huge amount of data because of limited participants and robust correlation among data.

Montaño& Cabrera-Loayza [25] explored Artificial Neural Network (ANN) to estimate the behavior of students utilizing historical data and obtaining outcomes in the early phases of student performance. But, it has a high complexity while increasing the number of hidden neurons and attributes. Deo et al. [26] introduced Extreme Learning Machines (ELM) to examine samples entrenched in constant analysis to devise the weighted rank and the test rank in engineering mathematics programs at an Australian provincial university. But, it needs more data like education gap, socio-economic status, activity on online platforms, etc. Dabhade et al. [27] presented the linear regression and support vector regression algorithms to estimate the educational quality of the last year's undergraduate learners of the particular institutions. But, these algorithms were prone to missing data and not suitable for a huge quantity of attributes mined from the student database.

Chi & Huang [28] suggested the Improved NB (INB)-based classifier model to estimate and advise the student achievement in the mid-term phase, which avoids the event of a huge fraction of omitted topics, thus guaranteeing the efficiency of student's training during the examinations. But, its accuracy was not efficient since it needs more attributes about the students and their learning scores. Matzavela and Alepis [29] developed DT learning using a predictive framework to classify the learner's talent level according to the weights of the DT. But, it needs more dynamic features that provide an effective m-learning platform in tertiary education. Ulloa-Cazarez et al. [30] developed a Multi-layer Adaptive Neuro-Fuzzy Inference System (MANFIS) to predict the learner's achievement in virtual post-graduation. This system has been modeled and evaluated by the database merged with the ranks attained by learners in different virtual post-graduation programs. But, the dataset utilized was limited to a single online educational course.

Tian et al. [31] concentrated on huge free virtual program learners' choice for special name confession in their display names as an analyst of their ultimate success at the closing stages of a program. But, it needs to analyze whether the complicated level of programs influences the student's performance on online learning platforms. Rodríguez-Hernández et al. [32] utilized ANN to categorize learner's training efficiency into high and low. But, the significant data defined by the higher class mean score was not considered. Also, each data about learner's socioeconomic situations was self-reported by the learners, which tends to be an inaccurate prediction.

III. PROPOSED METHODOLOGY

In this section, the presented system for predicting students' performance during the COVID-19 pandemic is explained briefly. The major goal of this work is to model and train sufficient data, as well as, design a prediction system with maximum accuracy. The presented system involves 4 key phases: (i) data acquisition, (ii) attribute selection, (iii) prediction and (iv) performance analysis. Initially, the database is acquired from Kaggle and the redundant attributes are discarded by the optimization algorithm. Then, an ICGAN-EDLSTM classifier is proposed to design a prediction model. At last, the efficiency of the presented algorithms is evaluated regarding different evaluation metrics. The phases included in this presented system are illustrated in Figure 1.

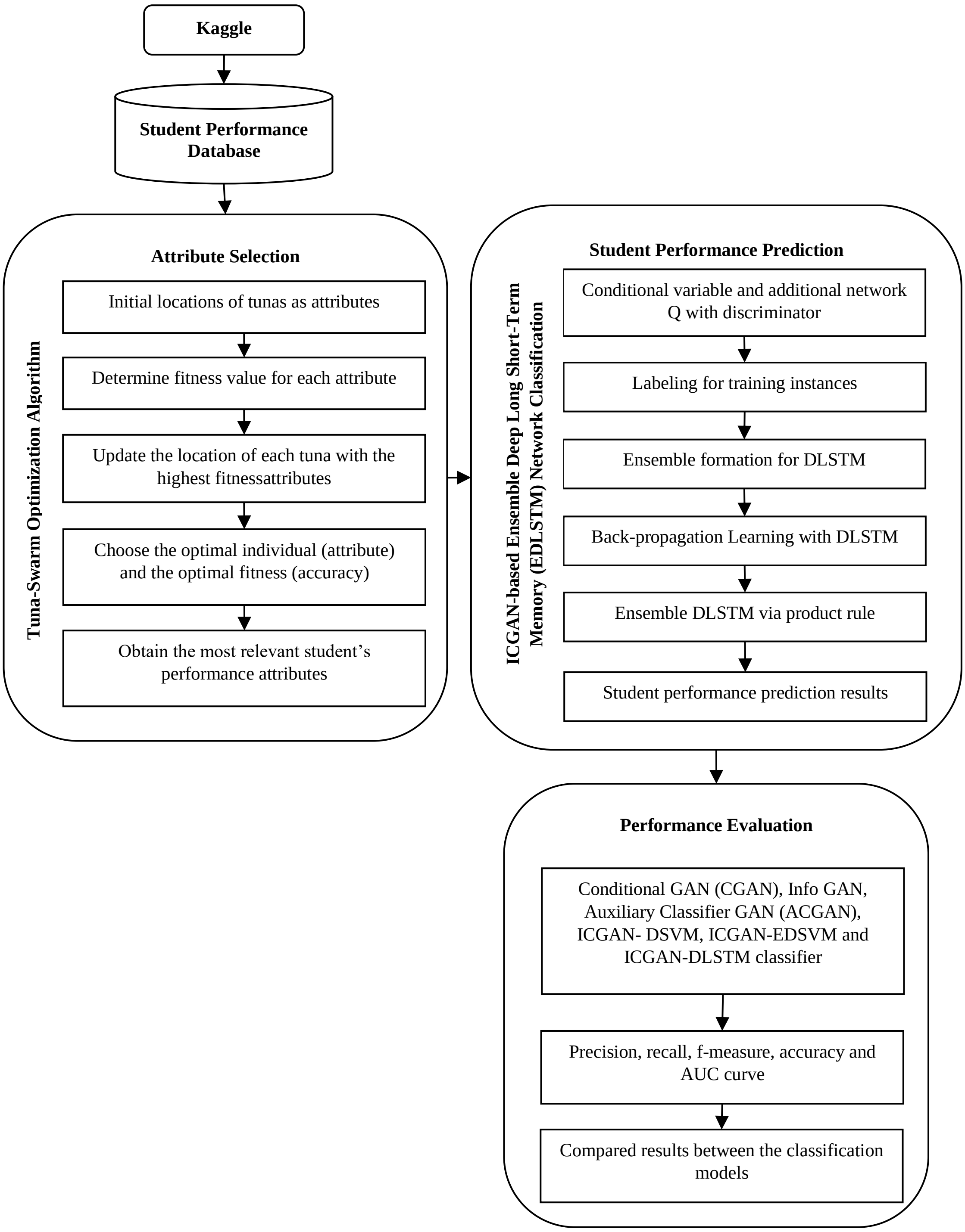


Figure 1. Proposed Student Performance Prediction Model

3.1 Data Acquisition

In this study, 3 different benchmark databases are considered for analysis.

Dataset 1 (Student performance school dataset): The dataset for student performance prediction was retrieved from recent work. It is comprised of two classes from 1044 students (i) Portuguese language class of 649 records; and (ii) Mathematics class of 395 records. The dataset has 33 attributes of which 9 of them are related to school tutoring and family tutoring. Parental cohabitation status, mother's education, mother's work, father's education, father's job, student's guardian, the standard of family relationships, school educational support, and family educational support are some of the characteristics. The remaining 29 attributes were gathered through a questionnaire, with the remainder coming from school records. Student's school, sex, age, home address type, family size, the reason for choosing this school, home-to-school travel time, weekly study time, number of previous class failures, extra paid classes within the course subject, extracurricular activities, attended nursery school, wants to pursue higher education, internet access at home, with romantic relations are some of these characteristics.

Dataset 2 (University Dataset): This dataset was created to analyze professors' and students' results. The dataset is designed to monitor student's progress through a course from the time they enroll to the time they complete it. Yes, of course. The dataset is made up of data from a fictional university. Information on its divisions, professors, student counseling, course offerings, course students were chosen for entry and student success on various exams. This dataset is about Professors' and Students' success over a period. Simply delve into your college memories, frame questions, and use this dataset to find answers. Once the dataset is collected then the next step is to perform feature selection for reducing the total number of features in the dataset for student performance prediction.

Dataset 3 (Covid-19 Go Away 2020 (C-19GA20) Database) [33]**:** The C-19GA20 dataset was generated digitally in April 2020 from school and university students aged 14 to 24. It includes data regarding learners' emotional stability, social life, opinions toward Covid-19, the influence of the Covid-19 outbreak on learners' academics and their expertise with virtual classrooms (See Table 1). This database comprises 5 main sets of attributes:

(a). Socio-demographic information: Learner's age, sex, the present location of the home and their learning degree

(b). Four entries for data about network access during quarantine: Gadget accessibility for exclusive usage, network bandwidth, 5 most important utilized web applications and screen time.

(c). Nine entries determined the effect of Covid-19 on the learner's social lives: Learner's present living conditions, number of people around them where they reside, their emotions about seeing their classmates, contacting their institutions of learning and activities that were conducted offline. During the quarantine, learners were questioned about their five major previous activities and how much time they spent on social sites.

(d). Six entries were utilized to assess their satisfaction with virtual classrooms during the quarantine: queries regarding emotionally close to their friends, obedience, scheduled training and the anxiety they experienced as a result of virtual classrooms during the quarantine.

(e). Eleven entries were used to collect extensive data regarding learner's emotional stability: How well they adjusted to stay-at-home guidelines, their overall emotions during the quarantine, experiences against Covid-19, their major worries about the curriculum, being upgraded and advised regarding Covid-19, as well as, the effect of online sites on their opinions.

Moreover, the learners were encouraged to write about how they believe the outbreak has impacted them as a human and influenced their thoughts, as well as, the learners were requested to submit a single-line statement for the nation during the quarantine.

Table 1. Online Education System – Review Database

S. No.	Attributes	Description	Category
1.	Timestamp	Date – Month – Year	Numeric
2.	Informed Consent for Data Sharing - Your responses and any other information that could be used to tie you to the study will be kept private by the research assistants. If university or government officials are checking on the integrity of this research, laws may force us to provide them with information. Personal data, such as names and addresses, will be protected and kept in a safe environment. You won't be mentioned by name in any study publications. We shall adhere to all ethical concerns of respondents.	I Agree	Nominal
3.	Year of birth (learners' age)	Years	Numeric

4.	Sex (learner's sex)	Male or Female	Binary
5.	Nationality	Indian	Binary
6.	State	Delhi / Haryana / Rajasthan / ….	Nominal
7.	City	Bangalore / Aluva / Indore / ….	Nominal
8.	Pin Code	110032 / 246149 / ….	Numeric
9.	Studying in	School, Graduate or Post Graduate	Nominal
10.	During this period do you have a smartphone/ computer for your exclusive use?	Yes or No	Binary
11.	Average number of hours per day that you spend on smartphone or computer (your screen time) during this lockdown?	Hours	Numeric
12.	Out of the following, choose 5 tools which you use most often these days.	WhatsApp, Instagram or Facebook	Nominal
13.	During the period of the current COVID-19 lockdown, you are:	At home with family or Other	Nominal
14.	Number of people you are staying with during this period	1 - 25	Numeric
15.	Do you miss seeing your college and school friends in person?	Number range (Extremely bad to Extremely good)	Numeric
16.	Do you spend more time chatting online with your friends now than when there was no lockdown?	Number range (Extremely bad to Extremely good)	Numeric
17.	Do you feel disappointed for events and opportunities (e.g., annual festival, farewell etc.) in college/ school you had been looking forward to?	Number range (Extremely bad to Extremely good)	Numeric
18.	Do you miss shopping in a market or mall etc.?	Number range (Extremely bad to Extremely good)	Numeric
19.	Other than attending online classes for your course, what are the 5 activities that you mostly indulge in during this period?	Spending time with family, Reading books, Sleeping, Cooking, Catching up on your fitness and weight management, Others	Nominal
20.	Other than for academic activities, how much time on an average do you spend per day on social media apps these days?	Hours	Numeric
21.	What is the one thing you will want do on the first day when lockdown is lifted?	Meet friends, Gym, Shopping, Others	Nominal
22.	Is the online teaching-learning helping you feel connected as a group?	Number range (Very bad to Very good)	Numeric
23.	Is the online teaching-learning helping you in maintaining a routine?	Number range (Very bad to Very good)	Numeric
24.	Is the online teaching-learning providing some hours of structured learning?	Number range (Very bad to Very good)	Numeric
25.	Do you look forward to these online teaching sessions?	Number range (Very	Numeric

		bad to Very good)	
26.	Do you find these classes too much burden in these difficult times?	Number range (Very bad to Very good)	Numeric
27.	In your opinion, what are the things in classroom teaching that cannot be substituted by online teaching?	Feeling, Attention, Classroom environment, Others	Nominal
28.	How well you have adapted to strictly stay-at-home restrictions?	Number scale (Extremely bad to Extremely good)	Numeric
29.	Which of the following best describes your overall mood these days?	Relaxed, Hopeful, Others	Nominal
30.	Tick the ones that describe your feelings towards COVID-19?	Stressed, Hopeful, Others	Nominal
31.	Which among the following are your prime concerns due to changes in academic schedule because of COVID-19?	Examination, Admission, Job offer, Others	Nominal
32.	Average number of hours per day that you spend on updating yourself on Covid-19 related news?	Hours	Numeric
33.	How do you keep yourself informed about COVID-19?	Television, Online news, Others	Nominal
34.	Do you believe all content you get on social media groups?	Number scale (Extremely disagree to Strongly agree)	Numeric
35.	Do you forward most content you get on social media groups?	Number scale (Extremely disagree to Strongly agree)	Numeric
36.	Do you verify the authenticity of the messages you forward on social media groups i.e., they are correct or have fake information?	Number scale (Extremely disagree to Strongly agree)	Numeric
37.	How has the pandemic COVID-19 changed you as a person? How has it changed your thinking process?	No, very much or Others	Nominal
38.	Your One Line message to the World during the Lockdown !!	Stay home stay safe, Be safe, or Others	Numeric

After acquiring the database, the attribute selection is conducted to minimize the overall number of attributes in the acquired database for achieving an effective prediction of students' performance.

3.2 Wrapper Attribute Selection

Wrapper attribute selection is applied by the AI approach, which utilizes a search scheme to find the accurate attribute space and select the most relevant attributes for enhancing the prediction accuracy. It includes the TSOA to choose the most relevant attributes for predicting student outcomes during the Covid-19 outbreak. This TSOA relies on the collaborative hunting nature of tuna swarm. Tuna, also known as Thunnini is a marine carnivorous fish. Tunas are major marine hunters that eat a wide range of midwater and surface species. They have a distinct swimming style (i.e., the fishtail form), wherein the torso continues rigid whilst the long, slender tail rolls rapid. Even if a particular tuna whirls rapidly, it is not as quick as the nimble tiny fish reaction. So, the tuna can utilize the crowd movement strategy for hunting [34]. They utilize their cleverness to discover and hit their prey. Such species have progressed different efficient and smart foraging policies.The initial policy is spiral hunting. If tuna eat, they whirl in a spiral arrangement to force their food into shallow water, wherein they may be simply hit. Another policy is parabolic foraging. All tunas follow the former one, constructing a parabolic arc to encircle its prey. According to these foraging policies, this TSOA is modeled and applied as the attribute selection method.The processes in this TSOA are described below:

a. *Initialization:* It initiates the task of optimization by creating initial inhabitants randomly in the hunt space as:

$$X_i^{int} = rand \cdot (l_u - l_l) + l_l, i = 1,2, \dots, N \tag{1}$$

In Eq. (1), X_i^{int} defines i^{th} initial individual, l_u and l_l denote the high and low margins of the hunt region, N represents the quantity of tuna inhabitants and $rand$ indicates the regularly distributed arbitrary vector between 0 and 1.

b. *Spiral hunting:* While sardines, herring and another tiny moving fish interacted with hunters, the complete group of fish grows a dense environment, which continually modifies direction, discovering it complex for hunters to focus on a victim. At this point, the tuna swarm is chasing the prey in a strict spiral shape. Although most of the fish in the crowd contain no vision, if a tiny group of fish swims strongly in specific direction, the adjacent fish will alter their direction one by one until they produce a huge crowd with an identical intention and instigate to forage. As well, circling after their prey, swarms of tuna exchange data with all others. All tunas pursue the earlier fish, therefore facilitating data distribution among adjacent tuna. According to these rules, the spiral foraging policy is defined by

$$X_i^{t+1} = \begin{cases} \alpha_1 \cdot (X_{best}^t + \beta \cdot |X_{best}^t - X_i^t|) + \alpha_2 \cdot X_i^t, & i = 1 \\ \alpha_1 \cdot (X_{best}^t + \beta \cdot |X_{best}^t - X_i^t|) + \alpha_2 \cdot X_{i-1}^t, & i = 2,3, \dots, N \end{cases} \tag{2}$$

$$\alpha_1 = a + (1 - a) \cdot \frac{t}{t_{max}} \tag{3}$$

$$\alpha_2 = (1 - a) - (1 - a) \cdot \frac{t}{t_{max}} \tag{4}$$

$$\beta = e^{kq} \cdot \cos(2\pi k) \tag{5}$$

$$q = e^{3\cos\left(\left((t_{max}+1/t)-1\right)\pi\right)} \tag{6}$$

In Eqns. (2) – (6), X_i^{t+1} denotes i^{th} individual of $t+1$ iteration, X_{best}^t represents the present optimal individual (prey), α_1 and α_2 denote the weight factors, which manage the affinity of individuals to travel towards the best individual and the earlier individual, a refers to the constant utilized to compute the scale at which the tuna pursues the best individual and the earlier individual in the primary stage, t is the present iteration, t_{max} denotes the maximum iterations and k indicates the arbitrary number regularly distributed from 0 to 1.

Every tuna has strong exploitation abilities for the hunt space around the prey if they forage swirly around the prey. On the other hand, if the best individual cannot locateprey, after the best individual to hunt is not favorable to cluster hunting. As a result, a random direction in the hunt region is created as an orientation point for spiral hunt. It allows every individual to explore a larger area and provides TSOA with global search capabilities. It is defined by

$$X_i^{t+1} = \begin{cases} \alpha_1 \cdot (X_{rand}^t + \beta \cdot |X_{rand}^t - X_i^t|) + \alpha_2 \cdot X_i^t, & i = 1 \\ \alpha_1 \cdot (X_{rand}^t + \beta \cdot |X_{rand}^t - X_i^t|) + \alpha_2 \cdot X_{i-1}^t, & i = 2,3, \dots, N \end{cases} \tag{7}$$

In Eq. (7), X_{rand}^t denotes the arbitrarily created orientation point in the hunt region.

In general, metaheuristic schemes accomplish significant global search in the premature steps and slowly change to accurate local search. So, TSOA alters the orientation points of spiral hunting from arbitrary individuals to best individuals as the iteration grows. To summarize, the absolute spiral hunting policy is defined by

$$X_i^{t+1} = \begin{cases} \alpha_1 \cdot (X_{rand}^t + \beta \cdot |X_{rand}^t - X_i^t|) + \alpha_2 \cdot X_i^t, & i = 1, \quad if\ rand < \frac{t}{t_{max}} \\ \alpha_1 \cdot (X_{rand}^t + \beta \cdot |X_{rand}^t - X_i^t|) + \alpha_2 \cdot X_{i-1}^t, & i = 2,3, \dots, N, \\ \alpha_1 \cdot (X_{best}^t + \beta \cdot |X_{best}^t - X_i^t|) + \alpha_2 \cdot X_i^t, & i = 1, \quad if\ rand \geq \frac{t}{t_{max}} \\ \alpha_1 \cdot (X_{best}^t + \beta \cdot |X_{best}^t - X_i^t|) + \alpha_2 \cdot X_{i-1}^t, & i = 2,3, \dots, N, \end{cases} \tag{8}$$

c. *Parabolic hunting:* Besides, tunas feed in a parabolic collaborative arrangement. With prey as a reference point, tuna develops a parabolic shape. As well, tuna hunt for prey through staring around them. Such 2 procedures are conducted simultaneously, with the premise that the selection chance for each is 50%. This parabolic hunting is defined by

$$X_i^{t+1} = \begin{cases} X_{best}^t + rand \cdot (X_{best}^t + X_i^t) + c \cdot p^2 \cdot (X_{best}^t + X_i^t), & if\ rand < \frac{1}{2} \\ c \cdot p^2 \cdot X_i^t, & if\ rand \geq \frac{1}{2} \end{cases} \tag{9}$$

$$p = \left(1 - \frac{t}{t_{max}}\right)^{\left(\frac{t}{t_{max}}\right)} \tag{10}$$

In Eq. (9), c denotes the arbitrary number with a value of either 1 or -1.

Tuna forage jointly using 2 foraging policies and discover their food. The population in the hunt space is initially created at random for the TSOA. In all iterations, all individuals select one of the 2 foraging policies at random or select to regenerate the position in the hunt region according to the chance z. In the whole optimization task, every individual of TSOA is constantly modified and determined until the termination conditions is satisfied, at which the optimum individual and the respective fitness value (maximum prediction accuracy) are obtained. The pseudocode for TSOA is presented in Algorithm 1.

Algorithm 1: Pseudocode for TSOA

Input: Population size (N) and highest iteration (t_{max})
Result: Position of prey (the most excellent individual) and its fitness range (prediction accuracy)

Begin

Initialize the arbitrary inhabitant of tuna swarms $X_i^{int}(i = 1, \ldots, N)$ (student performance attributes in the databases)

Set variables $a = 0.7$ and $z = 0.05$;

$\boldsymbol{while}(t < t_{max})$

Determine the fitness values (prediction accuracy) of tuna swarms (attributes related to the student performance);

Update X_{best}^t;

$\boldsymbol{for}(all\ tuna\ swarms)$

Update α_1, α_2 and p;

$\boldsymbol{if}(rand < z)$

Update the location X_i^{t+1} based on Eq. (1);

$\boldsymbol{elseif}(rand \geq z)$

$\boldsymbol{if}\left(rand < \frac{1}{2}\right)$

$\boldsymbol{if}\left(\frac{t}{t_{max}} < rand\right)$

Update the location X_i^{t+1} based on Eq. (7);

$\boldsymbol{elseif}\left(\frac{t}{t_{max}} \geq rand\right)$

Update the location X_i^{t+1} based on Eq. (2);

$\boldsymbol{elseif}\left(rand \geq \frac{1}{2}\right)$

Update the location X_i^{t+1} based on Eq. (9);

$\boldsymbol{end\ if}$

$\boldsymbol{end\ if}$

$\boldsymbol{end\ if}$

$\boldsymbol{end\ for}$

$t = t + 1;$

$\boldsymbol{end\ while}$

Return the best individual X_{best} (optimal attributes related to the student performance) and the best fitness value $F(X_{best})$ (i.e., maximum prediction accuracy)

End

Once the relevant attributes are chosen, the following phase is to conduct the student performance prediction.

3.3 Student Performance Prediction

The ICGAN is initially applied to create the training student performance data. It is integrated with the EDLSTM, which directs the student performance prediction. This ensemble learning technique is developed to increase the LSTM's efficiency for predicting the student's performance using the considered databases. The ensemble should build proper and fuse results from the LSTM classifier to design a robust ensemble classifier. To fuse many probability results from LSTM classifiers, a product rule is utilized. It creates a 2-layer LSTM classifier for each one-vs-all classification model. In the initial layer, a series of β-stabilized LSTM classifiers is segregated and trained independently. Then, the help vector activations from all classifiers in the initial layer are obtained to train the other β-stabilized LSTM classifier for the second layer design, which belongs to a similar one-vs-all classifier. Because the inputs to the second layer classifiers are created on activations of prototype samples instead of the fundamental functions, the results from the second layer might contain higher variations than the initial layer. It comprises many LSTM layers and an output LSTM layer in general.

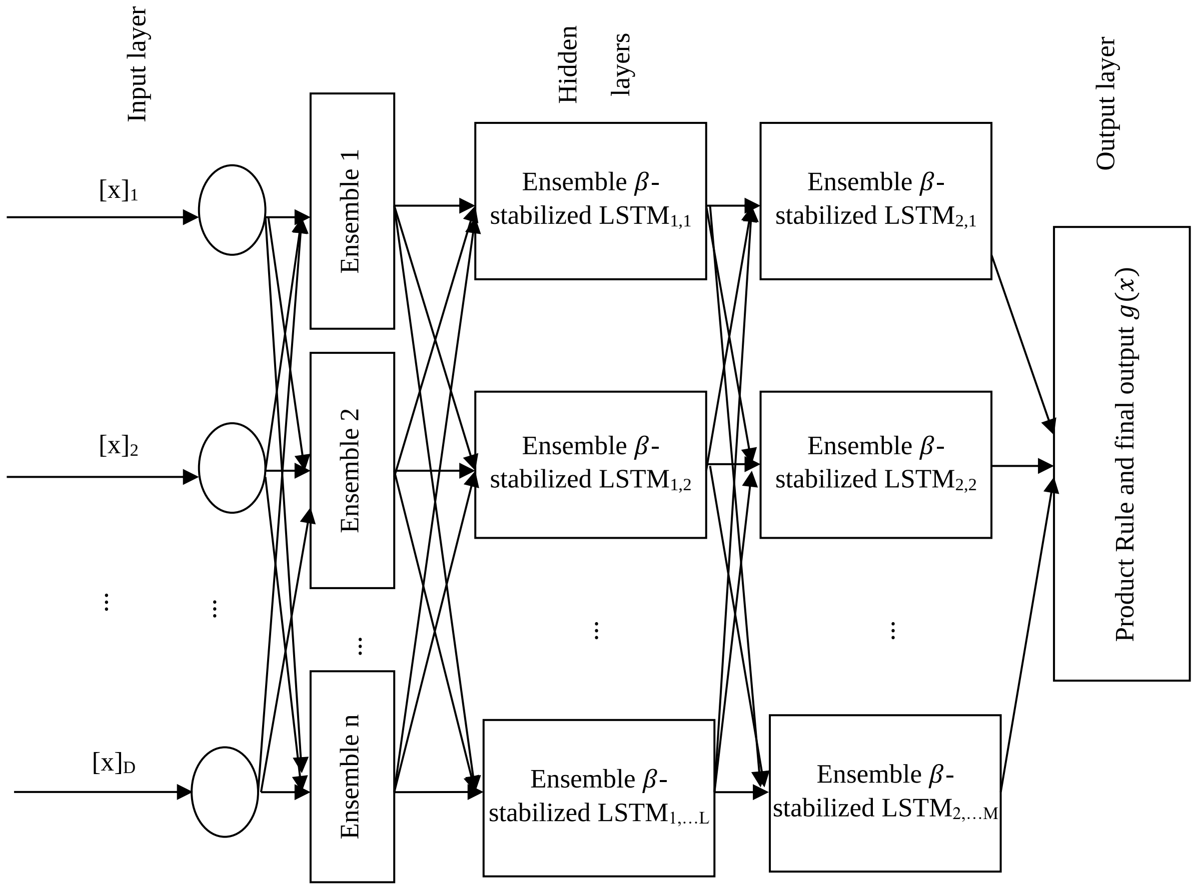


Figure 2. Conceptual Flow of EDLSTM Model

An ensemble DLSTM framework is portrayed in Figure 2, wherein D,L,M, N and P denote integers. Typically, the LSTM is a structure, which utilizes memory cell to preserve data. Figure 3 portrays the standard LSTM network, which has a 4-level configuration including connections in which h_{t-1} and h_t denote the results of the past and the ongoing cells, x_t defines the input of the ongoing cell, Sigmoid and $tanh$ indicate the activation factors, C_t defines the neuron condition at period t and f_t indicates the forgetting threshold for changing the chance of not remembering the last neuron condition via the sigmoid factor. Next, $tanh$ factor is employed to produce a new memory C_t' and it modifies how much new information is added to the neuron condition.

Additionally, o_t refers to the resultant threshold that predicts the outcome neuron conditions using the sigmoid factor and employs $tanh$ to analyze the neuron condition for obtaining the ultimate solution.

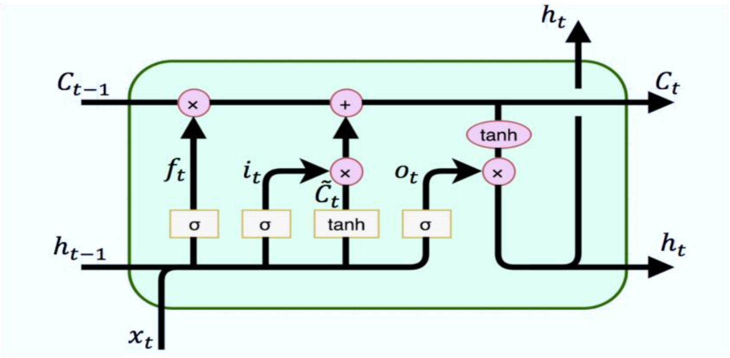


Figure 3. Standard LSTM Network Structure

It is performed based on the below formulas:

$$i_t = sigmoid(w_{xi}x_t + w_{hi}h_{t-1} + w_{Ci}C_{t-1} + b_i) \quad (11)$$

$$f_t = sigmoid\left(w_{xf}x_t + w_{hf}h_{t-1} + w_{Cf}C_{t-1} + b_f\right) \quad (12)$$

$$C_t = f_t \cdot C_{t-1} + i_t \cdot tanh(w_{xC}x_t + w_{hC}h_{t-1} + b_C) \quad (13)$$

$$o_t = sigmoid(w_{xo}x_t + w_{ho}h_{t-1} + w_{Co}C_t + b_o) \quad (14)$$

$$h_t = o_t \cdot tanh(C_t) \quad (15)$$

In Eqns. (11) – (15), w_{xi}, w_{hi}, w_{Ci} denote the weight vectors for the input gate, w_{xf}, w_{hf}, w_{Cf} characterize the weight vectors for the forgetting gate and w_{xo}, w_{ho}, w_{Co} denote the weight vectors for the output gate and w_{xC}, w_{hC} represent the weight vector for the neuron state vectors. Also, b_f, b_i, b_o and b_C are the offset constants of the input, forgetting, output gates and neuron state vectors, correspondingly. Similarly, C_t denotes the cell condition and h_t denotes the hidden condition that serves as the solution of the unit over t. The tensors i_t, f_t and o_t are the control gates.

In single LSTM layer, there are 3 gates and a major affine function. To enhance the LSTM network, an independent β-stabilizer is incorporated with all linear transform functions. It is considered that independent β-stabilizer may fine-tune the scale of all matrices individually and properly [35]. As a result, Eqns. (16) – (20) can be rewritten as:

$$i_t = sigmoid\left(e^{\beta_{xi}}w_{xi}x_t + e^{\beta_{hi}}w_{hi}h_{t-1} + e^{\beta_{ci}}w_{ci}c_{t-1} + b_i\right) \quad (16)$$

$$f_t = sigmoid\left(e^{\beta_{xf}}w_{xf}x_t + e^{\beta_{hf}}w_{hf}h_{t-1} + e^{\beta_{cf}}w_{cf}c_{t-1} + b_f\right) \quad (17)$$

$$c_t = f_t \cdot c_{t-1} + i_t \cdot tanh\left(e^{\beta_{xc}}w_{xc}x_t + e^{\beta_{hc}}w_{hc}h_{t-1} + b_c\right) \quad (18)$$

$$o_t = sigmoid\left(e^{\beta_{xo}}w_{xo}x_t + e^{\beta_{ho}}w_{ho}h_{t-1} + e^{\beta_{co}}w_{co}c_t + b_o\right) \quad (19)$$

$$h_t = o_t \cdot tanh(c_t) \quad (20)$$

The β-stablized LSTM is trained by the selected student's performance attributes based on the back-propagation scheme. During the training process, the back-propagation scheme needs to determine the gradient of fitness factor for x, w, b and β. The gradients with respect to x and w have less modifications,

$$\frac{\partial L}{\partial x} = e^{\beta}w^T\frac{\partial L}{\partial o} \text{ and } \frac{\partial L}{\partial w} = e^{\beta}\frac{\partial L}{\partial o}x^T \quad (21)$$

The gradients with respect to b remains unaltered,

$$\frac{\partial L}{\partial b} = \frac{\partial L}{\partial o} \quad (22)$$

The final challenge is how to fine-tune the control variable β. Using the chain rule,

$$\frac{\partial L}{\partial \beta} = \frac{\partial L}{\partial o}\frac{\partial L}{\partial \beta} = e^{\beta}\frac{\partial L^T}{\partial o}wx \quad (23)$$

Since

$$\frac{\partial L^T}{\partial x} = e^{\beta}\frac{\partial L^T}{\partial o}w \quad (24)$$

Obtain

$$\frac{\partial L}{\partial \beta} = \frac{\partial L^T}{\partial x}x \quad (25)$$

That is, the inner product of $\frac{\partial L}{\partial x}$ and x. The fine-tuning rule is as:

$$\beta = \beta - \eta\frac{\partial L^T}{\partial x}x \quad (26)$$

This refers to the range of β depends on the association between layer input and its gradient. It observes that β can be raised when scaling x up will enhance the prediction efficiency and vice versa. During initialization, each β value is assigned as 0 so $e^{\beta} = 1$, where the primary model remains similar with the one without control variable. Thus, the ICGAN-EDLSTM is trained to predict the student's outcome during pandemic with a high precision.

IV. RESULTS AND DISCUSSION

The efficiency of the presented deep learning classifier is evaluated by implementing it in MATLAB 2020a with the help of 3 different benchmark databases. In this comparative analysis, the efficiency of the proposed ICGAN-EDLSTM is compared with the CGAN, InfoGAN, ACGAN, ICGAN-DSVM, ICGAN-EDSVM and ICGAN-DLSTM classifiers. The system configurations are Intel CoreTM i7-11375H processor (12M Cache, up to 5 GHz processor) with 11th gen, 4 GB RAM, Windows 8.1 Pro, 64-bit OS and 1 TB hard disk.

4.1 Performance Evaluation Metrics

The efficiency of the classifiers on 3 benchmark databases is evaluated in the MATLAB tool. To assess the efficiency of various classification algorithms, precision, recall, f-measure, accuracy and Area Under Curve (AUC) are considered. The confusion matrix as shown in Table 2 is utilized to determine the True Positive (TP), True Negative (TN), False Positive (FP) and False Negative (FN).

Table 2. Confusion Matrix

Overall data instances		Classified label	
		Positive	**Negative**
Real label	**Positive**	**TP**	**FN**
	Negative	**FP**	**TN**

The evaluation metrics are defined as follows:

- **Precision:** It is proportion of exactly predicted instances from those predicted as positives.
  $$Precision = \frac{TP}{TP+FP} \quad (27)$$
- **Recall:** It is the fraction of exactly estimated positive instances from actual positives.
  $$Recall = \frac{TP}{TP+FN} \quad (28)$$
- **F-measure:** It is determined by
  $$F - measure = \frac{2\times(Precision \times Recall)}{Precision+Recall} \quad (29)$$
- **Accuracy:** It measures the classifier's ability to exactly predict both positive and negative instances over total instances tested.
  $$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (30)$$

- **AUC:**It is also known as the Receiver Operating Characteristic (ROC) curve. When a binary classification's discriminating threshold is different, this graph compares the True Positive Rate (TPR) and False Positive Rate (FPR) for that classification.

Table 3. Performance of comparison of classifiers of school dataset vs metrics

Methods	Precision (%)	Recall (%)	F-measure (%)	Accuracy (%)	AUC(%)
CGAN	79.2629	86.8172	85.4337	86.5000	93.4919
InfoGAN	82.5793	87.6464	86.9531	87.9500	93.8635
ACGAN	87.0895	89.3943	87.7554	89.2500	94.2063
ICGAN- DSVM	88.5366	92.6464	88.0893	91.3255	96.5172
ICGAN-EDSVM	89.7543	93.6237	89.5739	94.3215	96.6820
ICGAN-DLSTM	89.8830	97.7943	93.6210	97.6811	99.4932
ICGAN-EDLSTM	90.1880	98.1942	94.0210	98.0811	99.0931

Table 4. Performance of comparison of classifiers of university dataset vs metrics

Methods	Precision (%)	Recall (%)	F-measure (%)	Accuracy (%)	AUC (%)
CGAN	83.3593	82.5184	82.9367	82.1000	92.3266
InfoGAN	83.3829	85.3545	85.9779	85.3500	93.1607
ACGAN	88.6104	88.4935	86.3205	88.4000	93.4375
ICGAN- DSVM	88.3469	89.4726	88.8742	89.6510	93.9675
ICGAN-EDSVM	89.2582	93.1142	90.6679	93.0387	94.5844
ICGAN-DLSTM	92.9880	96.0192	94.4792	95.6461	98.2221
ICGAN-EDLSTM	93.3881	96.4191	94.8793	96.0460	98.6221

TABLE 5. PERFORMANCE OF COMPARISON OF CLASSIFIERS ON C-19GA20 DATABASE

Classifiers	Precision (%)	Recall (%)	F-measure (%)	Accuracy (%)	AUC (%)
CGAN	80.2487	79.4162	79.8324	79.4010	89.7565
InfoGAN	80.2720	82.2239	81.2480	82.6347	90.5881
ACGAN	85.4473	85.3315	85.3894	85.6695	90.8640
ICGAN- DSVM	85.1864	86.3008	85.7436	86.9142	91.3925
ICGAN-EDSVM	86.0886	89.9060	87.9973	90.2850	92.0075
ICGAN-DLSTM	87.5720	96.9430	89.7080	95.7640	99.4040
ICGAN-EDLSTM	89.9120	97.3630	90.1380	96.2640	99.8040

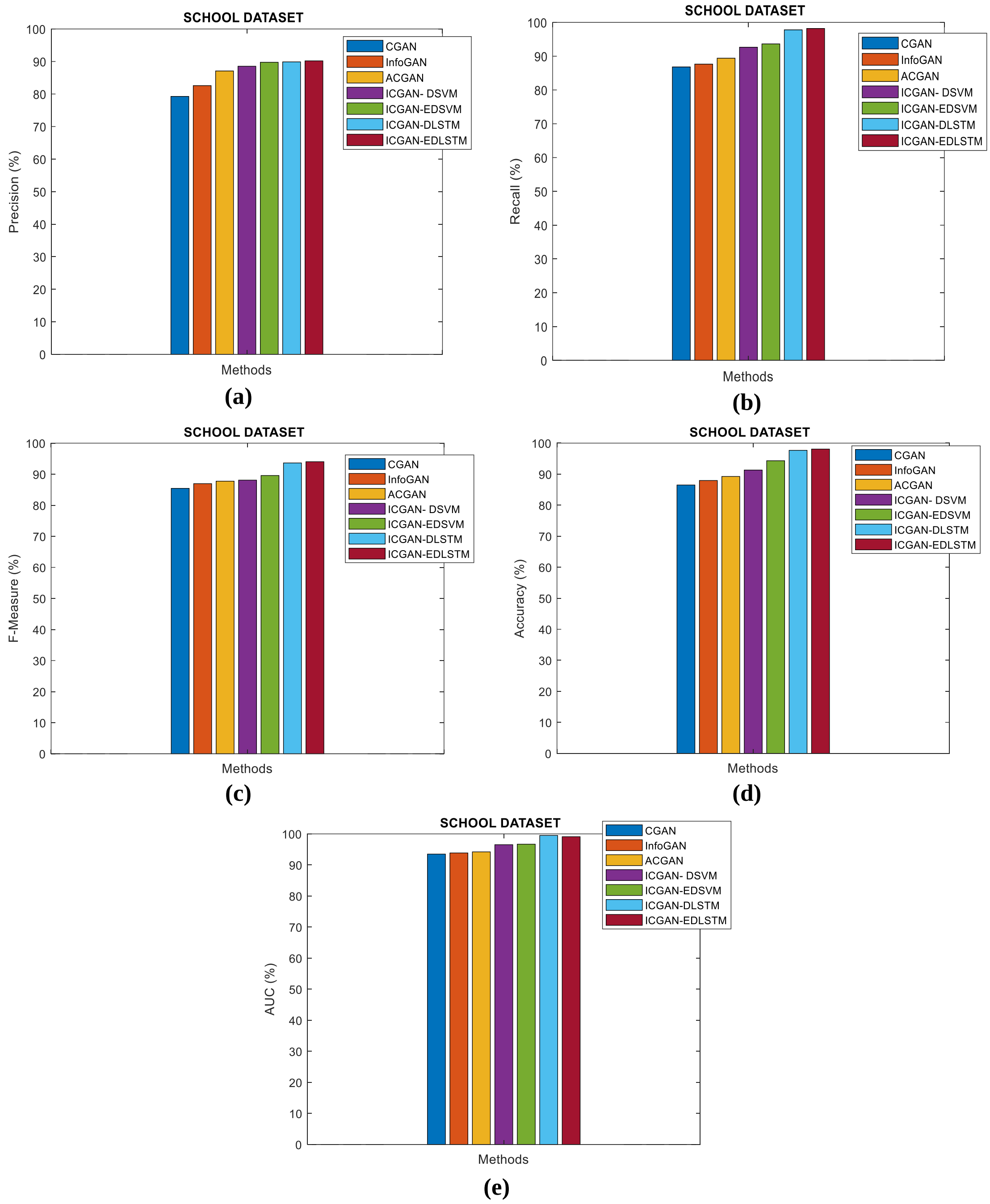


Figure 4. Comparison of various classification algorithms vs. School dataset. (a) Precision, (b) Recall, (c) F-measure, (d) Accuracy and (e) AUC

Figure 4 shows the performance values achieved by various classification algorithms such as CGAN, InfoGAN, ACGAN, ICGAN-DSVM, ICGAN-EDSVM, ICGAN-DLSTM and ICGAN-EDLSTM on the school dataset. Figure 4(a) shows the precision values of various classification algorithms on the school dataset. From the analyses, the proposed ICGAN-EDLSTM classifier gives the precision value of 90.188% for the school dataset, whereas the other classifiers such as CGAN, InfoGAN, ACGAN, ICGAN-DSVM, ICGAN-EDSVM and ICGAN-DLSTM gives lesser precision values of 79.2629%, 82.5793%, 87.0895%, 88.5366%, 89.7543% and 89.883% for the school dataset. Figure 4(b) shows the recall values of various classification algorithms on the school dataset. From the analyses, the proposed ICGAN-EDLSTM classifier gives the recall value of 98.1942% for the school dataset, whereas the other classifiers such as CGAN, InfoGAN, ACGAN, ICGAN-DSVM, ICGAN-EDSVM and ICGAN-DLSTM gives lesser recall values of 86.8172%, 87.6464%, 89.3943%, 92.6464%, 93.6237% and 97.7943%, respectively for the school dataset. Figure 4(c) shows the f-measure values of various classification algorithms on the school dataset. From the analyses, the proposed ICGAN-EDLSTM classifier gives the f-measure value of 94.0210% for the school dataset, whereas the other classifiers such as

CGAN, InfoGAN, ACGAN, ICGAN-DSVM, ICGAN-EDSVM and ICGAN-DLSTM gives lesser f-measure values of 85.4337%, 86.9531%, 87.7554%, 88.0893%, 89.5739% and 93.6210%, respectively for the school dataset. Similarly, Figure 4(d) shows the accuracy values of various classification algorithms on the school dataset. From the analyses, the proposed ICGAN-EDLSTM classifier gives the accuracy value of 98.0811% for the school dataset, whereas the other classifiers such as CGAN, InfoGAN, ACGAN, ICGAN-DSVM, ICGAN-EDSVM and ICGAN-DLSTM gives lesser accuracy values of 86.5000%, 87.9500%, 89.2500%, 91.3255%, 94.3215% and 97.6811%, respectively for the school dataset. Also, Figure 4(e) shows the AUC values of various classification algorithms on the school dataset. From the analyses, the proposed ICGAN-EDLSTM classifier gives the AUC value of 99.0931% for the school dataset, whereas the other classifiers such as CGAN, InfoGAN, ACGAN, ICGAN-DSVM, ICGAN-EDSVM and ICGAN-DLSTM gives lesser AUC values of 93.4919%, 93.8635%, 94.2063%, 96.5172%, 96.6820% and 99.4932%, respectively for the school dataset.

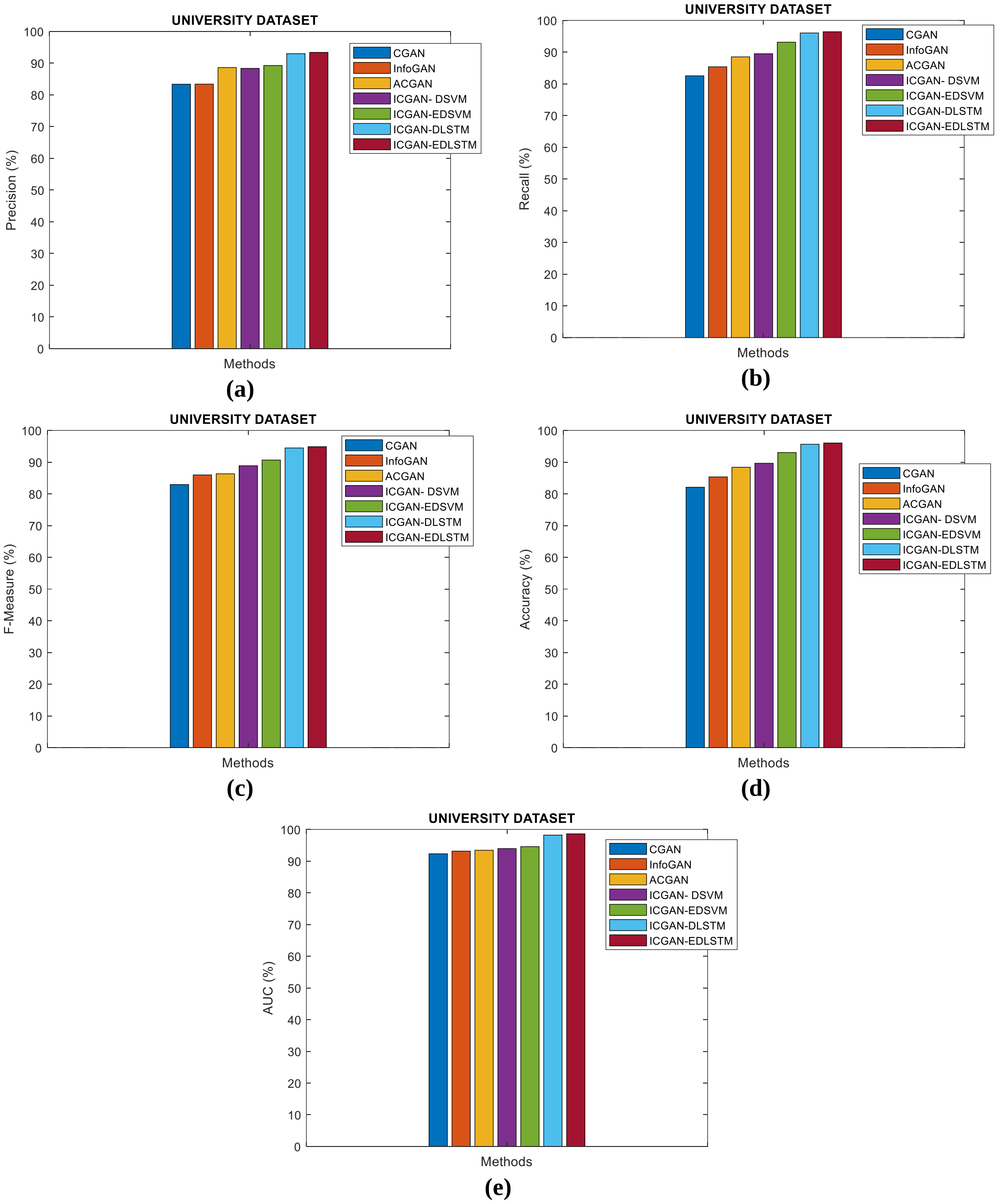
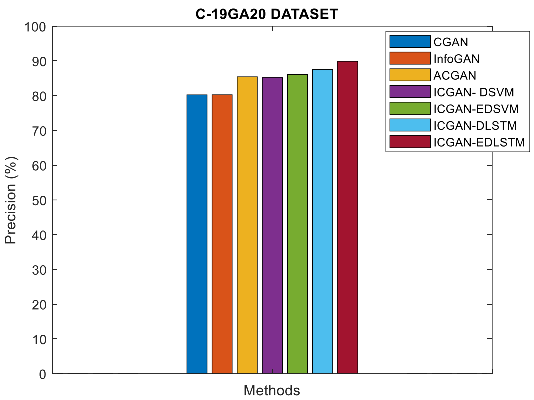
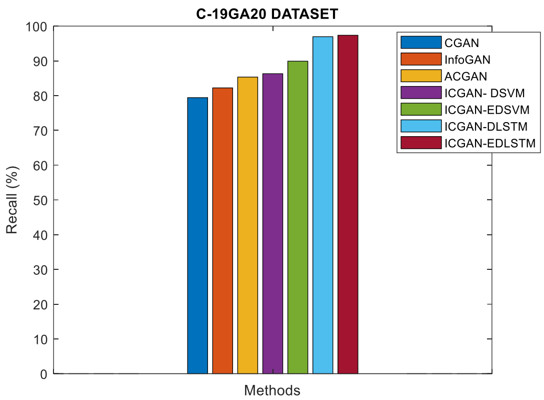


Figure 5. Comparison of various classification algorithms vs. University dataset. (a) Precision, (b) Recall, (c) F-measure, (d) Accuracy and (e) AUC

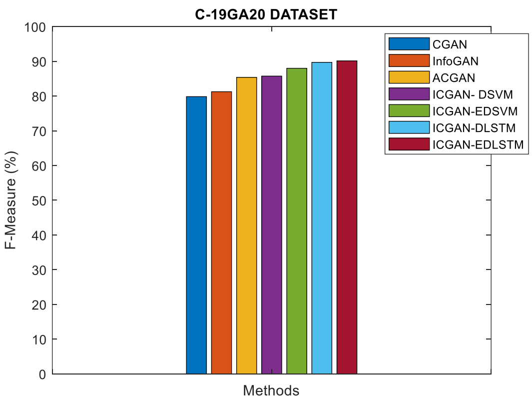
Figure 5 portrays the performance values achieved by various classification algorithms such as CGAN, InfoGAN, ACGAN, ICGAN-DSVM, ICGAN-EDSVM, ICGAN-DLSTM and ICGAN-EDLSTM on the university dataset. Figure 5(a) exhibits the precision values of the different classifiers on the university dataset. From the analyses, the proposed ICGAN-EDLSTM classifier gives the precision value of 93.3881% for the university dataset, whereas the other classifiers such as CGAN, InfoGAN, ACGAN, ICGAN-DSVM, ICGAN-EDSVM and ICGAN-DLSTM gives lesser precision values of 83.3593%, 83.3829%, 88.6104%, 88.3469%, 89.2582% and 92.9880% for university dataset. Figure 5(b) exhibits the recall values of various classification algorithms on the university dataset. From the analyses, the proposed ICGAN-EDLSTM classifier gives the recall value of 96.4191% for the university dataset, whereas the other classifiers such as CGAN, InfoGAN, ACGAN, ICGAN-DSVM, ICGAN-EDSVM and ICGAN-DLSTM gives lesser recall values of 82.5184%, 85.3545%, 88.4935%, 89.4726%, 93.1142% and 96.0192%, respectively for the university dataset. Figure 5(c) exhibits the f-measure values of various classification algorithms on the university dataset. From the analyses, the proposed ICGAN-EDLSTM classifier gives the f-measure value of 94.8793% for the university dataset, whereas the other classifiers such as CGAN, InfoGAN, ACGAN, ICGAN-DSVM, ICGAN-EDSVM and ICGAN-DLSTM gives lesser f-measure values of 82.9367%, 85.9779%, 86.3205%, 88.8742%, 90.6679% and 94.4792%, respectively for the university dataset. Similarly, Figure 5(d) exhibits the accuracy values of various classification algorithms on the university dataset. From the analyses, the proposed ICGAN-EDLSTM classifier gives the accuracy value of 96.0460% for the university dataset, whereas the other classifiers such as CGAN, InfoGAN, ACGAN, ICGAN-DSVM, ICGAN-EDSVM and ICGAN-DLSTM gives lesser accuracy values of 82.1000%, 85.3500%, 88.4000%, 89.6510%, 93.0387% and 95.6461%, respectively for the university dataset. Also, Figure 5(e) exhibits the AUC values of various classification algorithms on the university dataset. From the analyses, the proposed ICGAN-EDLSTM classifier gives the AUC value of 98.6221% for the university dataset, whereas the other classifiers such as CGAN, InfoGAN, ACGAN, ICGAN-DSVM, ICGAN-EDSVM and ICGAN-DLSTM gives lesser AUC values of 92.3266%, 93.1607%, 93.4375%, 93.9675%, 94.5844% and 98.2221%, respectively for the university dataset.



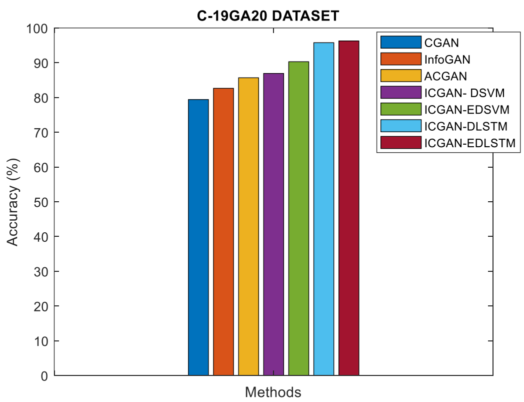
(a)



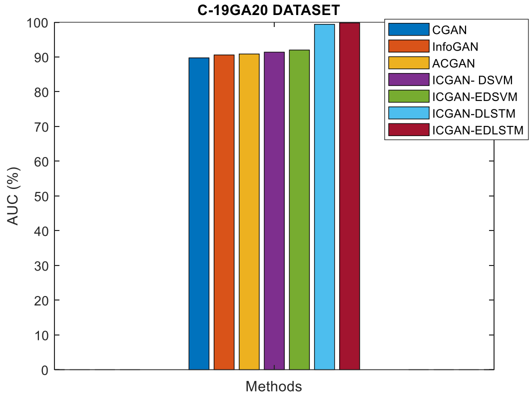
(b)



(c)



(d)



(e)

Figure 6. Comparison of various classification algorithms vs. C-19GA20 dataset. (a) Precision, (b) Recall, (c) F-measure, (d) Accuracy and (e) AUC

Figure 6 displays the performance values achieved by various classification algorithms such as CGAN, InfoGAN, ACGAN, ICGAN-DSVM, ICGAN-EDSVM, ICGAN-DLSTM and ICGAN-EDLSTM on the C-19GA20 dataset. Figure 6(a) illustrates the precision values of the different classifiers on the C-19GA20 dataset. From the analyses, the proposed ICGAN-EDLSTM classifier gives the precision value of 89.9120% for the C-19GA20 dataset, whereas the other classifiers such as CGAN, InfoGAN, ACGAN, ICGAN-DSVM, ICGAN-EDSVM and ICGAN-DLSTM gives lesser precision values of 80.2487%, 80.2720%, 85.4473%, 85.1864%, 86.0886% and 87.5720% for the C-19GA20 dataset. Figure 6(b) illustrates the recall values of the various classification algorithms on the C-19GA20 dataset. From the analyses, the proposed ICGAN-EDLSTM classifier gives the recall value of 97.3630% for the C-19GA20 dataset, whereas the other classifiers such as CGAN, InfoGAN, ACGAN, ICGAN-DSVM, ICGAN-EDSVM and ICGAN-DLSTM gives lesser recall values of 79.4162%, 82.2239%, 85.3315%, 86.3008%, 89.9060% and 96.9430%, respectively for the C-19GA20 dataset. Figure 6(c) illustrates the f-measure values of various classification algorithms on the C-19GA20 dataset. From the analyses, the proposed ICGAN-EDLSTM classifier gives the f-measure value of 90.1380% for the C-19GA20 dataset, whereas the other classifiers such as CGAN, InfoGAN, ACGAN, ICGAN-DSVM, ICGAN-EDSVM and ICGAN-DLSTM gives lesser f-measure values of 79.8324%, 81.2480%, 85.3894%, 85.7436%, 87.9973% and 89.7080%, respectively for the C-19GA20 dataset. Similarly, Figure 6(d) illustrates the accuracy values of various classification algorithms on the C-19GA20 dataset. From the analyses, the proposed ICGAN-EDLSTM classifier gives the accuracy value of 96.2640% for the C-19GA20 dataset, whereas the other classifiers such as CGAN, InfoGAN, ACGAN, ICGAN-DSVM, ICGAN-EDSVM and ICGAN-DLSTM gives lesser accuracy values of 79.4010%, 82.6347%, 85.6695%, 86.9142%, 90.2850% and 95.7640%, respectively for the C-19GA20 dataset. Also, Figure 6(e) illustrates the AUC values of various classification algorithms on the C-19GA20 dataset. From the analyses, the proposed ICGAN-EDLSTM classifier gives the AUC value of 99.8040% for the C-19GA20 dataset, whereas the other classifiers such as CGAN, InfoGAN, ACGAN, ICGAN-DSVM, ICGAN-EDSVM and ICGAN-DLSTM gives lesser AUC values of 89.7565%, 90.5881%, 90.8640%, 91.3925%, 92.0075% and 99.4040%, respectively for the C-19GA20 dataset.

V. CONCLUSION AND FUTURE WORK

The core contribution of this work is to predict the outcome of students in virtual classrooms during the pandemic. The advances are accomplished by creating the proposed predictive system, which would be utilized to predict the students' performance specifically during the Covid-19 outbreak. The proposed model has 4 primary stages: database acquisition, attribute selection, student performance prediction and analysis. In the initial stage, data associated with the students' learning through online classes are collected from Kaggle. In the second stage, irrelevant attributes in the collected database are removed by the TSOA, which is encouraged by the behavior of tuna swarms. This TSOA finds the best fitness tuna individuals for optimal selection of student performance attributes in the database. In the third stage, an ICGAN-EDLSTM was employed to create the prediction system, which supports educational institutions and academic professionals to estimate their student's performance in virtual learning and prevent them from dropout or getting an ineffective grades. At last, experimental findings are obtained based on the different evaluation metrics and it proves that the proposed ICGAN-EDLSTM classifier achieves maximum efficiency when compared to the other classification models. On the other hand, the LSTM needs more memory and time to train because of the sequential computation. So, the future work will focus on developing a Temporal Convolutional Network (TCN)-based ensemble classifier to solve the limitations of the LSTM across a wide range of attributes or datasets.

REFERENCES

[1] WHO. (2020). WHO Director-General’s opening remarks at the media briefing on COVID-19 -11 March 2020. https://www.who.int/director-general/speeches/detail/who-director-general-s-opening-remarks-at-the-media-briefing-on-covid-19---11-march-2020. Accessed 20 May 2022.

[2] Spinelli, A., &Pellino, G. (2020). COVID-19 pandemic: perspectives on an unfolding crisis. *Journal of British Surgery*, *107*(7), 785-787.

[3] UNESCO. (2020). COVID-19 impact on education. https://en.unesco.org/covid19/educationresponse. Accessed 20 May 2022.

[4] Alsammak, I. L. H., Mohammed, A. H., & Nasir, I. S. (2022). E-learning and COVID-19: predicting student academic performance using data mining algorithms. *Webology*, *19*(1), 3419-3432.

[5] Maatuk, A. M., Elberkawi, E. K., Aljawarneh, S., Rashaideh, H., & Alharbi, H. (2022). The COVID-19 pandemic and e-learning: challenges and opportunities from the perspective of students and instructors. *Journal of Computing in Higher Education*, *34*(1), 21-38.

[6] Pallavi, D. R., Ramachandran, M., &Chinnasamy, S. (2022). An empirical study on effectiveness of e-learning over conventional class room learning–a case study with respect to online degree programmes in higher education. *Recent trends in Management and Commerce*, *3*(1), 25-33.

[7] Mangione, G. R. J., & Cannella, G. (2021). Small school, smart schools: distance education in remoteness conditions. *Technology, Knowledge and Learning*, *26*(4), 845-865.

[8] Korać, D., Damjanović, B., &Simić, D. (2022). A model of digital identity for better information security in e-learning systems. *The Journal of Supercomputing*, *78*(3), 3325-3354.

[9] Alyahyan, E., &Düştegör, D. (2020). Predicting academic success in higher education: literature review and best practices. *International Journal of Educational Technology in Higher Education*, *17*(1), 1-21.

[10] Bansal, V., Buckchash, H., & Raman, B. (2022). Computational intelligence enabled student performance estimation in the age of COVID-19. *SN Computer Science*, *3*(1), 1-11.

[11] Dhawan, S. (2020). Online learning: a panacea in the time of COVID-19 crisis. *Journal of Educational Technology Systems*, *49*(1), 5-22.

[12] Aldowah, H., Al-Samarraie, H., &Fauzy, W. M. (2019). Educational data mining and learning analytics for 21st century higher education: a review and synthesis. *Telematics and Informatics*, *37*, 13-49.

[13] Rodrigues, M. W., Isotani, S., & Zarate, L. E. (2018). Educational data mining: a review of evaluation process in the e-learning. *Telematics and Informatics*, *35*(6), 1701-1717.

[14] Tarik, A., Aissa, H., & Yousef, F. (2021). Artificial intelligence and machine learning to predict student performance during the COVID-19. *Procedia Computer Science*, *184*, 835-840.

[15] Kuhfeld, M., Soland, J., Tarasawa, B., Johnson, A., Ruzek, E., & Liu, J. (2020). Projecting the potential impact of COVID-19 school closures on academic achievement. *Educational Researcher*, *49*(8), 549-565.

[16] Koob, C., Schröpfer, K., Coenen, M., Kus, S., & Schmidt, N. (2021). Factors influencing study engagement during the COVID-19 pandemic: a cross-sectional study among health and social professions students. *PlosOne*, *16*(7), 1-19.

[17] Chen, W., Brinton, C. G., Cao, D., Mason-Singh, A., Lu, C., & Chiang, M. (2018). Early detection prediction of learning outcomes in online short-courses via learning behaviors. *IEEE Transactions on Learning Technologies*, *12*(1), 44-58.

[18] Cazarez, R. L. U., & Martin, C. L. (2018). Neural Networks for predicting student performance in online education. *IEEE Latin America Transactions*, *16*(7), 2053-2060.

[19] Wang, G. H., Zhang, J., & Fu, G. S. (2018). Predicting student behaviors and performance in online learning using decision tree. In *Seventh IEEE International Conference of Educational Innovation through Technology*, pp. 214-219.

[20] Francis, B. K., & Babu, S. S. (2019). Predicting academic performance of students using a hybrid data mining approach. *Journal of Medical Systems*, *43*(6), 1-15.

[21] Li, Y., Gou, J., & Fan, Z. (2019). Educational data mining for students' performance based on fuzzy C-means clustering. *The Journal of Engineering*, *2019*(11), 8245-8250.

[22] Alshabandar, R., Hussain, A., Keight, R., & Khan, W. (2020). Students performance prediction in online courses using machine learning algorithms. In *IEEE International Joint Conference on Neural Networks*, pp. 1-7.

[23] Xu, Z., Yuan, H., & Liu, Q. (2020). Student performance prediction based on blended learning. *IEEE Transactions on Education*, *64*(1), 66-73.

[24] Jiang, P., & Wang, X. (2020). Preference cognitive diagnosis for student performance prediction. *IEEE Access*, *8*, 219775-219787.

[25] Montaño, H. S. A., & Cabrera-Loayza, M. C. (2020). Early prediction of dropout in online courses using artificial neural networks. In*IEEEXV ConferenciaLatinoamericana de Tecnologias de Aprendizaje*,pp. 1-6.

[26] Deo, R. C., Yaseen, Z. M., Al-Ansari, N., Nguyen-Huy, T., Langlands, T. A. M., & Galligan, L. (2020). Modern artificial intelligence model development for undergraduate student performance prediction: an investigation on engineering mathematics courses. *IEEE Access*, *8*, 136697-136724.
[27] Dabhade, P., Agarwal, R., Alameen, K. P., Fathima, A. T., Sridharan, R., &Gopakumar, G. (2021). Educational data mining for predicting students’ academic performance using machine learning algorithms. *Materials Today: Proceedings*, 1-8.
[28] Chi, D., & Huang, Y. (2021). Research on application of online teaching performance prediction based on data mining algorithm. In *IEEE International Conference on Consumer Electronics and Computer Engineering*, pp. 394-397.
[29] Matzavela, V., &Alepis, E. (2021). Decision tree learning through a predictive model for student academic performance in intelligent m-learning environments. *Computers and Education: Artificial Intelligence*, *2*, 1-12.
[30] Ulloa-Cazarez, R. L., García-Díaz, N., & Soriano-Equigua, L. (2021). Multi-layer adaptive fuzzy inference system for predicting student performance in online higher education. *IEEE Latin America Transactions*, *19*(01), 98-106.
[31] Tian, Y., Huang, W. N., Zhang, L., & Wang, T. (2021). Disclosing personal names in screen names predicts better final achievement levels in massive open online courses. *IEEE Access*, *9*, 50926-50938.
[32] Rodríguez-Hernández, C. F., Musso, M., Kyndt, E., &Cascallar, E. (2021). Artificial neural networks in academic performance prediction: systematic implementation and predictor evaluation. *Computers and Education: Artificial Intelligence*, *2*, 1-14.
[33] https://www.kaggle.com/datasets/720bbd5647878bf2694065a6adbbd85ba048068887e95db00c9c3bfb4cda8c5b
[34] Xie, L., Han, T., Zhou, H., Zhang, Z. R., Han, B., & Tang, A. (2021). Tuna swarm optimization: a novel swarm-based metaheuristic algorithm for global optimization. *Computational Intelligence and Neuroscience*, *2021*, 1-22.
[35] Liu, Q., Tan, T., & Yu, K. (2016). An investigation on deep learning with beta stabilizer. In *IEEE 13th International Conference on Signal Processing*, pp. 557-561.