J. Electrical Systems 20-7s (2024): 3213-3226

Zhilin Sun[1*]

Learning Behavior Analysis of College English Learners Based on Data Mining Technology

Journal of Electrical Systems

Abstract: **-** Learning English is crucial as global economic integration develops since it would facilitate greater communication and collaboration with other countries. Data mining must be used in online English education given the diverse range of English learning techniques available in today's modern world. Behavior analysis has several limitations, including the possibility for simplifying or generalization that might ignore unique learning preferences or cultural quirks. In this manuscript, Learning Behavior Analysis of College English Learners Based on Data Mining Technology (LBA-CEL-DMT-DTRSRN) is proposed. Initially the data is collected from Global Language Learning Popularity Dataset. Then the data is fed into pre-processing utilizing Non-Integer Order Generalized Filters (NIOF). The NIOF is used for data cleaning. Then the pre-processed data are given to Double Transformer Residual Super-Resolution Network (DTRSRN) for predict the Error of Students Learning Behavior. In general, DTRSRN does not express adapting optimization strategies to determine optimal parameters. In order to efficiently optimize DTRSRN and precisely anticipate the error of students' learning behavior, the Multiplayer Battle Game-Inspired Optimizer (MBGIO) was developed. The proposed (LBA-CEL-DMT-DTRSRN) approach is implemented in Python. The performance of proposed method examined utilizing performance metrics like Accuracy, Precision, Recall, F1-Score and ROC. In comparison to existing methods, such as Analysis of Students' Behavior in English Online Education Based on Data Mining (SB-EOEDM-FNN), IoT-enabled Personalized English Learning in Colleges using Big Data Analysis and Decision Support System (IOT-PELC-BDSVM), and College English Flipped Classroom Teaching Model Based on Big Data and Deep Neural Networks (CEFCT-BD-DNN), the proposed (LBA-CEL-DMT-DTRSRN) approach has 28.01%, 25.29%, and 21.05% higher accuracy, 26.35%, 21.05%, and 28.45% higher precision, and 23.78%, 26.54%, and 25.14% higher recall.

Keywords: Data Mining, Double Transformer Residual Super-Resolution Network, English, Global Language Learning Popularity Dataset, Multiplayer Battle Game-Inspired Optimizer, Non-Integer Order Generalized Filters, Online Education.

I. INTRODUCTION

The emergence of information and communication technology has ushered in a new era of personal knowledge. A major shift in people's lives, education, and cognitive processes has resulted from the growing use of computers and other electronic devices coupled with the quick expansion of computer networks (CN) [1-3]. The extensive use of computers and the Internet in education has led to the rapid emergence of new forms as well as the major modification of traditional teaching methods, competencies, and ideas. The primary means of acquiring information is now the Internet. The key issue that has to be resolved is how to leverage the online learning platform to enhance the teaching process. Making use of data mining and the Internet, creating a model connected to education may actively assist in platform decision-making [4-6]. Developing a platform for remote learning that provides a better learning experience, carrying out more accurate and efficient assessments of E-learning, and providing suggestions for learning English all depend on the analysis of E-learning behavior in online education [7-9]. For the user, session, and event datasets that were part of the "National Library Open Course" data collection, assessments were made of correlation, event rating, and network social networking. Using a mix of information gain (rate) analysis and correlation analysis, eight effective characteristic variables were chosen to generate the final feature set. Based on the findings, the research combines and assesses the information from the perspectives of effective learning behavior traits and markers, which might facilitate learning assessment and analysis in hybrid learning environments [10-12].The English online learning system included a variety of data mining techniques, which presented challenges for the field of data mining in education. The preparation stages were examined, and the data source was looked into. Also looked at were the benefits and drawbacks of various data mining methods [13, 14]. Some academics have also developed an online learning behavior analysis platform, a model of students' online learning behavior, and a big data-based approach to online learning behavior analysis. K-means clustering algorithm [15-17].Large prediction error, increased memory requirements, and low processing efficiency are some of the issues that still need to be

[1] [1] Lanzhou City University, Lanzhou, Gansu, 730070, China
*Corresponding author e-mail: sunzhilin@lzcu.edu.cn

resolved in order for the aforementioned approaches to be used practically. The approach to data mining behavior analysis of online English learning behavior is suggested by this research [18, 19].

While there are potential advantages to using data mining and internet platforms for education, there are also some disadvantages. First of all, it ignores the digital gap that denies many students their rights, especially those who come from underprivileged neighbourhoods, by assuming that everyone has access to technology and internet connectivity. Furthermore, the focus on online learning environments may override the value of in-person interactions and experiential learning, which are crucial for a number of courses and competencies. Concerns regarding data security and privacy are also raised by the use of data mining in decision-making, particularly when handling sensitive student data. Furthermore, the emphasis on uniform measures and algorithms might oversimplify the intricacies of customized learning requirements and learning behavior. Lastly, the statement ignores socioeconomic and cultural variables that affect educational achievements, which might skew the research and suggestions made using data mining tools.

It is important to develop a balanced strategy that prioritizes conventional techniques while utilizing technology to augment learning experiences in order to effectively tackle the issues presented by the increasing prevalence of online learning platforms in education. This means giving ethical and data privacy concerns first priority while maintaining the caliber and accessibility of online resources. Focus should be placed on closing the digital gap, giving teachers the assistance and training they need, and regularly assessing and enhancing online learning environments. Furthermore, it is imperative to acknowledge and accommodate cultural and linguistic variety, especially in English language learning environments, through the integration of culturally appropriate materials and the provision of multilingual education.By considering these aspects and addressing the pedagogical, technological, and economical barriers to online learning, stakeholders may create a more complete, efficient, and sustainable education system that meets the diverse needs of students in the digital age.

Major contribution of this paper a follows;

- In this research, LBA-CEL-DMT-DTRSRN is proposed.
- Initially the input data is collected from the Global Language Learning Popularity Dataset.
- The proposed LBA-CEL-DMT-DTRSRN method integrates multiple advanced techniques, including Non-Integer Order Generalized Filters (NIOF) for pre-processing.
- The pre-processed data undergoes prediction using Double Transformer Residual Super-Resolution Network (DTRSRN).
- Unlike traditional DTRSRN approaches, which lack optimization methods for computing optimal parameters, the proposed method incorporates Multiplayer Battle Game-Inspired Optimizer (MBGIO). MBGIO optimizes the weight parameters of DTRSRN.
- The efficiency of the suggested model is analysed with the existing methods like SB-EOEDM-FNN, IOT-PELC-BDSVM and CEFCT-BD-DNN models respectively.

Remaining portion of this work structured follows: Part 2: literature survey, Part3: describes suggested methodology, Part 4: illustrates results and discussion and Part 5: conclusion.

II. LITERATURE SURVEY

A few recent manuscript that address deep learning-based learning behavior analysis of college English learners using data mining technology are included below. A number of other works were also proposed in the literature.

Wang [20] has presented, Evaluation of SB-EOEDM-FNN. Learning English vital given the emergence of global economic integration and the need for improved communication and collaboration with countries worldwide. Data mining has to be used in online English education given the diverse range of English learning techniques available in this modern day. Data mining techniques were becoming more and more common in education due to their ongoing application and the advancement of online learning platforms. Traditional approaches suffer from low processing efficiency, high prediction error, and higher memory needs when faced with vast amounts of learning and student behavior data.

Wu [21] has presented, Big Data analysis and decision support systems(DSS) are used in IoT-enabled personalized English learning in colleges. The advancements in computer and Internet of Things (IoT) technology have accelerated the integration of these technologies into education in recent years. Conversely, this project aims to create a customized, Internet of Things-enabled collegiate English language learning environment through the use of big data analysis and a data mining algorithm-based DSS. Through the analysis of given data, educational analytics produces valuable learning outcomes that improve a multitude of individualized learning experiences in Chinese, French, and English.

Chang [22] has presented, A Tossed Classroom Education Model for College English Utilizing Deep Neural Networks and Big Data (DNN). In several educational sectors, the flipped classroom—a novel sort of mixed teaching style that depends on computer technology—has transformed traditional instruction and established a "learning first and teaching later" method. The quick advancement of information technology was the reason of this. In flipped classrooms, the knowledge transmission and internalization phases of traditional education were inverted, increasing student autonomy. However, studies examining the specific impacts of the tossed classroom model on college students' ability to study English independently were still in their infancy.

Wu [23] has presented, A mobile edge computing (MEC) environment's smart classroom learning model for English in college depends on data mining technology. Students' capacity to enhance their English abilities was impacted by a variety of challenges that persist in college English classrooms, such as the relatively confined learning environment and the lack of creativity and change in certain professors. Industry managers and decision-makers, educational academics, and practitioners of industry and application have all become more interested in the growth and use of big data technologies in education.

Chen et al. [24] have suggested,This study examines the requirements of users while downloading English vocabulary applications using data mining for online comments. The advent of sophisticated social media platforms has prompted the development of fresh mobile educational materials.such as English learning applications. After users have used these applications, online comments have become a valuable source of intellectual competition for businesses. These remarks may also assist companies in addressing problems with product creation and iteration as well as improving their understanding of their target market. Because of this, in order to gather user preferences, this work used hotspot mining and emotional evaluation to crawl the online user comments of three well-known APPs. The consumers' demands were then assessed using the K-means grouping method.

Feng et al. [25] have suggested Students' academic performance is analyzed and predicted via educational data mining. The development of intelligent technology was becoming increasingly widespread in the field of education. The rate at which educational data was expanding raises the possibility that traditional processing methods were flawed or had their limitations. To avoid unfavorable evaluation results and to keep an eye on students' future performance, The present research evaluates and predicts pupil academic achievement by heavily utilizing the pertinent hypotheses of convolution neural networks, discriminating neural networks, and grouping networks. First, this study suggests using a statistic never utilized in the K-means technique to maximize the clustering-number calculation.

Chen et al. [26] have suggested, Study on Information Mining for The combination Model Analysis and Performance Prediction Based on Students’ Behavior Attributes. It was crucial to raise the standard of school information management by using data mining technologies to evaluate student behavior in order to anticipate performance and other assessment metrics. A data mining approach based on student behavior has been established in an attempt to address the issues of inadequate management of information platforms and limited data analysis skills in colleges and universities. The GBDT, ANN, and K-means algorithms' properties were first examined, and then the three methods were integrated to create a composite prediction model. Five common data sets were merged at the same time to train the simulation. On the basis of the behavioural traits of the pupils, an analysis and prediction platform was then constructed. The assessment index system of students' behavior was built, and data collecting, modeling, analysis, and mining were achieved in conjunction with educational administration, the library, and management systems like "Campus All-in-one Card."

III. PROPOSED METHODOLOGY

In this section, LBA-CEL-DMT-DTRSRN is proposed. This process consists of three steps: Data Acquisition, pre-processing and prediction. In this proposed Learning Behavior Analysis of College English Learners undergo pre-processing to prepare them for further analysis. MBGIO involved optimizing hyper parameters by dynamically adjusting particle position within a predefined search space based on their performance and solutions. Then the prediction performance was evaluated on an independent testing dataset, aiming for both higher accuracy and overall model improvement. The final step involves employing a DTRSRN for predict the error of students learning behavior. Multiplayer Battle Game-Inspired Optimizer method was introduced for optimize the weight parameter of DTRSRN. By using data mining to analyze the behavior of college English learners, teachers may get insights into learning patterns and develop more effective teaching tactics. It identifies areas of struggle, facilitates personalized learning paths, enhances student engagement, and improves overall academic performance. The block diagram of proposed LBA-CEL-DMT-DTRSRN approach was represented in Fig 1. As a result, a thorough explanation of each step is provided below.

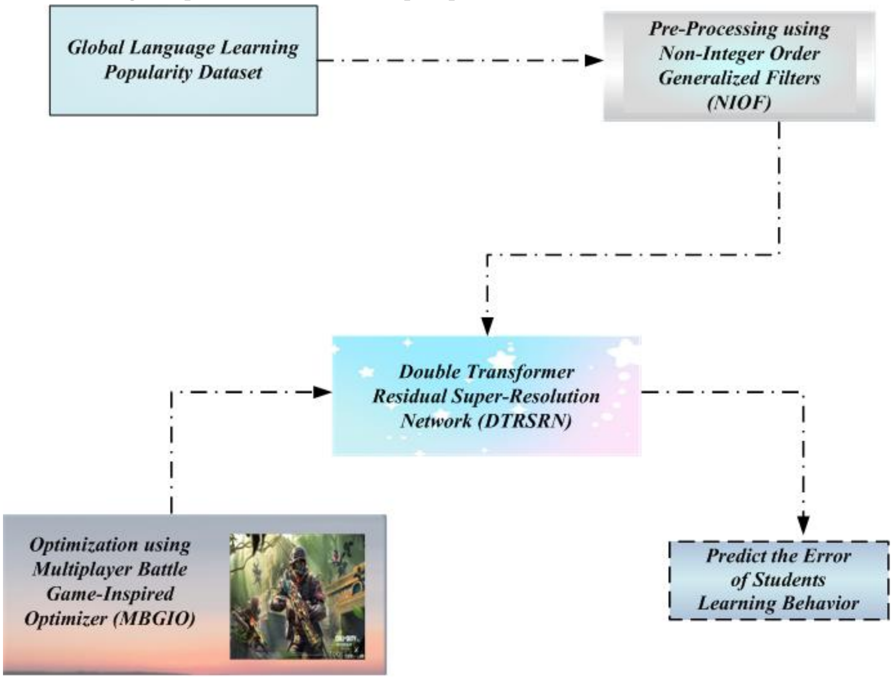

Figure 1: Block Diagram for Proposed LBA-CEL-DMT-DTRSRN

A. Data Acquisition

Information on language learners who used Duo lingo is available in the Duo lingo Language Report. To protect learner privacy, the data was combined based on language or nation. Country aggregations are predicated on the autonomous, self-governing, and globally acknowledged entities listed below. Learners under the age of 13 were not included in any analysis, and self-reported age and motivation data were used. Additionally, in order to preserve the privacy of learners, nations with fewer than 5,000 Duo lingo users are not included in the rankings [27]. With the aid of this time-series CSV file, investigate the changes in language popularity among nations! For the purpose of monitoring patterns, examining changes, and advancing research, it presents the two most often spoken languages in each nation over a period of years. Go deep into the data to learn about the intriguing dynamics of global language change.

B. Pre-processing using Non-Integer Order Generalized Filters (NIOF)

In this section, NIOF [28] is discussed. The input data is given to pre-processing. In pre-processing NIOF is used for data cleaning. When employing data mining technologies to analyze the learning behaviours of college-level English language learners, non-integer order generalized filters provide a number of benefits. First of all, by capturing fractional-order dynamics, they offer a more nuanced picture of complicated learning behaviours, enabling a more in-depth examination of how students interact with the learning resources. Teachers may now more clearly see subtle trends in student involvement, understanding, and development thanks to this increased granularity. Non-integer order filters are more adaptable for simulating a wide range of learning behaviours and can better handle time-varying and non-linear dynamics. Because of its adaptability, predictive models perform better, which helps teachers anticipate students' needs and design interventions that meet those requirements. Providing enhanced precision, adaptability, and effectiveness in data-driven educational research and practice, non-integer order generalized filters are a potent instrument for developing learning behavior analysis in college English teaching.

$$I_{BP,CO}(m) = E_0\left[\frac{1}{(\tau r)^{\alpha}+1}\right]^{\gamma} \tag{1}$$

Where α s stands for the integer order filter, $I_{BP,CO}$ stands for the transfer function, E_0 is a constant, and τ for the time constant. The pole frequency is m, while the laplacian operator is r. High-frequency components can get through the filter while low-frequency components are reduced .Thus it is given equation (2).

$$I'_{BP,CO}(m) = E_0 \frac{(\tau r)^{\alpha}}{(\tau r)^{\alpha} + 1} \tag{2}$$

where m is the knee frequency, τr is the slope transition value, and $I'_{BP,CO}$ represent the transfer function of the associated inverse filter, The integer order filter is called α. Slopes that are the reverse of those seen in low-pass and high-pass filters are produced by the inverse gain responses that are connected with them. Thus, it is given equation (3)

$$G_{BP,PL}(r) = E_0 \left(\frac{\tau r}{\tau r + 1} \right)^{\gamma} \tag{3}$$

The transfer functions for the power-law inverse are represented by G, the phase angle range is represented by γ, the scaled form of the pole frequency is represented by τ, and the gain response expressions are represented by E_o. The slope of the transition between the filter's two characteristic bands and the range that defines its phase response are also, the transfer function of a non-integer order high-pass filter is produced. The data is cleaned by this equation (4).

$$\left| G_{BP,PL}(\omega) \right| = E_0 \frac{(\omega \tau)^{\gamma}}{\left[1 + (\omega \tau)^3 \right]^{\gamma/2}} \tag{4}$$

Where, E_0 is represent the maximum gain and factorial-order filter; $\omega\tau$ is represent the time frequency; ω represents the asymmetric band-pass filter, and $G_{BP,PL}$ denotes the slope of the transition between the two filter characteristics, the roll-off of the prevalence response represents the knee level, which is dependent on both orders.

$$G_{BP,PL}(m) = E_0 \left[\frac{\tau r + 1}{(\tau r)^{\beta}} \right]^{\gamma} \tag{5}$$

Where, $(\tau r)^{\beta}$ is represent the located at a relatively higher frequency;The filter's transition from the stop-band to the pass-band is represented by (τr), E_o is the low frequency range, and m is the high frequency range. Non-integer order filters reduce dimensionality and computational complexity while maintaining pertinent information, making it easier to extract useful features from large-scale learning datasets. NIOF has successfully completed the data cleaning process.The pre-processed input data is given to the DTRSRN in order to forecast the Error of Students Learning Behavior.

C. Prediction using Double Transformer Residual Super-Resolution Network (DTRSRN)

In this section, DTRSRN is discussed [29]. DTRSRN is used predict the Error of Students Learning Behavior. By utilizing state-of-the-art deep learning techniques, the DTRSRN offers a comprehensive solution for learning behavior analysis among college English learners. Because of its double transformer design, DTRSRN is very good at extracting complex characteristics from input data, which helps the network identify minute trends in students' learning habits. The network reduces training difficulties by utilizing residual connections, which makes it easier to explore deeper topologies and boosts overall performance. By reconstructing high-resolution representations from low-resolution inputs, its super-resolution capabilities improve the granularity of analysis and capture more minute aspects of how students engage with English language learning resources. The transformer design, which is well-known for its effectiveness in sequence modeling, enables DTRSRN to capture long-range dependencies in student interactions' temporal sequences, which is essential for comprehending how learning behaviours gradually become sequential. Furthermore, DTRSRN is a good fit for assessing a variety of datasets produced by student interactions since it is scalable, flexible enough to adapt to different inputs, and robust enough to withstand noise and unpredictability. Essentially, DTRSRN shows itself to be an effective instrument for thorough learning behavior analysis and it is well-positioned to propel improvements in educational data mining in college English learning environments.

$$q_{lh} = \left\langle \frac{w_f^{\max}}{\left\| w_f^{\max} \right\|}, \frac{w_h}{\left\| w_h \right\|} \right\rangle \tag{6}$$

Where, q_{lh} represents the relationship between w_h *and* $w_f^{\max}$, $w_f^{\max}$ represents the max pooling and w_h represents the individual vector. Advanced data mining techniques are used in the Double Transformer Residual Super-Resolution Network for Learning Behavior Analysis of College English Learners in order to develope the resolution of learning performance data.

$$q_{zh} = \left\langle \frac{w_f^{avg}}{\left\| w_f^{avg} \right\|}, \frac{w_h}{\left\| w_h \right\|} \right\rangle \tag{7}$$

Where, q_{zh} represents the relationship between w_h *and* w_f^{avg}, w_f^{avg} represents the average pooling and w_h represents the individual vector. It efficiently ascertains patterns, trends, and difficulties in English learning by utilizing transformer architecture and residual learning. Teachers may better grasp students' cognitive levels, areas of engagement, and opportunities for growth with the use of this technology. For pupils who are having difficulty learning, it can tailor lessons, optimize teaching methods, and offer timely interventions.

$$y_h = [\alpha_r(w_h), \beta_r(q_{lh})] + [\alpha_r(w_h), \beta_r(r_{zh})] \tag{8}$$

In this case, s α *and* β tand for the global relationship and the embedded functionality of the feature itself, y_h represents the activation function, α_r denotes the mapping function, β_r denotes the super resolution image and link between the individual and average vectors is shown by the letter r_{zh}. The DTRSRN may contribute to educational research and policy formulation, offer individualized learning experiences, and ease performance evaluation and feedback—all of which have the potential to completely transform English language teaching.

$$H_{RQ}(h,i) = e\left(E_{KQ}(h', i'), V(h,i)\right) \tag{9}$$

super-resolution picture at (h,i) is represented by $H_{RQ}(h,i)$, the feature mapping function by $e\left(E_{KQ}(h', i')\right)$, the weight forecast segment of the pixel at (h,i) by $V(h,i)$, and the matching feature vectors in the low-resolution image by $E_{KQ}(h', i')$.Additionally, by offering insightful information about the variables influencing English language learning results in college settings, technology may support educational research. Large datasets of student learning behaviours may be analyzed by researchers using the system, and they can find trends and patterns that can guide the development of educational policies and practices.

$$u_{h,i} = \left(\frac{h}{r} - \left\lfloor \frac{h}{r} \right\rfloor, \frac{i}{r} - \left\lfloor \frac{i}{r} \right\rfloor, \frac{1}{r} \right) \tag{10}$$

Here i is the count of key points, r represent the local characteristics of various key points, $u_{h,i}$ represent the vector linked to the consistent pixel, and h represent the key point detection model. Personalized education is one important use of this technology. The system may offer customized recommendations and interventions to target areas of weakness and maximize learning results by studying the learning habits of individual students. Students may benefit from more effective and efficient learning experiences as a result of this individualized approach, which would eventually aid in their academic success.

$$V(h,i) = \varphi(u_{hi}; \theta)$$

(11)

Where φ represent the weight forecast network, θ represent the weight of the weight forecast network, u_{hi} represent the key point detection, and $V(h,i)$ represent the convolution kernel weight. Data mining is used by the DTRSRN to examine the behavior of college-level English language learners. For precise learning insights, it uses transformer models and improves resolution. Finally, DTRSRN is sued to predict the Error of Students

Learning Behavior. In this work, Multiplayer Battle Game-Inspired Optimizer is assigned to enhance DTRSRN. Here, MBGIO is assigned for turning weight parameter of DTRSRN.

D. Optimization using Multiplayer Battle Game-Inspired Optimizer (MBGIO)

In this section, Optimization using Multiplayer Battle Game-Inspired Optimizer (MBGIO) [30] is discussed. Here the proposed DTRSRN weight and bias parameters q_{lh} *and* y_h are optimized using MBGIO. Numerous data mining technologies are used in Multiplayer Battle Game-Inspired Optimizer for Learning Behavior Analysis of College English Learners. The first benefit of this strategy is that it makes use of the captivating and immersive qualities of multiplayer gaming to encourage students to actively engage in and stay committed to their English learning process. Gamification of the learning process increases the likelihood that students would show continuous interest and passion, which improves engagement and helps them retain language skills. Furthermore, teachers may obtain thorough insights into every student's learning style, preferences, and areas of strength and growth by utilizing data mining technologies. English language training may be made more successful by using a data-driven strategy that enables customized learning experiences based on each student's needs. In addition, the friendly rivalry and collaborative atmosphere that multiplayer games provide encourage peer-to-peer communication and knowledge exchange among students.

Step 1: Initialization

The starting population of MBGIO is generated randomly. Then the initialization is derived in equation (12).

$$y = \begin{bmatrix} y_1^1 & y_1^2 & y_1^1 & \cdots & y_1^D \\ y_2^1 & y_2^2 & y_2^3 & \cdots & y_2^D \\ \vdots & \vdots & \vdots & \vdots & \\ y_N^1 & y_N^2 & y_N^3 & \cdots & y_N^D \end{bmatrix} \tag{12}$$

Where y represent the uniform distribution that generates D is the dimension's initial value, then N represent the neutral for population size.

Step 2: Random Generation

Through MBGIO, the input fitness function acquired unpredictability upon setup.

Through MBGIO, the input fitness function acquired unpredictability upon setup.

Step 3: Fitness Function

The initialized parameters are determined by the best location that is currently available. Determine the individual's fitness value.

$$FitnessFunction = Optimizing\ [q_{lh}\ and\ y_h] \tag{13}$$

Where q_{lh} represents the increasing accuracy and y_h represents the lowering computational time.

Step 4: Movement Phase (q_{lh})

The Movement Phase is primarily in charge of the exploitation capabilities and directs people into possible regions by using the idea of the "safe zone." The primary objective is to determine a portion of the area as safe and then gradually reduce the safe prefetching and motivate people to group as the game proceeds, however the rules in all games may vary.

$$q_{lh} = \left(\left|Y_{best} - Y_{worst}\right| + eps\right) \times rand\,(0.8, 1.2) \tag{14}$$

Whereas eps is a small, non-negative value and q_{lh} represents the safety radius, the safe area's radius cannot be zero. $rand\ (0.8, 1.2)$ satisfies the uniform distribution and yields aarbitrary integer between 0 and 1.

$$y_{new}^k = \begin{cases} y_i^k + normal(\ ), & if\ rand(0,1) < 0.5 \\ y_i^k + \left(y_{best}^k - y_i^k\right) \times rand(0,1), & otherwise \end{cases} \tag{15}$$

Here $normal(\ )$ produces a casual integer that adheres to the traditional normal distribution, and y_{new} and y_{new}^k, correspondingly, denotes the relocated separate and its efficient value in the k^{th} dimension.

Step 5: Battle Phase (y_h)

The primary purposes of the fighting Phase are to allow for exploration and to imitate different fighting performances that rise when players arbitrarily cross paths with one another in the game. When playing against opponents of different skill levels, players may use diverse tactics, but their common goals are to avoid enemy attacks to reduce damage taken and to deal maximum damage to the opponent to win.

$$y_h = \begin{cases} y_i^k + rand(0,1) \times dir^k, & if\ rand(0,1) < 0.5 \\ y_{opponent}^k + rand(0,1) \times dir^k, & otherwise \end{cases}$$

(16)

Whereas $y_{opponen}$ represents the opponent person and y_i represents the original individual, dir represents the vector between the i^{th} individual and the randomly chosen opponent individual.

$$H_l^{(t)} = Y_i + dir \times \cos\ (2 \times \pi \times rand\,(0,1)) \quad (17)$$

Wherein the vector among the i^{th} individual and the opponent separate chosen at random is represented by dir . The uniform distribution is satisfied by $rand(0,1)$, which also yields a random integer between 0 and 1. Figure 2 shows Flow Chart for Multiplayer Battle Game-Inspired Optimizer.

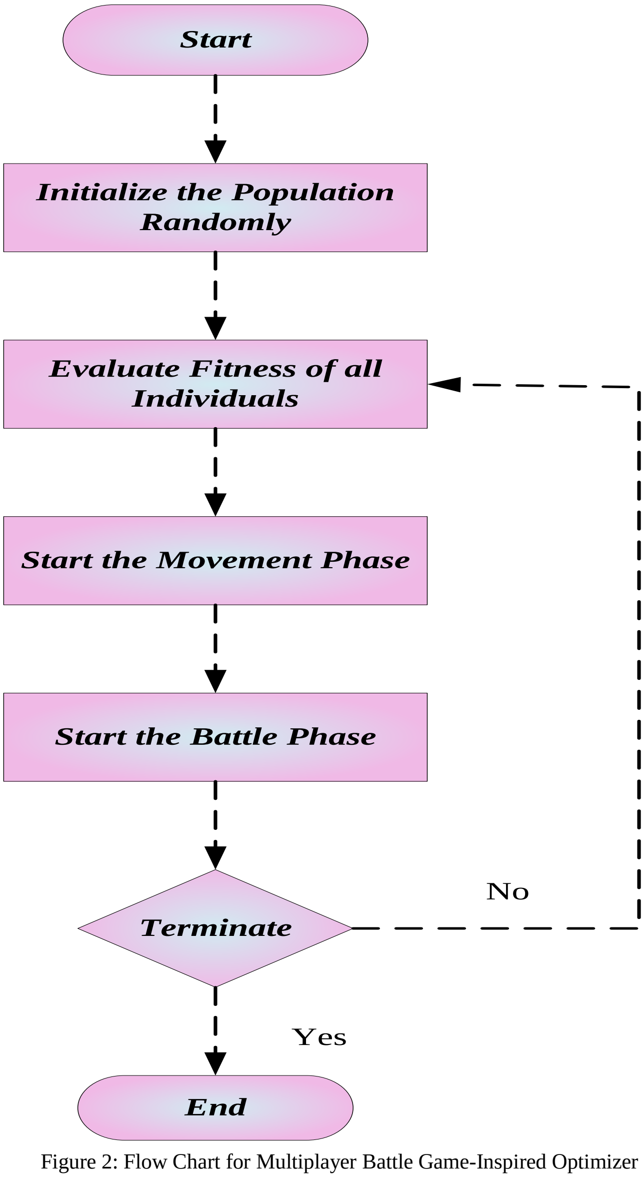

Figure 2: Flow Chart for Multiplayer Battle Game-Inspired Optimizer

Step 6: Termination Condition

Using MBGIO, the weight parameter (q_{lh} *and* y_h) of the Continual Spatio-Temporal Graph Convolutional Network is optimized at this stage. Step 3 is then iteratedly repeated until $y = y + 1$ is reached. Using data mining technology, DTRSRN is improved with MBGIO for Behavior Analysis of College English Learners.

IV. RESULT AND DISCUSSION

The behavior analysis of college English learners based on data mining technology is one of the experimental results of the suggested LBA-CEL-DMT-DTRSRN approach.The results of the proposed LBA-CEL-DMT-DTRSRN methodology are compared with the results of IOT-PELC-BDSVM.

A. Performance Measures

Selecting the best classifier requires taking this critical step. Accuracy, Precision, Recall, F1-score, and ROC are amongst the performance metrics that are evaluated in order to assess performance. The performance metric is deemed in order to scale the metrics. To scale the performance metric, you need the True Negative represent (TN), True Positive represent (TP), False Negative represent FN and False Positive represent (FP) sample

1) Accuracy

The percentage of samples (both positive and negative) relative to the total samples is called accuracy, and it is reported by the equation (18).

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{18}$$

2) Precision

Precision estimation how many positive labels had expected with high accuracy, its expressed equation (19)

$$\Pr ecision = \frac{TP}{(TP + FP)} \tag{19}$$

3) Recall

Recall is computed by dividing the total count of true positive and false negative predictions by the total number of true positives. The model's capacity to collect all pertinent instances is measured. It is shown in equation (20),

$$\mathrm{Re}call = \frac{TP}{TP + FN} \tag{20}$$

4) F1-Score

The F1 score is a widely used statistic to evaluate the model's performance in binary classification problems. It is the vocal mean of recall and accuracy. It is shown in equation (21),

$$F1 - score = 2 \times \frac{\Pr ecision \times \mathrm{Re}call}{\Pr ecision + \mathrm{Re}call} \tag{21}$$

5) ROC

The false negative to true positive region ratio is known as the ROC. given in equation (22),

$$ROC = 0.5 \times \left(\frac{TP}{TP + FN} + \frac{TN}{TN + FP} \right) \tag{22}$$

Here, TP represent true positive, TN signifies true negative, FP represent false positive and FN denotes false negative.

B. Performance Analysis

Figure 3 to 7 shows imitation of suggested LBA-CEL-DMT-DTRSRN technique. Then the suggested LBA-CEL-DMT technique is likened with current SB-EOEDM-FNN, IOT-PELC-BDSVM and CEFCT-BD-DNN methods respectively.

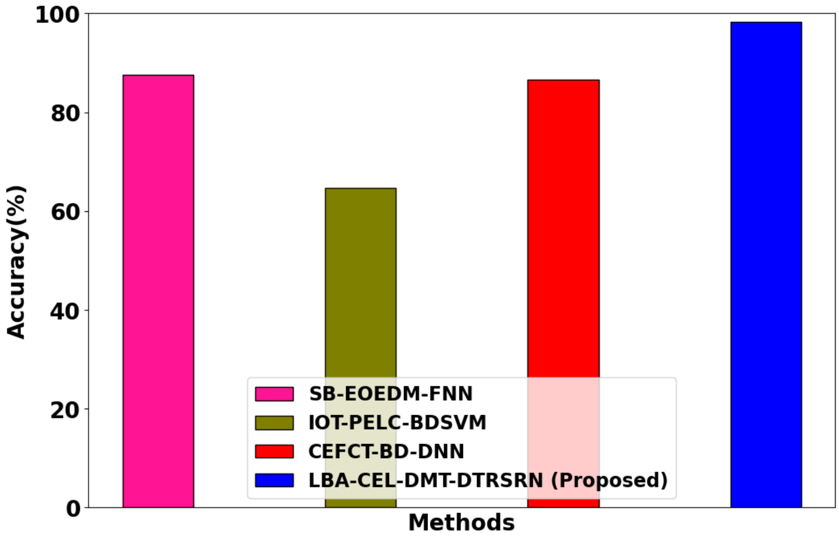

Figure 3: Performance analysis of Accuracy

Figure 3 presents the accuracy analysis and performance comparison of several methods: LBA-CEL-DMT-DTRSRN, SB-EOEDM-FNN, IOT-PELC-BDSVM, and CEFCT-BD-DNN. The suggested LBA-CEL-DMT-DTRSRN approach, in contrast, yields much greater accuracy rates. It performs better than SB-EOEDM-FNN by 28.01%, IOT-PELC-BDSVM by 25.29% and CEFCT-BD-DNN by 21.05%, to be exact. By utilizing cutting-edge methods such as LBA-CEL-DMT-DTRSRN, networks can become more efficient and reliable by improving their capacity to adapt to complicated data patterns. This supports current infrastructure and opens the door for more advanced technical advancements, promoting advancement and creativity in a variety of fields. This suggests that, when compared to the other techniques considered in the investigation, the LBA-CEL-DMT-DTRSRN method has greater accuracy.

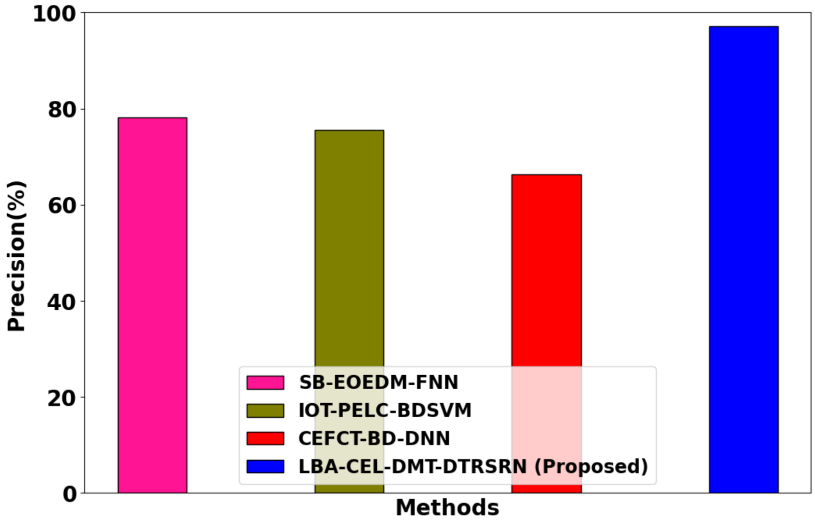

Figure 4: Performance analysis of Precision

Figure 4 compares the accuracy and precision of several approaches. Comparing the suggested LBA-CEL-DMT-DTRSRN approach to three existing methods SB-EOEDM-FNN, IOT-PELC-BDSVM, and CEFCT-BD-DNN reveals a considerable increase in precision. To be more precise, LBA-CEL-DMT-DTRSRN outperforms SB-EOEDM-FNN, IOT-PELC-BDSVM, and CEFCT-BD-DNN techniques by 26.35%, 21.05%, and 28.45%, respectively. This implies that, in comparison to the other approaches examined, the LBA-CEL-DMT-DTRSRN method makes predictions that are more accurate. This is especially important for activities when memory is not as important as precision, such anomaly detection or medical diagnosis, where false positives might have serious repercussions. In this situation, accuracy refers to the proportion of relevant examples among the retrieved instances.

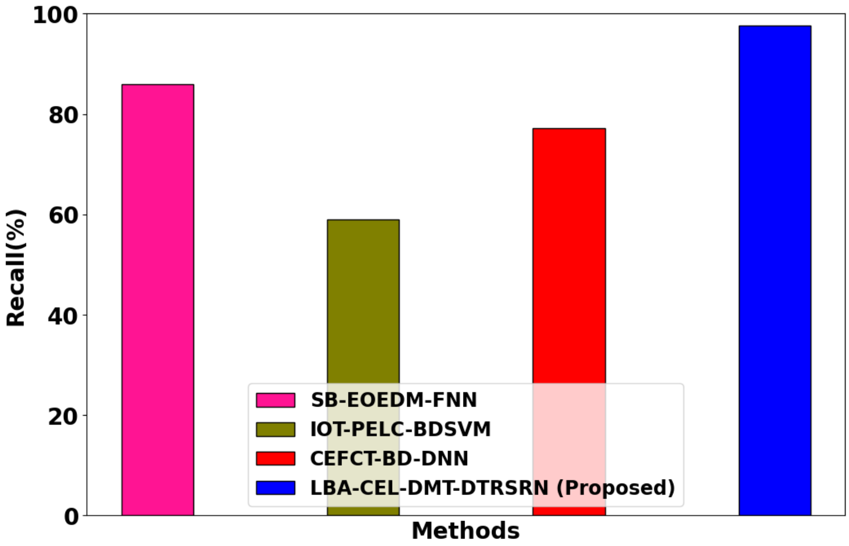

Figure 5: Performance analysis of Recall

Figure 5 shows the performance of several approaches in a particular situation by calculating the Recall metric. The LBA-CEL-DMT-DTRSRN strategy that has been suggested performs noticeably better than the other approaches, with recall rates that are significantly higher. In particular, the LBA-CEL-DMT-DTRSRN approach performs better than the SB-EOEDM-FNN, IOT-PELC-BDSVM, and CEFCT-BD-DNN methods by 23.78%, 26.54%, and 25.14%, respectively. This indicates that the LBA-CEL-DMT-DTRSRN approach outperforms its competitors by a significant margin in accurately identifying pertinent cases in the dataset. These notable enhancements demonstrate the effectiveness and promise of the suggested approach in improving recall performance within the specified parameters.

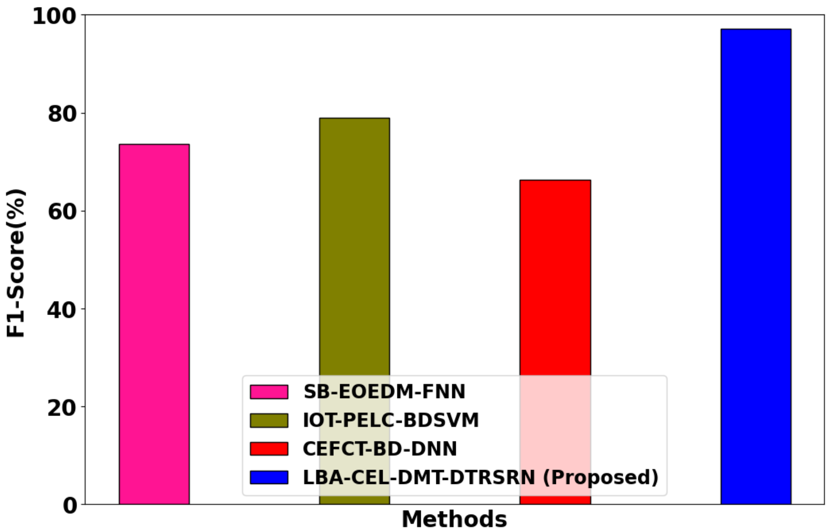

Figure 6: Performance analysis of F1-Score

The F1-scores obtained using the different procedures are displayed in Figure 6, where the suggested LBA-CEL-DMT-DTRSRN method is the most noticeable. It performs much better than the SB-EOEDM-FNN, IOT-PELC-BDSVM, and CEFCT-BD-DNN approaches. In particular, compared to the previously described techniques, LBA-CEL-DMT-DTRSRN generates F1-scores that are 26.35%, 21.05%, and 28.45% higher, respectively. A measure that strikes a compromise between recall and accuracy, the F1-score sheds light on a model's overall effectiveness in classification tasks. When compared to other methods, the LBA-CEL-DMT-DTRSRN approach significantly outperforms others, indicating that it is effective in reliably identifying data points. The significant variations observed in F1-scores highlight the efficacy and possible superiority of the suggested LBA-CEL-DMT-DTRSRN technique in tackling the given problem.

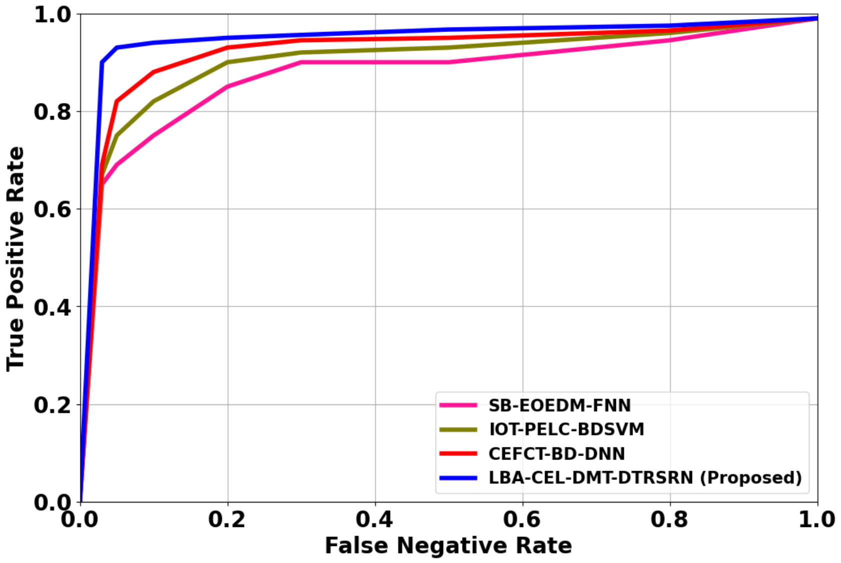

Figure 7: Performance analysis of ROC

Figure 7 uses the Receiver Operating Characteristic (ROC) to compare how well various approaches function. Compared to previous methods, the suggested LBA-CEL-DMT-DTRSRN technique shows a dramatic improvement with much better ROC values. In particular, the LBA-CEL-DMT-DTRSRN technique yields a 26.35% improvement in ROC compared to SB-EOEDM-FNN. Likewise, it demonstrates a significant 21.05% improvement over IOT-PELC-BDSVM. Moreover, the LBA-CEL-DMT-DTRSRN approach shows a notable improvement of 28.45% in ROC when compared to CEFCT-BD-DNN. These findings highlight the LBA-CEL-DMT-DTRSRN method's superiority and efficacy in classifying performance, as demonstrated by the method’s higher ROC values in a variety of comparison studies.

C. Discussion

This work is the initial step toward a data mining-based learning behavior analysis of college-level English learners. The pre-processing unit uses the Global Language Learning Popularity Dataset to enhance the behavior analysis of college English learners. The data’s are predicted using the DTRSRN. Accuracy, precision, recall, F1-Score, and ROC are among the metrics that were utilized to assess the developed LBA-CEL-DMT-DTRSRN's performance. When the suggested LBA-CEL-DMT-DTRSRN is put up against existing methods like SB-EOEDM-FNN, IOT-PELC-BDSVM and CEFCT-BD-DNN, it performs better, achieving greater accuracy 26.35%, 21.05% and 28.45%. Analyzed over an average recall 23.15%, 26.38% and 20.36% value of the recommended technique is 20.36% which is comparable. The assessment metrics of precision and accuracy area greater for the LBA-CEL-DMT-DTRSRN approach in comparison with earlier methodologies.

V. CONCLUSION

In this section, the effective implementation of LearningThe article describes the behavior analysis of college-level English learners using data mining technology and an optimized double transformer residual super-resolution network inspired by multiplayer battle games.The proposed NIOF was used to data cleaning from the Global Language Learning Popularity Dataset. The proposed LBA-CEL-DMT-DTRSRN method is applied in Python. The presentation of suggested LBA-CEL-DMT-DTRSRN approach contains 28.01%, 25.29% and 21.05% higher accuracy,26.35%, 21.05% and 28.45% higher precision when analysed to the existing methods like SB-EOEDM-FNN, IOT-PELC-BDSVM and CEFCT-BD-DNN methods respectively. Further compared to its competitors, the suggested LBA-CEL-DMT-DTRSRN strategy achieves a significant computational time decrease with reductions of 20.15%, 22.63% and 19.21% with higher accuracy. Its ability to improve Error of Students Learning Behavior prediction is validated by this. Future research may concentrate on improving algorithm efficiency through the investigation of methods such as memory optimization or parallel processing. Further experimentation with hybrid techniques or various data mining algorithms may enhance forecast accuracy. Contextual elements and real-time data streaming integration for dynamic analysis should improve the method's resilience. It would also be helpful to look into methods for scalability to handle larger datasets and automatic parameter adjustment. The actual implementation of the approach in online English instruction might be enhanced by working with educators to customize insights and interventions based on behavioural patterns.

Acknowledgement

Gansu Provincial Philosophy and Social Science General Project, Research on the Accessibility and Realization Mechanism of High-Quality Basic Education in Rural Gansu Province（2022YB093）

REFERENCE

[1]Shi, L. & Di, X., (2023). A recognition method of learning behaviour in English online classroom based on feature data mining. *International Journal of Reasoning-based Intelligent Systems*, *15*(1), pp.8-14.

[2]Cenka, B.A.N., Santoso, H.B. & Junus, K., (2022). Analysing student behaviour in a learning management system using a process mining approach. *Knowledge Management & E-Learning*, *14*(1), pp.62-80.

[3]Saiz-Manzanares, M.C., Rodriguez-Diez, J.J., Diez-Pastor, J.F., Rodriguez-Arribas, S., Marticorena-Sanchez, R. & Ji, Y.P., (2021). Monitoring of student learning in learning management systems: An application of educational data mining techniques. *Applied Sciences*, *11*(6), p.2677.

[4]Dabhade, P., Agarwal, R., Alameen, K.P., Fathima, A.T., Sridharan, R. & Gopakumar, G., (2021). Educational data mining for predicting students' academic performance using machine learning algorithms. *Materials Today: Proceedings*, *47*, pp.5260-5267.

[5]Lee, C.A., Tzeng, J.W., Huang, N.F. & Su, Y.S., (2021). Prediction of student performance in massive open online courses using deep learning system based on learning behaviours. *Educational Technology & Society*, *24*(3), pp.130-146.

[6]Trakunphutthirak, R. & Lee, V.C., (2022). Application of educational data mining approach for student academic performance prediction using progressive temporal data. *Journal of Educational Computing Research*, *60*(3), pp.742-776.

[7]Xin, Y., (2021). Analyzing the quality of business English teaching using multimedia data mining. *Mobile Information Systems*, *2021*, pp.1-8.

[8]Peng, N., (2022). Research on the effectiveness of English online learning based on neural network. *Neural Computing and Applications*, *34*(4), pp.2543-2554.

[9]Yin, X., (2021). Construction of student information management system based on data mining and clustering algorithm. *Complexity*, *2021*, pp.1-11.

[10]Hong, X., (2022). Interactivity of English online learning based on neural network. *Neural Computing and Applications*, *34*(5), pp.3349-3364.

[11]Yu, J., (2021). Academic Performance Prediction Method of Online Education using Random Forest Algorithm and Artificial Intelligence Methods. *International Journal of Emerging Technologies in Learning*, *15*(5).

[12]Kukkar, A., Mohana, R., Sharma, A. & Nayyar, A., (2023). Prediction of student academic performance based on their emotional wellbeing and interaction on various e-learning platforms. *Education and Information Technologies*, *28*(8), pp.9655-9684.

[13]Kanetaki, Z., Stergiou, C., Bekas, G., Troussas, C. & Sgouropoulou, C., (2022). Evaluating remote task assignment of an online engineering module through data mining in a virtual communication platform environment. *Electronics*, *11*(1), p.158.

[14]Liu, L. & Tsai, S.B., (2021). Intelligent recognition and teaching of English fuzzy texts based on fuzzy computing and big data. *Wireless Communications and Mobile Computing*, *2021*, pp.1-10.

[15]Tang, Y. & Quan, Z., (2022). College Classroom English Teaching Methods and Practice Based on Big Data Technology. *Wireless Communications & Mobile Computing*.

[16]Yuzhong, H., (2021). Students' emotional analysis on ideological and political teaching classes based on artificial intelligence and data mining. *Journal of Intelligent & Fuzzy Systems*, *40*(2), pp.3801-3809.

[17]Yu, Z., Yu, L., Xu, Q., Xu, W. & Wu, P., (2022). Effects of mobile learning technologies and social media tools on student engagement and learning outcomes of English learning. *Technology, Pedagogy and Education*, *31*(3), pp.381-398.

[18]Hai-tao, P., Ming-qu, F., Hong-bin, Z., Bi-zhen, Y., Jin-jiao, L., Chun-fang, L., Yan-ze, Z. & Rui, S., (2021). Predicting academic performance of students in Chinese-foreign cooperation in running schools with graph convolutional network. *Neural Computing and Applications*, *33*, pp.637-645.

[19]Qiu, F., Zhang, G., Sheng, X., Jiang, L., Zhu, L., Xiang, Q., Jiang, B. & Chen, P.K., (2022). Predicting students' performance in e-learning using learning process and behaviour data. *Scientific Reports*, *12*(1), p.453.

[20]Wang, C., (2021). Analysis of students' behavior in English online education based on data mining. *Mobile Information Systems*, *2021*, pp.1-10.

[21]Wu, L., (2023). IoT-enabled Personalized English Learning in Colleges using Big Data Analysis and Decision Support System.

[22]Chang, H., (2021). College English flipped classroom teaching model based on big data and deep neural networks. *Scientific Programming*, *2021*, pp.1-10.

[23]Wu, J., (2022). Intelligent classroom learning model of college English based on data mining technology in mobile edge computing environment. *Wireless Communications & Mobile Computing (Online)*, *2022*.

[24]Chen, T., Peng, L., Yang, J. & Cong, G., (2021). Analysis of user needs on downloading behavior of English vocabulary APPs based on data mining for online comments. *Mathematics*, *9*(12), p.1341.

[25]Feng, G., Fan, M. & Chen, Y., (2022). Analysis and prediction of students' academic performance based on educational data mining. *IEEE Access*, *10*, pp.19558-19571.

[26]Chen, L., Wang, L. & Zhou, Y., (2022). Research on data mining combination model analysis and performance prediction based on students' behavior characteristics. *Mathematical Problems in Engineering*, *2022*, pp.1-10.

[27]https://www.kaggle.com/datasets/surakarthikeya/global-language-learning-popularity-2020-2023

[28]Nako, J., Psychalinos, C., Elwakil, A.S. & Minaei, S., (2023). Non-Integer Order Generalized Filters Designs. *IEEE Access.*

[29]Zhu, F., Sun, C., Wang, C. & Zhu, B., (2023). A double transformer residual super-resolution network for cross-resolution person re-identification. *The Egyptian Journal of Remote Sensing and Space Science*, *26*(3), pp.768-776.

[30]Xu, Y., Zhong, R., Zhang, C. & Yu, J., (2024). Multiplayer battle game-inspired optimizer for complex optimization problems. *Cluster Computing*, pp.1-25.