Journal of Data Acquisition and Processing
ISSN: 1004-9037
https://sjcjycl.cn/
DOI: 10.5281/zenodo.776872

CIRCLE MAP TUNA-SWARM OPTIMIZATION ALGORITHM (CMTSOA) AND ENSEMBLE MULTIPLE RECURRENT DEEP LEARNING (EMRDL) CLASSIFIER FOR COGNITIVE ANALYSIS OF STUDENT PERFORMANCE PREDICTION

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. Email: s.sathipriya@gmail.com
[2] Assistant Professor, Department of Computer Science, Nallamuthu Gounder Mahalingam College Pollachi, Tamil Nadu - 642001, India. Email: a.kanagaraj@gmail.com

ABSTRACT
The abundance of data in educational databases makes it more difficult to predict students' success. Tutoring in the classroom and at home offers students individualised support and constructive criticism of their learning. Methods that correctly predict a student's performance in a particular cognitive analysis or course can help identify students who are at risk of failing a course during a pandemic and enable their educational institution to take appropriate action. These days, combining classifiers is proposed as a novel approach for enhancing prediction performance. The Ensemble Multiple Recurrent Deep Learning (EMRDL) algorithm has been put forth in this paper as a way to forecast student success under supportive learning conditions during the Coronavirus Disease (COVID-19) epidemic, which includes tutoring from families and schools. First, benchmark repository records are gathered. In order to select the most pertinent features and reduce the dataspace, the Circle Map Tuna-Swarm Optimization Algorithm (CMTSOA) is developed. A more precise feature search is made possible by the TSOA, circle map function, which has been added for random number generation. Following that, the EMRDL classifier is introduced to improve the categorization outcomes by fusing several models, including Ensemble Deep Long Short-Term Memory (EDLSTM), Weight Recurrent Multilayer Perceptron Network (WRMLP), and Gated Recurrent Unit (GRU). MLP has been introduced by adding delayed connections among nearby nodes of a hidden layer, an RMLP network is able to be built. Through connections established through a series of nodes, GRU is intended to conduct student performance prediction. By using majority voting, various classifiers are merged. By integrating the proposed system in Matrix Laboratory R 2020a (MATLAB R2020a) and using three distinct benchmark databases, the results of the proposed system and state-of-the-art classifiers is assessed (School, University, and C-19GA20). Metrics such as Precision, Recall, Accuracy, and Area Under Curve (AUC) are used to evaluate the performance of these models.

Keywords: *Education data mining, Virtual learning, Circle Map Tuna-Swarm Optimization Algorithm (CMTSOA), Deep Learning, Ensemble Multiple Recurrent Deep Learning (EMRDL), Ensemble Deep Long Short-Term Memory (EDLSTM), Weight Recurrent Multilayer Perceptron Network (WRMLP), and Gated Recurrent Unit (GRU).*

1. INTRODUCTION

In the current era of big data and artificial intelligence, learning analytics [1], [2], and supportive learning [3], [4] have emerged as new study areas to aid students' learning. Because students acquire skills and information that they can use to benefit the community, education is essential for the sustainable development of society. In the very young field of educational data mining (EDM), which focuses on learning latent patterns in a variety of educational contexts, such as student knowledge analysis [5], student learning behaviour analysis [6], teacher curriculum planning, and course time arrangement, latent patterns are latent patterns that are not explicitly observed. All studies are involved have as their ultimate objective improving student learning outcomes [7, 8, 9], in addition to other objectives like lowering educational expenses.

Each year, numerous student advances to higher levels or graduate. The analysis of the interconnected negative effects on students who barely passed or failed the course is covered in great depth in many research works. These problems have led to an increase in the ability to predict student success in terms of global development. Since it completely relies on the educational process that results in the production of a generation capable of taking on the responsibility of guiding towards growth in all area of life [10].The effectiveness of educational institutions, raising future generations in accordance with the many stages of life in every nation is also reflected in how well students perform on assessments of their performance. Many machine learning methods have been proposed in the literature and tested using real-world datasets. For feature extraction, researchers have examined the student population's heterogeneity [11]. Four popular machine learning techniques were used to create prediction models: JRip, sequential minimal optimization, C4.5, and Naive-Bayes (NB). The prediction accuracy for all algorithms is the same at 80%. Yet, prior research mainly relies on feature engineering. The process of choosing, modifying, manipulating, and creating new variables from unprocessed data using domain expertise is known as feature engineering. The fact that feature engineering uses features that are customised for a particular dataset results in models that are rigid when applied to new datasets.

A direct consequence is that features need to be rebuilt for each course. Recent superior feature selection methods are using the influence of optimization algorithms for choosing a subset of relevant features to get better classification results [12, 13]. Existing works [14]–[16] possessed a common design of analyzing the optimal feature vector from the dataset. Taking the review articles [17] into account, to the best of knowledge, there has no consideration on the prediction of students' performance under shadow education environment, that is school tutoring and family tutoring. Until now, machine learning algorithms tended to adopt deep learning techniques due to the typically small amount of data in educational datasets. The current educational systems are unsuccessful as they do cannot place the necessary importance on predicting student achievement. The current educational systems suffer from a lack of efficiency as a result of not placing enough emphasis on predicting student achievement. Raising educational efficiency is a result of identifying the lessons that the student will be interested in and tracking his activity in the classroom. In comparison to other standard classifiers, deep learning outperforms them, according to the researchers [10,18]. Data mining can be utilised to get better and more accurate outcomes in learning management systems. It is possible to forecast student performance using machine learning and deep learning algorithms, as well as to quickly identify any risk factors kids may be exposed to them. By combining the predictions of several learners; ensemble learning approaches improve overall accuracy. The most popular learning strategies are stacking, boosting, and bagging. The capacity of layered classification is used to guarantee personalised teaching and learning assignments. The learner's abilities, skills, interests, and requirements can be used as the classification criteria. To overcome learning disabilities and ensure success, each student's class follows a distinct learning route [19, 20–21].

To increase the adaptability and automation of prediction models, two key contributions are presented in this study's details: 1) A circular map tuna-swarm optimization method (CMTSOA) system has been developed for the purpose of removing features, and 2) the Ensemble Multiple Recurrent Deep Learning (EMRDL) framework has been developed for the purpose of predicting student performance. Individual classifiers like Ensemble Deep Long Short-Term Memory (EDLSTM), Weight Recurrent Multilayer Perceptron Network (WRMLP), and Gated Recurrent Unit (GRU) are used to results the EMRDL classifier. Majority voting is used to integrate the results of these distinct classifiers. In the end, metrics such as precision, recall, f-measure, accuracy, and AUC are used to evaluate the prediction algorithms' performance.

2. LITERATURE REVIEW

New machine learning approach was created by Xu et al. [22] to forecast student performance in degree programmes. Two main features define the proposed method. First, a bilayered structure with a number of base predictors and a series of ensemble predictors is created to make predictions based on the changing performance statuses of the students. Second, a data-driven strategy using latent factor models and probabilistic matrix factorization is proposed to find course relevance, which is crucial for developing effective base predictors. Demonstrate that the proposed strategy outperforms benchmark approaches through thoroughly simulations on a dataset of undergraduate students gathered over three years at the University of California, Los Angeles. Farissi and Dahlan [23] proposed combining a classification method with a Genetic Algorithm (GA) feature selection strategy. The ideal answer is rather challenging to create because nearly all previous feature selection strategies use local search throughout the process. Because of this, GA is used as a technique of features selection using Random Forest (RF) based ensemble classification to increase classification accuracy value of student academic performance prediction. It can also be applied to datasets with high dimensionality and imbalanced class. Behavioral, academic, and demographic data are the three main categories of the Kaggle repository's dataset that is used for experiments. The AUC is used to assess how well the proposed strategy performs. According to the experiment's findings, the proposed approach performs impressively when it comes to predicting students' academic achievement.

In order to categorise student based on their past performance, Surenthiran et al. [24] proposed a Deep Belief Neural Network (DBNN) with Atom Search Optimization (ASO) optimization. The cognitive divergence approach used by the Restricted Boltzmann Machine (RBM) in DBNN employs several cascaded RBMs. An educational dataset that is available for use in the public is initially used to develop the proposed methodology. Preparation takes place in the first phase, while classification happens in the second. To improve the learning rate parameters in the second phase is a difficult task for DBNN. Student level prediction is carried out automatically by optimization using ASO. Performance measures, which are assessed to evaluate the proposed work, produced an improved result over earlier algorithms. The proposed model performs better than average, with a reduced error value of under 20.00% and a 90.00% accuracy level. Kim et al [25] proposed a GritNet for predicting student performance which builds on the Bidirectional Long Short Term Memory (BiLSTM). GritNet routinely beats the conventional Logistic-Regression (LR), according on results from graduation predictions made for actual Udacity students. Improvements are particularly noticeable in the initial few weeks, when making correct forecasts is the most difficult. This is mostly due to the models' continued reliance on feature engineering to lower input dimensions, which seems to restrict the development of improved deep learning models. Results comparison was done using student two Udacity Nanodegree (ND) programs: from the ND-A and ND-B. A dataset of students from three institutions in Assam, India, was used by Hussain et al. [26] for their deep learning (Sequential Neural Model (SNN) and Adam optimization technique). In order to forecast the students' results, the study contrasted AdaBoost and other classification techniques such the Artificial Immune Recognition System (AIRS) v2.0. As part of a statistical analysis to determine the best classification algorithms, performance metrics like as precision, recall, F-Score, accuracy, and Kappa Statistics were calculated. The 10140 student records in the dataset were used in this investigation.

A technique for assessing teacher performance that combines stacking and voting ensemble was put forth by Ahuja and Sharma [27]. Two datasets were taken into consideration; one was obtained from the Teaching Assistant Evaluation machine learning repository at the University of California, Irvine (UCI), while the other was gathered from university students. RF, Decision Tree (DT), Ridge Classifier, Gradient Boosting (GB), Logistic Regression (LR), Ada Boost Classifier, Linear Discriminant Analysis (LDA), Quadratic Discriminant Analysis (QDA), Linear Kernel Support Vector Machine (LKSVM), NB, and proposed model using Python. According to the data, a professor's success is greatly influenced by his or her capacity to inspire students to acquire the material, stimulate involvement in class discussions, and connect lectures with real-world situations. Asselman et al. [28] proposed an ensemble learning techniques as a highly effective machine learning paradigm has been employed to develop numerous new solutions across a variety of sectors. In order to improve the predictive validity of student performance, a new Performance Factors Analysis (PFA) has been introduced based on various models—RF, AdaBoost, and XGBoost. According to the experimental results, the scalable XGBoost beat the other models that were considered and significantly enhanced performance prediction compared to the original PFA algorithm. Convolutional Neural Network (CNN) captures local dominant features and alleviates the curse of dimensionality, while Recurrent Neural Network (RNN) retrieves the semantic connection between features.

Xiong et al. [29] introduced a hybrid deep learning model with this CNN-RNN configuration. The trials' findings show that the hybrid CNN-RNN prediction model outperforms deep learning model by 3.16%, with hybrid CNN-accuracy RNN's increasing from 73.07% to 79.23%.

Recent research [30] examined a variety of student features, evaluations, and interactions with the online learning environment. It is proposed to use a RNN-Gated Recurrent Unit (GRU) to fit both static and sequential data, with the data completion mechanism also being used to fill in the gaps in the stream of data. Experimental findings using the Open University Learning Analytics Dataset (OULAD) demonstrate that simple approaches such as GRU and simple RNN outperform the comparatively complicated LSTM model in terms of performance. The findings also show that different models perform well at different times, leading to the proposed combined model's achievement of over 80% prediction accuracy for at-risk students at the conclusion of the semester. In a learning management system, Hidalgo et al. [31] looked into the possibilities of deep learning and meta-learning to predict student performance. 500 students enrolled in an online master's programme served as the dataset on which the developed predictive model was tested. It demonstrates the model's capacity for autonomous hyperparameter optimization. The autonomous model's performance was comparable to that of the conventionally constructed model, which has important advantages in terms of efficiency and scalability. In order to predict student academic performance, Adejo and Connolly [32] proposed a new system that empirically investigates and contrasts the use of various data sources, various classifiers, and ensembles of classifiers by three different classification algorithms (DT, Artificial Neural Network (ANN), and SVM). A novel hybrid model that may be used to predict student performance has been developed as a result of the study is more accurate and effective. In general, this study contributes to understanding of how ensemble approaches can be used to forecast student performance utilising learner data. Also, it increases performance accuracy and lowers the rate of classification error.

3. PROPOSED METHODOLOGY

These classifiers include Ensemble Deep Long Short-Term Memory (EDLSTM), WRMLP, and GRU has been integrated in EMRDL approach. The four main stages of the proposed system are data gathering, feature selection, prediction, and performance analysis. The database is initially obtained via UCI and Kaggle. Second, the Circle Map Tuna-Swarm Optimization Algorithm (CMTSOA) has been used to eliminate unnecessary features. The creation of a prediction model is then proposed using an EMRDL classifier. Finally evaluation measures have been introduced to assess the results of classifiers. Figure 1 provides an illustration of the proposed system.

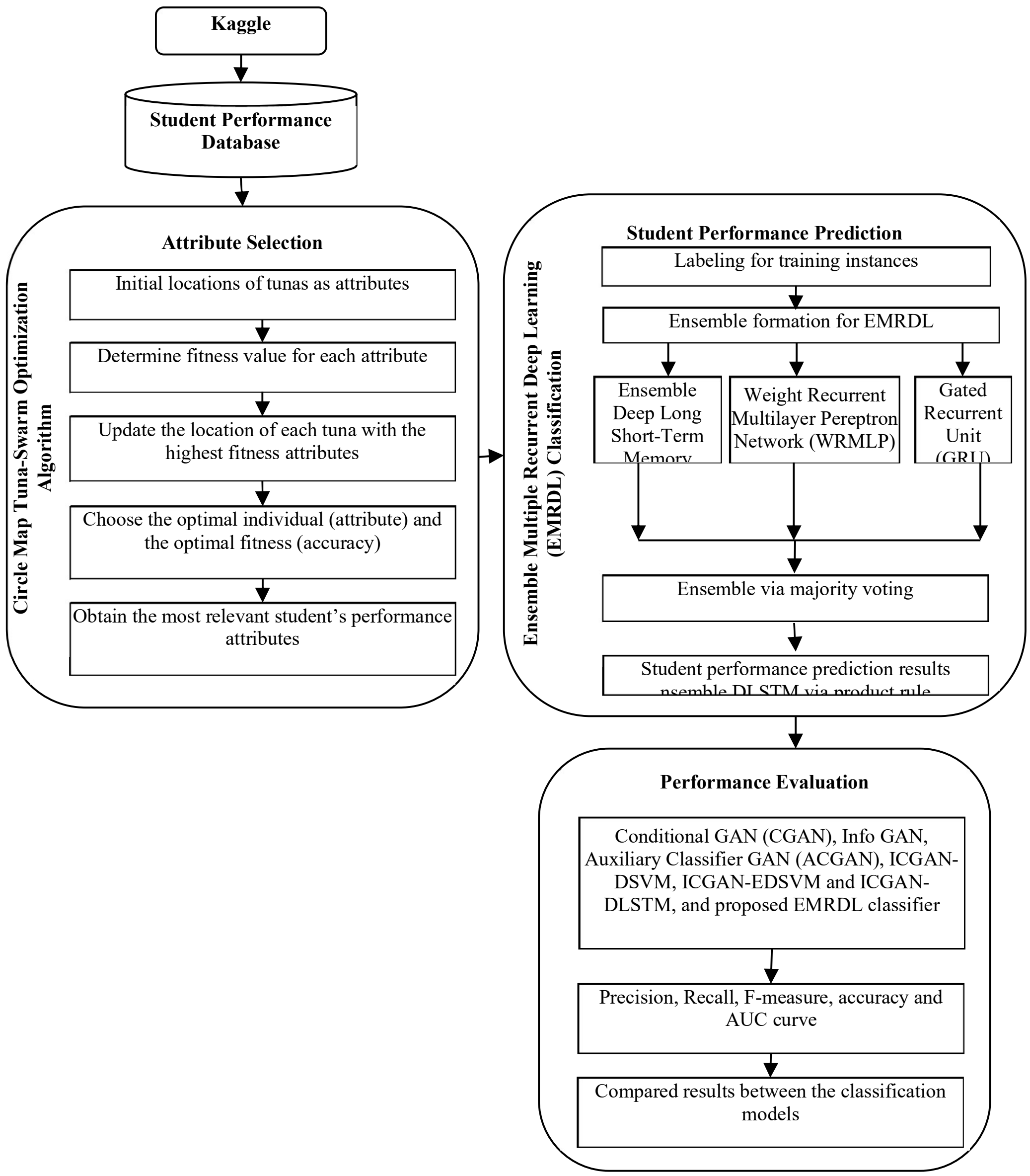


Figure 1. Proposed Student Performance Prediction Model

a. Data Acquisition

Three separate benchmark databases are taken into consideration for examination in this study.

The dataset for predicting student performance was taken from recent studies. It is titled Dataset 1 (Student Performance School Dataset). It consists of two courses totaling 1044 students: one for Portuguese language (649 records), and the other for mathematics (395 records). The dataset contains 33 variables, nine of which are connected to tutoring in schools and in families. The remaining 29 features were obtained from school records and the remainder from a questionnaire.

University Dataset 2, Dataset 2: To examine the findings from academics and students, this dataset was developed. Details on the school's divisions, teachers, student counselling, course offerings, how students were admitted to certain courses, and how well students performed on particular exams. After the dataset has been gathered, the following step is to execute feature selection in order to lower the dataset's overall feature count.

Covid-19 Go Away 2020 (C-19GA20) Database, Dataset 3: Digital data for the C-19GA20 dataset was created in April 2020 using school and university students between the ages of 14 and 24. It contains information on students' mental stability, social life, attitudes towards Covid-19, how the Covid-19 outbreak has affected their academic performance, and their experience with virtual classrooms [33]. Following database acquisition, attribute selection is carried out to reduce the total number of attributes in the database in order to produce an accurate prediction of students' performance.

b. Wrapper Attribute Selection

The AI approach applies wrapper attribute selection, which uses a search strategy to locate the precise attribute space and choose the most pertinent attributes to improve prediction accuracy. The most pertinent features for forecasting student outcomes during the Covid-19 outbreak are chosen using the CTSOA. Tuna swarms' cooperative hunting style is crucial to CTSOA. As a result, tuna can use the swarm movement technique to hunt [34]. Spiral hunting is the first rule. While tuna consume, they spin in a spiral pattern to push their prey into shallow water where they can easily be struck. All tunas make a parabolic arc to encircle their prey in the same manner as the first one. These foraging policies dictate how this CTSOA is simulated and used as a technique of attribute selection. This is a description of the procedures in this CTSOA.

- *Initialization:* It starts the optimization process by introducing first occupants into the search space.

- *Spiral hunting:* As sardines, herring, and other small, moving fish interacted with hunters, the entire school of fish developed a dense habitat that continuously changed direction, making it difficult for hunters to concentrate on a victim. The tuna swarm is currently pursuing its target in a precise spiral pattern. Although the majority of the fish in crowd lack eyesight, the nearby fish will gradually change their orientation until they form a large crowd by the same objective and begin to forage. Tuna swarms communicate with one another as they circle after their prey. All tunas follow the earlier fish, making it easier for nearby tuna to share information.

- *Random number generation:* Creating random numbers: A circle map that projects a random integer onto a circle. A nonlinear iterated map called the Circle map is given by an equation (1),

  - $$\theta_{n+1} = \left(\theta_n + \Omega - \frac{K}{2\pi}\sin(2\pi\theta_n)\right) mod\ 1 \quad (1)$$

  where K is the coupling strength, andΩ is a constant that represents the fixed angular progression of the sinusoidal oscillator.

- Hunting in a parabola Moreover, tunas collaborate in a parabolic manner to feed. Tuna assumes a parabolic shape using its prey as a point of reference. Tuna also scan the area around them in search of prey. These two procedures are run concurrently on the assumption that each has a 50% chance of being chosen.

Tuna cooperatively forage utilising two foraging techniques to find food. With the TSOA, the population in the hunt space is originally generated at random. All individuals choose one of the two foraging policies at random or choose to regenerate their location in the hunt region according to chance z during each iteration. Every TSOA person is continuously changed and determined throughout the whole optimization process until the termination requirements are met, at which point the ideal individual and the corresponding fitness value (highest prediction accuracy) are obtained.

c. Ensemble classifier

By merging many classifiers, including EDLSTM, WRMLP, and GRU, the EMRDL classifier improves classification results. Several models are developed and combined in an ensemble manner to predict student performance. The performance of the classifier model is enhanced by the use of ensemble. The technique entails creating numerous classifiers independently and returning the average of each classifier's output. As there is less chance of inaccuracy, the combined output is superior to an individual output. With majority vote, the results are combined.

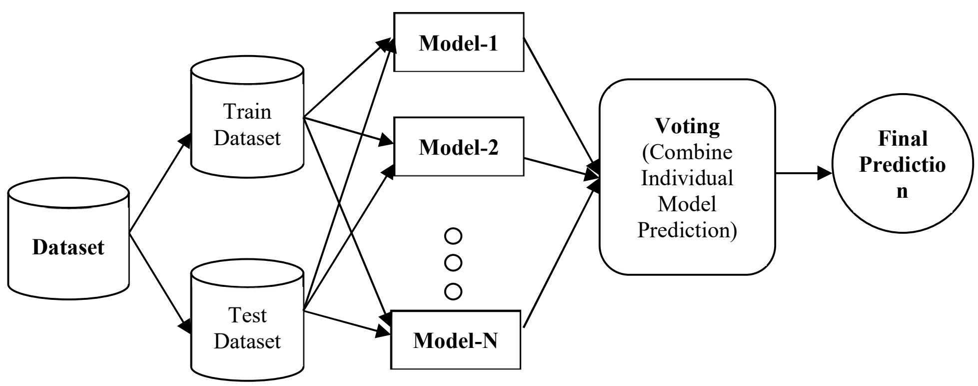


Figure 2. Ensemble Multiple Recurrent Deep Learning (EMRDL) Classifier

i. Ensemble Deep Long Short-Term Memory (EDLSTM)

To create a reliable ensemble classifier, the ensemble should properly generate and combine LSTM classifier findings. A product rule is used to combine numerous probability outcomes from LSTM classifiers. For each one-vs-all classification model, a 2-layer LSTM classifier is produced. A number of β-stabilized LSTM classifiers are divided and trained individually in the first layer. The other LSTM classifier for the second layer design, which is a similar one-vs.-all classifier, is then trained using the assistance vector activations from every classifier in the initial layer. There are three gates and a significant affine function in a single LSTM layer. Every linear transform functions have an independent β -stabilizer to improve the LSTM network. It is believed that an independent -stabilizer can correctly and independently adjust the scale of each matrix [35]. You can rewrite equations (2-6) as follows:

$$i_t = sigmoid\left(e^{\beta_{xi}} w_{xi} x_t + e^{\beta_{hi}} w_{hi} h_{t-1} + e^{\beta_{ci}} w_{ci} c_{t-1} + b_i\right) \tag{2}$$

$$f_t = sigmoid\left(e^{\beta_{xf}} w_{xf} x_t + e^{\beta_{hf}} w_{hf} h_{t-1} + e^{\beta_{cf}} w_{cf} c_{t-1} + b_f\right) \tag{3}$$

$$c_t = f_t \cdot c_{t-1} + i_t \cdot tanh\left(e^{\beta_{xc}} w_{xc} x_t + e^{\beta_{hc}} w_{hc} h_{t-1} + b_c\right) \tag{4}$$

$$o_t = sigmoid\left(e^{\beta_{xo}} w_{xo} x_t + e^{\beta_{ho}} w_{ho} h_{t-1} + e^{\beta_{co}} w_{co} c_t + b_o\right) \tag{5}$$

$$h_t = o_t \cdot tanh(c_t) \tag{6}$$

Based on the back-propagation scheme, the LSTM that has beenβ stabilised is trained using the performance features of the chosen student. The back-propagation scheme must figure out the gradient of fitness factor for variables β during the training procedure.

ii. Recurrent Multilayer Perceptron (RMLP)

Multilayer Perceptron (MLP) has been introduced by adding delayed connections among the hard by nodes of a hidden layer including the originating node; a consequential RMLP network is able to be built. RMLP belong to a broad type of nonlinear systems with features for student success prediction.

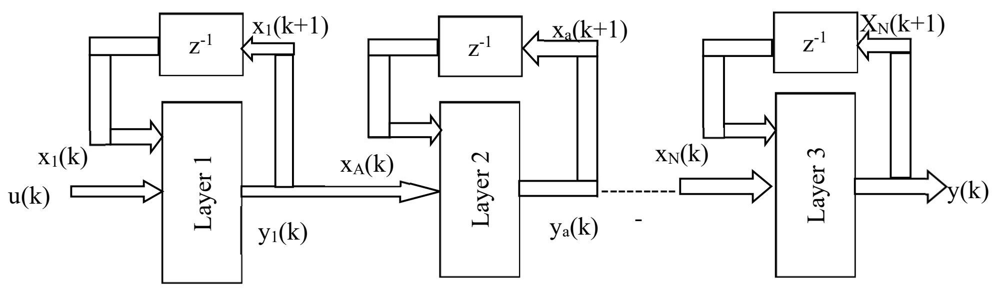


Figure 3. RMLP architecture with l layers

RMLP network architecture is shown in Figure 3. Figure 3 depicts the final RMLP construction with L layers. The layers are all typical MLP layers. The delayed layer outputs are fed back into the layer at the same time as the layer outputs are passed forward towards the inputs of the following layer. Consequently, the layer output at time k-1 designed for a particular layer serves as the state variable at with the purpose of moment. All layer values [i(k)] together make up the network's overall state. No links between layers occur repeatedly. Assume that the input vector is $(u(k) \in \mathbb{R}^k)$ the hidden layer's output at instant k is $x(k) \in \mathbb{R}^q$and the output vector is $y(k) \in \mathbb{R}^k$.Consider the RMLP network, which has L layers and N(l) nodes, for $l = 1, \ldots, L$. The network's input goes towards the first layer, which serves as a buffer, and the network's output comes from the final layer. The subsequent equations (7-8), can be used towards describing the ith node placed at thel^{th} layer of an RMLP system.

$$z_{[l,i]}(n) = \sum_{j=1}^{N(l)} w_{[l,j][l.i]} x_{[l,j]}\,(n-1) + \sum_{j=1}^{N(l)} w_{[l-1,j][l.i]} x_{[l-1,j]}\,(n) + b_{[l,i]} \tag{7}$$

$$x_{[l,i]}(n) = F_{[l]}\left(z_{[l,i]}(k)\right) \tag{8}$$

where $z_{[l,i]}(n)$ represents the internal state variable of the ith node by lth layer; $x_{[l,i]}(n)$is the output of the ith node by lth layer,$b_{[l,i]}(n)$is the bias towards the ith node by l^{th} layer,$w_{[l,j][l',i]}$ is the weight between the jth node by l^{th}layer and the ith node by l^{th} layer; and $F_{[l]}(.)$ represents the discriminatory function associated by the l^{th}layer, a hyperbolic tangent for this study. RMLP network outputs are functions of the present and past network internal states of the present network inputs.

iii. Gated Recurrent Unit (GRU)

A reset-gate and an update-gate student performance variable which are internal states used towards evaluate the long-term dependency and keep student performance data .It is added to the simple RNN described above by Gated Recurrent Unit (GRU). Using the following formulae, a GRU's forward step can be calculated (9-12),

$$u_i^{(t)} = sigmoid(W_{xu}x_i^{(t)} + W_{hu}h_i^{(t-1)} + b_z) \tag{9}$$

$$r_i^{(t)} = sigmoid(W_{xr}x_i^{(t)} + W_{hr}h_i^{(t-1)} + b_r) \tag{10}$$

$$\hat{h}_i^{(t)} = tanh(W_{x\hat{h}}x_i^{(t)} + b_{\hat{h}_1} + r_i^{(t)}o(\, W_{h\hat{h}}h_i^{(t-1)} + b_{\hat{h}2}) \tag{11}$$

$$h_i^{(t)} = \left(1 - u^{(t)}\right)oh_i^{(t-1)} + u^{(t)}o\hat{h}^{(t)} \tag{12}$$

where $y_i^{(t)}$ is the candidate output, W and b are the weight and bias of classifier, and o is the element-wise product; $z_i^{(t)}$ and $r_i^{(t)}$are the update-gate and reset-gate vectors, respectively.

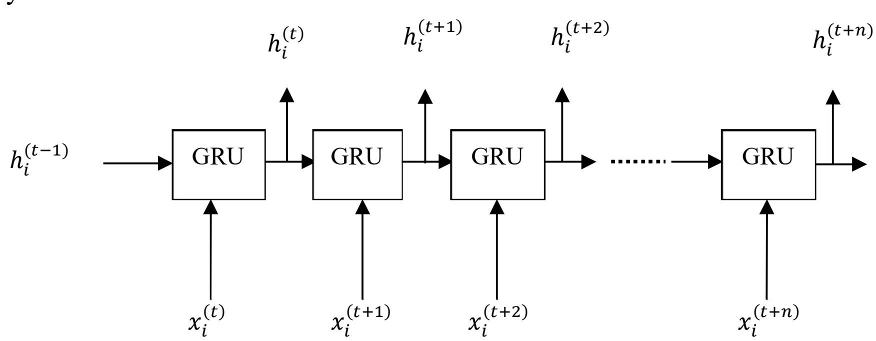


Figure 4. Gated recurrent unit (GRU) architecture

A GRU network is illustrated in Figure 4 as it develops over time. GRU in an extended network each and every one distributes the similar parameters, it must be noted. GRU has been introduced to update-gate u and reset-gate r which regulate the data flow from one time step to the next and be able to capture long-term.

d. Majority Voting

Every individual classifier, such as EDLSTM, WRMLP, and GRU, casts a vote for a class and the majority in majority voting. The simplest ensemble algorithm, voting, is frequently highly effective. It functions by adding more sub-models like EDLSTM, WRMLP, and GRU. The majority voting mechanism is used to combine the predictions made by each sub-model. The majority voting algorithm's operation is shown in Figure 5. It is a meta-classifier for merging conceptually similar or dissimilar machine learning classifiers like EDLSTM, WRMLP, and GRU for classification by majority voting. Using classification models like EDLSTM, WRMLP, and GRU, forecast the final class label as the one that has been predicted most frequently. Use the majority vote of each classifier C_j. to forecast the class label y in this case. The following equation can be used to summarise the ensemble's result in majority voting (13),

$$\hat{y} = mode\{C_1(x), C_2(x), C_3(x)\} \tag{13}$$

Where$\hat{y}$ the classification models $C_1(x), C_2(x), C_3(x)$include EDLSTM, WRMLP, and GRU.

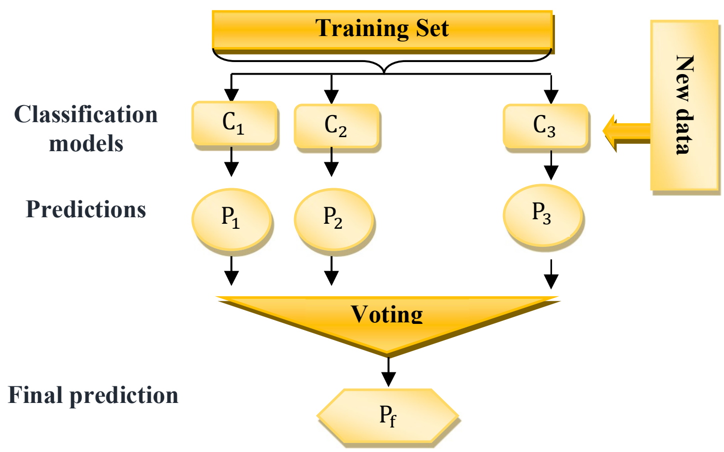


Figure 5. Majority Voting Algorithm

By using ensemble learning, various classifiers' predictions are combined. Voting by majority is the most straightforward but most used combo method. Although there are other ways to use majority voting, the most popular one takes into account the most votes, i.e. assigning an instance to the base class that the majority of base classifiers agree on. In this kind of voting, each classifier receives one vote, which has a value of 1.

4. Results and Discussion

Using three different benchmark databases, the performance of the proposed deep learning classifier is assessed through its implementation in MATLABR2020a. CGAN, InfoGAN, ACGAN, ICGAN-DSVM, ICGAN-EDSVM, and ICGAN-DLSTM classifier efficiency are all analysed in this comparison analysis along with the proposed ICGAN-EDLSTM classifier. Intel CoreTM i7-11375H processor (12M Cache, up to 5 GHz processor), 4 GB RAM, Windows 8.1 Pro, 64-bit OS, and 1 TB hard drive are the system options.

Performance Metrics: The MATLAB application assesses the effectiveness of the classifiers on three benchmark databases. Precision, recall, f-measure, accuracy, and AUC are taken into account to rate the effectiveness of different categorization algorithms. To determine the True Positive (TP), True Negative (TN), False Positive (FP), and False Negative (FN) values, the confusion matrix provided in Table 1 is used.

TABLE 1. CONFUSION MATRIX

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

The following is a definition of the evaluation metrics:

Precision: Precision is the percentage of precisely predicted events compared to those projected as positives. Equation(14) describes it,

$$Precision = \frac{TP}{TP + FP} \tag{14}$$

Recall: It is the fraction of exactly estimated positive instances from actual positives. It is described by equation (15),

$$Recall = \frac{TP}{TP + FN} \tag{15}$$

F-measure: Harmonic mean of recall and precision is the F-measure. Equation(16) describes it,

$$F - measure = \frac{2 \times (Precision \times Recall)}{Precision + Recall} \tag{16}$$

- **Accuracy:** This metric assesses how well the classifier predicts both positive and negative instances across all examples examined. Equation(17) describes it,

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{17}$$

- **AUC:** The ROC curve, or Receiver Operating Characteristic, is another name for it. This graph compares the True Positive Rate (TPR) and False Positive Rate (FPR) for each classification when the discriminating threshold for that classification is varied.
-

Table 2. Metrics Comparison of Classifiers VS School Dataset

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

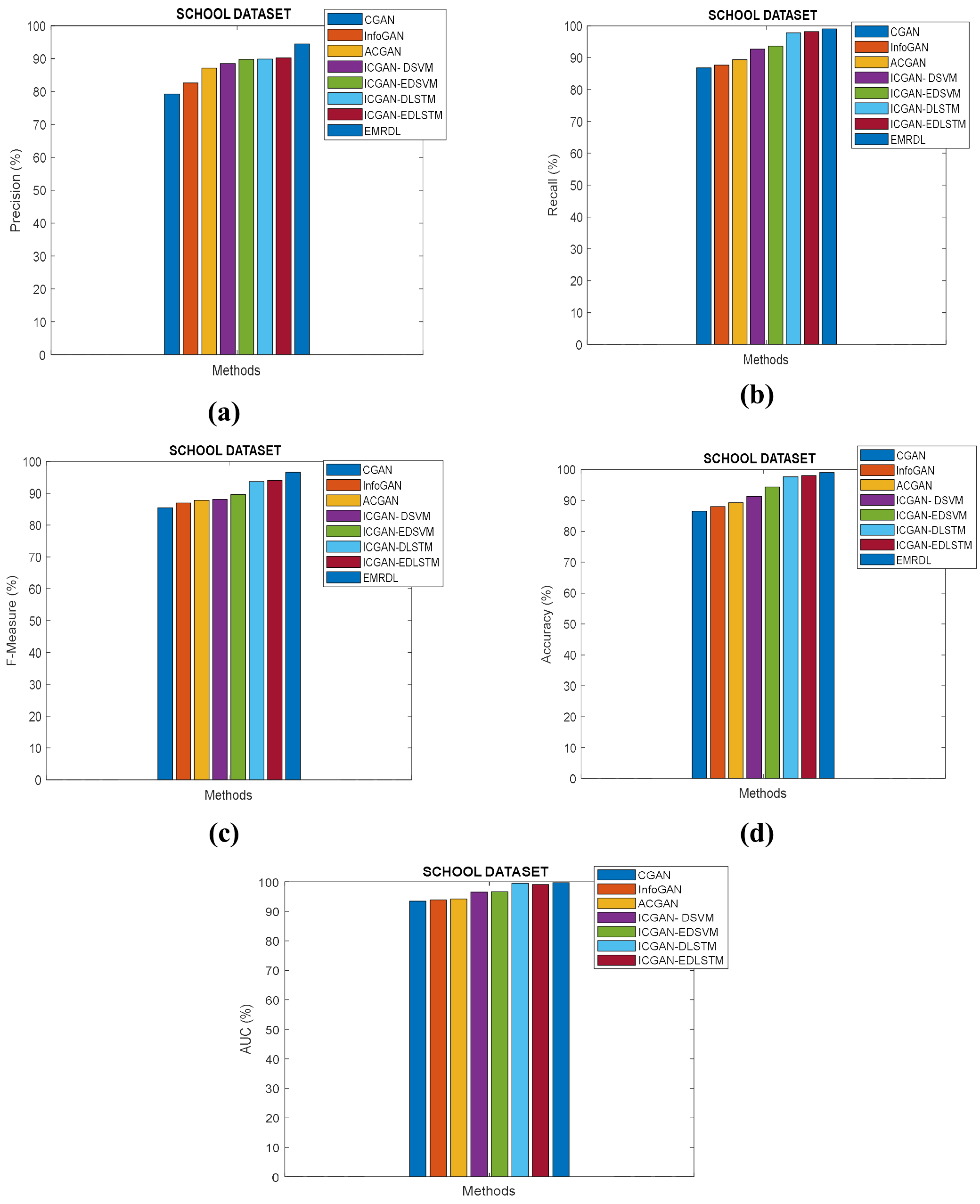
ICGAN-EDSVM	89.7543	93.6237	89.5739	94.3215	96.6820
ICGAN-DLSTM	89.8830	97.7943	93.6210	97.6811	99.4932
ICGAN-EDLSTM	90.1880	98.1942	94.0210	98.0811	99.0931
EMRDL	94.4500	98.9900	96.5600	99.0100	99.6700

Table 3. Metrics Comparison of Classifiers vs University Dataset

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
EMRDL	96.6600	98.0300	97.7800	97.8900	99.0700

Table 4. Metrics Comparison of Classifiers vs c-19ga20 Database

Classifiers	**Precision (%)**	**Recall (%)**	**F-meas0ure (%)**	**Accuracy (%)**	**AUC (%)**
CGAN	80.2487	79.4162	79.8/324	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
EMRDL	93.2300	98.7800	94.6700	98.0600	99.9300



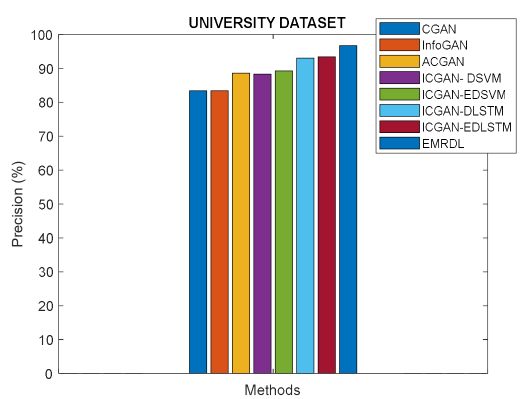
(e)

Figure 6. Analysis of Classification Algorithms vs. School Dataset. (a) Precision, (b) Recall, (c) F-measure, (d) Accuracy and (e) AUC

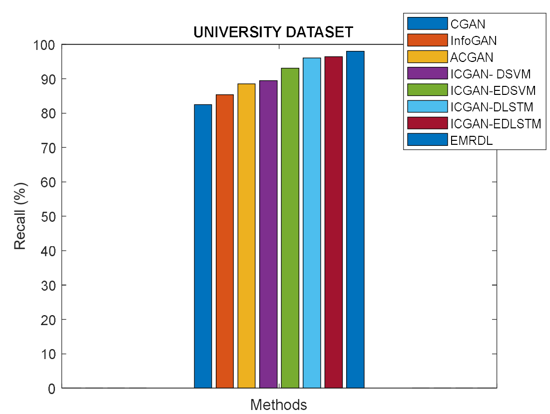
Using the school dataset, the performance values of several classification algorithms are shown in Figure 6. Using the school dataset, Figure 6(a) displays the precision values for several classification techniques. ACGAN, CGAN, InfoGAN, ACGAN, ICGAN-DSVM, ICGAN-EDSVM, ICGAN-DLSTM, and ICGAN-EDLSTM yield less precise results for the school dataset, with precision values of 79.2629%, 82.5793%, 87.0895%, 88.5366%, 89.7543%, 89.883%, and 90.1880%, respectively, according to the analysis. Using the school dataset, Figure 6(b) displays the recall values for several categorization techniques.

The f-measure values for several classification techniques on the school dataset are displayed in Figure 6(c). CGAN, InfoGAN, ACGAN, ICGAN-DSVM, ICGAN-EDSVM, ICGAN-DLSTM, and ICGAN-EDLSTM offer lower f-measure values of 85.4337%, 86.9531%, 87.7554%, 88.0893%, 89.5739%, 93.6210%, and 94.0210%, respectively, for the school dataset than the proposed EMRDL classifier, according to the analyses. Using the school dataset, Figure 6(d) displays the accuracy scores of various categorization systems. The proposed EMRDL classifier, according to the analyses, has an accuracy value of 99.0100% for the school dataset, while other classifiers like CGAN, InfoGAN, ACGAN, ICGAN-DSVM, ICGAN-EDSVM, ICGAN-DLSTM, and ICGAN-EDLSTM have lower accuracy values of 86.5000%, 87.9500%, 89.2500%, 91.3255%, 94.3215%, 97.6811% and 98.0811%, respectively for the school dataset.

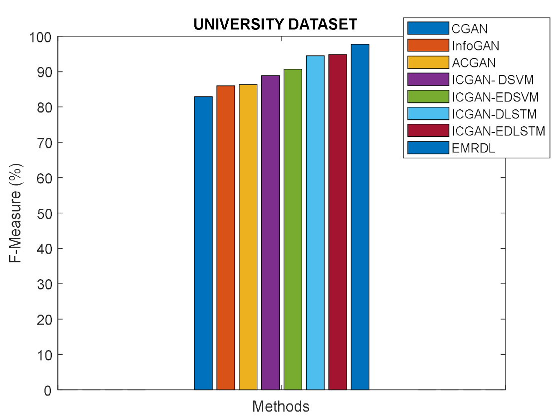
Also, Figure 6(e) displays the AUC values of different classification algorithms on the dataset from the school. According to the analysis, the proposed EMRDL classifier has an AUC value of 99.6700% for the school dataset, compared to classifiers like CGAN, InfoGAN, ACGAN, ICGAN-DSVM, ICGAN-EDSVM, ICGAN-DLSTM, and ICGAN-EDLSTM that have lower AUC values of 93.4919%, 93.8635%, 94.2063%, 96.5172%, 96.6820%, 99.4932% and 99.0931%, respectively for the school dataset.



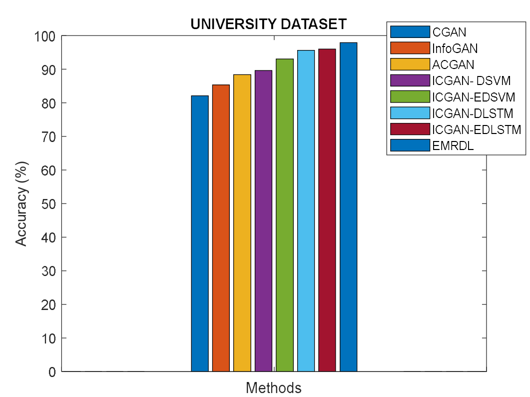
(a)



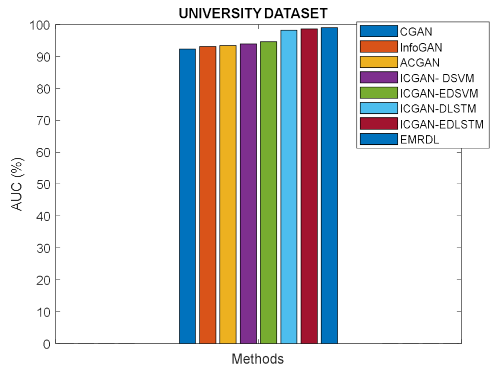
(b)



(c)



(d)



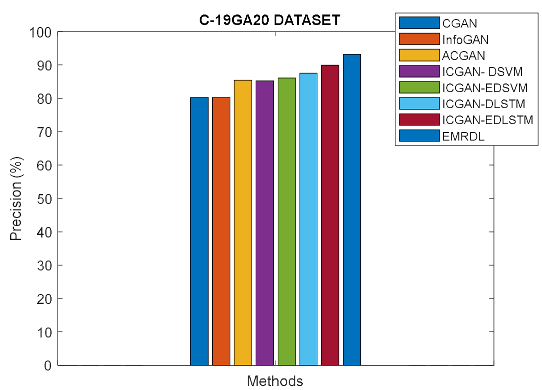
(e)

Figure 7. Analysis of classification Algorithms vs. University Dataset. (a) Precision, (b) Recall, (c) F-measure, (d) Accuracy and (e) AUC

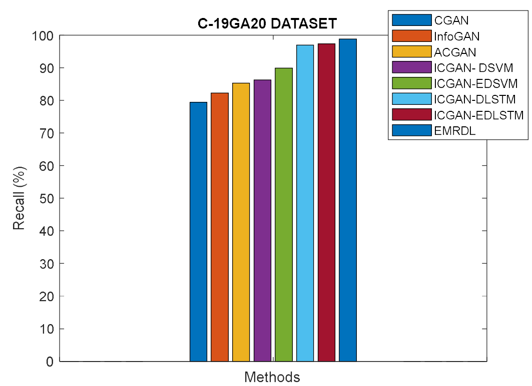
The results of several classification methods on the university dataset are shown in Figure 7. Using the university dataset, Figure 7(a) displays the precision values of the various classifiers. From the analyses, the proposed EMRDL classifier gives the precision value of 96.6600% for the university dataset, whereas the other classifiers such as CGAN, InfoGAN, ACGAN, ICGAN-DSVM, ICGAN-EDSVM, ICGAN-DLSTM and ICGAN-EDLSTM gives lesser precision values of 83.3593%, 83.3829%, 88.6104%, 88.3469%, 89.2582%, 92.9880% and 93.3881% for university dataset. Using the university dataset, Figure 7(b) shows the recall values for several classification techniques. From the analyses, the proposed EMRDL classifier gives the recall value of 98.0300% for the university dataset, whereas the other classifiers such as CGAN, InfoGAN, ACGAN, ICGAN-DSVM, ICGAN-EDSVM, ICGAN-DLSTM and ICGAN-EDLSTM gives lesser recall values of 82.5184%, 85.3545%, 88.4935%, 89.4726%, 93.1142%, 96.0192% and 96.4191% respectively for the university dataset. The f-measure values of various classification algorithms on the university dataset are shown in Figure 7(c). From the analyses, the proposed EMRDL classifier gives the f-measure value of 97.7800% for the university dataset, whereas the other classifiers such as CGAN, InfoGAN, ACGAN, ICGAN-DSVM, ICGAN-EDSVM, ICGAN-DLSTM and ICGAN-EDLSTM gives lesser f-measure values of 82.9367%, 85.9779%, 86.3205%, 88.8742%, 90.6679%, 94.4792% and 94.8793% respectively for the university dataset.

From the analyses, the proposed EMRDL classifier gives the recall value of 98.0300% for the university dataset, whereas the other classifiers such as CGAN, InfoGAN, ACGAN, ICGAN-DSVM, ICGAN-EDSVM, ICGAN-DLSTM and ICGAN-EDLSTM gives lesser recall values of 82.5184%, 85.3545%, 88.4935%, 89.4726%, 93.1142%, 96.0192% and 96.4191% respectively for the university dataset. Figure 7(c) exhibits the f-measure values of various classification algorithms on the university dataset. From the analyses, the proposed EMRDL classifier gives the f-measure value of 97.7800% for the university dataset, whereas the other classifiers such as CGAN, InfoGAN, ACGAN, ICGAN-DSVM, ICGAN-EDSVM, ICGAN-DLSTM and ICGAN-EDLSTM gives lesser f-measure values of 82.9367%, 85.9779%, 86.3205%, 88.8742%, 90.6679%, 94.4792% and 94.8793% respectively for the university dataset. For the university dataset, Figure 7(d) displays the accuracy values of various classification systems.

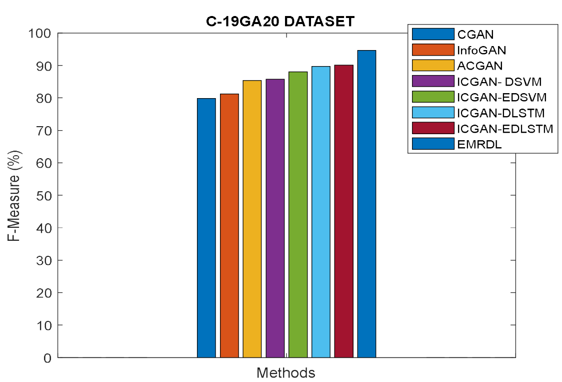
According to the analyses, the proposed EMRDL classifier has an accuracy value of 97.8900% for the university dataset, while other classifiers like CGAN, InfoGAN, ACGAN, ICGAN-DSVM, ICGAN-EDSVM, ICGAN-DLSTM, and ICGAN-EDLSTM have lower accuracy values for the same dataset of 82.100%, 85.3500%, 88.4000%, 89.6510%, 93. Using the university dataset, Figure 7(e) also shows the AUC values of various classification systems. According to the analyses, the proposed EMRDL classifier achieves an AUC value of 99.0700% for the university dataset, while other classifiers like CGAN, InfoGAN, ACGAN, ICGAN-DSVM, ICGAN-EDSVM, ICGAN-DLSTM, and ICGAN-EDLSTM achieve lower AUC values of 92.3266%, 93.1607%, 93.4375%, 93.9675%, 94.5844%, and 98.6221% respectively for the university dataset.



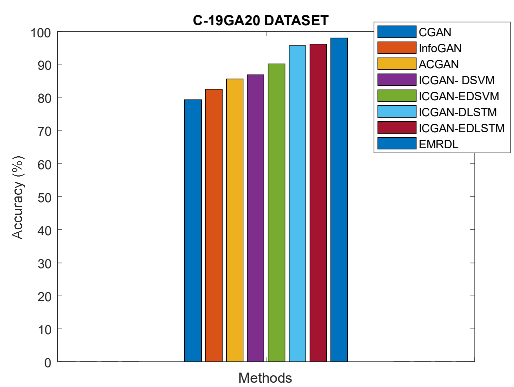
(a)



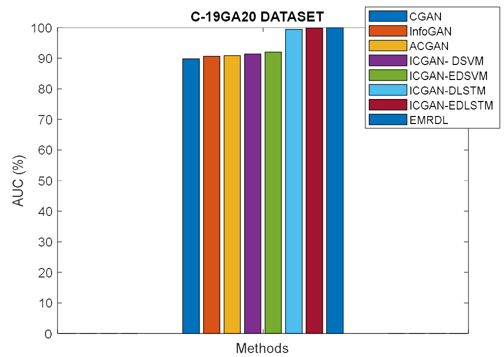
(b)



(c)



(d)



(e)

FIGURE 8. ANALYSIS OF CLASSIFICATION ALGORITHMS VS. C-19GA20 DATABASE. (a) Precision, (b) Recall, (c) F-measure, (d) Accuracy and (e) AUC

Using the C-19GA20 dataset, Figure 8 shows the performance values attained by various classification methods including CGAN, InfoGAN, ACGAN, ICGAN-DSVM, ICGAN-EDSVM, ICGAN-DLSTM, ICGAN-EDLSTM, and EMRDL. Using the C-19GA20 dataset, Figure 8(a) shows the precision values for the various classifiers. From the analyses, the proposed EMRDL classifier gives the precision value of 93.2300% for the C-19GA20 dataset, whereas the other classifiers such as CGAN, InfoGAN, ACGAN, ICGAN-DSVM, ICGAN-EDSVM, ICGAN-DLSTM and ICGAN-EDLSTM gives lesser precision values of 80.2487%, 80.2720%, 85.4473%, 85.1864%, 86.0886%, 87.5720% and 89.9120% for the C-19GA20 dataset. Using the C-19GA20 dataset, Figure 8(b) shows the recall values for several classification techniques.According to the analysis, the proposed EMRDL classifier has a recall value of 98.7800% for the C-19GA20 dataset, compared to classifiers like CGAN, InfoGAN, ACGAN, ICGAN-DSVM, ICGAN-EDSVM, ICGAN-DLSTM, and ICGAN-EDLSTM that have recall values of 79.4162%, 82.2239%, 85.3315%, 86.3008%, 89. Using the C-19GA20 dataset, Figure 8(c) shows the f-measure values of various classification techniques. From the analyses, the proposed EMRDL classifier gives the f-measure value of 94.6700% for the C-19GA20 dataset, whereas the other classifiers such as CGAN, InfoGAN, ACGAN, ICGAN-DSVM, ICGAN-EDSVM, ICGAN-DLSTM and ICGAN-EDLSTM gives lesser f-measure values of 79.8324%, 81.2480%, 85.3894%, 85.7436%, 87.9973%, 89.7080% and 90.1380% respectively for the C-19GA20 dataset. From the analyses, the proposed EMRDL classifier gives the recall value of 98.7800% for the C-19GA20 dataset, whereas the other classifiers such as CGAN, InfoGAN, ACGAN, ICGAN-DSVM, ICGAN-EDSVM, ICGAN-DLSTM and ICGAN-EDLSTM gives lesser recall values of 79.4162%, 82.2239%, 85.3315%, 86.3008%, 89.9060%, 96.9430% and 97.3630% respectively for the C-19GA20 dataset. Figure 8(c) illustrates the f-measure values of various classification algorithms on the C-19GA20 dataset. From the analyses, the proposed EMRDLclassifier gives the f-measure value of 94.6700% for the C-19GA20 dataset, whereas the other classifiers such as CGAN, InfoGAN, ACGAN, ICGAN-DSVM, ICGAN-EDSVM, ICGAN-DLSTM and ICGAN-EDLSTM gives lesser f-measure values of 79.8324%, 81.2480%, 85.3894%, 85.7436%, 87.9973%, 89.7080% and 90.1380% respectively for the C-19GA20 dataset. Using the C-19GA20 dataset, Figure 8(d) shows the accuracy values of various classification systems. From the analyses, the proposed EMRDL classifier gives the accuracy value of 98.0600% for the C-19GA20 dataset, whereas the other classifiers such as CGAN, InfoGAN, ACGAN, ICGAN-DSVM, ICGAN-EDSVM, ICGAN-DLSTM and ICGAN-EDLSTM gives lesser accuracy values of 79.4010%, 82.6347%, 85.6695%, 86.9142%, 90.2850%, 95.7640% and 96.2640% respectively for the C-19GA20 dataset. The AUC values of different classification methods on the C-19GA20 dataset are also shown in Figure 8(e).CGAN, InfoGAN, ACGAN, ICGAN-DSVM, ICGAN-EDSVM, ICGAN-DLSTM, and ICGAN-EDLSTM offer lower AUC values of 89.7565%, 90.5881%, 90.8640%, 91.3925%, 92.0075%, 99.4040%, and 99.8040% for the C-19GA20 dataset, respectively, whereas the proposed EMRDL classifier gives the AUC value of 99.4040% and 99.8040% respectively for the C-19GA20 dataset.

5. CONCLUSION AND FUTURE WORK

In this paper, predict the outcome of students in virtual classrooms during the pandemic. The proposed model has four major phases: data collection, feature selection, prediction and performance analysis. In the initial stage, dataset associated with the students' learning through online classes are collected from Kaggle, and UCI. In the second stage, redundant features are discarded by the Circle Map Tuna-Swarm Optimization Algorithm (CMTSOA). It includes the CTSOA to choose the most relevant attributes for predicting student outcomes during the Covid-19 outbreak. CTSOA relies on the collaborative hunting nature of tuna swarm, and circle map has been introduced for random number generation. CMTSOA finds the best fitness tuna individuals for optimal selection of student performance features in the database. In the third stage, EMRDL classifier 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 grade.Ensemble Deep Long Short-Term Memory (EDLSTM) classifier should build proper and fuse results from the LSTM classifier to design a robust ensemble classifier. RMLP network can be constructed by starting from the MLP architecture by adding delayed connections between the neighboring nodes of a hidden layer, together with the originating node itself. Gated Recurrent Unit (GRU) has been implemented depending on reset-gate and update-gate to evaluate the long-term dependency. Results of the classifiers are combined via majority voting. The combined output is better than an individual output since error is reduced. It empirically compared the performance and efficiency of ensemble classifier that make use of single classification methods. At last, experimental findings are obtained based on the different evaluation metrics and it proves that the proposed EMRDL classifier achieves maximum efficiency when compared to the other classification models.In the future work, hybrid model has been introduced for feature selection, and ensemble model can be used for reducing the irrelevant attributes which gives higher accuracy and efficient in performance.

6. References

1. Moubayed, M. Injadat, A. Bou Nassif, H. Lutfiyya, and A. Shami, ''Elearning: Challenges and research opportunities using machine learning & data analytics,'' IEEE Access, vol. 6, pp. 39117–39138, 2018.

2. Pardo A., J. Jovanovic, S. Dawson, D. Gašević, and N. Mirriahi, ''Using learning analytics to scale the provision of personalised feedback,'' Brit. J. Educ. Technol., vol. 50, no. 1, pp. 128–138, 2019.
3. Batanero C., L. de-Marcos, J. Holvikivi, J. R. Hilera, and S. Oton, ''Effects of new supportive technologies for blind and deaf engineering students in online learning,'' IEEE Trans. Educ., vol. 62, no. 4, pp. 270–277, 2019.
4. Wang C., H.-C.-K. Hsu, E. M. Bonem, J. D. Moss, S. Yu, D. B. Nelson, and C. Levesque-Bristol, "Need satisfaction and need dissatisfaction: A comparative study of online and face-to-face learning contexts," Comput. Hum. Behav., vol. 95, pp. 114–125, 2019.
5. Yeung, C.-K., and Yeung, D.-Y. (2018). "Addressing two problems in deep knowledge tracing via prediction-consistent regularization," in Proceedings of the Fifth Annual ACM Conference on Learning at Scale (London), 1–10.
6. Juhaňák, L., Zounek, J., and Rohlíková, L. (2019). Using process mining to analyze students' quiz-taking behavior patterns in a learning management system. Comput. Hum. Behav. 92, pp.496–506.
7. Liu, Q., Tong, S., Liu, C., Zhao, H., Chen, E., Ma, H., et al. (2019). "Exploiting cognitive structure for adaptive learning," in in Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining (Anchorage, AK), pp.627–635.
8. Liu, Q., Wu, R., Chen, E., Xu, G., Su, Y., Chen, Z., et al. (2018). Fuzzy cognitive diagnosis for modelling examinee performance. ACM Trans. Intell. Syst. Technol. 9, pp.1–26.
9. Wang, F., Liu, Q., Chen, E., Huang, Z., Chen, Y., Yin, Y., et al. (2020). "Neural cognitive diagnosis for intelligent education systems," in Proceedings of the AAAI Conference on Artificial Intelligence, Vol. 34, pp.6153–6161.
10. Li, S. and Liu, T., 2021. Performance prediction for higher education students using deep learning. Complexity, vol.2021, no. 9958203, pp.1-10.
11. Helal S., J. Li, L. Liu, E. Ebrahimie, S. Dawson, D. J. Murray, and Q. Long, 2018, ''Predicting academic performance by considering student heterogeneity,'' Knowl.-Based Syst., vol. 161, pp. 134–146.
12. Turabieh, H., Azwari, S.A., Rokaya, M., Alosaimi, W., Alharbi, A., Alhakami, W. and Alnfiai, M., 2021. Enhanced Harris Hawks optimization as a feature selection for the prediction of student performance. Computing, vol.103, no.7, pp.1417-1438.
13. Allam, M. and Nandhini, M., 2022. Optimal feature selection using binary teaching learning based optimization algorithm. Journal of King Saud University-Computer and Information Sciences, vol.34, no.2, pp.329-341.
14. Fernandes E., M. Holanda, M. Victorino, V. Borges, R. Carvalho, and G. V. Erven, 2019, ''Educational data mining: Predictive analysis of academic performance of public school students in the capital of Brazil,'' J. Bus. Res., vol. 94, pp. 335–343.
15. Polyzou A. and G. Karypis, 2019, ''Feature extraction for next-term prediction of poor student performance,'' IEEE Trans. Learn. Technol., vol. 12, no. 2, pp. 237–248.

16. Xu X., J. Wang, H. Peng, and R. Wu, 2019, ‘‘Prediction of academic performance associated with Internet usage behaviors using machine learning algorithms,’’ Comput. Hum. Behav., vol. 98, pp. 166–173, Sep. 2019.
17. Widyahastuti F. and V. U. Tjhin, 2018, ‘‘Performance prediction in online discussion forum: State-of-the-art and comparative analysis,’’ Procedia Comput. Sci., vol. 135, pp. 302–314.
18. Aslam, N., Khan, I., Alamri, L. and Almuslim, R., 2021. An Improved Early Student’s Academic Performance Prediction Using Deep Learning. International Journal of Emerging Technologies in Learning (iJET), 16(12), pp.108-122.
19. Zian, S., Kareem, S.A. and Varathan, K.D., 2021. An empirical evaluation of stacked ensembles with different meta-learners in imbalanced classification. IEEE Access, 9, pp.87434-87452.
20. Kalyani B. S., D. Harisha, V. RamyaKrishna, and S. Manne, “Evaluation of students performance using neural networks,” International Conference on Intelligent Computing, Information and Control Systems, Springer, Berlin, Germany, pp. 499–505, 2019.
21. Okubo F., T. Yamashita, A. Shimada, and H. Ogata, “A neural network approach for students’ performance prediction,” in Proceedings of the Seventh International Learning Analytics & Knowledge Conference, pp. 598-599, New York, NY, USA, March 2017.
22. Xu, J., Moon, K.H. and Van Der Schaar, M., 2017. A machine learning approach for tracking and predicting student performance in degree programs. IEEE Journal of Selected Topics in Signal Processing, 11(5), pp.742-753.
23. Farissi, A. and Dahlan, H.M., 2020, Genetic algorithm based feature selection with ensemble methods for student academic performance prediction. In Journal of Physics: Conference Series (Vol. 1500, No. 1, p. 012110). IOP Publishing.
24. Surenthiran, S., Rajalakshmi, R. and Sujatha, S.S., 2021. Student performance prediction using atom search optimization based deep belief neural network. Optical Memory and Neural Networks, 30(2), pp.157-171.
25. Kim, B.H., Vizitei, E. and Ganapathi, V., 2018. GritNet: Student performance prediction with deep learning. arXiv preprint arXiv:1804.07405.
26. Hussain, S., Muhsin, Z., Salal, Y., Theodorou, P., Kurtoğlu, F. and Hazarika, G., 2019. Prediction model on student performance based on internal assessment using deep learning. International Journal of Emerging Technologies in Learning, 14(8), pp.4-22.
27. Ahuja R. and S. Sharma, “Stacking and voting ensemble methods fusion to evaluate instructor performance in higher education,” International Journal on Information Technology, vol. 13, pp. 1–11, 2021.
28. Asselman, A., Khaldi, M. and Aammou, S., 2021. Enhancing the prediction of student performance based on the machine learning XGBoost algorithm. Interactive Learning Environments, pp.1-20.
29. Xiong, S.Y., Gasim, E.F.M., Ying, C.X., Wah, K.K. and Ha, L.M., 2022. A Proposed Hybrid CNN-RNN Architecture for Student Performance Prediction. International Journal of Intelligent Systems and Applications in Engineering, 10(3), pp.347-355.

30. He, Y., Chen, R., Li, X., Hao, C., Liu, S., Zhang, G. and Jiang, B., 2020. Online at-risk student identification using RNN-GRU joint neural networks. Information, 11(10), pp.1-12.
31. Hidalgo, Á.C., Ger, P.M. and Valentín, L.D.L.F., 2022. Using Meta-Learning to predict student performance in virtual learning environments. Applied Intelligence, 52(3), pp.3352-3365.
32. Adejo, O.W. and Connolly, T. (2018), "Predicting student academic performance using multi-model heterogeneous ensemble approach", Journal of Applied Research in Higher Education, Vol. 10 No. 1, pp. 61-75.
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.
36. Zia, M., Hussain, M. and Jaffar, M. (2018). A novel spontaneous facial expression recognition using dynamically weighted majority voting based ensemble classifier. Multimedia Tools Application, 77, pp. 25537-25567.
37. Moustafa, S., ElNainay, M., Makky, N. and Abougabal, M. (2018). Software bug prediction using weighted majority voting techniques. Alexandria Engineering Journal, 7, pp. 2763-2774.