East African Scholars Journal of Engineering and Computer Sciences
Abbreviated Key Title: East African Scholars J Eng Comput Sci
ISSN: 2617-4480 (Print) & ISSN: 2663-0346 (Online)
Published By East African Scholars Publisher, Kenya

Volume-7 | Issue-7 | Oct-2024 | DOI: https://doi.org/10.36349/easjecs.2024.v07i07.001

Original Research Article

WEKA: A Tool for Shaping Effective Student Learning Strategies for College School Leavers

Aiyegbusi Akinyemi Edward[1*], Adeyemi Toyin Sidikat[1], Agoi Adeolu Moses[1], Kehinde Esther Olabisi[1]

[1]Department of Computer Science, College of Information and Technology Education, Lagos State University of Education, Lagos Nigeria

Article History
Received: 15.08.2024
Accepted: 23.09.2024
Published: 14.10.2024

Journal homepage:
https://www.easpublisher.com

Quick Response Code

Abstract: As college school leavers transition into higher education or the workforce, they face numerous challenges in adapting to new learning environments and expectations. Effective learning strategies are crucial to ensuring their success in these new settings. This paper explores how WEKA, a powerful data mining tool, can be leveraged to shape and enhance student learning strategies for college school leavers. By analysing educational data, WEKA can provide insights into the most effective learning methods, helping educators tailor their approaches to meet the specific needs of these students. The paper discusses the application of WEKA in educational data mining, highlights its capabilities, and presents case studies demonstrating its effectiveness in shaping learning strategies.
Keywords: WEKA, Data Mining, Student Learning, School Leavers, Mining Tools.

INTRODUCTION

The transition from secondary education to higher education or the workforce is a significant milestone in a student's life. College school leavers often face difficulties in adjusting to new academic demands, time management, and self-directed learning. To support these students, educators, and institutions must develop and implement effective learning strategies that cater to their unique needs. In recent years, data mining tools like WEKA have emerged as valuable resources in the field of educational data mining (EDM) (Romero & Ventura, 2010). WEKA, developed by the University of Waikato, is a collection of machine-learning algorithms for data mining tasks. It provides a comprehensive suite of tools for data analysis, visualization, and model evaluation, making it an ideal choice for educators seeking to improve student learning outcomes (Peña-Ayala, 2014). The use of effective study strategies is important for academic achievement, yet research indicates that students often use relatively ineffective learning strategies (McDaniel, Einstein & Een, 2021). Cognitive psychological research from the last decades has shown that learning strategies that create desirable difficulties during learning, e.g., practice testing, are most effective for long-term learning outcomes. However, there is a paucity of research on how to effectively translate these insights into training students in higher education (Wei, Shi, MacLeod & Yang 2022). Various scenarios to improve the education system have been carried out. One of them is the application of educational data mining to gain the student's academic achievement, also decreasing the risk of drop out (Triayudi, Widyarto & Rosalina, 2022; Baker & Siemens, 2014). Cognitive psychological research from the last decades has shown that learning strategies that create desirable difficulties during learning, e.g., practice testing, are most effective for long-term learning outcomes. However, there is a paucity of research on how to effectively translate these insights into training students in higher education (Felicitas *et al.*, 2020). Mavo & Mcgrath, (2021) review strategies for effective design in online instruction. The authors explore the traditional debate between advocates and critics of online education and discuss effectiveness in retention, engagement, and overall academic performance.

Research of Zhou & Wang, (2023) suggests that student engagement is a key determinant of students' learning effectiveness. More engaged students are expected to have positive academic and non-academic outcomes, including educational achievement, persistence and completion of schoolwork, and physical or psychological well-being. However, engaging students in online classes is usually more challenging than in the traditional face-to-face environment, where

***Corresponding Author:** Aiyegbusi Akinyemi Edward
Department of Computer Science, College of Information and Technology Education, Lagos State University of Education, Lagos Nigeria

students are more likely to participate actively. Therefore, the following research questions guide the study:

- Is there any significant difference between the use of Tree Classifier called J48 pruned tree, Decision Table Classifier, K-Nearest Neighour Classifier, Multilayer-Perceptron called Neural Network in predicting secondary school students learning outcomes?
- Which of the algorithms is the best predictor among them?

LITERATURE REVIEW

Background on WEKA

WEKA (Waikato Environment for Knowledge Analysis) is an open-source software platform that offers a variety of machine learning algorithms and data preprocessing tools. Initially developed in the late 1990s, WEKA has since become one of the most widely used tools in the field of data mining and machine learning (Hall *et al*., 2009). Its user-friendly interface, extensive documentation, and robust functionality have made it a popular choice among researchers, educators, and data scientists. This work shows the use of WEKA, a tool that implements the most common machine learning algorithms, to perform a Text Mining analysis on a set of documents. Applying these methods requires initial steps where the text is converted into a structured format. Both the processing phase and the analysis of the transformed dataset, using classification and clustering algorithms, can be carried out entirely with this tool, rigorously and simply. The work describes the construction of two classification models starting from two different sets of documents. These models are not meant to be good or realistic, but just illustrate how WEKA can be used for a Text Mining analysis (Holmes, Donkin & Witten 1994). Triayudi, Widyarto & Rosalina, (2022) classified all possibilities related to predicting students' academic performance, the data was collected from three private universities in Jakarta, where it consists of academic, social, and economic information, as the demographics of 350 students with 23 attributes. Educational Data Mining applied a WEKA's tool that takes a rule in this study, while the classification phases applied PART, BayesNet, Random Tree, and J48 as their methods of classification. Attributes that have a significant influence on the classification process are selected as a medium of classification. Srivastava, Sharma & Kumar, (2020) explained that WEKA (Waikato Condition for Information Investigation) is a broad suite of Java class libraries that acknowledge many top-level PC based knowledge and information mining checks. It gives executions of reenacted knowledge estimations that you can without a great deal of a stretch apply to a dataset. It moreover combines a game plan of devices for developing datasets. According to Saleh *et al*., (2020) Data mining is characterized as searching for useful information through very large data sets. Some of the key and most common techniques for data mining are association rules, classification, clustering, prediction, and sequential models. WEKA is a potent artificial intelligence data platform. One may simply study data with WEKA AI software, which also provides tools for data visualization and analysis methods. It is dependable and offers the necessary performance and scalability. It can be used in the cloud, on-site, or bursting between platforms (Wei *et al*., 2022). Three hundred students' socioeconomic, demographic, and academic data were included in the data set, which included twenty-four variables. The J48, PART, Random Forest, and Bayes Network Classifiers were the four classification techniques applied. WEKA was the data mining tool utilized. Using the technique, the highly influential attributes were chosen. The students' final semester results in our dataset are most influenced by the internal assessment attribute in the ongoing evaluation process. Based on accuracy and classifier errors, the results demonstrated that random forest performs better than the other classifiers. The optimal rules were also shown after the Apriori method was utilized to locate the association rule mining among all the attributes.

Zhou & Wang, (2023) evaluated the data of computer engineering students using classification, a data mining technique, to determine whether students require academic counseling in the field. The classification model was constructed taking into account five attributes. The model's classifier, a decision tree, was selected. Cross-validation was used to compare the accuracy of the decision tree algorithms Random Tree, RepTree, and J48. Random Tree produced the highest accuracy of 75.188%. The categorization model was created using the data mining program Waikato Environment for Knowledge Analysis (WEKA). The Research Project Grade Predictor application was created using Visual C#, and its algorithm made use of the classification rules that were taken out of the decision tree. The program made it simple for research teachers or advisers to identify students who are expected to perform poorly and who require further attention.

Felicitas *et al*., (2020) designed an intervention program aiming to create awareness about, foster reflection on, and stimulate practice of effective learning strategies. In a first examination of the pilot intervention (N=47), we tested the effects of the intervention on metacognitive knowledge and self-reported use of effective learning strategies during self-study, using a control-group mixed-methods design. The intervention program had positive effects on knowledge about effective learning strategies and increased the use of practice testing. Qualitative interview results suggested that to sustainably change students' learning strategies, we may consider tackling their uncertainty about effort and time, and increase availability of practice questions.

Donatella & Martina, (2021) showed the use of WEKA, a tool that implements the most common machine learning algorithms, to perform a Text Mining analysis on a set of documents. According to them, applying those methods requires initial steps where the text is converted into a structured format. Both the processing phase and the analysis of the transformed dataset, using classification and clustering algorithms, can be carried out entirely with this tool, in a rigorous and simple way. The work describes the construction of two classification models starting from two different sets of documents. These models were not meant to be good or realistic, but just illustrate how WEKA can be used for a Text Mining analysis.

Key Features of WEKA

1. Extensive Algorithm Library:

WEKA includes a wide range of machine learning algorithms, including classification, regression, clustering, association, and feature selection methods. This diversity allows users to experiment with different approaches to identify the most effective models for their data.

2. Data Preprocessing Tools:

WEKA provides tools for data cleaning, normalization, and transformation, which are essential steps in preparing data for analysis. These preprocessing capabilities ensure that the data is suitable for the chosen algorithms.

3. Visualization Capabilities:

WEKA offers various visualization options, enabling users to explore and interpret their data effectively. Visualization helps in understanding patterns and trends that may not be immediately apparent from the raw data.

4. Evaluation and Validation:

WEKA allows users to evaluate the performance of their models using various metrics, such as accuracy, precision, recall, and F1-score. Additionally, it supports cross-validation techniques, which help in assessing the generalizability of the models.

5. Open-Source and Extensible:

As an open-source tool, WEKA is freely available for use and modification. Its extensibility allows users to develop and integrate custom algorithms or tools, further enhancing its utility.

WEKA in Educational Data Mining

Educational Data Mining (EDM) is an interdisciplinary field that applies data mining techniques to educational data to gain insights into teaching and learning processes. EDM aims to improve academic outcomes by analyzing data from various sources, such as student performance, behavior, and demographics. WEKA has proven to be a valuable tool in this field, allowing educators to mine educational data and extract actionable insights.

Applications of WEKA in Shaping Learning Strategies

1. Identifying At-Risk Students:

WEKA can be used to analyze historical student data to identify patterns associated with academic success or failure. By applying classification algorithms, educators can predict which students are at risk of underperforming or dropping out. Early identification allows for timely interventions, such as personalized tutoring or counseling, to support these students.

2. Personalizing Learning Experiences:

By analyzing student data, WEKA can help educators understand the learning preferences and needs of individual students. Clustering algorithms can group students based on their learning styles, enabling educators to design personalized learning experiences that cater to the strengths and weaknesses of each group.

3. Optimizing Curriculum Design:

WEKA can analyze data on student performance in different courses or modules to identify which elements of the curriculum are most effective. Regression analysis can reveal the impact of specific instructional strategies on student outcomes, allowing educators to refine their curriculum to enhance learning effectiveness.

4. Improving Assessment Techniques:

WEKA's data analysis capabilities can be used to evaluate the effectiveness of different assessment methods. For example, educators can analyze the correlation between assessment types (e.g., quizzes, projects, exams) and student performance to determine the most accurate predictors of student success. This information can guide the development of more effective assessment strategies.

5. Supporting Decision-Making:

WEKA's ability to handle large datasets and provide meaningful insights makes it an invaluable tool for educational decision-makers. By analyzing data on student demographics, performance, and engagement, administrators can make informed decisions about resource allocation, program development, and policy implementation.

Identifying At-Risk Students

In a study conducted at a university in New Zealand, WEKA was used to analyze historical student data to identify factors that contributed to academic failure. The researchers applied various classification algorithms, including decision trees and logistic regression, to predict student outcomes based on factors such as attendance, assignment submission rates, and previous academic performance. The results showed that WEKA could accurately identify at-risk students, allowing the university to implement targeted interventions that significantly improved retention rates. For predicting heart problems, Saleh *et al*., (2020) postulated that data extraction algorithms like K-star, J48, SMO, Naïve Bayes, MLP, Random Forest, Bayes Net, and REPTREE are used for this study (Weka 3.8.3) software. The results of the predictive accuracy, the ROC curve, and the AUC value are combined using a standard set of data and a collected dataset. By applying different data mining algorithms, the patient data can be used for diagnosis as training samples. Data mining functions and techniques are used to identify the level of risk factors to help the patients in taking precautions in advance to save their life.

Personalizing Learning Experiences

A high school in Australia implemented WEKA to analyze student data and develop personalized learning plans. The school used clustering algorithms to group students based on their learning preferences, such as visual, auditory, or kinesthetic learning styles. Teachers then tailored their instructional methods to match the preferences of each group. The implementation of these personalized strategies resulted in improved student engagement and academic performance, demonstrating the effectiveness of WEKA in shaping learning strategies.

Optimizing Curriculum Design

A community college in the United States used WEKA to analyze data on student performance in various courses. By applying regression analysis, the college identified specific instructional methods and course materials that were most strongly correlated with student success. The findings led to a revision of the curriculum, with an emphasis on incorporating more interactive and hands-on learning activities. As a result, the college observed a significant improvement in student outcomes, particularly in courses that had previously been challenging for students. Srivastava, Sharma & Kumar, (2020) survey the preprocess a dataset, feed it into a learning plan, and separate the subsequent classifier and its presentation-all without framing any program code whatsoever utilizing Weka. Weka is wholeheartedly accessible on the Internet and goes with another substance on information mining and man-made insight which reports and absolutely clarifies all the figuring it contains. Applications made utilizing the Weka class libraries can be run on any PC with an Internet investigating limit.

Challenges and Considerations

While WEKA offers numerous benefits for shaping effective student learning strategies, there are also challenges and considerations that educators must be aware of:

1. **Data Quality**: The accuracy and reliability of the insights derived from WEKA depend on the quality of the data used. Incomplete, inaccurate, or biased data can lead to misleading conclusions, which may negatively impact student learning outcomes.
2. **Technical Expertise**: Using WEKA effectively requires a certain level of technical expertise in data mining and machine learning. Educators and administrators may need training or support to fully utilize the tool's capabilities.
3. **Ethical Considerations**: The use of student data for analysis raises ethical concerns related to privacy and consent. Educators must ensure that data is used responsibly and that students' rights are protected.
4. **Scalability**: While WEKA is a powerful tool, its effectiveness can be limited by the size and complexity of the dataset. For large-scale implementations, more advanced or specialized tools may be required.

METHODOLOGY

This study employed Ex post facto research design which compares groups of already existing characteristics based on a dependent variable. The study used Students' Mock Examination results and WASSCE results of graduates from secondary schools. The population of this study comprised of 122,498 students, made of 58,971 male students and 63,527 female students, that is, the total number of all secondary school students' who have written both Mock examination and WASSCE in Ogun State in 2021/2022 academic session. The researcher ensure samples were selected from all the local government areas of the total population of secondary schools in Ogun State. The first sample technique used was the purposive sampling technique. This was necessarily to identify and select only the schools that their students wrote computer studies in Mock and WASSC examination for the research year which was 2021 /2022 academic session. To arrive at the expected 1200 participants, an extra pellet was picked from 10 schools. The SSCE Result Dataset Collection Template was validated by the researcher's supervisor to ensure content and construct validity of the instrument. In order to ensure face validity of the instrument used, copies of the template were given to teachers that were not selected for the study as well as researcher's colleagues to seek their understanding of the semantical structure of the instrument before it was finally used to collect data for the study. In order to determine the extent to which the instrument measure what was supposed to measure, the instrument was rated and subjected to test-retest method where (20) twenty copies of the instrument (template) were used to collect Mock and WASSCE results from selected schools within the population. The rated scores obtained were tested for reliability and internal consistence using Chronbach alpha analyses. The Chronbach Alpha reliability coefficients computed value were found to be 0.871. The raw data collected via SSCE Result Dataset Collection Template were collated and transformed into readable format useful for the analysis. The datasets from SSCE Result Dataset Collection Template were subjected to software application packages WEKA (using Tree Classifier called J48 pruned tree, Decision Table Classifier, K-

Nearest Neighbour Classifier, Multilayer-Perceptron called Neural Network).

RESULTS

Presentation of Demographic Information of the Students

Table 1: Gender difference of the students

	Frequency	Percent	Cumulative Percent
Male	523	43.6	43.6
Female	677	56.4	100.0
Total	1200	100.0	

Table 1 shows gender difference of sampled students that sat for both Mock and WASSC Examinations in the year 2021/2022 academic session. Out of 1,200 sampled, 523 (44%) students were male while the remaining 677 (56%) students were female. This reveals that female students were more than male counterparts.

Table 2: School Location of the students

	Frequency	**Percent**	**Cumulative Percent**
Rural	465	38.8	38.8
Urban	735	61.3	100.0
Total	1200	100.0	

Table 2 explains the school location of the students that sat for both Mock and WASSC Examinations in the year 2021/2022 academic session. It is revealed that 465 (39%) students were from rural location while 735 (61%) students' schools were located at urban settlement. The statistics show that students from urban were more than those from rural locations.

Table 2: School Type of the students

	Frequency	**Percent**	**Cumulative Percent**
Private Schools	375	31.25	31.25
Public Schools	825	68.75	100.0
Total	1200	100.0	

Table 3 gives detailed statistics of the number of students that participated in the Mock and WASSC examinations in the year 2021/2022 academic session based on school type. The frequency of the table shows that 375 (31%) students were from private schools while 825 (69%) students' schools were from public schools. The statistics shows that students from public schools were more than those from private schools

Use of WEKA for Prediction

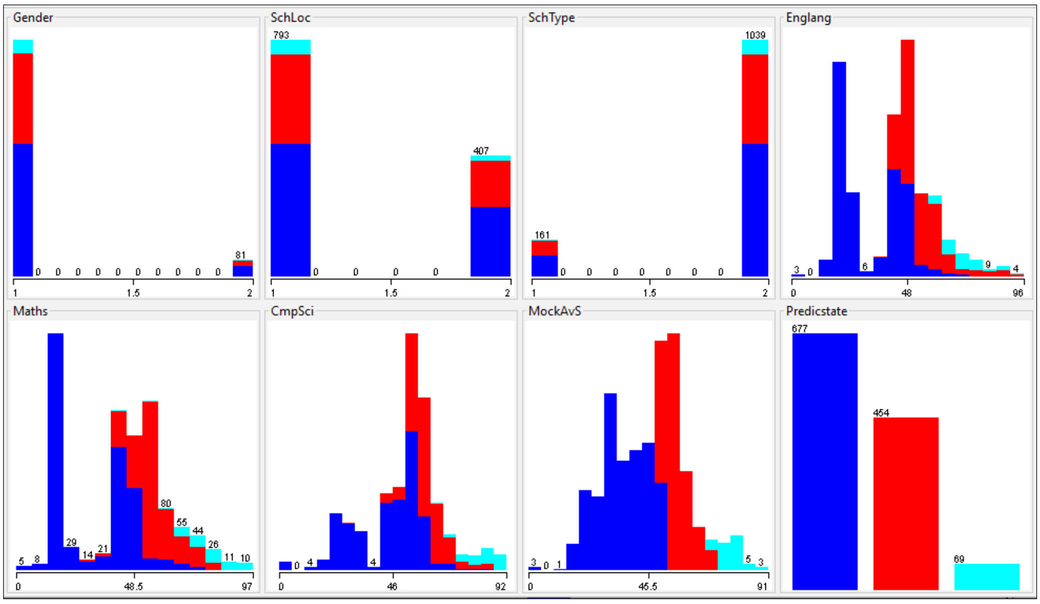

Figure 1: All attributes used in WEKA

A. WEKA (Using Tree Classifier called J48 pruned tree)

It was shown that 8 attributes and 1200 instances tested, the attributes include: Gender, SchLoc, SchType, Englang, Maths, CmpSci, MockAvS, and Predict state. It produced 5 number of Leaves and 9 size of the tree (See fig.4.6). The time taken to build model was 0.03 seconds while the time taken to test the model on test split was 0.48 seconds.

Summary of prediction

Correctly Classified Instances	357	99.1667 %
Incorrectly Classified Instances	3	0.8333 %
Kappa statistic	0.9839	
Mean absolute error	0.0085	
Root mean squared error	0.0663	
Relative absolute error	2.4024%	
Root relative squared error	15.9363%	
Precision value	0.992	
Recall	0.992	
F-Measure	0.922	

Confusion Matrix

```
a     b      c   <-- classified as
202   3      0 |  a = Evidence to pass with poor grade
0     143    0 |  b = Evidence to pass with average grade
0     0     12 |  c = Evidence to pass with good grade
```

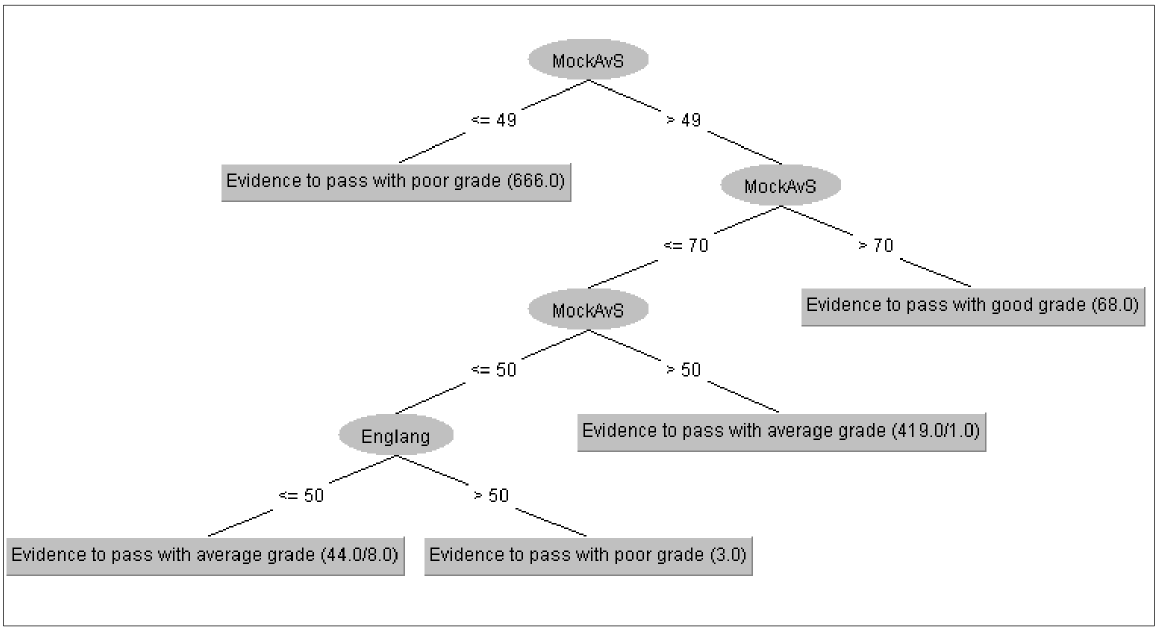

Figure 2: Tree model example

B. WEKA (Using Decision Table Classifier)

The Number of training instances was 1200 with 17 rules. It takes 0.22 seconds to build the model and 0.46 seconds to test model on test split.

Summary of Prediction

Correctly Classified Instances 357 99.1667 %
Incorrectly Classified Instances 3 0.8333 %
Kappa statistic 0.9839
Mean absolute error 0.0312
Root mean squared error 0.0849
Relative absolute error 8.8202 %
Root relative squared error 20.4067 %
Precision value = 0.992,
Recall = 0.992,
F-Measure = 0.992

Confusion Matrix

a b c <-- classified as
202 3 0 | a = Evidence to pass with poor grade

0 143 0 | b = Evidence to pass with average grade
0 0 12 | c = Evidence to pass with good grade

C. WEKA (Using K-Nearest Neighour Classifier)

The time taken to build model was 0.01 seconds while the time taken to test the model on test split was 0.62 seconds

Summary of Prediction

Correctly Classified Instances 351 97.5 %
Incorrectly Classified Instances 9 2.5 %
Kappa statistic 0.9512
Mean absolute error 0.0181
Root mean squared error 0.1289
Relative absolute error 5.0997 %
Root relative squared error 30.9886 %
Precision value = 0.975,
Recall = 0.975,
F-Measure = 0.974

Confusion Matrix

a b c <-- classified as
204 1 0 | a = Evidence to pass with poor grade
5 138 0 | b = Evidence to pass with average grade
0 3 9 | c = Evidence to pass with good grade

D. WEKA (Using Multilayer-Perceptron called Neural Network)

The time taken to build model was 1.35 seconds while the time taken to test the model on test split was 0.46 seconds

Summary of Prediction

Correctly Classified Instances 357 99.1667 %
Incorrectly Classified Instances 3 0.8333 %
Kappa statistic 0.9839
Mean absolute error 0.0051
Root mean squared error 0.0444
Relative absolute error 1.4368 %
Root relative squared error 10.6862 %
Precision value = 0.992,
Recall = 0.992,
F-Measure = 0.992

Confusion Matrix

a b c <-- classified as
202 3 0 | a = Evidence to pass with poor grade
0 143 0 | b = Evidence to pass with average grade
0 0 12 | c = Evidence to pass with good grade

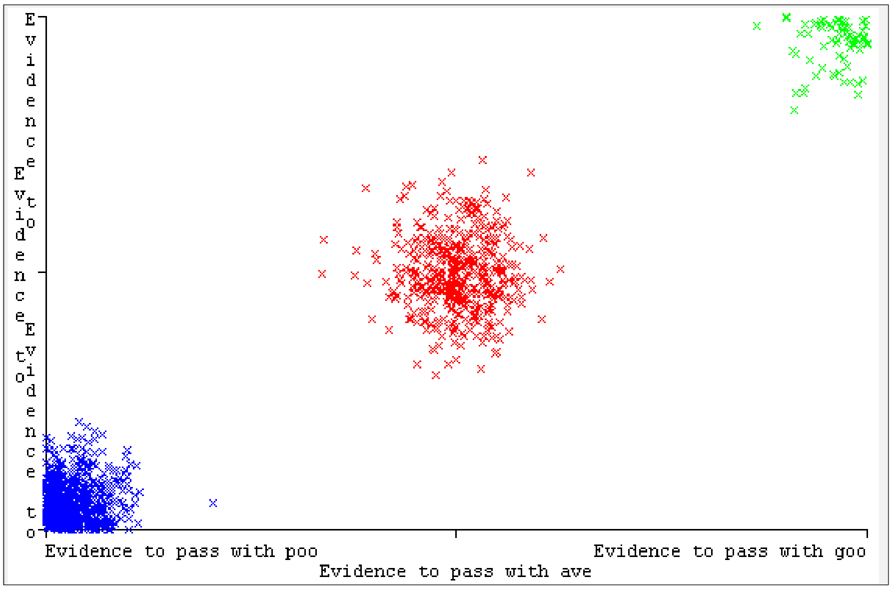

Figure 3: Graphical classification sample

Fig. 3 shows clear classification of the predictions: Evidence to pass with good grade (Green colour), Evidence to pass with average grade (Red colour), and Evidence to pass with poor grade (Blue colour).

Table 4: Comparison of the results and decision within WEKA software algorithms

Name	*Prediction Accuracy %*	*Precision value*	*Recall value*	*F-Measure*	*MAE*	*RMSE*	*RAE %*	*RRSE %*
Tree Classifier	99.1667	0.992	0.992	0.992	0.0085	0.0663	2.4	15.94
Decision Table Classifier	99.1667	0.992	0.992	0.992	0.0312	0.0849	8.82	20.41
K-Nearest Neighour Classifier	97.5	0.975	0.975	0.974	0.0181	0.1289	5.10	30.99
Multilayer-Perceptron	99.1667	0.992	0.992	0.992	0.0051	0.0444	1.44	10.69
Final Decision	98.75	98.77	98.77	98.75	0.016	0.081	4.44	19.51

Generally, the dataset ratio splits were based on 70% training set and 30% testing set. The accuracies obtained across four classifiers showed that all algorithms can predict the students learning outcomes correctly well as shown in the table 4.

Research Question Testing

- Is there any significant difference between the use of Tree Classifier called J48 pruned tree, Decision Table Classifier, K-Nearest Neighbour Classifier, Multilayer-Perceptron called Neural Network in predicting secondary school students learning outcomes?

The statistical difference from table 4 show that the observed difference is marginal. Therefore, there is no significant difference between the use of Tree Classifier called J48 pruned tree, Decision Table Classifier, K-Nearest Neighbour Classifier, Multilayer-Perceptron called Neural Network in predicting secondary school students learning outcomes." Is not rejected.

- Which of the algorithms is the best predictor among them?

The statistical results from table 4 show that Decision Table Classifier and, K-Nearest Neighbour Classifier, Multilayer-Perceptron called Neural Network performed better with 99%. So, the two classifiers were considered the best(s)

Discussion of Findings

The research question on that says "Is there any significant difference between the use of Tree Classifier called J48 pruned tree, Decision Table Classifier, K-Nearest Neighbour Classifier, Multilayer-Perceptron called Neural Network in predicting secondary school students learning outcomes?" proved that the use of WEKA for predicting secondary school students learning outcomes is a great advantage and opportunity. This study is in line with the work of Zhou & Wang, (2023) that said "The use of WEKA made it simple for research teachers or advisers to identify students who are expected to perform poorly and who require further attention". Also, Srivastava, Sharma & Kumar, (2020) claimed that the use of WEKA gives executions of reenacted knowledge estimations that you can without a great deal of a stretch apply to a dataset. Similarly, among the four algorithms used decision tree as well as neural network performed well with 99% accuracy. This showed that one model is likely to performed better on a dataset than other.

Conclusion

WEKA is a versatile and powerful tool that can significantly enhance the effectiveness of student learning strategies for college school leavers. By leveraging WEKA's data mining capabilities, educators can gain valuable insights into student behavior, preferences, and performance, allowing them to tailor their instructional methods to meet the unique needs of each student. The case studies presented in this paper demonstrate the potential of WEKA to improve educational outcomes and support the successful transition of college school leavers into higher education or the workforce. However, to fully realize the benefits of WEKA, educators must address challenges related to data quality, technical expertise, and ethical considerations. As educational institutions continue to embrace data-driven decision-making, tools like WEKA will play an increasingly important role in shaping the future of education. By harnessing the power of data, educators can create learning environments that are more personalized, effective, and equitable, ultimately helping students achieve their full potential.

References

- Baker, R. S., & Siemens, G. (2014). *Educational Data Mining and Learning Analytics. In Cambridge Handbook of the Learning Sciences* (pp. 253-272). Cambridge University Press.
- Donatella, M., & Martina, R. (2021). Text categorization with WEKA: A survey, *Machine Learning with Application, 4*(1), Doi: https://doi.org/10.1016/j.mlwa.2021.100033.
- Felicitas, B., Mirjam, G. A., oude, E., Pauline, A., & Anique, B. H. de. B. (2020) Fostering Effective Learning Strategies in Higher Education – A Mixed-Methods Study. *Journal of Applied Research in Memory and Cognition, 9*(2), 186-203, Doi: https://doi.org/10.1016/j.jarmac.2020.03.004
- Hall, M., Frank, E., Holmes, G., Pfahringer, B., Reutemann, P., & Witten, I. H. (2009). The WEKA Data Mining Software: An Update. *ACM SIGKDD Explorations Newsletter, 11*(1), 10-18.
- Holmes, G., Donkin, A., & Witten, I. H. (1994). WEKA: a machine learning workbench. *Proceedings of ANZIIS '94 - Australian New Zealnd Intelligent Information Systems Conference*, 357-361.
- Mavo Navarro, J., & Mcgrath, B. (2021). Strategies for Effective Online Teaching and Learning: Practices and Techniques with a Proven Track of Success in Online Education. Doi:10.4018/978-1-7998-8275-6.ch029.
- McDaniel, M. A., Einstein, G. O., & Een, E. (2021). Training College Students to Use Learning Strategies: A Framework and Pilot Course. *Psychology Learning & Teaching, 20*(3), 364-382. https://doi.org/10.1177/1475725721989489
- Peña-Ayala, A. (2014). Educational data mining: A survey and a data mining-based analysis of recent works. *Expert Systems with Applications, 41*(4), 1432-1462.
- Romero, C., & Ventura, S. (2010). Educational Data Mining: A Review of the State of the Art. IEEE Transactions on Systems, Man, and Cybernetics, *Part C (Applications and Reviews), 40*(6), 601-618.

- Saleh, B., Saedi, A., Al-aqbi, A., & Salman, L. (2020). *A Review Paper: Analysis of Weka Data Mining Techniques for Heart Disease Prediction System*.
- Srivastava, S., Sharma, Y., & Kumar, S. (2020). Using Of WEKA Tool in Machine Learning: A Review. 29. 8604-8614.
- Triayudi, A., Widyarto, W. O., Rosalina, V. (2022). Analysis of Educational Data Mining Using WEKA for the Performance Students Achievements. In: Triwiyanto, T., Rizal, A., Caesarendra, W. (eds) Proceedings of the 2nd International Conference on Electronics, Biomedical Engineering, and Health Informatics. Lecture Notes in Electrical Engineering, vol 898. Springer, Singapore. https://doi.org/10.1007/978-981-19-1804-9_1
- Wei, Y., Shi, Y., MacLeod, J., & Yang, H. H. (2022). Exploring the Factors That Influence College Students' Academic Self-Efficacy in Blended Learning: A Study from the Personal, Interpersonal, and Environmental Perspectives. *Sage Open*, *12*(2). https://doi.org/10.1177/21582440221104815
- Zhou, Z., & Wang, S. (2023). *Strategies for Online Student Engagement*.

Cite This Article: Aiyegbusi Akinyemi Edward, Adeyemi Toyin Sidikat, Agoi Adeolu Moses, Kehinde Esther Olabisi (2024). WEKA: A Tool for Shaping Effective Student Learning Strategies for College School Leavers. *East African Scholars J Eng Comput Sci*, *7*(7), 61-69.