International Journal of Scientific and Academic Research (IJSAR)

E-ISSN : 2583-0279

DOI: 10.54756/IJSAR.2025.6.1

Vol. 5 (6) October - 2025

Analyzing Teaching Performance of Instructors Using Data Mining Techniques

Fernandez. Mary Joy M.[1] & Maquiling. Kryzza J.[2]

Department of Graduate School

University of Immaculate Conception

Davao City, Philippines

ABSTRACT

To identify trends into instructor effectiveness, this paper utilizes data mining techniques on student evaluations of teaching quality. PCA and clustering were employed to identify distinct response patterns and underlying constructs in a set of student-faculty ratings on multiple performance dimensions. Three categories of students with different patterns in evaluation emerged using K-means clustering presenting consistent and inconsistent evaluations of teachers. PCA further reduced dimensionality by extracting a single criterion and examining the main variation in the data in a single predominant component that reflected standard evaluation biases. The results provide an evidence-based perspective on teaching quality, and identify specific strengths, and opportunities for improvement. This provides the basis for faculty development and evidence-based decision making in higher education.

Keywords: Clustering Analysis, Educational Data Mining, Evidence-Based Decisions, Faculty Development, Principal Component Analysis, Student Evaluation, Teaching Performance.

1.0 INTRODUCTION

1.1 Background and Context

One of the most critical activities in academic organizations is the assessment of teaching effectiveness since it affects the quality level of the teaching and students' attainment of it. Student evaluations of teaching (SET) are the main instrument used to evaluate teaching effectiveness, and their outcomes affect decisions regarding faculty development, promotions, and teaching assignment [1]. However, notwithstanding the precious information which could be obtained from SETs, analysis of great amounts of information collected from students' evaluations is often a difficult task. Most methods concentrate on single metrics or global behaviors, leaving behind subtle structures in data [2].

Recent progress in the field of educational data mining (EDM) provides promising tools to tackle this issue. Utilizing machine learning algorithms to mine student evaluation data, institutions could discover underlying patterns and valuable insights that conventional frameworks fail to detect, such as clustering and predictive modeling for better understanding of teaching performance patterns that can be discovered through students feedback [14]. Several factors such as course difficulty, instructor personality, and student engagement can have considerable impact on SET outcomes [9]. An examination of this kind with EDM can be useful to disclose such factors and offers a more detailed view on teaching quality.

The purpose of this study is to apply data mining technology, clustering and predictive modeling specifically, to consider the evaluation data of SET and derive the implication of SET on teaching performance. By examining trends in student ratings over a variety of dimensions, we discover typical patterns and deviances in the way that instructors are viewed [13]. The work also employs visualization via heatmaps and boxplots to aid in the interpretation of student ratings distribution [16]. The purpose is to identify teaching strengths and weaknesses that provide useful feedback for type of instruction and educational decision making.

In this way, the found study adds to a wider literature on educational analytics and provides an evidence-based way of how HEIs can develop data-informed strategies to improve quality of teaching and make informed decisions [15]. This study efficiently uses SET data via data mining in that it can discover the major trend of teaching performance effectively with implications for teaching improvement.

2.0 METHOD

The study applied unsupervised learning models (K-means clustering and PCA) to student evaluations of teaching performance. K-means cluster analysis applied to student responses to form distinct clusters for evaluation patterns and optimal number of clusters was determined using the Elbow Method. PCA was used to reduce the dimension of the data and discovering significant components which explained the maximum variability in the data as well as extracted underlying themes for the evaluations. These unsupervised learning techniques enabled us to understand evaluation trends without the need of labeled data. The R programming language was used for the analysis, along with a number of R packages: tidyverse for data manipulation, cluster for clustering analysis, factoextra for clustering and PCA results visualization, ggplot2 to create heatmaps and boxplots, stats to carry out PCA.

2.1 Data Collection and Preparation

The student course evaluation is being conducted on ninth and tenth weeks of each semester by the Dean in the College of Information and Communication technology of South East Asian Institute of Technology, Inc. These assessments are intended only for programming-related courses and are targeted for assessing the instructors performance in these courses.

The dataset contains student evaluations of teaching (SET) scoreings, ranking instructors on dimensions A-M, then dimensions N-R apply to the course and the section. The assessments were developed to measure aspects of the quality of teaching and learning in programming subjects, including the teacher's preparedness, knowledge of the subject, responsiveness and ability to communicate the course material and relation to theoretical and practical programming.

In the dataset, each row represents one student's ratings and is a view of the instructor's and the course's success in all these 18 criteria specific to the student. The ratings were recorded on a numerical Likert scale ranging from 1 (poorer performance) to 5 (better performance) in the condiciones. The variables reported present a dichotomy in terms of inference. The first group of attributes are the information from evaluation questionnaire questionnaire is where the students to anonymously 19 question about the instructor (A – M) and the course (N – R). The rates are numbers such as four-point Likart scale on 1 to 5 scale, denoting the more is the value the more a subject performs in each criterion. The questions of this form are presented in Table 1.

Table 1: Questions of the Evaluation Form

	Instructor
A	The instructor is prepared for each class.
B	The instructor demonstrates knowledge of the Subject.
C	The instructor has completed the whole course.
D	The instructor provides additional material apart from the textbook.
E	The instructor gives citations regarding current situations
F	The instructor communicates the subject matter effectively.
G	The instructor shows respect towards students and encourages class participation.
H	The instructor maintains an environment that is conductive to learning.

DOI: 10.54756/IJSAR.2025.6.1

I	The instructor arrives on time.
J	The instructor leaves class on time.
K	The instructor is fair in examination.
L	The instructor returns the quizzes, examinations, etc. in a reasonable amount of time.
M	The instructor was available during the specified office hours for consultations.
	Course
N	The subject matter presented in the course has increased your knowledge of the subject.
O	The syllabus clearly states course objectives, requirements, procedures and grading criteria.
P	The course integrates theoretical course concepts with real-world applications.
Q	The assignments and exams covered the materials presented in the course.
R	The course material is modern and updated.

2.2 Modeling and Evaluation Using Data Mining Techniques

The modeling part of the research was conducted in the R software with some of the major R packages (Kelley & Lai, 2012), including used in the research: tidyverse, stats, factoextra. Modeling and evaluation were based on data mining principles for systematic analysis of student evaluations of teaching performance. After the dataset was generated according to the study's methodology, several steps including the dimensionality reduction and clustering were conducted in order to identify important patterns and knowledge within the dataset.

PCA analysis was used for dimensionality, and to show the variables with the highest significance. Clustering analysis was conducted by K-means and elbow method was applied to identify the optimal number of clusters to analyze.

To evaluate models, visualization methods including scree plots and cluster plots were created, using the factoextra package in R, and within the clustering, crossvalidation was implemented to create clusters.

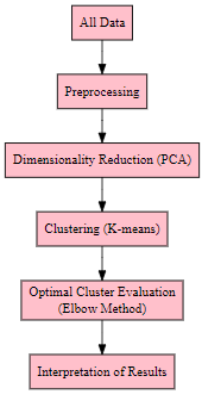

Figure 1: The Data Analysis Workflow for Teaching Performance Evaluation

The Silhouette Coefficient was selected as a clustering criterion to assess clustering performance. The Silhouette Coefficient is a measure of how close each point in one cluster is to the points in the neighboring clusters and thus provides a way to assess parameters such as number of clusters visually. It is calculated as:

$$S(i) = \frac{b(i) - a\,(i)}{\max(a(i), b(i))}$$

DOI: 10.54756/IJSAR.2025.6.1

where a(i) is the matrix of the mean distance between i and the all other points in the same cluster (intra-cluster distance); b(i) is the matrix of the mean distance between i and all the points in the nearby cluster (inter-cluster distance). The value for the coefficient lies in the range of −1to 1, where the value ≈ 1 stands a good clustered data point than the others, close to its own cluster. A score of 0 indicates that this point is on the exact same distance from 2 different clusters, if scores are negative, the point might be assigned to the wrong cluster. The Silhouette Coefficient yields an overall indication of how well clustered the data is, higher average Silhouette Coefficient values indicate that the objects are well matched to the cluster to which they are assigned [21].

All points were scored with this metric and used to evaluate the clustering performance as well as interpretation of the meaningless of the student-evaluation cluster. An increased average Silhouette Coefficient value indicates more distinct clusters, and stable clusters that are suitable for subsequent analysis [21].

3.0 RESULTS AND DISCUSSION

3.1 Result

3.1.1 Descriptive Analysis

The description analysis was also done by enabling the use of mean for all question sets (A-M) and (N-R) by their instructor-related and course related questions respectively across student ratings (S1-S596). Once separated the instructor questions from the course questions, the neural mean for each question was calculated in order to determine the neural average rating that they obtained.

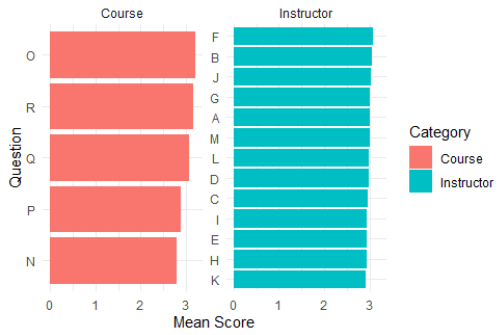

Figure 2: Mean Scores by Category

The plot shows average scores of questions concerning courses and instructors aggregated and compared by categories. RED bars: mean scores of the students The mean course-related questions' scores, represented by red bars, are approximately in the same range for the entire set of the criteria (N, O, P, Q, and R), with the lowest standard deviation, The combination of the questions concerning the course yielded the typical student's mean score for the initial data decomposition analysis that follows, indicating that the students rated these types of criteria (i.e., structure, content, organization etc.) in the same way. In contrast, there is more variability in the mean scores for teacher-related questions, indicated as blue bars. This could indicate that for certain elements of teaching performance such as engagement, preparedness, or feedback, students have different expectations or standards.

Clustering

The clustering was formed by K-means clustering of the dataset based on student responses. Initially, the student rating columns (S1 to S596) were extracted and the missing values were imputed as the mean of the questions. Next, the data were transposed such that each row represented a particular student, and columns represented each of the evaluation questions. The data points were normalized before clustering to ensure similar scaling on questions. By the Elbow method, to use three number of clusters. From the standardized responses, the K-means technique clustered groups of students based on similarity in standardized responses which resulted in three different clusters that describe groups of students that share similar evaluation patterns by which individual perceptions can be distinguished or patterned trends in the feedback data can be observed

DOI: 10.54756/IJSAR.2025.6.1

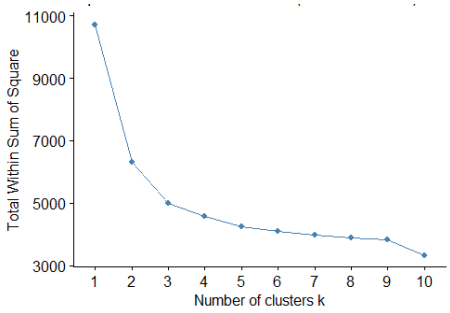

Figure 3: Optimal Number of Clusters (Elbow Method)

The Elbow Method plot for finding the number of clusters (k) to use in the clustering analysis. The number of clusters (k) is on the x-axis, and the Total Within Sum of Squares (WSS) reflecting the compactness of the clusters is indicated by the y-axis.

Total WSS vs K The plot of total WSS against the number of clusters shows a strong decrease of WSS for each added cluster up to 4, suggesting that the cluster WSS themselves are that: their number of centers in the ideal value is at most 3. This shows that points are closely clustered within their respective clusters when we have 3 clusters. But after three clusters, the decrease of WSS slows down, and adding more clusters will not bring much improvement of cluster compactness. The "elbow" point, at which the curve bends from steep curve to shallow curve, occurs at k = 3. This indicates that the optimal trade-off between compactness and simplicity arise when choosing three clusters to represent in the model and it is enough to capture the structure of the data without making things more complicated.

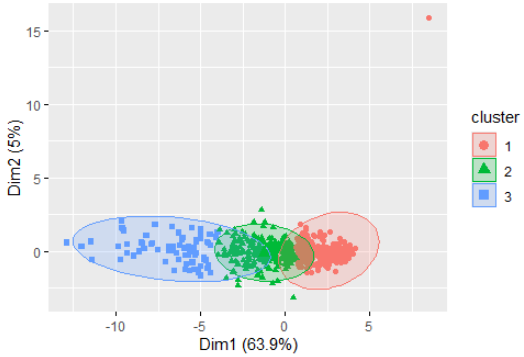

Figure 4: Clustering of Students Based on Response Patterns

Clustering of student data clusters analysis of student data show three obvious groups of student responses, each group representing a distinctive pattern of evaluation scores. Cluster 1 in Red circles: These are students whose response seems to deviate from the others, indicating a different view and experience with the course and the instructor. Cluster 2 (green triangles) partially overlaps with Cluster 1 and Cluster 3, implying these students have medium or mixed evaluation responses. This class can consist of students with various scores which are not too high and too low. Finally, cluster 3 (blue squares) is more isolated at the left, identifying students who respond uniformly with students in their group but differently from the other groups of students. This division may suggest a pattern of contrasting evaluations, such as overall higher or lower ratings on certain components of the course or instruction. The Horizontal axis (Dim1) that explains 63.9% of the inertia, would probably reflect the principal factor that accounts for such breathing patterns, and the Vertical axis (Dim2) that explains 5% of the inertia being indicative of a further disturbance.

Factor Analysis or Principal Component Analysis

Principal Component Analysis (PCA) Data from the student evaluation dataset was first converted to all numeric and scaled, as PCA treats each question as equal no matter of the scale. We then took only the student-rating

DOI: 10.54756/IJSAR.2025.6.1

columns, and transposed this data to perspectives where each question is a variable, before running a principal components analysis (PCA) with prcomp() in R, a dimension reduction tool that identifies the principal axis (i.e., principal component) that explains the most variability in the data. By the plot of above PCA results, it could observe how question related to each other via the their contribution of the principal component. These components were interpreted through scree plot and variable plot, in which Dim1 explained most variance, unveiled a main pattern, and Dim2 showed additional but less important variability. This data reduction collapses the data to larger contours that can shed light on the correlatively linked patterns of the student answers.

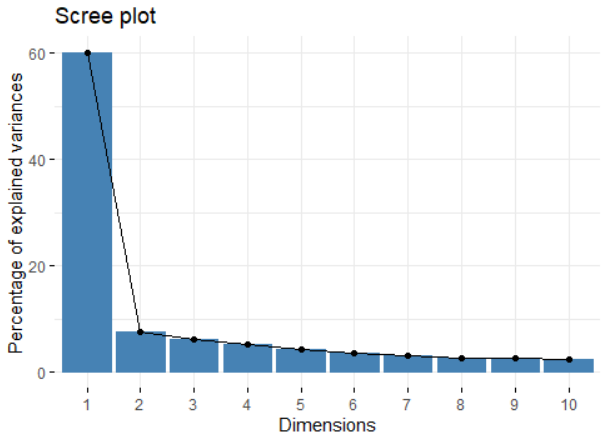

Figure 5: Scree Plot

The scree plot indicates the proportion of variance that is accounted for by each principal component (dimension) in your data. The first component accounts for about 60% of the variance – a large proportion, meaning it is able to capture the main pattern or theme in the data. This rapid decrease in explained variance between the first and second component indicates that the first one is the most meaningful, while the remaining ones contribute much less individually. The "elbow" forming in the scatter plot about the second or third component indicates that it is the first component, or possibly only the first few components, that are interpretable with the largest patterns in the data. Terms beyond the first few don't add much new information, so it can probably be safely disregarded without losing too much descriptive power. This indicates that the student evaluations are structured that can be summarize effectively by concentrating on the first principal component, or perhaps the first two or three components.

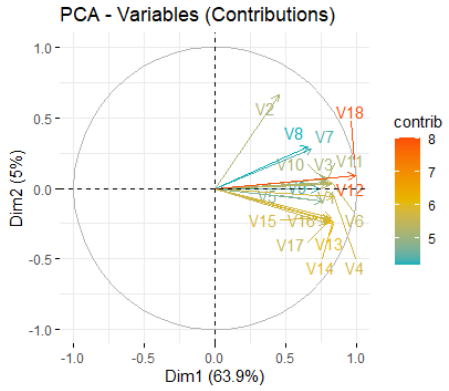

Figure 6: PCA Plot

A positive gradient is used to represent the contribution of a variable in PCA plot, with warm colors (orange/red) indicating high, and cool colors (blue/yellow) indicating low contribution. Variable V2 is the one that has the most influence, since it is the most red and has the longest arrow, explaining the most variance along the two dimensions Dim1 and Dim2. Like most of the variables fall onto the x-axis (Dim1) and explain 63.9% of variance. However, features such as V4, V6, V15, and V17 strongly indicate that they correlate and reflect the similar styles in data. The dimension along the y-axis (Dim2) and it explains 5% of the variance, captures slight (and perhaps more

DOI: 10.54756/IJSAR.2025.6.1

subtle) trends supported by variables such as V2, V7, V8 that express other patterns not shown by Dim1.

Closer clustering of, e.g., V14 and V15 (V17) in Figure 2d implies that they are positively correlated and likely measure correlated properties of the data. On the other hand, long distinct arrows such as V2 and V18 make unique contributions and they describe separate data dimensions that are different from the major structures. Such grouping and separation can aid to recognize which factors play an important role on the variance in the data, in this case, V2 and V18 have considerable importance.

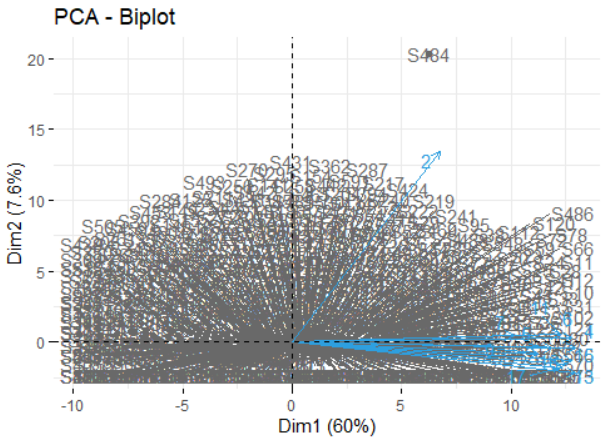

Figure 7: PCA - Biplot

A biplot of the PCA shows the relationship of each user response and every evaluation question in the two main directions of the PCA. Dim1 (explaining 60% of the variance) reflects the main feature in student evaluations, i.e., most students and questions are oriented along this dimension. This would indicate that a great deal of the variance in student responses can be accounted for due to a predominant theme, which is believed to be due to like perceptions or similar response tendencies for the overall group of items. Dim2 which accounts for only 7.6% of the variance contributes much less and points to specific differences, in particular to Question 2 that emerges as an outlier highly in-line with Dim2. This suggests that Question 2 reflects a special reaction to the evaluation product which is different to the general trend of the other questions. Consider further the range of the student responses across Dim1 to reinforce the observation of a common evaluation pattern, and the location of Question 2 across Dim2 to indicate that the former question possibly serves as a different criterion or probes distinct responses that differ from all other questions in the evaluation.

3.2 Discussion

In this study, Data mining was used, more specifically Principal Component Analysis(PCA) and K -means clustering, to assess the teaching performance based on student's opinion from evaluations. The results identified three main clusters of students with contrasting patterns of assessment. These clusters offer an understanding of differing student views of teaching effectiveness that go beyond the most and least consistent responses.

PCA also revealed a major component explaining more than 60% of the variance, indicating a general agreement among students on some aspects of teaching. That is, some criteria for judging teaching differ by individual, but others are widely accepted standards of teaching quality. For example, items related to teacher readiness and expertise tended to load closely supporting their status as central aspects of teaching effectiveness.

The clustering analysis also revealed different student opinions among the clusters, highlighting the importance for instructors to respond to diverse student needs and expectations. For instance, one cluster exhibited a high consistent mean rating on nearly all questions, which could be the students who were VERY satisfied with their learning. Conversely, a second cluster displayed a broad range of profile values or polarized profile values, which could be indicative of students' troubles or in unusual situations.

DOI: 10.54756/IJSAR.2025.6.1

Using methods of data mining showed the possibility of extracting subtle patterns from student evaluations that descriptive statistics might fail to detect. These findings can contribute to evidence-based approaches to faculty development and could inform interventions aimed at, for example, promoting classroom engagement and matching course content with real-world application.

4.0 CONCLUSION

The use of data mining (PCA, K-means clustering) is also made a lot easier in teaching performance analysis as it provides a strong methodology for higher education institutes that are proscribing in providing quality teaching. Hence, the application of these approaches is able to uncover patterns in student evaluations that otherwise are hidden and, for example, the finding of three different types of students indicates diversity in student perceptions and encourages instructors to consider different teaching strategies for those different needs. Clustering allows institutions to move beyond averages and look at specific trends in student feedback.

PCA emphasize fundamental teaching competencies (study) Details of the ICA factors, including instructor preparedness and content knowledge emerged as staples of student satisfaction, suggesting that these attributes are regarded as important for effective teaching regardless of discipline. This might inform programs for faculty development at institutions, programs that can focus not just on anything, but on the important things for students. Another suggestion for improvement was to better integrate real-world examples into course content—the change of these classes from the theoretical to the applied.

This paper also contributes to the larger field of AI in education through the use of data mining. It holds that data-driven approach provides insight into and guidance for improving teaching effectiveness. The approach to data analysis utilized in this study provides a template for other institutions to conduct similar analyses of their teaching evaluation data. This kind of progress is helpful not just in strengthening teaching, but also in enabling schools to meet the unique needs and interests of their students.

5.0 ACKNOWLEDGEMENTS

I wish to thank South East Asian Institute of Technology, Inc. for supporting during the work of this research. Special thanks are owed to Kane, Jember and Rose who provided indispensible assistance in executing the data collection. I also extend my sincere thanks to the participants who provided their feedback in order to undertake this research. Finally, I want to thank my family and friends, they have been supporting and cheering me up during the process.

REFERENCES

[1] Aggarwal, C. C. (2015). *Data mining: The textbook. Springer*.

[2] Bholowalia, P., & Kumar, A. (2014). EBK-Means: A clustering technique based on elbow method and K-means in WSN. *International Journal of Computer Applications*, 105(9), 17-24. [1] M. Sigala, A. Beer, L. Hodgson, and A. O'Connor, *Big Data for Measuring the Impact of Tourism Economic Development Programmes: A Process and Quality Criteria Framework for Using Big Data*. 2019.

[3] Brusilovsky, P., & Millán, E. (2016). User models for adaptive hypermedia and adaptive educational systems. *In The adaptive web* (pp. 3-53). Springer. https://doi.org/10.1007/978-3-540-72079-9_1.

[4] Chen, L., Chen, P., & Lin, Z. (2020). Artificial intelligence in education: A review. IEEE Access, 8, 75264-75278. https://doi.org/10.1109/ACCESS.2020.2988510.

[5] Dawson, S., Joksimovic, S., Poquet, O., & Siemens, G. (2020). Increasing the impact of learning analytics. *Journal of Learning Analytics*, 7(3), 1-12. https://doi.org/10.18608/jla.2020.7399.

[6] Ferguson, R., & Clow, D. (2017). Where is the evidence? A call to action for learning analytics. *Proceedings of the Seventh International Learning Analytics & Knowledge Conference* (pp. 56-65). https://doi.org/10.1145/3027385.3027395.

[7] Greenacre, M. (2017). *Correspondence analysis in practice* (3rd ed.). CRC Press.

[8] Hwang, G. J., & Chang, H. F. (2021). A review of opportunities and challenges of chatbots in education.

DOI: 10.54756/IJSAR.2025.6.1

Interactive Learning Environments, 29(5), 765-777. https://doi.org/10.1080/10494820.2020.1710171.

[9] Jolliffe, I. T., & Cadima, J. (2016). Principal component analysis: A review and recent developments. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 374(2065), 20150202. https://doi.org/10.1098/rsta.2015.0202.

[10] Karimi, M., & Yousefi, A. (2021). Predicting student performance using learning analytics techniques: *A systematic review*. *Education and Information Technologies*, 26, 491-512. https://doi.org/10.1007/s10639-020-10256-7.

[11] Kassambara, A. (2017). *Practical guide to cluster analysis in R: Unsupervised machine learning.* STHDA.

[12] Luan, J., & Zhao, C. M. (2018). Predictive analytics in higher education: *An overview*. *New Directions for Institutional Research*, 179, 5-15. https://doi.org/10.1002/ir.20268.

[13] Macfadyen, L. P., & Dawson, S. (2016). Mining LMS data to develop an early warning system for educators. *Computers & Education*, 54(2), 588-599. https://doi.org/10.1016/j.compedu.2016.09.009.

[14] Romero, C., & Ventura, S. (2020). Educational data mining: A review of the state of the art. *IEEE Transactions on Systems, Man, and Cybernetics, Part C*, 40(6), 601-618. https://doi.org/10.1109/TSMCC.2020.2053532.

[15] Siemens, G., & Baker, R. S. J. d. (2020). Learning analytics and educational data mining: Towards communication and collaboration. *Proceedings of the 10th International Conference on Learning Analytics and Knowledge* (LAK '20), 252-254. https://doi.org/10.1145/2330601.2330661.

[16] Wilkinson, L., & Friendly, M. (2012). The history of the cluster heat map. *The American Statistician*, 63(2), 179-184. https://doi.org/10.1198/tas.2009.0033.

[17] Wickham, H., & Grolemund, G. (2016). *R for data science: Import, tidy, transform, visualize, and model data.* O'Reilly Media, Inc.

[18] You, J. W. (2021). Clustering LMS data to identify student behaviors: Enhancing predictive analytics for online learning. *Computers & Education*, 165, 104110. https://doi.org/10.1016/j.compedu.2021.104110.

[19] Zhang, Z., Kim, H. J., Lonjon, G., & Zhu, Y. (2018). Balance diagnostics after propensity score matching. *Annals of Translational Medicine*, 7(1), 16. https://doi.org/10.21037/atm.2018.12.10.

[20] Zimek, A., & Schubert, E. (2017). Outlier detection. *In Encyclopedia of Database Systems* (pp. 1-5). https://doi.org/10.1007/978-1-4899-7993-3_807654o.

[21] Shutaywi, M., & Kachouie, N. N. (2021). Silhouette Analysis for Performance Evaluation in Machine Learning with Applications to Clustering. *Journal Name, Volume*(Issue), Pages. https://doi.org/DOI7.

DOI: 10.54756/IJSAR.2025.6.1