Information Sciences Letters
An International Journal

http://dx.doi.org/10.18576/isl/120346

Predicting Student's Performance in Online Education through Deep Learning Model

Mohammed H. R. Khalaf [1,*] *and Zeinab M. Abdel Azim* [2]

[1]Department of Educational Technology, Faculty of Education, Alexandria University, Alexandria, Egypt
[2]Department of Management Information Systems, Egyptian Institute of Alexandria Academy for Administration and Accounting, Alexandria, Egypt

Received: 27 Jul. 2022, Revised: 2 Nov. 2022, Accepted: 5 Nov. 2022.
Published online: 1 Mar. 2023.

Abstract: This epidemic has prompted the development of Education 4.0, virtual learning, and the demand to adapt educational practices to meet the needs of younger demographics. A rising epidemic has necessitated the shutdown of campuses where education programs are now being carried out online in educational institutions all over the globe. The report includes a study on the effectiveness and perceptions of students toward digital learning during the pandemic. A Convolutional Neural Network (CNN) and Particle swarm optimization model, which forecasts the student's learning rates, are used to tackle this issue. This study will categorize student performance into low, medium, and high grades to forecast student achievement. The Kaggle student's performance assessment database is utilized to gather the student information logs, which are then pre-processed to eliminate noise and redundant data. The CNN derives features based on the student's attention and arbitrary patterns sequencing by examining the pre-processed information. Then, utilizing the Minimum Redundancy Maximum Relevance (mRMR) approach, the retrieved characteristics are evaluated. The lowest one that treats each characteristic individually is chosen as the greatest feature by mRMR. CNN uses stochastic Gradient Descent (SGD) to calculate the characteristic weights, which are then modified for improved extracting features. Finally, the CNN-WOA method forecasts the final academic achievement forecast outcome. Studies revealed that the suggested approach outperforms existing ones in terms of accuracy, precision, recall, and F-score while requiring less computing time.

Keywords: Deep Learning, Convolutional Neural Network, Particle swarm optimization, Stochastic Gradient Descent.

1 Introduction

Education has seen an increase in the popularity of online learning as a result of its accessibility and scheduling adaptability. However, it has struggled to maintain a steady stream of graduates and ensure curriculum improvement over time [1]. A worrying aspect of online assignment continuance is disengagement. In the academic sector, forecasting academic achievement is crucial because the impact on academic position enables improvement for enhanced efficiency. Because it can improve the learning process in developing individuals who require further help and prevent individuals from quitting during examinations, the student's ability prediction is a crucial area of investigation. Massive open online courses used mostly in distant education systems have become more popular due to information and communications technology advances. One of the most popular electronic learning sites is MOOC. The Massive Open Online Courses delivers the curriculum using digitized tool resources in various formats, including text format, sound recording, and audio-visual. Most students find that watching video tutorials allows them to understand information better than reading regular text materials [2].

Teachers can analyze academic achievement using documented user information in web-based educational experiences like massive open online courses, digital electronics education, design suite, and learning management systems. However, instructors could find it challenging to analyze individual records. Web-based focus on important like massive open online courses and learning management systems are common; they focus on providing course work from many institutions and give opportunities for students worldwide. In addition, they offer execution, writing, content

*Corresponding author e-mail: dr_mhragab@alexu.edu.eg

development, management, and identity [3]. Knowledge about a person's self - esteem all through the entire course can be given to teachers by forecasting a competency level in a course or sessions via, for instance, tests, tasks, examinations, and workshop events Investigators have used a variety of statistical approaches to data collected from conventional and virtual colleges to achieve this objective. Analysts mainly use a pupil's academic background and survey details to forecast their achievement [4].

The last few centuries have seen significant advancements in educational technologies, which have been demonstrated to be quite helpful throughout the pandemic. Many online venues assist in online learning. However, colleges found it difficult to map actual instructional operations online. Students and instructors also had various logistical, technological, economic, and sociological issues. The online education platform is shown in Fig. 1.

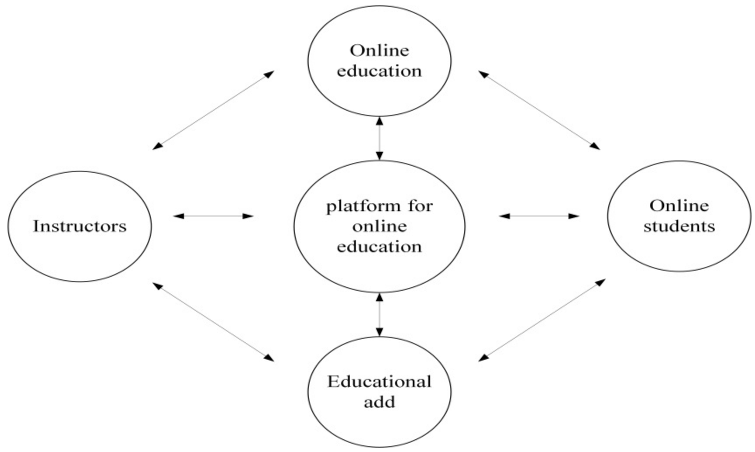

Fig. 1: Online education platform.

The educational environment has evolved into a beginner on web-based learning networks like Chalkboard, submit assignments, Moodle, and web-based educational standards. The advantages of the network are used in online learning settings to increase students' technological adaptability and advance established teaching methods. It offers a variety of remote academic courses for every respective school. As a result, it has drawn many students who do not want to be on the faculty for many reasons. Open courseware platforms house academic materials that teachers can use to provide components to their pupils and boost their academic abilities. The online course offers great potential for faculty members to monitor and record students' learning behavior and improve instructional interconnection [5]. It also offers a consistent method for students to continuously monitor class notes, coursework, application forms, documents of examinations, periodic or daily practice tests, and other operations. While online learning environments have several amenities, they nevertheless have significant problems, such as poor academic motivation that affects how well they function academically and how the outcomes turn out. Compared to older schooling, the lower graduation incidence also emerges as a significant issue. The student's performance is shown in Fig. 2.

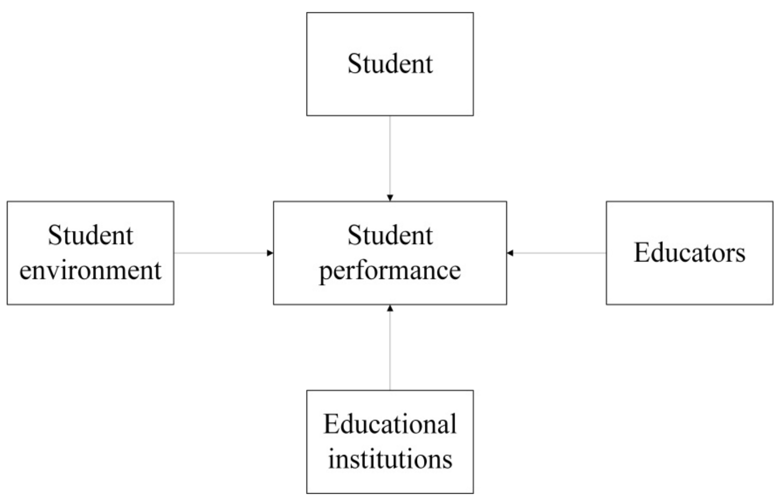

Fig. 2: Student's performance toward online education.

The ongoing inclusion of data to the data sets pertaining to individuals' academic performance in academic institutions makes it difficult to predict how students will do throughout their learning process. A structure for teaching contains a vast amount of illuminating knowledge. The database could include information about the participants, teachers, graduate courses, resources, etc. In deciding the appropriate, academic mining techniques are used to find the models [6]. Electronic learning provides academic benefits such as virtual learning and online service, eliminating learners' need to attend in-person classes. Electronic learning is less costly than conventional instructional strategies, and more individuals can sign up for online classes. Moreover, there is no direct involvement between learners and educators when learning online. Consequently, electronic learning presents various difficulties. To start, it is challenging for teachers to evaluate a track's performance. Second, compared to conventional instructional strategies, the participation rate for participants in electronic learning sessions is substantially greater. Third, judging an achievement is challenging. Finally, it can be challenging to identify at-risk pupils in brand-new modules. Last but not least, educators want to forecast their learners' performance in test format [7].

The most widely used technique in the artificial neural networks sector of educational data mining. Most of these worries were addressed throughout the last ten years with the development of deep artificial neural networks, even though there have occasionally been problems with artificial neural networks, particularly when collecting characteristics humans can understand from the expected outcomes. DL, which developed upon ML and is defined by several statistical levels, allows the system to learn from instances, processes, or occurrences instead of using the characteristics manually as has traditionally been done. In the secondary end, instruction analytics is classified continuously with ideas of teaching data analysis, assisting organizations in their financial health by lowering turnover rates, enhancing learning consequences by taking students ' learning behavior into consideration, placing more emphasis on initiatives for teachers that possibly result in providing a secure institute, and getting exercise an appropriate allocation of resources procedure [8]. The phenomenon of learning analytics has helped universities retain learners over time, which has led to an increase in student achievement. Retaining individuals became a regular strategic target for organizations. Various methods are accessible to analyze students' effectiveness in online learning. Deep learning is the method most frequently employed in the educational industry to assess students' progress. It is a growing field of study that emphasizes numerous deep learning techniques, such as selecting features, predictions, and classifications. This research presents a novel technique for assessing student progress in online learning and educational behavior, the convolutional neural network with whale optimization (CNN-WOA). Database collection, pre-processing, extraction of features, and classifications are the broad categories these fall into. The following section covers the succinct overview from Section 2, followed by Section 3 discussion of similar works. Section 4 explanation of the proposed tactics and strategies. Section 5 presents the anticipated research after Section 6 presents the comprehensive empirical data with appropriate analysis and conclusions.

2 Related Works

Ahmed et al. [9] proposed a deep learning model for predictive learning. Since most students see the same online lecture videos, analyzing the learning behavior of MOOC aficionados has emerged as a problem in educational analytics. It is beneficial to thoroughly analyze these behaviors, investigate different learning styles for students, and forecast their achievement using MOOC course videos. This study uses clickstream data from videos to analyze a temporally sequential classification task and forecast learner performance, a crucial challenge in decision-making, to better serve students. This study uses a deep neural network (LSTM) on a collection of implicit variables from video devices to forecast students' weekly performance and give teachers the tools to design prompt appropriate interventions. Results demonstrate that the suggested model's accuracy rate ranges from 82% to 93% over the twelve weeks. The suggested LSTM model works 93% more accurately in datasets from actual courses than baseline ANNs, SVM, and Logistic Regression.

Hajra et al. [10] used deep learning to predict students' performance. This study adds to the body of knowledge by identifying students who are likely to be unsuccessful early on, identifying those who are most likely to drop out of a module, and identifying key characteristics that help some students perform better than others. After the module's start, the results show that demographic factors and students' clickstream activity substantially impacted their performance. First, before modules start, students' involvement in the educational environment has little effect on how well they perform. This study evaluates the efficiency of the deep learning model in predicting student outcomes early on, enabling prompt university involvement to put corrective procedures for student assistance and counseling into place. These studies will make it easier for institutions to set up committees to support students' needs and benefits, assisting the institution in upholding its standards of politeness and efficiency. A weakness of our research was the failure to detect a distinct tendency for such students to the imbalanced class issue in "qualitative difference" instances. However, geography and demographic factors frequently have a big impact on performance. The results show that the methodologies used to evaluate student initial assessment are efficient. Such data-driven research is necessary to support higher education in creating a framework for deep learning and assisting in their judgment.

The student's performance was compared using a machine learning algorithm by Vijayalakshmi [11]. To analyze student status and make improvements for higher results, it is critical to predict student outcomes in the education field. Academic data mining uses data mining principles and techniques in education. Machine learning algorithms are now quite helpful in practically all industries. Many scientists have only employed machine learning algorithms. In this paper, we developed a Deep Neural Network-based approach for predicting student performance. The Kaggle dataset is used to build the model, test it, and evaluate the accuracy of various techniques, including Decision Tree (C5.0), Naive Bayes, Random Forest, Support Vector Machine, K-Nearest Neighbour, and Deep Neural Network in R. Deep Neural Network beat six algorithms with an accuracy of 84%.

Mustaq et al. [12] used machine learning to predict the learning of students. Since it may enable teachers to identify students who require further help and prevent pupils from quitting before final examinations, the student's outcome prediction is a crucial research topic. This study's goal is to forecast the challenges that learners will face in the following session of a digital design course. We used machine learning techniques to examine the data logged by the analog and digital instruction and design suite (DEEDS), a TEL system. Among the machine learning algorithms, are ANN, SVM, logistic regression, Naïve Bayesian classifiers, and decision trees. Students can complete digital design activities with varying difficulty using the Actions system while documenting input data. The average length of time, the overall number of activities, the average idle time, the average number of keystrokes, and overall similar activities for every exercise during private counseling in the electronic design course served as the original study input parameters. The session grades for any given student served as the study's output variable. We then used the information from the previous session to train algorithms for machine learning and the information from the subsequent session to test the methods. To assess the models' capabilities, we rank-fold cross-validation and calculated the transmitter operating characteristics and root median square error metrics.

Sadiq et al. [13] proposed a deep learning method for the internal assessment of students. Academic system research advances the way that learning is organized. The education organization can provide the student with extra assessments with the help of the student's predicted achievement, which also advances the growth of the educational system inside the company. Through an evaluation of an applicant, the outcome of the applicant may be predicted based on the experimental data. As a result, tutors may spend more time working with applicants with low internal assessment scores to help them perform better on the final exam. Kids whose skills received low internal evaluations may be categorized as at-risk pupils. The modeling approach may be employed to raise such candidates' test scores in the same way as it may be used to alert such pupils and parents. The instructors may act in real-time by checking the grades students receive on their assessments. Therefore, the assessment may remain a crucial component of a certain course. The use of machine learning techniques to forecast student performance has proven useful for identifying underachievers and enabling tutors to implement corrective actions earlier, even at the start of an educational year, using only internal assessment results from prior academic sessions to help the group most in need.

Sofia et al. [14] used deep learning to support online learning. University and school face-to-face instruction have been completely suspended due to the coronavirus (Covid-19) epidemic, forcing students and instructors to take quick lessons in online courses and technology. LMS, such as Moodle, Blackboard, and Google Classroom, as well as teleconferencing platforms (e.g., Zoom, WebEx, MS Teams), are being adopted and heavily used as online learning environments during this unprecedented crisis (OLEs). However, as these media just serve as a platform for online engagement, there is a need for efficient tools to support teachers and serve as higher cognitive cues for students to predict their conduct in OLEs. Here, we demonstrate for the first time how Deep Learning techniques can be applied to manage data collected from LMS users and create a brand-new predictive model, called DeepLMS, that can anticipate the QoI with LMS. When tested on QoI data from one database before the Covid-19 outbreak and two databases during the pandemic, DeepLMS produces mean testing Root Mean Square and maximum correlation coefficients between underlying data and anticipated QoI values. DeepLMS-tailored QoI forecasting increases engagement rates in online courses. It gives teachers an extra evaluation pathway in contrast to the material assessment, enhancing the overall picture of the student's desire and involvement in the learning process.

Deep learning-based prediction of student performance was discussed by Byung et al. [15]. It is difficult to anticipate students' future success as they engage with online education. This is known as forecasting student performance. It may be essential to make accurate early projections of a student's performance to provide timely instructional strategies during a course. However, there hasn't been much prior research that has looked at this issue from a deep-learning approach. The student achievement person who thinks is reformulated as a consecutive event forecasting model in this research, and a new deep learning-based system called GritNet is proposed. GritNet relies on the BLSTM. It is difficult to anticipate students' future success as they engage with online education. This is known as forecasting student performance. It may be essential to make accurate early projections of a student's performance to provide timely instructional strategies during a course. However, there hasn't been much prior research that has looked at this issue from a deep-learning approach. The student achievement person who thinks is reformulated as a consecutive event forecasting model in this research, and a new deep learning-based system called GritNet is proposed.

GritNet relies on the BLSTM.

Jabeen et al. [16] proposed a deep-learning approach for predicting educational data. Sentiment analysis identifies and classifies user sentiments from a portion of text into several sentiments. For instance, users' attitudes toward a certain topic or object are influenced by their emotions, such as happiness, sadness, anger, positivity, negativity, or neutrality. Sentiment analysis is one of the most active study fields in language processing, information extraction, and text analytics. It is significant in various domains, such as administrative and social sciences, and education, where student input is crucial to determining how successful learning tools are. With the growth of educational institutions, online learning platforms have drawn in many learners by providing cost-free courses. Thousands of students enroll in these enormous online courses every year, and their opinions of the course material and educational quality are further evaluated. Also, make blog ideas to enhance instruction by expressing favorable or unfavorable opinions. To do sentiment analysis on educational data, this study provides a framework based on the deep Learning method. To forecast the best model, we concentrated on the efficiency and durability of the training data set in this study. It is acknowledged that MLP and SVM are the superior models.

3 Proposed Methodology

Various methods are accessible to analyze students' effectiveness in online learning. Deep learning is the method most frequently employed in the educational industry to assess students' progress. It is a growing field of study that emphasizes numerous deep learning techniques, such as selecting features, predictions, and classifications. This research presents a novel technique for assessing student progress in online learning and educational behavior, the convolutional neural network with whale optimization (CNN-WOA). Database collection, pre-processing, extraction of features, and classifications are the broad categories these fall into. The suggested system architecture procedure is shown in Fig. 3.

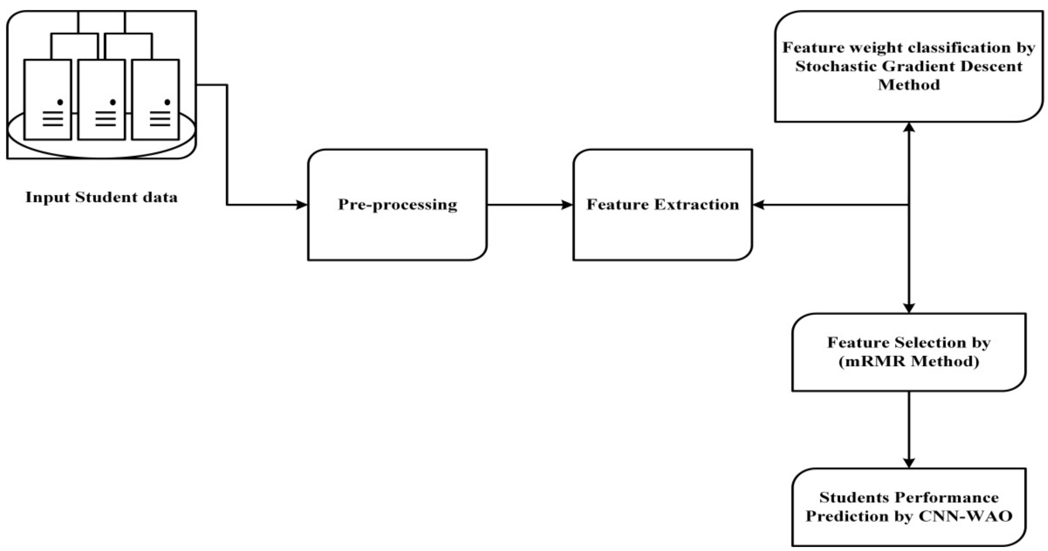

Fig. 3: Workflow of the Proposed System.

3.1 Database Collection

Information from the Kaggle repository was used to analyze the progress of students. A database is divided in two halves with a proportion of 70:30 for the purpose of assigning testing and training samples.

3.2 Pre-processing

Information preparation is necessary to create an efficient deep-learning forecasting model because the effectiveness of classifiers is slowed down by original information. The student performance prediction analysis database is presented in a skewed version with lots of redundant information. The duplicate information must be eliminated to have a clear input for the deep learning model. This system carried out the subsequent pre-processing stages:

- The gathered information is cleaned of distortion, and duplicate data is eliminated.
- The category parameters are converted to a numerical format, and some qualities have multiple entries for a single parameter.

3.3 Feature Extraction

Convolutional Neural Network determines the periodicity levels by extracting the characteristics from the student's already examined database and determining the attention rates for every subject. Students' attention and persistent traits for various motifs are estimated and categorized after extracting the characteristics.

Algorithm: Extraction of Features depending on Student's Attention
Input: Students report S_r
Output: Functional Load f_l
Step 1: Assessment of students' interests s_a Accessibility to the computation frequency characteristic $$C_f = \frac{\sum s_a .non\ like == Feature\ assessment}{Overall\ characteristic\ access + dterm\ repeated\ query}$$
Step 2: Determine the accessing rating for the learner $$L_a = \frac{\sum s_a.\ service == features}{overall\ characteristic\ access} \times s_a(generated\ accessibility\ weighting)$$
Step 3: Compute relevance weightage score $$if\ s_a > C_f - Ranking\ table\ follows(service + performance\ of\ the\ students)$$ $$Service\ interest\ rate\{C_{f1}, C_{f2}, \dots \dots\}$$
Step 4: Feature weight-based f_l students recognized ranking list is returned. According to various performing records, the aforementioned system assesses every student's attention characteristic. Analysis and characteristic extraction are done on the students' performances and behavior characteristics.

3.4 Feature Selection

By choosing the most advantageous characteristics to be evaluated, the Maximum Relevance — Minimum Redundancy algorithm lowers computation expenses. Those attributes are selected on the basis of their resemblance in Eq. 1, which would be dependent on the greatest correlations of the parameters.

$$G(p,q) = \sum_{P\in p} \sum_{Q\in q} (P,Q) log\left(\frac{u(P,Q)}{u_1(P)u_2(Q)}\right) \tag{1}$$

The fundamental objective of the Maximum Relevance — Minimum Redundancy approach is to choose the best features and weaken the undesirable characteristics. This approach analyses every characteristic independently from the database and makes utilization of the local data among them to calculate the degree of similarities among the 2 characteristics, P and Q, employing the $G(p,q)$ statistic. The marginal distribution functions for a randomized variables are determined by adding the P and Q probabilities.

Every probabilistic function has a k-size vectors definition$\left(r_j = \left[r_j^1, r_j^2, r_j^3, \dots \dots, r_j^n\right]\right)$. x_{st} stands for the selectable characteristics, while K stands for the classes' identities. To get the best characteristics, Eqs. (2) and (3) must be met.

$$Max\ A, A = \frac{1}{|x_{st}|} \sum_{r_j \in x_{st}} G(r_j, K) \tag{2}$$

Whereas x_{st} sets the highest fitting parameters on the basis of which the undesired factors are eliminated, r_j represents the similarity measurement among the parameters. The large percentage of dependent parameters and duplication between the factors are reduced using Eq. (3). The greatest relevancy $(Max\ A, A)\ and\ (Min\ B, B)$ the minimal redundancies are determined by Eqs. (2) and (3). Equations (4) and (5) are used to do the optimisation simultaneously.

$$Min(B,B) = \frac{1}{|x_{st}|^2} \sum_{r_j, r_i \in x_{st}} G(r_j, r_i) \tag{3}$$

$$Max(A,B) = A - B \tag{4}$$

$$Max(A,B) = \frac{A}{B} \tag{5}$$

Until the Convolutional Neural Network learning procedure is optimized, the Stochastic gradient descent approach adjusts the features weighting at each iteration. The best characteristics are provided for Maximum Relevance — Minimum Redundancy to choose from via Convolutional Neural Network, whose learning coefficients rates increases, and thus loops the procedure until the optimum characteristics are chosen from the classification model.

3.5 CNN technique utilizing deep learning for assessing performance of students

Employing deep learning, in order to forecast student progress, Convolutional Neural Network was implemented. Since Convolutional Neural Network trains on hierarchical representations, deep learning boosts recognition effectiveness. Since it does not require any prior knowledge, this could feature extraction. As a result of its benefits, an automatic assessment system has been created to enable non-specialists to design their subsequent Convolutional Neural

Network architecture without any existing understanding of this topic. The supervised learning architecture consists of a variety of hidden units and permits the generalization of characteristics by having a number of stages. Convolution, subsampling, and fully connected layers are three of the deep learning network's hidden layers, which are extensively discussed in the following subcategories. Here, the hidden layer's convolutional and subsampling layers both aid in the feature extraction process from low to extremely high levels of inputs information characteristics.

3.5.1 Convolution layer

A filtering or kernels of dimension $c \times c$ is initially used to convolve an input data with a size of $u \times v$. Every block in the input sequence is combined with the filters separately, as well as the result of this phase is a brand-new data. The output results created by the convolution of the input data and the filters utilized in this phrase typically has characteristics. In generally, the inputs to the size of the characteristic mapping $i \times i$ are expressed as the outputs data features obtained by the kernels of the convolving layers. The Convolutional Neural Network architecture is made up of numerous convolutional layers, with the succeeding convolutional layer's inputs and outputs serving as its classification model. A filter called an R filter is combined with the inputs within every convolution operation. Additionally, the number of feature maps used in the convolution technique is equal to the thickness of the convolution layer that was formed. Every one of the characteristic mappings is identified as a particular characteristic of the input data at a specified position. As a result, $C_{i(k)}$, which includes an extracted feature, is the outcome for the z-th phase in the convolution layer. It is statistically expressed as follows:

$$C_{i(k)} = A_{i(k)} \sum_{j=1}^{a_{i(k-1)}} M_{i,j}^{(k-1)} \times C_{i(k)} \tag{6}$$

Where $A_{i(k)}$ illustrates a bias matrix and $M_{i,j}^{(k-1)}$ is the convolutional kernels or filter size of the $c \times c$ that connects the thinner $(k-1)$ layer extracted features with the corresponding layer in the $k-th$ data mapping. Outcome $C_{i(k)}$ the characteristic mappings are included in the layers. The inputs region is the convolution layers during initial $C_{i(k)}$ of Eq. (6). The characteristic mapping is obtained by the kernel that is displayed in the networking framework. The activation functions are used to change the outcome convolution layer nonlinearly after the convolution operation.

$$Y_{i(k)} = Y(C_{i(k)}) \tag{7}$$

The inputs region is the convolutional layers during initial ($C_{i(k)}$ of Eq. (7). The characteristic mapping is obtained by the kernels that are displayed in the networking framework. The activating functions $Y_{i(k)}$ are used to change the outcome convolutional layers nonlinearly after the convolutional operation.

$$Y_{i(k)} = maxi(0, Y_{i(k)}) \tag{8}$$

The interactions and non-linear impacts that the model of deep learning must adjust to are reduced by Eq. 8. Rectified linear activation functions have a replacements function that, in the event of a negative outcome result, substitutes it with zero while simultaneously preceding a positive outcome significance. Owing to its erroneous variance, the activating functional executes training more quickly than the other processes. It slows down in the saturated area, having an impact on disappearing gradients issues as the weights upgrades eventually completely disappear.

3.5.2 Sub-sampling layer

The main goal of this layers is to reduce the spatially complexity of the data map that was retrieved from the convolutional layers before it. A filter of size $s \times s$ is chosen for this task; the sub-sampling processing is done among the characteristic mappings and masks. Aggregate sampling could be done in a number of ways, including sum pooled, averaged pooling, and max pooling. The most effective of them is maximum pooling, which uses the highest benefit of every block as the output data. This layer aids in tolerating the convolutional layer's effectiveness when the input data are rotated and translated.

3.5.3 Fully connected Layer

The categorization of student's performance is carried out in this layer depending on the retrieved characteristic from the previous convolutional layers. Here, the standard feed-forward neural network with one or even more hidden units is used to interconnect every layer with the preceding stage of each and every neuron. In this output units, the soft max layer mechanism is applied.

$$y_{i(k)} = FC(k_{i(k)}) \tag{9}$$

Where:

$$k_{i(k)} = \sum_{i=1}^{m_{i(k-1)}} r_{i,j(k)} y_{i(k-1)} \tag{10}$$

Thus, $r_{i,j(k)}$ the fully connected layers tune the loads to create a representations of the students performance records, and the transferring characteristic is designated as FC, which stands for non-linearity. The output results are then used to designate the categories. The key strategy used in the suggested study to enhance the effectiveness of CNN architecture utilizing deep learning is parameterization tuning. The CNN architecture is depicted in Fig. 4. Consequently, a unique Whale Optimization Algorithm that is based on the operation of the whale and its method of organizing feed has been developed in order to optimize the convolutional neural network parameters. Additionally, these training variables that result to the lightest convolutional neural network dependent on WOA features and display the framework developed for histological performance classification are modified in order to produce the highest efficiency measurement for convolutional neural network. The following sections have addressed parameters optimization using Whale Optimization Algorithm.

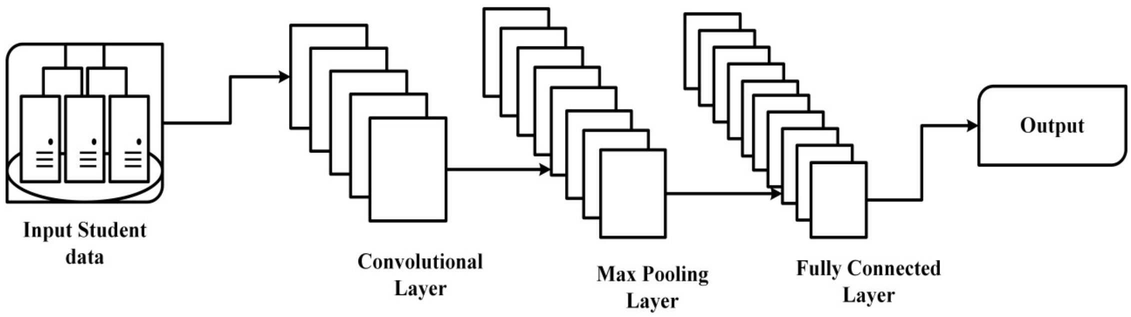

Fig. 4: Workflow of CNN.

4 Factors to optimize

The suggested deep learning-based WOA accelerates the training process by carefully choosing the variable values for students' performance. The humpback whales use three operations to pursue their predators: looking for it, encircling it, and creating a bubbles trap to catch it. This is the basis of the whale optimization method. The main steps of the whale optimization approach are described in great detail in Fig. 3.

4.1 Mathematical modelling

This segment illustrates the mathematical depiction of enveloping the food, spiraling bubble-net feeding operations, and hunting for predators.

4.2 Deployment

The first result is created at randomness during the initialization stage of the suggested procedure. For example, using the suggested optimization technique, the optimum performance created by the CNN variables is chosen for the histopathology data of student's records following pre-processing. The amount of feature mappings, padding, kind of pooling, number of kernels, and number of whales—or, rather, the community of whales—are all initialized at randomly in this case. Consequently, the following is how the randomized variable in the searching area is depicted:

$$W(h) = (w_1, w_2, \dots\dots, w_u) \tag{11}$$

Thus, $\boldsymbol{W}$ designates the whale's initial density, while $\boldsymbol{u}$ stands for the amount of connectivity layers to be optimised.

1.3 Fitness calculation

For automatic student progress predictions, the fitness functions are created to maximize reliability and produce the best categorization measurement. It is then assessed using the following expression:

$$\boldsymbol{F_{i(W(h))}} = \boldsymbol{Maximum(Accuracy)} \tag{12}$$

4.4 Reposition the present approach to encircling the target

In this stage, whales begin to attack after observing their prey's location, and they subsequently circle their target. Then, the whale regarded to be the best whale is discovered as the greatest option. The other whales would continue towards such a best whale following learning its location. The equations below show how whales update their information:

$$\vec{P} = \left|\vec{G}.\overrightarrow{W_{best}}\,(h) - \overrightarrow{W(h)}\right| \tag{13}$$

$$\overrightarrow{w(h+1)} = \overrightarrow{W_{best}}\,(h) - \vec{A}.\vec{P} \tag{14}$$

Here,h stands for the most recent rendition., $\overrightarrow{W_{best}}$ determines the ideal response, $\overrightarrow{W(h)}$ relates to the present situation, $\vec{A}$ and $\vec{G}$ indicates a vector of coefficients, $|A \times P|$ the exact location is indicated. The coefficients matrices are additionally expressed numerically as follows:

$$\vec{A} = 2\vec{A}.\vec{q} - \vec{A} \tag{15}$$

$$\vec{G}= 2.\ \vec{q} \tag{16}$$

Where $\vec{A}$ is a list of linear recurrence from $2\ to\ 0$, $\vec{q} \in (0,1)$ for both the periods of extraction and research.

4.5 Stage of Exploitation

This stage is often referred to as the bubble-net offensive strategy. There are two methods have emerged:

- Shrinking encircling mechanism and its $\vec{A}$ Quantity is reduced in order to achieve this efficiency. Here, $\vec{A}$ is employed for the lowering of a variety of $\vec{A}$. In the absence of that, it is asserted that in the frequencies between $[-A, A]$, $\vec{A}$ is a position wherein a is accidentally lowered from $2\ to\ 0$. The position of the discovering agent may change for any number of reasons for $\vec{A} \in [-1,1]$ is mathematically given by the following equation:
$$\vec{A} = 2\vec{A}.\vec{q} - \vec{A} \tag{17}$$

- Spiral updating location is determined by calculating the relationship among the location of the whale and its food as described in the following:

$$\vec{W}(h+1) = \overrightarrow{P_{dist}}.Exp_{ctd}.\cos(2\pi d) + \overrightarrow{W_{best}}\,(h) \tag{18}$$

Where $\vec{P} = \left|\vec{G}.\overrightarrow{W_{best}}\,(h) - \overrightarrow{W(h)}\right|$ the optimal idea found so far in this, indicated as the range between the y-th whale and its victim, is designed to take values from $[-1, 1]$, m is depicted as the form of the exponential spirals. When undertaking optimisation, the position of the whale with a likelihood of 70percent of the overall should be determined by choosing any diminishing or spiraling encircling models, and then the corresponding mathematical expression should be used:

$$\vec{W}(h+1) = \begin{cases} \overrightarrow{W_{best}}(h) - \vec{A}.\vec{P}, if\ Q < 0.5 \\ \overrightarrow{P_{dist}}.Exp_{ctd}.\cos(2\pi d) + \overrightarrow{W_{best}}\,(h), if\ Q \geq 0.5 \end{cases} \tag{19}$$

Where $Q \in [0,1]$ The humpback whales create bubble netting by randomly selecting their food.

4.6 Exploration Stage

Hunting the victim is another name for this stage. The next equation describes the statistical structure of the exploratory step in more detail.

$$\vec{P} = \left|\vec{G}.\overrightarrow{W_{random}} - \vec{W}\right| \tag{20}$$

$$\overrightarrow{w(h+1)} = \left|\overrightarrow{W_{random}} - \vec{A}.\vec{P}\right|$$

The following is a representation of the currently population's randomized locations $\overrightarrow{W_{random}}$. The fitness computation is assessed during the improve phase of every option to determine which is the best one available. A series of fresh ideas are discovered based on the optimal results produced, and the fitness factor is computed to continue the procedure of upgrading the optimal answer described above.

4.7 Conditions for termination

Due to the whale's hunting habit, it meets all of CNN's strict criteria. The predictive algorithm is validated as a consequence of locating the optimum fitness functional or optimization method. The predictions system generated for the optimum performance framework is highly suited to estimate unseen information because the target functionality is to enhance the precision of training information.

4.8 Evaluation stage

Following the completion of the training procedure, the suggested models were put to the trial utilizing a selection of data in the testing stage. Precise, recall, F1-score, and reliability have all been calculated for those data for each of the categories. The aggregate precision is calculated by adding together all of those measurements. The most intriguing aspect of the suggested technique is that the Convolutional Neural Network structure extracts a document's characteristics directly, allowing the network to identify particular characteristics inside the logs and recognize them wherever in the data. It until data is processed, the procedures would be performed. 70 percent of total of the information is evaluated during the assessment stage in order to gauge how well Convolutional Neural Network performs. For instance, in this project, student logs are examined, and progress is assessed based on the percentage of accurately projected records.

5 Results and Discussions

By running a number of tests on the Kaggle students' progress assessment database, the suggested CNN-WOA approach is examined. In order to predict students' progress across several classes using CNN-WOA, this method incorporates Python. The Kaggle database is divided into training and assessment sets, with the training dataset making up 70percent of total of the information. To assess its effectiveness, which is reported in table 1, the quantity of the trainable variables is employed. The features were constructed for the simulation platform with graphics processing unit (GPU) capabilities, and Version of windows 64-bit was utilized to run the operating systems. Performance measures for the algorithms include Accuracy, Recall, Precision, and F1-score. The previous techniques SPDN, BiLSTM, and SVM, which are shown in Fig. 5, are significantly less accurate than the suggested CNN and WOA, which yields 98.5percentage effectiveness.

Table 1. Evaluating the proposed CNN- WOA performance

	Proposed CNN-WOA	Bi-LSTM	SVM	SPDN
Accuracy	98.66	89.67	94.77	71.55
Precision	97.80	90.00	90.33	73.66
Recall	93.44	87.99	84.66	70.77
F1-score	96.34	91.33	78.77	68.67

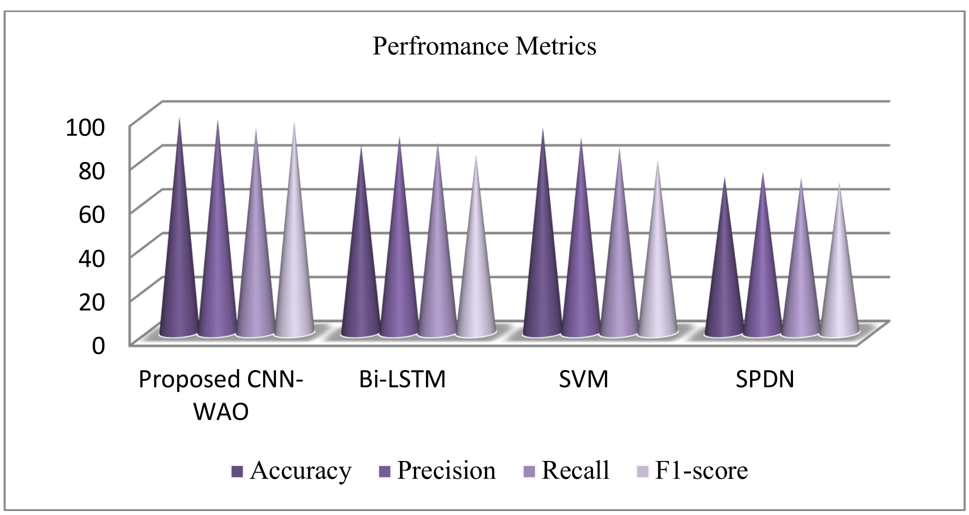

Fig. 5: Performance Metrics.

The most popular assessment measures, including precision, recall, F-measure, and accuracy, are used to analyse the system 's effectiveness. The following are some of the evaluation metrics taken into account for the suggested system's assessment process:

- Precision

The suggested system's efficiency can be determined with precision. The proportion of appropriate photographs to total information that were found is known as the precision proportion.

$$P_r = \frac{T^P}{T^P + F^P} \tag{21}$$

- Recall

The recall is calculated as the proportion of retrieved significant data to all of the database's relevance student's record.

$$R_c = \frac{T^P}{T^P + F^N} \tag{22}$$

- F1- Measure

The F1-score, which assesses a test's correctness, could be calculated as follows:

$$F1\ score = 2 * {}^{P_r * R_c}/_{P_r + R_c} \tag{23}$$

- Accuracy

The percentage of accurately classified information occurrences over all data instances is known as accuracy.

$$Accuracy = \frac{T^N + T^P}{T^N + F^P + T^P + F^N} \tag{24}$$

6 Conclusions

The massive amounts information kept in the ecosystems of academic records and digital learning database make it increasingly hard to anticipate students' progress on an online system. There are numerous methods available for assessing academic performance. Deep learning, also known as learning analytics is the method that is most frequently used to assess students' progress. It is a growing field of research that concentrates on several data mining approaches like categorization, predictions, and extraction of features. In this article, suggest the CNN-WOA approach for forecasting student performance and behavior in students. Nearly each element of the social and economic lives has undergone significant upheaval as a result of COVID-19 and the ensuing regulatory changes. Governments have used lockdowns and social segregation policies to reduce human movement and the transmission of illness. Many schools switched from using physical lectures to online platforms as a result of the severe interruptions to teaching caused by the COVID-19 issue in order to reduce the educational losses brought on by the disturbances. These interruptions have a significant negative impact on students' desire to learn and lead to a decline in educational outcomes. Predicting a person's interest level and improving the efficiency of low-ranking pupils can help tackle this problem. The student record is processed using the CNN-WOA based extraction of features, which extracts useful features. The finest characteristics are then chosen using the mRMR featured choice method, which eliminates the fewest characteristics. For categorization, the characteristics are sent to CNN-WOA, which divides the outcome into low, moderate, and excellent performance measures based on the student's log. The outcome of this prediction helps several educational institutions and employees to recognize underperforming kids.

References

[1] J. L. Rastrollo-Guerrero, J. A. Gómez-Pulido, and A. Durán-Domínguez, Analyzing and predicting students' performance by means of machine learning: A review, *Applied sciences.*, **10(3),** 1042, 2020.

[2] A. A. Mubarak, H. Cao, and W. Zhang, Prediction of students' early dropout based on their interaction logs in online learning environment, *Interactive Learning Environments.*, **30(8),** 1414–1433, 2022.

[3] S. Hussain, Z. F. Muhsion, Y. K. Salal, P. Theodoru, F. Kurtoğlu, and G. C. Hazarika, Prediction model on student performance based on internal assessment using deep learning, *International Journal of Emerging Technologies in Learning*, **14(8),** 2019.

[4] V. Vijayalakshmi and K. Venkatachalapathy, Comparison of predicting student's performance using machine learning algorithms, *International Journal of Intelligent Systems and Applications.*, **11(12),** 34–45, 2019.

[5] M. Elloumi, M. A. Ahmad, A. H. Samak, A. M. Al-Sharafi, D. Kihara, and A. I. Taloba, Error correction algorithms in non-null aspheric testing next generation sequencing data, *Alexandria Engineering Journal.*, **61(12),** 9819-9829, 2022.

[6] P. Chakraborty, P. Mittal, M. S. Gupta, S. Yadav, and A. Arora, Opinion of students on online education during the COVID-19 pandemic, *Human Behavior and Emerging Technologies.*, **3(3)**, 357-365, 2021.

[7] M. Hussain, W. Zhu, W. Zhang, S. M. R. Abidi, and S. Ali, Using machine learning to predict student difficulties from learning session data, *Artificial Intelligence Review.*, **52(1)**, 381-407, 2019.

[8] M. Tsiakmaki, G. Kostopoulos, S. Kotsiantis, and O. Ragos, Transfer learning from deep neural networks for predicting student performance, *Applied Sciences.*, **10(6)**, 2145, 2020.

[9] A. A. Mubarak, H. Cao, and S. A. M. Ahmed, Predictive learning analytics using deep learning model in MOOCs' courses videos, *Education and Information Technologies.*, **26(1)**, 371-392, 2021.

[10] A. A. Sewisy, M. H. Marghny, and A. I. Taloba, Fast Efficient Clustering Algorithm for Balanced Data, *International Journal of Advanced Computer Science and Applications.*, **5(6)**, 123-129, 2014.

[11] Research Scholar, Computer Science and Engineering, Annamalai University, Tamilnadu, INDIA, V. Vijayalakshmi, and K. Venkatachalapathy, Comparison of predicting student's performance using machine learning algorithms, *International Journal of Intelligent Systems and Applications.*, **11(12)**, 34-45, 2019.

[12] M. Hussain, W. Zhu, W. Zhang, S. M. R. Abidi, and S. Ali, Using machine learning to predict student difficulties from learning session data, *Artificial Intelligence Review.*, **52(1)**, 381-407,2019.

[13] S. S. Ismail, R. F. Mansour, R. M. Abd ElAziz and A. I. Taloba, Efficient E-Mail Spam Detection Strategy Using Genetic Decision Tree Processing with NLP Features, *Computational Intelligence and Neuroscience.*, 2022.

[14] A. Elhadad, F. Alanazi, A. I. Taloba, and A. Abozeid, Fog Computing Service in the Healthcare Monitoring System for Managing the Real-Time Notification, *Journal of Healthcare Engineering.*, 2022.

[15] A. I. Taloba, R. M. Abd El-Aziz, H. M. Alshanbari and A. A. H. El-Bagoury, Estimation and Prediction of Hospitalization and Medical Care Costs Using Regression in Machine Learning, *Journal of Healthcare Engineering.*, 2022.

[16] J. Sultana, N. Sultana, K. Yadav, and F. AlFayez, *Prediction of sentiment analysis on educational data based on deep learning approach*, In 2018 21st Saudi computer society national computer conference., 1-5, 2018.