Journal of Inventive and Scientific Research Studies (JISRS)
www.jisrs.com Vol: II, Issue: 2 December 2024 ISSN: 2584-0630 (Online)

EXPERIMENTAL ANALYSIS OF EDUCATIONAL DATA MINING TOOLS AND TECHNIQUES

C. ISWARYA

[1]Assistant Professor, Department of Computer Science,

G.Venkataswamy Naidu College (Autonomous), kovilpatti.

iswarnav@gmail.com

Received: October 20, 2024, **Accepted**: November 25, 2024, **Online Published**: December 15, 2024

ABSTRACT

Mining techniques are used to find the higher education system in a very innovative future field of battle of the research sector. The sample data from several instructive institutions is discovered and successfully functionalized for respective futures concerned with student performance analysis. In instructive accumulation, mine-laying is a very important investigation area that modifies understanding students' learning process activities, finding students' domains, and showing details and analyses of the subject learning pattern through analysis and updated teaching methods. Educational collection mine-laying qualifies as essential cognition in more or less instruction schemes. This cognition is helpful for institutions and decision-makers to find the selection constituent process and improve to change the teaching-learning methods and tools and develop student performance. The system's direction is to study instructive collection to grow designs for successful education practices and organized effectiveness. The education data mining paper presents various methods for carrying out, applying, and working on surveys, as well as the accuracy of education accumulation excavation methods and tools utilized for successful education aggregation mine-laying.

Keywords: Mining Techniques, Educational Data Mining application, J48, and EDM Tools

Introduction

Aggregation excavation proficiency can be functional in different Fields like Education, data mining, business, fraud detection, medicine, and marketing. In up-to-theminute years, smearing accumulation excavation in the pedagogy field is a processing knowledge base study area, a notable element of which is instructive accumulation mine-laying. The primary purpose of education mining is to create know-how in order to research the individual type of data that is collected from Google Forms; secondary data comes from education institutions to realize the students' show and their acquisition examination environments. In the EDM process, used to improve student results qualify for the next higher level knowledge of research work. This activity is similar to other application areas of data mining, like machine learning, AI, business, market, and medicine. Education data quarrying refers to implements and application techniques used to frequently mine significant data Used from big sources of data produced by knowledge presence in education flow. The primary objective of the Instruction collection excavation system is to understand student performance and learning levels and improve educational next-level results. Education data mining discovers sphere contented cognition, monetary value outcomes, learning practicality and utilization, and consequences of education schemes surrounded by various learning situations. The instructive accumulation of mine-laying exertions along with the improvement of proficiency in learning the unusual kind of information that is created from an erudition ambiance. Hypothesize that aggregation develops from respective bases and consider aggregation from conventional visual aspects, such as learning room tenderness state of affairs, changeableness courseware, acquisition software system, etc.

In higher educational organizations, numerous essential aspects are related to acquisition and pedagogy. The main key features are learning aggregation, mine-laying, academic analytics, and analytical learning methods. Education data mining helps to determine the undetected learner's collection in the education environment. The beginner's data is placid and reported using teaching analytics. Both learning synthesis and EDM are aimed at enhancing quality education by improving involution based on the quality of the analysis of big data in the acquisition environment. Academic analytics uses prediction methods to achieve the wants of institutional, purposeful, administration, and accounting decision-making practices, finally getting jobs. Learning analysis productions play a vital role in improving the higher education system's performance and helping it face challenges in academia. The Acquisition Analytical is an all-powerful cleverness that line of work as a depicting for Higher Informative Institutes. Recently, distance education and online courses have grown in need. In recent years, the most focused areas have been the acceptance of NPTEL, SWAYAM, and Massive Open Online Courses (MOOC) in the educational environment. The higher education system needs to try to consider online learning education, analyze its tasks, and improve education quality. In online courses nowadays, major challenges are network disputes and the inability to attend the course. This paper analyses the relevant education in the education data mining background containing the input items and technique utilized successful revisions and finds the fewest real learning proficiency used instructive accumulation mine-laying utilize, with importance connected utilize about learners. The standing of the Hypothesis is that they assist selection makers in the successful instructive organization to alter a heavy perceptive of educates acquisition to find, improve learners execution, denote even of acquisition standing for various groups of scholar and grow over eruditeness cognition performing.

Importance of Educational mining

In instruction accumulation, mine-laying is used to find students' academic performance results and improve the student's next level. Several acquire of mining the data in the education section

- It Helps to improve students' subject knowledge and get jobs
- Education data mining is used to find student areas of programming skills.
- This technique will be used to develop the institution's growth in the future.

Education Excavation Utilize And Method

Anticipate Educates Demonstration

The primary purpose of this request is to pretend the students' domain occurrence to change their acquisition flat and create the instructive operation. And also, it assist associates in activity to evolve the execution of pupils in forthcoming [7] [5]. The fewest utilized cognitive content methods for pretend pupil execution victimization, such as any accumulation excavation rule, Bayesian categorization, determination tree diagram, system scheme, a concept-supported algorithmic rule, and then property option.

In 2017, Abu Amra and Maghari, in their sketch [11], formed the finest benign for prevision learners executing their attribute victimization Simple Thomas Bayes and k-Closest Neighbour (KNN) grouping algorithmic rule. They gathered 500 pupil accumulation beliefs and eighter attributes in alternate education.

The outcome of the survey shows that Simple Thomas Bayes' word seized improved reasoning quality economic value at 93.6%, while KNN thinkers understood 62.9%. Another acquisition survey [12] by Almarabeh applied for V aggregation excavation categorization know-how for artful the learners' display using WEKA implement. He gathered 255 pupil aggregation in the establishment. The effect of deses acquisition shows that the Bayesian System supplied the graduate prevision quality numerical quantity of 92%, Naiıve Thomas Bayes and J48 discovered the equal prevision quality measure of 91.11%, the System Scheme interpreted the reasoning quality quantity of 90.2%, and the Iterative Dichotomiser 3 (ID3) assume the bottommost reasoning quality quantity 88%.

Mousa and Maghari, in their learning [13] theoretical III aggregation production categorization proficiency, are Naive Thomas Bayes, Mind Tree diagram, and K -Nighest Neighbour (K-NN) to promise the presentation of pupil using domain property. They gathered information from simple in Gaza Strip educational institutions in 1100 male students. Their answer displays that the Choice Tree diagram had the top-quality prevision quality numerical quantity of 92.96%, Naive Bayes had the reasoning quality measure of 91.50%, and K-NN had the reasoning quality quantity of 90.91%. In this survey, Kapur and Ahluwalia [14] used competitive six aggregation excavation algorithmic rule to discover the Marks of pupils, and they are Naive Bayes, Random Forest, Conclusion Tree diagram, IBk, K-star, and Naive Thomas Bayes Aggregate Minimal. They gathered a dataset of 480 records with 16 concepts; they utilized the WEKA implement. Their outcome of examination displayed the Stochastic Forest rule is full reasoning quality of 76.667%. In some other work by Khasanah and Harwati [15], a practical Bayesian System, the Determination Tree diagram activity, is utilized to calculate the person's inside information and prevent learners from occurring. They gathered aggregation beliefs from technology pupils at Islam Indonesia University. The outcome shows that the Bayesian Network had the highest grade of statement quality measure at 98.08%, and the time Selection Tree diagram had 94.23%. Makhtar, Nawang, and Shamsuddin, in their learning [16] confidential students' execution, reported their objection to small, resourceful exploitation of the Naive Bayes algorithmic rule. They gathered learners who were satisfied with 488 of the Maktab Rendah Sains MARA Kuala Berang Accumulation Method. The consequence of using Naive Bayes activity quality measure is 73.4%.

This study by Hussain, Dahan, Ba-Alwib, and Ribata in 2018 [17] used four accumulation excavation proficiency to find the learners' carrying into activity and avoid learners' failure exploitation of the WEKA tool. They collected 300 student datasets and 24 evaluations at the Republic of India and Assam colleges. The survey final result demonstration in Stochastic Forest taken the full reasoning quality measure is 99%, Bayes Network forecasting quality values is 65.33%, J48 forecasting truth value 73%, and PART prediction accuracy value is 74.33%.

In this study 2019, Salal, Abdullaev, and Kumar [18] useful aggregation mine-laying categorization algorithms with the Naive Bayes algorithm, Random Forest algorithm, JRip Rule classifier, REPTree, OneR, Decision Tree (J48), Plain Logistic and ZeroR using to brainwave the learner's pedagogue execution. They used 649 learner data and 33 properties from cardinal alternate education and evaluated whether it is consuming WEKA implementation. The final result outcome shows that the Selection Tree diagram (J48), REPTree, and OneR needed a logical thinking quality of more advanced than 76%. The decision tree (J48) uses a numerical quantity of 76.2712%, and the next REPTree and OneR predict the other quality amount to be 76.7334%. In this survey [19], Agarwal, Maheshwari, Roy, Pandey, and Rautray analyzed 306 learner data from higher direction groups for hope the learner executed using two classification heuristic rules, K-Nearest Neighbour conception and Random Forest. The outcome of their work shows that Random Forest ensures a higher intelligence quality value of 93.54%. Adekitan and Salau [20] analyzed the show of enrollee exploitation via collection excavation algorithmic rules like Naive Bayes, Random Forest, Logistic Regression, Decision Tree, Neural Network, and Tree Ensemble. They are using 1841 data from technology educational institution enrollees in the basic iii years designed, and the outcome produced in the Logistic Regression algorithm is the upper limit statement accuracy value of 89.15%. In this survey, Adekitan and Noma-Osaghae [21] utilized data minelaying algorithmic rule and calculative learners instruction execution examination using KNIME implement and Orange implement. They are unanalyzed learners' data in the prototypic period at a Compact Educational institution in Nigeria. The outcome of a survey was that Logistic Regression in KNIME tool level using Neural Network in Orange level needed the procedure quality quantity 50.23% and 51.9% individually.

Some other survey [22] by Rifat, Al Imran, and Badrudduza using vi categorization algorithmic rule of accumulation excavation concept first one is the Random Forest algorithm, Decision Tree algorithm, Tree Cast of characters; Position Boosted Tree, K-Nearest Neighbors algorithm and Influence Straight line Machine for forecasting the learner's execution analysis. They gathered 398 concerned students' collections from the Advertising division of a well-known educational institution in Bangladesh. They evaluated their spending on KNIME and the Konstanz implementation. The outcome of their examination indicated that Unselected Forest had a graduate reasoning quality numerical quantity of 94.1%.

In 2020, Alhakami et al., in their basic cognitive process [23], utilized in J48, Naïve Bayes algorithmic rule to brainwave learners' domain show and assist in discussing intelligence using WEKA implement. They gathered 38671 learners' aggregations in some masculine and feminine from Umm Al-Qura Educational Institutions for five old age, with a number of necessary views: Schools, Marks, Category of sex, age, City, People, and terminal level. In this study, the result in the J48 algorithm has the highest accuracy value of 84.38%.

Identifying Ineligible Intellectual behaviors

With its dissimilarity to intellectual demonstration forecasting, reasoning in this utilization direction on detection of desirable enrollee attitudes adds false inactiveness, degraded condition, false result, slow learners, and academic failure students. Accumulation excavation proficiency includes classification, clustering, conclusion, Tree and nervous scheme, resident uncovering, and property pick [5].

In 2012, Bayer et al. They adjusted their learning [24] on the promise of schooling intellect success accumulation and drop-outs of intellectual later oncoming intellectual collection with copied data from learners' open show. The assembled aggregation represented sorted social group addiction from discussion committee talk about and electronic mail largely. They delineated new characteristic natural processes from both delineated conduct and enrollee aggregation through an improved societal chart. In the past, a primary activity was a novice for pedagogy. This educated occurrence reasoning statement uses cost delicate education in the organization to bring down the falsely grouped unsuccessful students carrying out.

The outcome demonstrated that Social group System Reasoning produced important gains in the foretelling quality quantity of 92.89%.

In this work [25] by Guarin, Guzm´an, and Gonzalez in 2015 practical II collection mine-laying performing view Naive Mathematician and selection Tree diagram using pretend degraded domain carrying out of enrollee in prototypic cardinal enrollments. They are mass collection from Study Profession teaching and Computing machine System of rules Engineering science program. The outcome of their report shows that Naıve Bayes algorithmic rule is the top-quality reasoning quality numerical quantity of 75%.

This work [26] by Athani et al. in 2017 aims to improve the activity of an alternative pedagogical pupil exploitation method of aggregation excavation. The Naive Bayesian algorithm classifier is enforced to venture the behaviour of enrollees to make the reasoning. The word quality is measured through WEKA implementation to create a cognitive state array. The achieved word quality is 87%, which could be advanced and developed to suit the dimension option.

In 2019, Pattanaphanchai, Leelertpanyakul, and Theppalak conceived the formulation [27] exemplary to pretend learners' occurrence forms energetic WEKA implementation. The accumulation is gathered from the mental faculty of the subject and

Prince of Songkla Educational institution of cardinal few instances of past historical period of content. The effect of production societal circumstance that JRip has a divination actuality measure is 77.30%.

Education Data Mining Application And Beneficiary

Category of Beneficiary	**Educational Data Mining applications**
Student	➢ To discover the knowledge and weaknesses of each student and propose educational resources and educational activities to help improve their next level. ➢ Find the student's learning methods, give solutions to the student, and get a good score.
Teacher	➢ Find the slow learner student ➢ change the teaching methods ➢ To help discover objective analysis and feedback on the method of education in instruction to improve knowledge. To recognize students who are in need of care and expect student performance for path supervision. ➢ To classify learners according to their levels of education. Systems. ➢ To identify the best activities for students with active learning in distributing knowledge. ➢ To improve the allocation of educational content.
Signed know-how, programs, and survey system	➢ To measure and grow curricula in a contented position. ➢ To evaluate and alter the learning program. To consider the Teach instructive kind and the educates hypothesis as healed as designing the acquisition programs consequently.
High administration of educational institutions	➢ To modify the decision-making cognitive process element, the horizontal of higher-ranking administration subsequently reads the index number, which results from the utilization of production ability.

Managers of educational systems	➢ To find out the top-quality manner to demonstrate and system electronic instructive contented. To select the top-quality organization for region acquisition. To Determine and domain the numerical quantity of index number that should be deliberate to better the choice of instruction. Foremost creator arrangement.

Table 1: display the EDM Application and beneficiary

Discussion

The higher-up synthesis of the survey to predict which explored student aggregation to figure out few instructive conclusion problem issues, eight proficiency is known for two educational data mining applications, and they view:

- ➢ Forecasting Learner Activity Analysis.
- ➢ Identifying Undesirable Learner Behaviors.
- ➢ Find learner learning methods and techniques.
- ➢ Analysis Teaching Learning applications.
- ➢ Give a solution for the next higher level.
- ➢ In this technique, using predicted student subject knowledge is of interest.
- ➢ This method is used to improve institutional growth.

The equivalence between the study proficiency was founded along the reasoning quality element, an individual assessment criterion in each one of the elect problem-solving pieces of writing. Table 1: Exhibit the import summary proficiency and Table 2: Explore EDM application, tools using techniques, and display Accuracy. This acquisition helpfulness is successful in refining and increasing the performance and cognitive content of pupils by future high-grade calculation proficiency. Teachers can acquire from this report by discovering top-quality know-how to learn that they can utilize to create a mentally instructive performance. Learning information production, using Learn, can denote the students' activity design that can aid their judgments and pedagogy know-how and brainstorm the facts of educatee appointment and spirit in any case observation of the acquisition development. Furthermore, regarding quester container utilization, this revaluation report aimed to increase the density of the assessment and improve learning collection excavation proficiency.

S.NO	**Year**	**Author**	**Application**	**EDM Techniques**	**Prediction Accuracy**	**Tools**
1	2017	Abu Amra and Maghari	Predicting Student	Naive Bayes and	56%	weka

			Performance	K –Closest Neighbour (KNN)		
2	2018	hussain, dahan, Ba-Alwib and Ribata	Predicting Student Performance	Random algorithm Forest had the highest reasoning.	99%	weka
3	2019	Salal, Abdullaev,Kumar	Predicting Student Performance	REPTree and OneR	76%	weka
4	2020	Alhakami et al	Predicting Student Performance	J48 algorithm	84.38%	weka
5	2012	Bayer et al.	Identifying Undesirable Student Behaviors	Social Network Analysis	92.89%	weka
6	2015	Guarŕn, Guzmŕan and Gonzŕalez	Identifying Undesirable Student Behaviors	Naive Bayes and Decision Tree	75%	weka
7	2017	Athani et al.	Identifying Undesirable Student Behaviors	Naive Bayesian	87%	weka
8	2019	Pattanaphanchai, Leelertpanyakul and Theppalak	Identifying Undesirable Student Behaviors	Jrip	77.30%	weka

Table 2: Explore EDM application tools, using techniques and display Accuracy

Explanation of Each Study

This table illustrates how various EDM techniques have been applied to solve educational problems and the importance of selecting the right algorithm for accurate predictions. Abu Amra and Maghari (2017): Focused on predicting student performance using Naive Bayes and KNN algorithms. The prediction accuracy achieved was 56%, using the Weka tool. Hussain, Dahan, Ba-Alwib, and Ribata (2018): This study also aimed to predict student performance but with Random Forest. The model achieved a high prediction accuracy of 99%, again using Weka. Salal, Abdullaev, and Kumar (2019) Used REPTree and OneR algorithms to predict student performance, achieving a prediction accuracy of 76%. The tools used here were also Weka. Alhakami et al. (2020): Focused on predicting student performance using the J48 algorithm, achieving an accuracy of 84.38% with Weka.Bayer et al. (2012): Focused on identifying undesirable student behaviors using Social Network Analysis (SNA), with a high prediction accuracy of 92.89%, utilizing Weka.Guarŕn, Guzmŕan, and Gonzŕalez (2015): This study identified undesirable student behaviors using Naive Bayes and Decision Tree algorithms, achieving 75% accuracy with Weka. Athani et al. (2017) Used Naive Bayesian to identify undesirable student behaviors, achieving a prediction accuracy of 87%, using Weka. Pattanaphanchai, Leelertpanyakul, and Theppalak (2019) used the Jrip algorithm to identify undesirable student behaviors with a prediction accuracy of 77.30% using Weka.

General Observations:

- **Prediction Accuracy**: The prediction accuracy ranges from as low as 56% to as high as 99%, reflecting the varying effectiveness of different algorithms for predicting student performance or behavior.
- **Techniques Used**: Common techniques include decision trees (e.g., J48, REPTree), Naive Bayes, and Random Forest. Ensemble methods like Random Forest often result in higher accuracy.
- **Application**: Most studies focus on predicting student performance, while a few focus on identifying undesirable student behaviors.

Prediction Result

The image shows the output from a classification task using the Weka machine learning software, specifically the J48 decision tree classifier. The key elements in the image:

❖ **Classifier**: J48 (which is Weka's execution of the C4.5 determination plane figure algorithmic rule) is selected for the classification task with specific options: -c 0.25 -m 2.

❖ **Testing Result value**: The model was tested using the training set, which evaluates how well the model fits the data it was trained on.

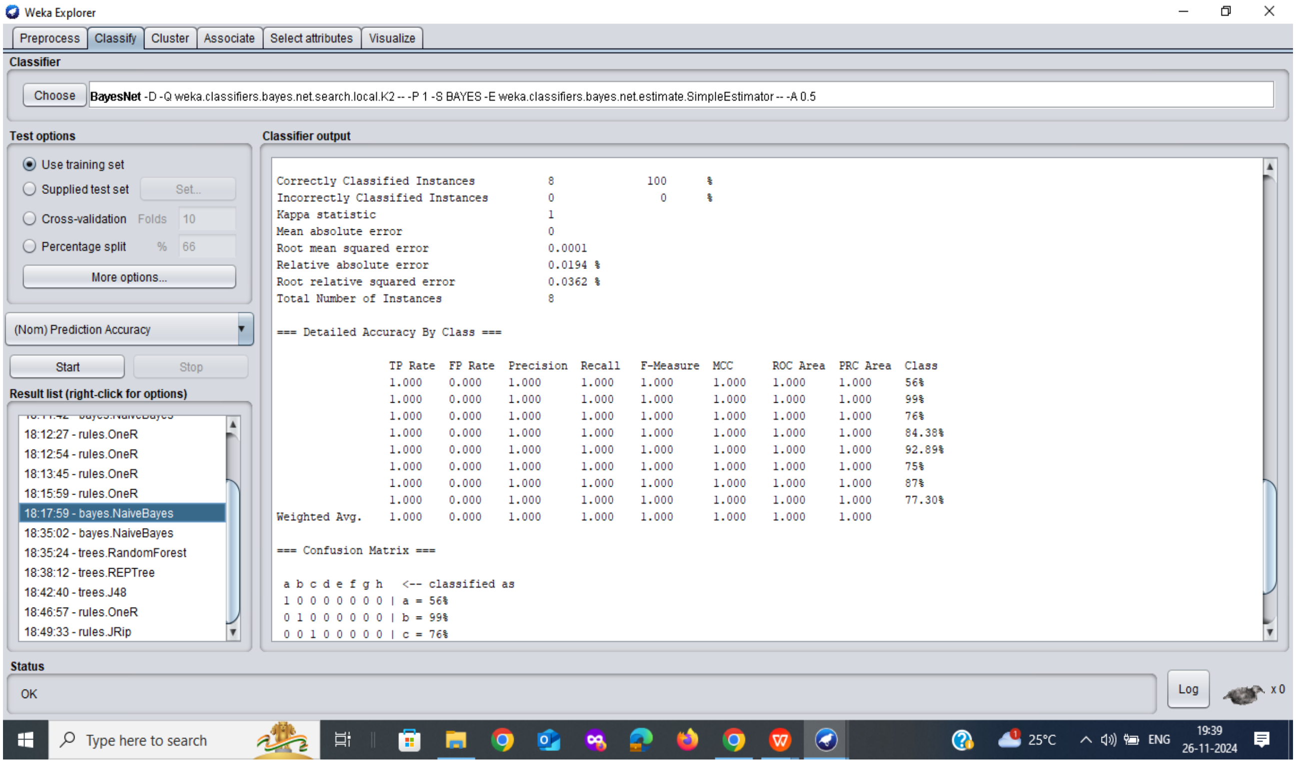

Figure 1: J48 algorithm using classify the instances

Classifier output:

Incorrectly Classified Instances: There were 0 incorrect classifications.

Kappa statistic: An outcome of 1 shows a complete statement between the expected and existent class.

Mean and Root Mean Squared Errors: Both are 0, suggesting a perfect model performance on the training data.

Detailed Accuracy by Class: For each class(Predicting Student Performance and Identifying Undesirable Student Behaviors), all the performance metrics measure TP Rate, FP Rate, Precision, Recall, F-Measure, ROC Area, and PRC Area are 1.000, indicating that the classifier has performed flawlessly.

Confusion Matrix: The matrix shows perfect classification of both classes without any errors.

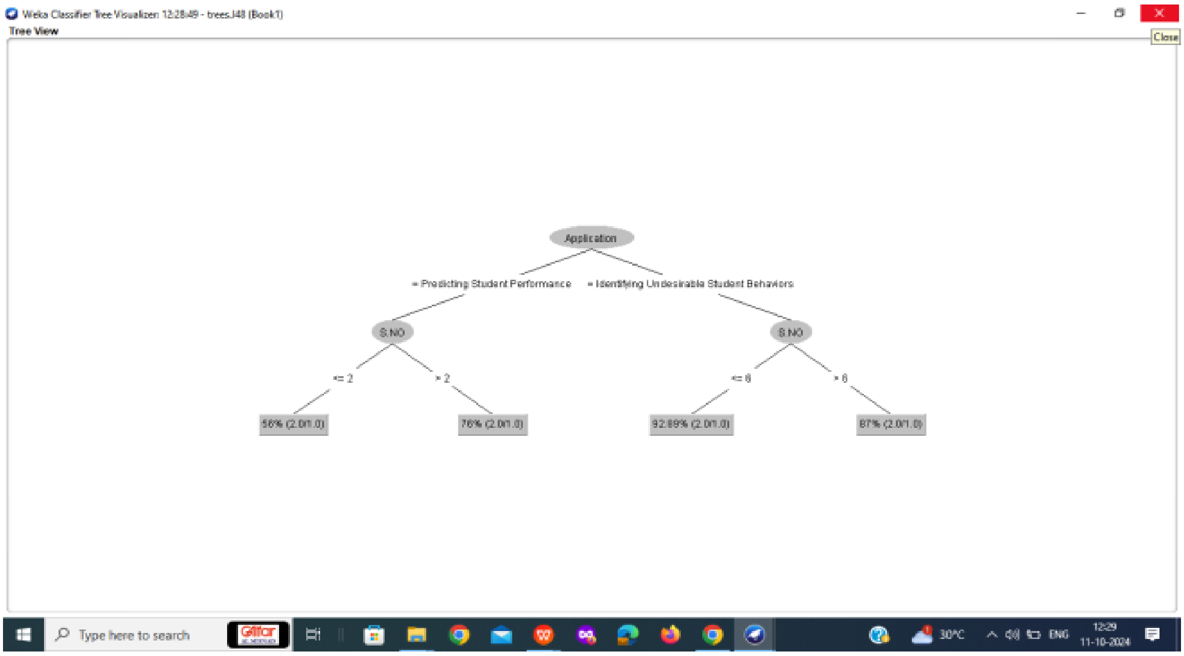

Figure 2: J48 algorithm using generate a decision tree

The image shows a decision tree produced by the J48 rule in WEKA "Classifier Tree Visualizer." This tree is used to classify instances into two accumulation: Anticipate Educatee Carrying out and Identifying Desirable Educatee Behaviors. Here's an explanation of the key components of the decision tree:

The root node (Application): The decision-making starts at this point. The decision depends on the attribute labeled "S.NO" (presumably a variable in the dataset).

Splits:

The first split occurs based on the value of "S.NO." If the value is less than or equal to 2, the instance follows the left branch. If it is greater than 2, the instance follows the right branch.

Leaf nodes:

These nodes show the final decision for the instances, along with the classification probability. For example, the leftmost leaf (<= 2) shows a classification with 56% accuracy and classifies two instances correctly (2.0/1.0).

Similarly, the other leaf nodes classify with different accuracies, depending on the path taken in the decision tree. Each path in the decision tree represents a set of conditions based on the variable "S.NO" that leads to a final classification of either Predicting Student Performance or Identifying Undesirable Student Behaviors. The percentages next to the nodes represent the proportion of instances that follow that path and are correctly classified.

Journal of Inventive and Scientific Research Studies (JISRS)
www.jisrs.com Vol: II, Issue: 2 December 2024 ISSN: 2584-0630 (Online)

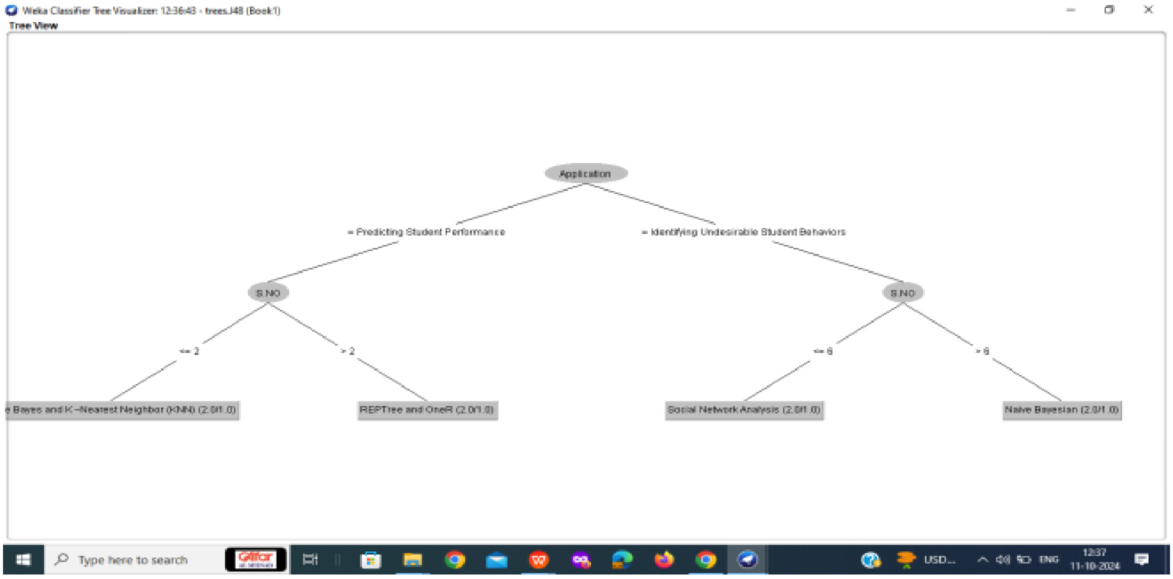

Figure 3: Generate decision tree j48 to find predicting student performance

The image displays another decision tree visualization created using the J48 algorithm in Weka, which classifies instances into two main collections: Anticipate Educatee Show and Characteristic Desirable Student Behaviors. Below is an explanation of the decision tree:

Root Node (Application): The starting point for classification is based on the attribute "S.NO" (presumably a feature in the dataset). The decision on this feature determines the further splits.

Splits:

If "S.NO" is less than or equal to 2, the left branch leads to a leaf node that applies Naive Bayes and K-Nearest Neighbor (KNN) methods. This branch classifies two instances correctly with the path value of "2.0/1.0."If "S.NO" is greater than 2, the right branch continues to another split based on the same "S.NO" attribute. If "S.NO" is less than or equal to 6, it leads to a leaf node that applies Social Network Analysis, which also classifies two instances correctly. If "S.NO" is greater than 6, the leaf node applies the Naive Bayesian method and correctly classifies another two cases.

Leaf Nodes: The tree ends at several leaf nodes where different algorithms are applied, such as Naive Bayes and KNN, for values less than or equal to 2.REPTree and OneR for values greater than 2. Social Network Analysis for values less than or equal to 6 in the right branch. Naive Bayesian for values greater than 6 in the right branch. Each path in the decision tree represents a sequence of conditions based on the value of "S.NO" and leads to different models or methods applied for classification. This decision tree indicates that a combination of devices' basic cognitive process know-how, such as

Naive Bayes, KNN, and Social Network Analysis, is used to classify student performance and behaviors effectively.

Algorithm Metrics Results

Top Performers: Naive Bayes, Random Forest, and OneR all achieved 100% accuracy, with Naive Bayes and OneR having perfect error metrics (0), indicating ideal performance.Moderate Performer: J48 showed 50% accuracy, which is much lower than the top performers, and also exhibited some errors, though not as high as REPTree or Jrip.Poor Performers: REPTree and Jrip performed very poorly, with 12.5% accuracy and very high error values across multiple metrics. Their Kappa statistic of 0 suggests their predictions were essentially random.

Correctly Classified Instances	Incorrectly Classified Instances	Kappa statistic	Mean absolute error	Root mean squared error	Relative absolute error	Root relative squared error	Instances	algorithm
100	0	1	0	0	0.0001	0.0194	0.0362	Naive Bayes Classifier
100	0	1	0	0.0847	0.1386	38.7143	41.9228	Random Forest
12.5	87.5	0	0.2188	0.3307	100	100	8	REPTree
50	50	0.4286	0.125	0.25	57.1429	75.5929	8	J48
100	0	1	0	0	0	0	8	OneR
12.5	87.5	0	0.2188	0.3307	100	100	8	Grip

Conclusion and Future Work

Lately, with the development of accumulation excavation applications in learning geographic areas, this article discerned the fewest effectual proficiency for each one of these Instruction data mining utilizes. The standing of this assessment report lies connected to the agreed assessment criteria for the questioning of the various proficiency for each education data mining request. Prevision correctness is used as a fact to measure the efficiency of the calculated proficiency. This physical element that the effectual method in cardinal utilization does not necessarily mean normal information technology will be stunning along with some other usage. Hence, the study should be promoted for each of the EDM later in the day to more accurately denote the most effective method. Naive Bayes and OneR provided the best performance with perfect accuracy and low error metrics. Random Forest also achieved 100% accuracy but had slightly higher errors than Naive Bayes.J48 performed moderately well with 50% accuracy. Still, REPTree and Jrip were ineffective, showing very poor classification accuracy and high error rates.

References

[1] K. Karthikeyan and K. PALANIAPPAN, "On Improving Student Performance Prediction in Education Systems Using Enhanced Data Mining Techniques," International Journal of Advanced Research in Computer Science and Software Engineering, vol. 7, no. 5, 2017.

[2] B. Bakhshinategh, O. R. Zaiane, S. ElAtia, and D. Ipperciel, "Educational Data Mining Applications and Tasks: A Survey of the Last 10 Years," Education and Information Technologies, vol. 23, no. 1, pp. 537–553, 2018.

[3]I. A. A. Amra and A. Y. Maghari, "Students Performance Prediction Using Known and Naïve Bayesian," in 2017 8th International Conference on Information Technology (ICIT), pp. 909–913, IEEE, 2017.

[4] H. Almarabeh, "Analysis of Students' Performance by Using Different Data Mining Classifiers," International Journal of Modern Education and Computer Science, vol. 9, no. 8, p. 9, 2017.

[5] H. Mousa and A. Y. Maghari, "School Students' Performance Predication Using Data Mining Classification," vol. 6, no. 8, 2017.

[6] B. Kapur, N. Ahluwalia, and R. Sathyaraj, "Comparative Study on Marks Prediction Using Data Mining and Classification Algorithms," International Journal of Advanced Research in Computer Science, vol. 8, no. 3, 2017.

[7] A. U. Khasanah et al., "A Comparative Study to Predict Student's Performance Using Educational Data Mining Techniques," in IOP Conference Series: Materials Science and Engineering, vol. 215, p. 012036, IOP Publishing, 2017.

[8] M. MAKHTAR, H. NAWANG, and S. N. WAN SHAMSUDDIN, "Analysis on Students' Performance Using Naïve Bayes Classifier.," Journal of Theoretical & Applied Information Technology, vol. 95, no. 16, 2017.

[9] S. Hussain, N. A. Dahan, F. M. Ba-Alwib, and N. Ribata, "Educational Data Mining and Analysis of Students' Academic Performance Using Weka," Indonesian Journal of Electrical Engineering and Computer Science, vol. 9, no. 2, pp. 447–459, 2018.

[10] Y. Salal, S. Abdullaev, and M. Kumar, "Educational Data Mining: Student Performance Prediction in Academic," IJ of Engineering and Advanced Tech, vol. 8, no. 4C, pp. 54–59, 2019.

[11] K. Agarwal, E. Maheshwari, C. Roy, M. Pandey, and S. S. Rautray, "Analyzing Student Performance in Engineering Placement Using Data Mining," in

Proceedings of International Conference on Computational Intelligence and Data Engineering, pp. 171–181, Springer, 2019.

[12] A. I. Adekitan and O. Salau, “The Impact of Engineering Student’s Performance in the First Three Years on their Graduation Result Using Educational Data Mining,” Heliyon, vol. 5, no. 2, p. e01250, 2019.

[13] A. I. Adekitan and E. Noma-Osaghae, “Data Mining Approach to Predicting the Performance of the First-Year Student in a University Using the Admission Requirements,” Education and Information Technologies, vol. 24, no. 2, pp. 1527–1543, 2019.

[14] M. R. I. Rifat, A. Al Imran, and A. Badrudduza, “Educational Performance Analytics of Undergraduate Business Students,” International Journal of Modern Education and Computer Science, vol. 11, no. 7, p. 44, 2019.

[15] H. Alhakami, T. Alsubait, and A. Aljarallah, “Data Mining for Student Advising,” International Journal of Advanced Computer Science and Applications, vol. 11, no. 3, 2020.

[16] J. Bayer, H. Bydzovsk´a, J. G´eryk, T. Obsivac, and L. Popelinsky, “Predicting Drop-out from Social Behaviour of Students.,” International Educational Data Mining Society, 2012.

[17] C. E. L. Guar´ın, E. L. Guzm´an, and F. A. Gonzalez, “A Model to Predict Low Academic Performance at a Specific Enrollment Using Data Mining,” IEEE Revista Iberoamericana de Tecnologias del Aprendizaje, vol. 10, no. 3, pp. 119–125, 2015.

[18] S. S. Athani, S. A. Kodli, M. N. Banavasi, and P. S. Hiremath, “Student Academic Performance and Social Behavior Predictor Using Data Mining Techniques,” in 2017 International Conference on Computing, Communication and Automation (ICCCA), pp. 170–174, IEEE, 2017.

[19] J. Pattanaphanchai, K. Leelertpanyakul, and N. Theppalak, “The Investigation of Student Dropout Prediction Model in Thai Higher Education Using Educational Data Mining: A Case Study of Faculty of Science, Prince of Songkla University,” Journal of University of Babylon for Pure and Applied Sciences, vol. 27, no. 1, pp. 356–367, 2019.