International Journal of Multiphysics
Volume 18, No. 3, 2024
ISSN: 1750-9548

Prediction of children's learning effectiveness using data mining technology

Shuai Xin[1], Yaqi Guo[1*], Nannan Ju[1]

1 Departmentof teacher education, Weifang Engineering Vocational College, Qingzhou, Shandong, 262500, China

*Corresponding author e-mail: anna111211@sina.com

Abstract- This study aimed to predict children's learning effectiveness using data mining (DM) technology. The expansion of academic institutions is happening very rapidly since both the public and private sectors are opening up fresh institutions. Medium- and relatively low-risk learners nevertheless continue to confront unemployment. Thus, a novel fine-tuned seagull-optimized weighted k-nearest neighbor (FSOA-KNN) strategy was used in this work to increase the effectiveness of the kids' learning. Three hundred students participated in this study, and their features were gathered and examined. To improve the prediction performance, gathered data samples are used for pre-processing procedures. The proposed approach is put into practice and its effectiveness is evaluated using metrics for accuracy, recall, f-measure, and precision. The study results found that the proposed model has provided an accuracy of 98.7%, which helps in forecasting children's learning efficiency. Additionally, this article aids in identifying the students that require extra guidance or counseling from a teacher who provides high-quality instruction.

Keywords- Education, children, learning, prediction, data mining (DM), fine-tuned seagull-optimized weighted k-nearest neighbor (FSOA-KNN)

1. Introduction

Education is considered one of the most important achievements and prerequisites that a person must meet in order to have access to the truth, progress, recognition as an individual and factual information. It is considered to be one of the cornerstones of human existence [1]. A student goes through multiple phases of education: primary, secondary and higher education. After completing these stages, the student may enroll in colleges, universities or other educational institutions.

Given that education teaches individuals to think, distinguish between good and bad, and make judgments, it aids in knowledge acquisition and information acquisition. Education helps individuals successfully integrate into society, where people become prosperous.

Monitoring teaching and learning activities in an educational setting on a continuous basis is necessary to give students a high-quality education [2]. Because of the abundance of data in educational databases, this is typically challengingnevertheless, big data and machine learning, two recent technological advances, provide solutions for all of these issues. Big data refers to a collection of approaches and strategies that demand novel combinations to extract significant hidden values on a huge scale from a variety of intricate datasets. Although it can detect patterns in data analysis to reveal hidden information and ease decision-making, this technology is very helpful.

Decision-making procedures use a greater amount of data due to the rapid development of various information technologies[3]. The difficulties in storing, handling and evaluating the data grew along with the volume and complexity of the data that was gathered. An exponential rise in demand led to the development of data warehouses, new data management systems that surpassed the capabilities of basic relational database systems. Data mining is a new word for the data analysis field. To put it simply, the tedious, repeated process of finding new patterns in large-scale data sources is known as data mining. Associations, trends, linkages, natural groupings, to improve evidence-based decision-making.

A plethora of knowledge is now easily accessible due to internet technology. Technology has been applied in a variety of ways in the aforementioned situations to increase learning [4]. The world's current focus is on utilizing data mining techniques to better understand learning trends and the effects of alterations in learning environments and methods.Decision makers can use this knowledge to improve the quality of education, as there are large databases at their disposal and the fields of Educational Data Mining (EDM) and Learning Analytics (LA) are fast developing.

Data mining can be applied to extract interesting and pertinent information from data. It encompasses various activities, including clustering, classification and prediction, and association rule mining [5]. Predictions may have biases since the provided data may not be typical of all children. Models that predict might not adapt well to various populations, for instance, if the dataset mostly includes children from particular geographic or demographic origins.This study's primary goal is to improve children learning efficacy by implementing a fine-tuned seagull-optimized weighted k-nearest neighbor (FSOA-KNN) approach. The remaining research falls under the following categories: Section 2deliberates related works. Methodology in Section 3. The outcomes of our techniques will be evaluated in section 4. A summary of the study's findings is provided in section 5.

2. Related work:

Study [6] pursued monitoring students' academic progress, identified the danger, and solved a categorized problem (successful or unsuccessful).Using SPSS Modeler 18, the experiment employed classification models derived from supervised learning methods.The highest specificity value was attained by the C&RT model, which performed the best.

The Orange software was used for DM (Data Mining)procedure and it used Decision Tree (DT) rules to identify the traits of students performing at low, moderate, and high levels in the study [7]. It found that the most important predictor was the students' prior academic performance, along with a few demographic and psychological characteristics.

Study [8] proposed that the methodology labeled student history data was utilized to train the regression model and decision tree classifier after the obtained data had been preprocessed to improve the quality of the data.The outcomes obtained demonstrated the value and efficacy of machine learning technologies in predicting student achievement.

Five machine learning (ML) techniques were used in the study [9] such as SVM, DT, NB, KNN, and the Federated Learning model. Research presented there showed how students made optimal use of learning data from several institutions while safeguarding data privacy, which promoted ML and EDM for the estimation of academic achievement.

Study [10] evaluated the ML model's performance and it compared the use of several data mining algorithms.The DM approach to performance analysis verified the effectiveness and correctness of the learning model, producing accurate and genuine forecasts.

A study [11] provided the impact of educational data mining (EDM) and learning analytics (LA) on adaptive learning. The empirical data supports the main goals of the possible integration of LA/EDM into general strategic planning for education.

The FA-BiGRU-consideration prototype produced the best prediction forecast and performed comparably in variant method analysis when related to individual models like the Back Propagation Neural Network,RF, Linear Regression (LR), GRUand ablation tests [12]. It means that the study has a great deal of potential to change the traditional education sector, guarantee the country's talent pool continues to grow, and offer data references for bettering educational activities.

The past performance in similar courses to predict a student's success in a given course [13]. Massive data sets can be mined for hidden patterns using a range of data mining techniques. The experiment instructed that student data provide pertinent information and therefore comprise among others, SVM, ID3, C&RT and Random Forest.

In order to identify learning practices that have enhanced student performance, a Study [14] examined the association between various factors, such as gender, parents' qualifications, exam preparation, and students' exam results. The results show that student's chances of passing exams were significantly impacted by their participation in test preparation courses.

Study [15] gathered learning data from a LMS to illustrate the significance of student behavioral aspects. The included dataset underwent feature analysis, followed by data preprocessing. It was a crucial phase in the process of discovering new information. The accuracy of the suggested model was increased by the use of ensemble methods. Table 1 illustrate the summary of existing methods

Table 1: Comparative study of applied technologies

Reference	Applied technology	Discovering	Benefits	Limitation
Tosun et al. [6]	SPSS Modeler 18,Classification models (C&RT)	Academic progress tracking was identified and resolved, with the best specificity using the C&RT approach.	Efficient for supervised learning, particular outcomes	Requires knowledge of SPSS Modeler; may not translate to other software
Roslan et al. [7]	Orange software, Decision Tree (DT)	Traits of low, moderate, and high student performance were identified.	Results that are comprehensible and highlight key indicators	Restricted to decision tree functionality and struggle to manage intricate interactions
Yousafzai et al. [8]	Regression model, Decision Tree classifier	predicted academic success using historical data	Effectively uses past data and enhances the quality of the data	The accuracy of preprocessing relies on the quality of historical data, and it can take a while.
Farooq et al. [9]	SVM, DT, NB, KNN, Federated Learning model	Calculated academic performance from information from several colleges	Encourages data privacy and skillfully incorporates learning data	Federated learning setup complexity and inconsistent data quality throughout institutions

Srimani et al. [10]	Various data mining algorithms	assessed ML model performance and confirmed the learning model's efficacy	Thorough study that yields precise forecasts	Results may be impacted by algorithm choice, which calls for knowledge of data mining techniques.
Agus et al [11]	Educational Data Mining (EDM), Learning Analytics (LA)	EDM/LA integration with strategic planning for schooling	Promotes instructional activities and supports adaptive learning	Infrastructure is needed for data integration due to implementation complexity.
Yin et al. [12]	FA-BiGRU prototype, comparison with other models	generated the most accurate estimates and was competitive in technique analysis	Improvements in forecast accuracy and the possibility of improving the education sector	Resource-intensive for training; the evaluation of the model could change based on the use case and dataset.
Sajja et al. [13]	SVM, ID3, C&RT, Random Forest	Employed data mining methods to forecast students' course achievement	discovers latent patterns in data sets and makes predictions using machine learning	Large datasets may need the use of powerful computing resources due to algorithm-specific biases.
Karmagatri et al. [14]	Association analysis between various factors	Examined the effects of variables (parents' qualifications, gender, etc.) on exam outcomes	Reveals strategies for preparing for exams and elements that affect academic success.	Restricted to correlations; may not prove causation
Ajibade et al. [15]	Feature analysis, ensemble methods (LMS data)	Enhanced precision in forecasting student conduct through behavioral factors	Improves model correctness and uses feature analysis to find fresh information	Ensemble approaches can be difficult to use, and feature selection and preprocessing have a big impact on the final product.

3. Methodology

This study passes through the data collections of children and the data is preprocessed with Min-Max normalization algorithm, then open source software WEKA is approached for assessment of generated predictive modelsaccuracy and followed with Fine-tuned seagull optimized K-nearest Neighbor (FSOA-KNN) as proposed model to predict children’s learning effectiveness with the parameters of accuracy, precision, recall and F-measure of predictivemodel.Figure 1 shows the flow of the proposed model

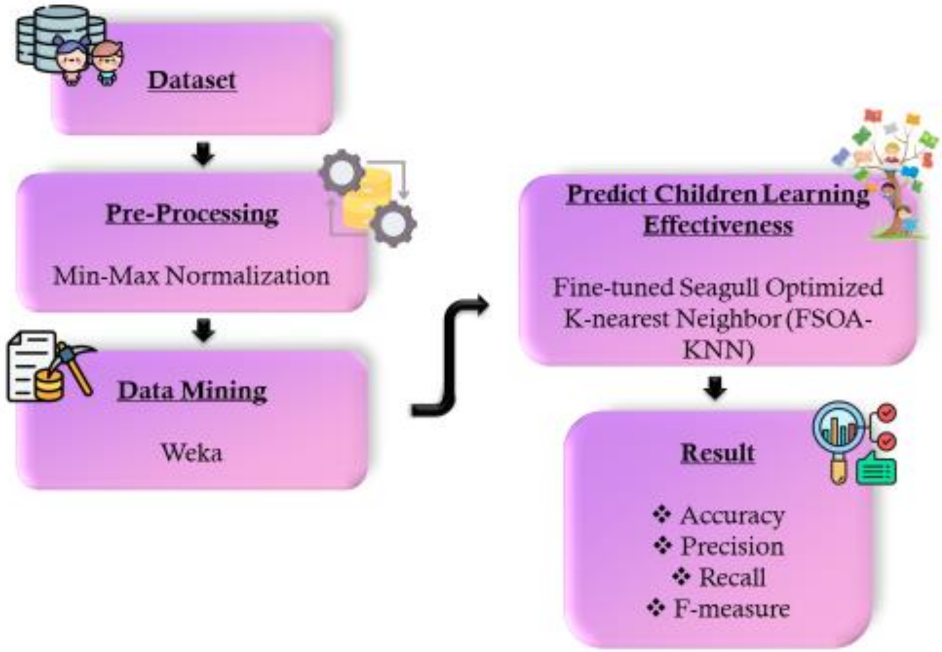

Figure1: Flow of methodology

3.1. Dataset

On the sample selection for children's learning effectiveness. Only the fields necessary for data mining were chosen in this step. A few variables that were derived were chosen. However, some data for the variables was taken straight out of the database. The data collection of around 300 students was initially gathered, all the predictor and response variables as shown in Table 2.

Table2: Dataset description

Variable	Narrative	Value
Sex	Student sex	Male – 150 Female-150
SG	Student grades	O – 90% -80%, A – 79% - 60%, B – 59% - 40%, C – below 40%,
Fsize	Size of student's family	1, 2, 3, >3
Fstat	Familial status of students	Joint or individual family
FAR	Family Annual revenue	High, average, or poor
Pqual	Parents' qualifications	Educated (schooling grade), uneducated, or well-educated (> 1 degree)
FP	Father's Profession	Government employer, businessman, or Day laborer
MP	Mother Profession	Homemakers or working women.

Sample domain values in Table 2 were defined as upcoming points:

- HSG: High School Grades - Students' academic progress. State board students take six courses totaling 100 marks. All students receive their grades using O is 90% to 80%, A is 79% to 60%, B is 59% to 40%, and C is below 40%.In the mapping that follows.
- SG: Student's grade in primary schooling. State board students take five subjects totaling 100 points each. Every student has a grade assigned to them.
- FSize: Family size is more than the number of members in the family; it's a representation of the complexities of resource allocation, interpersonal dynamics, and familial dynamics that shape people's values and goals.
- FStat: Family status is more than a picture; it tells a story about the ups and downs of support systems in the family, overcoming hardship and how socioeconomic status affects educational outcomes.
- FAR: Family annual income is a compass that shows the path of economic possibilities and limitations, affecting social mobility, access to high-quality healthcare, and education.

3.2. Preprocessing

3.2.1 Min-Max normalization:

It is a technique that keeps the original data's associations intact. It is a straightforward method which able to precisely organize the data into a predetermined border using a preset boundary. The data set of students from different schools was preprocessed by Min-Max normalization, a process that uses the concepts like mean and standard deviation to produce the range or normalized values of data from the original unstructured data.

$$B = \left(\frac{B - \min\ value\ of\ B}{\max value\ of\ B - \min\ value\ of\ B}\right) * (C - D) + D \tag{1}$$

3.3. Implementation of Mining Model:

Weka is an open-source program used in data mining applications that carries out a vast array of machine learning methods. WEKA Explorer was used to load this file. The classify panel allows the user to visualize incorrect predictions or the model itself, estimate the accuracy of generated predictive model, and apply classification techniques, such as Fine-tuned Seagull Optimized Weighted KNN, to the resulting the dataset.

3.4. Prediction of children's learning effectiveness

In this section, we predict the children's learning effectiveness using Fine-tuned seagull-optimized weighted KNN. The novelty of using a Fine-tuned Seagull Optimization Algorithm in conjunction with Weighted K-Nearest Neighbors lies in the dual process selecting the best features and determining the optimal weights for classification. This hybrid method improves thepredictive performance ofmaking a useful tool for teacher and successfully raising student learning outcomes.

3.4.1 K-Nearest Neighbor:

The neural network algorithm is one the most basic and traditional methods of classification. As computing power has increased, it has gained importance and emerged as one of the most often-used categorization strategies. Because of its distance-based methodology, NN is more appropriate for use with numerical datasets with categorical datasets. The query is allocated to the class label of the point in the training set that is closest to the query, as determined by applying the fundamental logic of neural networks to the algorithm training set. The query is included in a class based on the majority of tags of the closest k-neighbor in KNN, which is an extension of NN. The Euclidian distance calculation is the most well-known.

$$Euclidean_{j,i} = \sqrt{\sum_{l=1}^{m} \left(w_{jl} - w_{il}\right)^2} \tag{2}$$

3.4.2 Weighted K- Nearest Neighbor:

The W-KNN algorithm was initially presented by using the distance-weight function, close neighbors in the W-KNN are given a higher weight than their distant neighbors. The following Equation displays the x_j'weight of the NN in query w' in iteration j. The class to which the query will be assigned is then chosen by voting based on the l value, which determines the class label. The neighbor with the shortest distance is given more weight than the one with a greater distance. If the weights of the closest and furthest neighbors are 1, and 0, respectively, then the remaining neighbors' weights are scaled linearly based on their separation from one another.

$$x_j' = \begin{cases} \frac{c(w',w_l^{MM})-c(w',w_j^{MM})}{c(w',w_l^{MM})-c(w',w_J^{MM})}, if\ c(w',w_l^{MM}) \neq c(w',w_J^{MM}) \\ 1 \qquad if\ c(w',w_l^{MM}) = c(w',w_J^{MM}) \end{cases} (3)$$

3.4.3. Fine-tunedSeagull Optimization Algorithm (FSOA):

The benefits of the SOA method are its quick convergence speed, low computational cost, and ability to solve large-scale constrained problems. It has a lot of advantages over other optimization algorithms. The equation illustrates that the global optimization search process of SOA is linear. The global search capacity of SOA cannot be fully leveraged due to this linear search strategy. As a result, we provide a nonlinear search control formula, represented by Equation, which can enhance the algorithm's speed and accuracy by focusing on the stage of the seagull group exploration process. Figure2: Display the flow of Fine-tuned Seagull Optimization Algorithm (FSOA)

$$B = e_d \times \frac{1}{f^{4.\left(\frac{s}{Max_{iteration}}\right)^4}} (4)$$

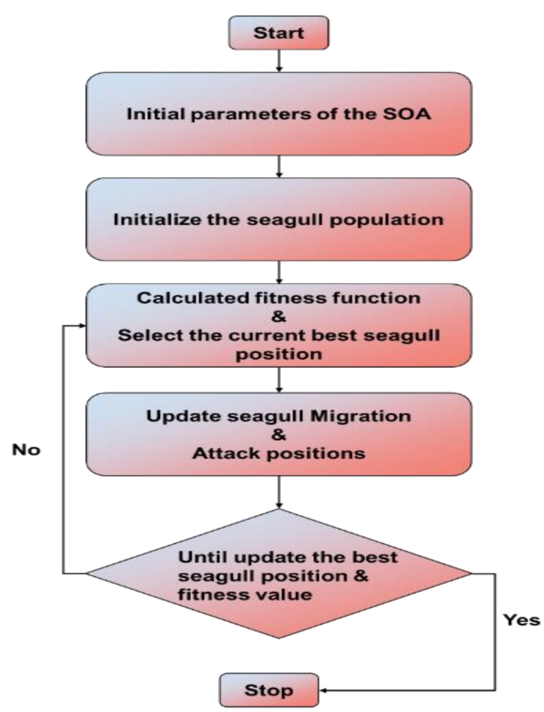

Figure2: Flow of Fine-tuned Seagull Optimization Algorithm (FSOA)

3.4.4. Fine-tuned Seagull-optimized weighted KNN algorithm

We accurately predict children's learning effectiveness by utilizing a perfectly calibrated Fine-tuned Seagull-optimized weighted KNN algorithm, allowing customized teaching tactics for the best possible results.In this case, KNN lays the foundation by offering a baseline classification or regression framework. Next WKNN enhances the procedure by assigning weighted affects to neighboring data points. FSOA then adjusts the parameters of hybrid model to maximize its performance on a variety of dataset. The best features of FSOA, WKNN and NKNN are combined to create the hybrid approach, a symbiotic fusion designed to maximize the capability of each constituent algorithm. By combining this methods meets better unique learning requirement of every student while optimizing academic achievement.

The proposed FSOA was based on the migratory behavior of seagulls and their assault prey. The coding of this algorithm was inspired by the tactics used by a flock of migrating seagulls to catch their meal as they fly from one location to another. To prevent searching agents from colliding with each other in FSOA, an extra parameter "N" is used to determine the new search agent's position. This study used this approach for the learning effectiveness ($\vec{F_s}$) given as

$\vec{F_s} = \mathrm{P}x\vec{\mathrm{M}}_s(i)$ (5)

$\vec{M_s}$ represents the seagull's present role, and “t” denotes the learning effectiveness iteration at that time. It is possible to model the crash prevention parameter "P" as

$P = E_c - \left(i * (F_c/\mathrm{m}ax \cdot Iter)\right)$ (6)

Here, our collision avoidance parameter is a sequentially reductive attribute from F_c to 0, which we set to 2. Using the following equation, the search agents try to approach the ideal student position once the avoidance occurrences in the collision mechanism are finished.

$\vec{N_s} = Ax(\vec{P}_{bs}(\mathrm{i}) - \vec{M}_s(\mathrm{i}))$ (7)

In order to attain the propensity of equilibrium between the phases of exploitation and exploration, and the parameter "P" is randomly assigned and can be computed as

$P = N^2 * 2 * \mathrm{rand}()$ (8)

Subsequently, the following positions of each iteration of the algorithm will be updated:

$\vec{R_s} = \left|\vec{\mathrm{E}}_s + \vec{N_s}\right|$ (9)

While migrating, seagulls frequently modify their frequency and targeting angle based on past experiences. Seagull migration behavior can be visualized in three dimensions as

$s' = x * \mathrm{Cos}(j)$ (10)

$T' = x * \mathrm{Sin}(j)$ (11)

$U' = x * \mathrm{j}$ (12)

The spiral movement radius of the seagulls is represented by "x," and an element between 0 and 2 is chosen at random for "j." The remaining agents in the search will have their positions updated in accordance with the optimal solution once it has been saved.

$\vec{M_s}(k) = \left(\vec{N_s} * s' * T' * U'\right) + \vec{P}_{bs}(k)$ (13)

A graphic illustration of the steps involved in FSOA optimization is shown in Figure 2. Although FSOA has been successfully used for other technical optimization problems, the authors are not aware of any literature that reports on its application to learning optimization problems. The clever behavior of seagulls motivated the authors to employ the FSOA searching approach for the study of learning behavior.

4. Result analysis

To evaluate children's learning using a range of measure, such as recall, accuracy, precision and F-Measure. Existing methods such as SVM, NB, ID3, and KNN [15].

Accuracy analyzes the proportion of incidents that are consistently classified efficiently. Figure 3 and table 3illustrates the result of the accuracy. Compared to the existing method NB (79.2%), ID3 (95.1%), SVM (91.8%), and KNN (96.8%) Our proposed method was higher FSO-WKNN (97%).The proposed approach, FSO-WKNN, has significantly improved children's learning effectiveness prediction when compared to current methods.

Table 3: Outcome value of accuracy

Method	Accuracy (%)
NB	79.2
ID3	95.1
SVM	91.8
KNN	96.8
FSO-WKNN [Proposed]	97

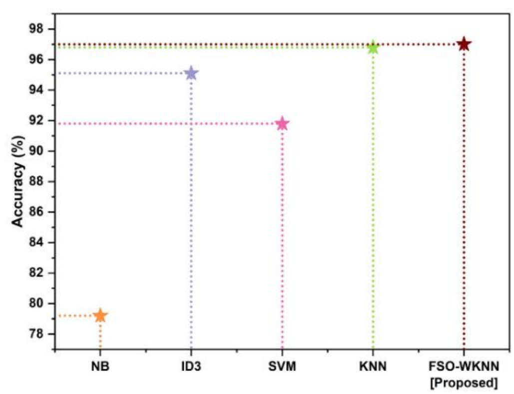

Figure 3: outcome of accuracy

Precision shows that the percentage of accurate, well-made forecasts among all predicted events.Figure 4 and table 4illustrates the result of the precision. Compared to the existing method NB (77.6%), ID3 (94.6%), SVM (90%), and KNN (95.5%) Our proposed method was higher FSO-WKNN (96.5%). The proposed approach, FSO-WKNN, has significantly improved children's learning effectiveness prediction when compared to current methods.

Table 4:Represent the precision comparison values

Method	Precision (%)
NB	77.6
ID3	94.6
SVM	90
KNN	95.5
FSO-WKNN [Proposed]	96.5

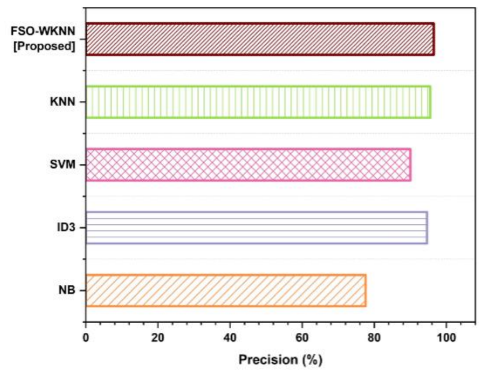

Figure 4: Outcome of precision

Recalls show the percentage of true positives that are effectively acknowledged out of all real positives.Figure 5 and table 5 illustrates the result of the recall. Compared to the existing method NB (75.7%), ID3 (93.1%), SVM (91.2%), and KNN (95.1%) Our proposed method was higher FSO-WKNN (96%). The proposed approach, FSO-WKNN, has significantly improved children's learning effectiveness prediction when compared to current methods.

Table 5:Represent the recall comparison value

Method	Recall (%)
NB	75.7
ID3	93.1
SVM	91.2
KNN	95.1
FSO-WKNN [Proposed]	96

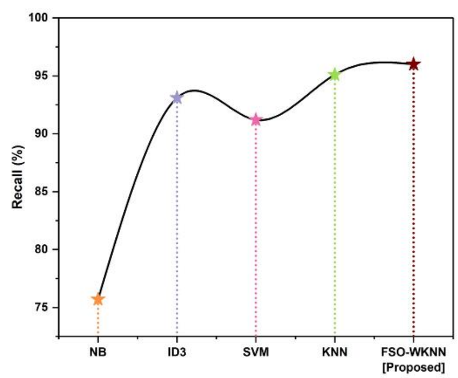

Figure 5: outcome of recall

F-measure is the precision and recall harmonic average that balances the two measurements and is especially useful for unbalanced datasets. Figure 6 and table 6 illustrates the result of the F-Measure. Compared to the existing method NB (75.1%), ID3 (91.3%), SVM (86.8%), and KNN (94.6%) Our proposed method was higher FSO-WKNN

(95.5%).The proposed approach, FSO-WKNN, has significantly improved children's learning effectiveness prediction when compared to current methods.

Table 6: Comparisons of F-Measure values

Method	F-Measure
NB	75.1
ID3	91.3
SVM	86.8
KNN	94.6
FSO-WKNN [Proposed]	95.5

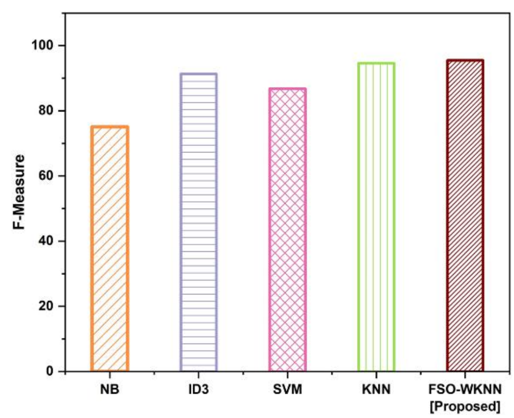

Figure 6: Outcome of F-Measure

4.1 Discussion

In this session we discuss about the findings and limitations of existing methods in table 7.

Table 7: Findings and Limitations of existing methods

Method	Limitation
Naive Bayes	Feature independence is assumed, however this may not be the case for all datasets.
ID3 (Decision Tree)	It is prone to overfitting with noisy data.
Support Vector Machine	Sensitive to parameter tuning, may not perform well with overlapping classes
k-Nearest Neighbors	Computationally costly for large datasets and susceptible to noise and extraneous information

To overcome these limitation we proposed Fine-tuned Seagull Optimized Algorithm K-Nearest Neighbor (FSOA-KNN) for children's learning effectiveness prediction.

5. Conclusion

Prediction of children's learning effectiveness entails evaluating elements such as cognitive capacity, socioeconomic background, teaching techniques, and individual attributes. This study uses data mining technologies to present a novel method for predicting children's learning effectiveness: FSOA-KNN. Using pre-processing we are able to anticipate learning efficiency with greater accuracy than previous approaches by utilizing the data collected from 300 students. The experimental results display that the suggested strategy is effective, outperforming conventional methods in phrases of accuracy (97%), precision (96.5%), recall (96%), and F-Measure (95.5%). This research improves educational practices and promotes improved learning outcomes.

Limitation and future scope

When working with huge datasets, the FSOA-KNN algorithm's computational complexity may cause scalability problems. The size of the dataset may affect the amount of time and money required for training and prediction. To improve prediction accuracy and applicability in a variety of educational settings, future developments may incorporate deep learning techniques to improve feature representation and scalability.

Acknowledgement

Weifang Enginnering Vocational College Horizontal Project"Technical services for the development of a cognitive ability training system for preschool children" (Number:20230430).

Reference:

1. Alghamdi, A.S. and Rahman, A., 2023. Data mining approach to predict success of secondary school students: A Saudi Arabian case study. Education Sciences, 13(3), p.293.
2. Lynn, N.D. and Emanuel, A.W.R., 2021, March. Using data mining techniques to predict students' performance. a review. In IOP Conference series: materials science and engineering (Vol. 1096, No. 1, p. 012083). IOP Publishing.
3. Şen, B., Uçar, E. and Delen, D., 2012. Predicting and analyzing secondary education placement-test scores: A data mining approach. Expert Systems with Applications, 39(10), pp.9468-9476.
4. Srimani, P.K. and Kamath, A.S., 2012. Data Mining Techniques for the Performance Analysis of a Learning Model-A Case Study. International Journal of Computer Applications, 53(5).
5. Pal, A.K. and Pal, S., 2013. Data mining techniques in EDM for predicting the performance of students. International Journal of Computer and Information Technology, 2(06), pp.2279-0764.
6. Tosun, S. and Kalaycıoğlu, D.B., 2024. Data mining approach for prediction of academic success in open and distance education. Journal of Educational Technology and Online Learning, 7(2), pp.168-176.
7. Roslan, M.H.B. and Chen, C.J., 2023. Predicting students' performance in English and Mathematics using data mining techniques. Education and Information Technologies, 28(2), pp.1427-1453.
8. Yousafzai, B.K., Hayat, M. and Afzal, S., 2020. Application of machine learning and data mining in predicting the performance of intermediate and secondary education level students. Education and Information Technologies, 25(6), pp.4677-4697.
9. Farooq, U., Naseem, S., Mahmood, T., Li, J., Rehman, A., Saba, T. and Mustafa, L., 2024. Transforming educational insights: Strategic integration of federated learning for enhanced prediction of student learning outcomes. The Journal of Supercomputing, pp.1-34.

10. Srimani, P.K. and Kamath, A.S., 2012. Data Mining Techniques for the Performance Analysis of a Learning Model-A Case Study. International Journal of Computer Applications, 53(5).
11. Agus, R. and Samuri, S.M., 2018. Learning analytics contribution in education and child development: A review on learning analytics. Asian Journal of Assessment in Teaching and Learning, 8, pp.36-47.
12. Yin, C., Tang, D., Zhang, F., Tang, Q., Feng, Y. and He, Z., 2023. Students learn performance prediction based on feature extraction algorithms and attention-based bidirectional gated recurrent unit networks. Plos one, 18(10), p.e0286156.
13. Sajja, G.S., Pallathadka, H., Phasinam, K. and Ray, S., 2022. Using Classification Data Mining for Predicting Student Performance. ECS Transactions, 107(1), p.10217.
14. Karmagatri, M., Kurnianingrum, D., Suciana, M.R. and Utami, S.A., 2023, October. Predicting Factors Related to Student Performance Using Decision Tree Algorithm. In 2023 5th International Conference on Cybernetics and Intelligent Systems (ICORIS) (pp. 1-6). IEEE.
15. Ajibade, S.S.M., Dayupay, J., Ngo-Hoang, D.L., Oyebode, O.J. and Sasan, J.M., 2022. Utilization of ensemble techniques for prediction of the academic performance of students. Journal of Optoelectronics Laser, 41(6), pp.48-54.

Appendix

Abbreviation	**Full form**
C&RT	Classification and regression tree
ID3	Decision tree
DM	Data Mining
EDM	Educational Data Mining
LA	Learning Analytics
SPSS	Statistical Package for the Social Sciences
DT	Decision Tree
SVM	Support Vector Machine
NB	Naïve Bayes
KNN	K-nearest Neighbors
ML	Machine Learning
RF	Random Forest
BP	Back Propagation Neural Network
GRU	Gate Recurrent Unit
LMS	Learning Management System