Department of Digital Business
Journal of Artificial Intelligence and Digital Business (RIGGS)
Homepage: https://journal.ilmudata.co.id/index.php/RIGGS
Vol. 4 No. 4 (2026) pp: 13734-13740
P-ISSN: 2963-9298, e-ISSN: 2963-914X

Collaboration in Learning and Curriculum Management Using Data Mining to Improve the Quality of Education in the Digital Era

Hamria, Adelia Dwi Putri
Universitas Ichsan Sidenreng Rappang
rhiyafatmawatihamka@gmail.com, adeliaadelia244@gmail.com

Abstract

In the rapidly growing digital era, a major challenge faced by the education system is adapting the curriculum to the fast-paced technological advancements. This study aims to analyze the effectiveness of the curriculum through the application of data mining, focusing on the use of clustering algorithms to identify patterns in student data and curriculum. The data used in this study were collected from 200 students and 50 teachers using a digital-based curriculum, which included information on student satisfaction, exam performance, and feedback on teaching materials. Through the application of clustering algorithms, this study identified three main clusters, each of which showed different curriculum needs. The first cluster indicates that students with low satisfaction and low exam performance require a more adaptive and technology-based curriculum. The second cluster, which shows high satisfaction and high exam performance, requires the maintenance and development of the existing curriculum. Meanwhile, the third cluster, which had high satisfaction but low exam performance, showed the need for a more personalized teaching approach and more effective exam evaluation techniques. The results of this study provide valuable insights for designing a curriculum that is more responsive to educational needs in the digital era. By utilizing data mining, particularly clustering, education systems can better understand and tailor to the diverse needs of students, creating more effective and relevant learning experiences in line with technological advancements. This research also contributes significantly to the design of data-driven curricula that can improve the quality of education in the digital age.

Keywords: Data Mining, Curriculum Effectiveness, Digital Education, Clustering Algorithms, Educational Analysis

1. Introduction

Education in the digital era today faces a big challenge in adapting the curriculum to the rapid development of technology.[1] The integration of technology in education involves not only the use of hardware and software, but also changes in learning methods and curriculum evaluation.[2] An effective curriculum is expected to produce students who are competent, creative, and ready to compete in an ever-changing world. However, how to assess the effectiveness of the curriculum in the midst of rapid technological developments.[3]

Data mining emerged as a potential solution to analyze the effectiveness of the curriculum.[4] With its ability to unearth patterns and trends from big data, data mining can provide in-depth information about how the curriculum is implemented and how well it adapts to the needs of students and the development of the world of work. Therefore, this study aims to explore the role of data mining in analyzing the effectiveness of curriculum in the digital era.[5]

Many previous studies have examined the use of data mining in various aspects of education, including in student performance analysis, learning personalization, and curriculum evaluation. that data mining can be used to analyze students' learning habits and provide recommendations related to curriculum that can improve learning outcomes.[6] Meanwhile, it shows that clustering algorithms can help identify segments of students with different learning needs, thus allowing for more targeted curriculum development. [7]

However, most of these studies have not focused on the use of data mining to analyze the effectiveness of the overall curriculum in the context of digital education.[8] This research aims to fill this gap by exploring how data mining techniques can be used to evaluate and develop curriculum in the digital era.[9]

This research has an important contribution to the development of science in the fields of education and technology.[10] First, this study offers a new view of how data mining can be leveraged to evaluate curriculum effectiveness.[11] Second, this research provides an empirical basis for the development of a curriculum that is more responsive to the changing needs of the world of education and the world of work in the digital era.[12] Third, this study introduces various data mining methods and techniques that can be used by education practitioners to analyze curriculum data more efficiently.[13]

Although curriculum has an important role to play in achieving educational goals, it is often difficult to evaluate the extent to which it is effective in creating optimal learning outcomes.[14] Therefore, the problem raised in this study is how data mining can be used to analyze and measure the effectiveness of the curriculum in the digital era. The focus of this research is to identify methods that can be used to collect and analyze educational data in order to evaluate and improve the existing curriculum.[15]

2. Methodology

2.1 Proposal (Constructive Step)

This study uses a quantitative approach by collecting educational data from various sources, including test result data, student satisfaction surveys, and other relevant data. The steps proposed in this study include:

1. Data Collection: Data is collected from educational institutions that use digital-based curriculum.
2. Data Preprocessing: The collected data is then processed to ensure its quality, including data cleansing and normalization.
3. Application of Data Mining Algorithms: Techniques such as clustering, classification, and association rule mining are applied to explore patterns and relationships in data that can demonstrate the effectiveness of the curriculum.
4. Evaluation of Results: The results of data analysis are used to evaluate the curriculum and provide recommendations for improvement.

Diagram of Research Flow for Digital-Based Curriculum Evaluation

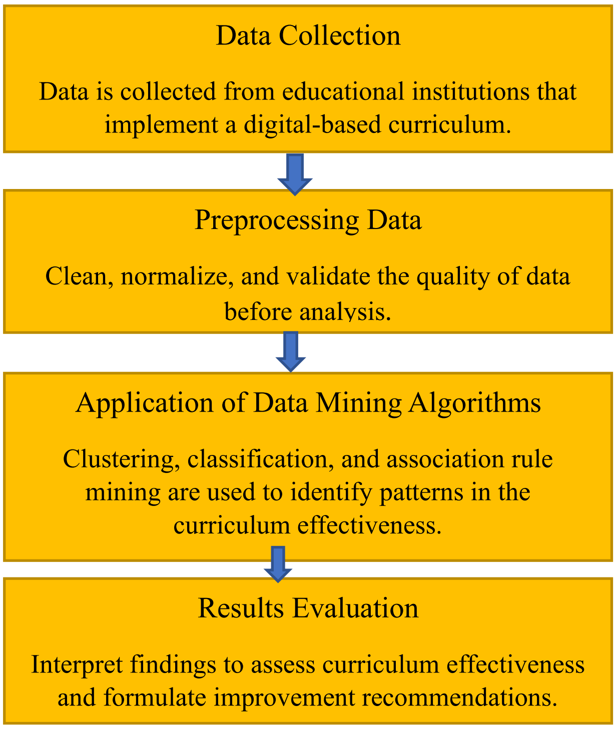

Figure 2.1 Proposal (Constructive Steps)

DOI: https://doi.org/10.31004/riggs.v4i4.5511

2.2 Theory Development & Solution Implementation

This study adopts the theory of learning analytics and educational data mining to develop a curriculum analysis model. The implementation of the solution is carried out by using data mining software such as WEKA or RapidMiner to implement the appropriate algorithm. The results of this analysis are expected to provide insight into the effectiveness of the curriculum and areas that need improvement.

3. Results and discussion

3.1 Test Data

The data used in this study consisted of survey results of 200 students and 50 teachers in several schools that used a digital-based curriculum. The data includes information about student satisfaction, exam performance, and feedback

Table 3.1 Test Data

Yes	Information	Number of Students	Number of Teachers	Data Description
1	Student Satisfaction	200	-	Measuring student satisfaction with the digital curriculum
2	Exam Performance	200	-	Data about students' exam results
3	Feedback on Materials	200	50	Feedback from students and teachers about teaching materials

3.2 Test Execution

The test is carried out by applying clustering algorithms to identify patterns in student and curriculum data. The results show that there are groups of students who need a more adaptive and technology-based curriculum.

To perform testing with clustering algorithms, here are the steps that can be taken to identify patterns in student data and curriculum, with the goal of segmenting students based on their needs for a more adaptive and technology-based curriculum:

1. Data Preparation

1. **Data Used**: The data used in the test was the result of a survey of 200 students and 50 teachers that included student satisfaction, exam performance, and feedback on the teaching material.
2. **Preprocessing Data**:
   a. The data needs to be cleaned first, by removing incomplete or irrelevant data.
   b. Data normalization (for example, scaling on test satisfaction scores or performance to match).

2. Test Results

Interpretation of Clustering Results:

1. **Cluster 1: Students with low satisfaction and low exam performance**

   Explanation: This cluster represents students who are dissatisfied with the curriculum and also perform poorly on exams. The low satisfaction suggests that the curriculum might not meet their learning preferences or needs, which could lead to a lack of motivation or engagement. Their low exam performance indicates that the teaching methods or materials may not be effective for their learning style. To address this, a more adaptive and technology-based curriculum could be introduced. This could involve

DOI: https://doi.org/10.31004/riggs.v4i4.5511

incorporating interactive, personalized learning tools that cater to different learning paces, such as online learning platforms, gamified content, or artificial intelligence-driven assessments. These students may also benefit from a curriculum that is more flexible and adaptable to their learning needs.

2. **Cluster 2: Students with high satisfaction and high exam performance**

   Explanation: This cluster represents students who are highly satisfied with the curriculum and perform well in exams. This suggests that these students are benefiting from the current curriculum structure, and their learning needs are being met effectively. However, while their performance is good, the curriculum may need to be regularly updated to ensure that it continues to challenge these students and encourages further growth. This group might require new opportunities for deeper learning, advanced topics, or innovative teaching techniques to maintain their high performance and engagement. Additionally, they may benefit from enrichment activities or specialized learning tracks that align with their strengths and interests, ensuring that they are continuously stimulated and prepared for future academic challenges.

3. **Cluster 3: Students with high satisfaction but low exam performance**

   Explanation: This cluster represents students who are satisfied with the curriculum but are not performing well on exams. This paradox suggests that while these students enjoy the curriculum and are perhaps motivated by the learning experience, there may be underlying issues with how their learning is being assessed. The gap between satisfaction and exam performance could indicate that the exam evaluation techniques are not aligning with the way these students learn or demonstrate their understanding. These students may benefit from more personalized teaching approaches, such as alternative assessment methods (e.g., project-based assessments, oral exams, or portfolios) that better reflect their knowledge. Additionally, a focus on improving their exam preparation strategies, study techniques, or targeted support in certain subject areas could help improve their performance while maintaining their satisfaction with the curriculum.

Here are example cases for each of the three clusters:

1. **Cluster 1: Students with low satisfaction and low exam performance**

   Case Example:
   Student A is enrolled in a digital-based curriculum for a high school mathematics course. Despite consistent effort, **Student A** shows a lack of engagement with the content and performs poorly on exams, scoring below average. The student expresses frustration with the pace and style of the lessons, finding them too theoretical and not aligned with their preferred learning style. After conducting surveys, it becomes clear that **Student A** and other students in this cluster are not motivated by traditional lecture-based teaching methods. They report that they prefer interactive learning tools, such as simulations, videos, and real-time problem-solving tasks. Their low satisfaction correlates with their low exam performance, as they struggle to connect with the material and feel disconnected from the traditional learning methods.

   Solution:
   The curriculum is adapted to include more hands-on learning experiences, integrating technology such as interactive math software or gamified learning platforms. Personalized learning paths and digital tools are introduced to cater to individual needs, allowing students like **Student A** to progress at their own pace.

2. **Cluster 2: Students with high satisfaction and high exam performance**

   Case Example:
   Student B excels in a digital-based chemistry course, consistently scoring high marks on exams and expressing satisfaction with the content, teaching methods, and overall learning experience. The student appreciates the well-structured lessons, which include a mix of video lectures, quizzes, and discussions. **Student B** enjoys how the curriculum challenges them with complex problems and encourages independent research. Their exam scores reflect their deep understanding of the subject, and they feel motivated by the learning experience.

DOI: https://doi.org/10.31004/riggs.v4i4.5511

Solution:
Although **Student B** is performing well, the curriculum is regularly reviewed and updated to ensure it remains challenging and enriching. To keep **Student B** engaged, opportunities for deeper exploration of advanced topics are added, such as special projects or more in-depth assignments. A mentorship program is also introduced to offer high-achieving students like **Student B** more personalized academic guidance and pathways for future learning.

3. **Cluster 3: Students with high satisfaction but low exam performance**

Case Example:
Student C is a student in a digital history course who enjoys the content and is highly satisfied with the engaging and visually appealing course materials, such as interactive timelines, discussion boards, and multimedia presentations. However, despite their high satisfaction with the curriculum, **Student C** consistently scores below average on exams. They have expressed that they often find it difficult to prepare for the exams, which primarily focus on traditional multiple-choice and essay questions. **Student C** performs well in discussions and group activities but struggles with the exam format, which they feel does not reflect their understanding of the material.

Solution:
In response to the mismatch between satisfaction and exam performance, the curriculum is adjusted to provide more diverse forms of assessment. Alternatives such as project-based assessments, collaborative presentations, or creative assignments are introduced, allowing **Student C** to demonstrate their knowledge in ways that better align with their strengths. Additionally, targeted workshops and study sessions are held to help students like **Student C** prepare for exams, focusing on time management, exam strategies, and specific areas where they are struggling.

1. Rumus Clustering K-Means

In the *K-Means algorithm*, the main goal is to minimize the distance between the data point and the center of the cluster (centroid). This process uses the following formula to calculate the distance between the data point and the centroid:

$$\boldsymbol{J} = \sum_{i=1}^{n} \sum_{j=1}^{k} \mathbf{b1}_{(c_i=j)} \left| \boldsymbol{x_i} - \boldsymbol{\mu_j} \right|^2$$

Where:

1. J is a cost function that must be minimized.
2. n is the sum of data (for example, 200 students).
3. is the specified number of clusters (for example, 3 clusters).
4. C_i is the cluster label for the i data.
5. X_i is a data point (e.g., student satisfaction and exam performance).
6. μ_j is the centroid cluster j.

2. Interpretation of Clustering Results

After clustering is complete, here is the interpretation of the results that can be based on the clusters that are formed:

1. **Cluster 1 (Low Satisfaction and Low Performance):** For students in Cluster 1, it is necessary to adjust the curriculum to be more adaptive and technology-based. This can be achieved by providing more interactive materials or using digital tools to increase students' interest and understanding.
2. **Cluster 2 (High Satisfaction and High Performance):** Students in this cluster indicate that the existing curriculum is already sufficiently effective, but there may need to be further improvement to maintain their

performance. For this group, the focus is on further development or a greater challenge to sustain their results.

3. **Cluster 3 (High Satisfaction but Low Performance):** Students in this cluster are satisfied with the curriculum, but their exam performance is low. This may indicate a need to delve into more effective teaching or exam evaluation techniques, or it may be the need for a more personalized approach.

3. Cluster Quality Measurement

To evaluate the quality of clustering results, we can use several metrics: Silhouette Score: Used to measure how well objects are grouped in their own cluster compared to other clusters. Silhouette values range from -1 to 1, with higher values indicating better clustering.

The Silhouette formula for each data point i is:

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

4. Evaluation and Use of Clustering Results

The results of clustering can be used to provide recommendations to educators and curriculum policymakers to adjust the materials or teaching methods based on the needs of students grouped into different clusters. Suppose we have 3 data points with distances that have been calculated as follows:

1. Point i (e.g., student 1) is in cluster 1.
2. $\mathbf{a_{(i)}=2.5a_{(i)} = 2.5a_{(i)}=2.5}$ (the average distance of point iii to other points in the same cluster).
3. $\mathbf{b(i)=4.0b_{(i)} = 4.0b_{(i)}=4.0}$ (the average distance of point iii to points in the nearest cluster).

So the silhouette value for point i is:

$$\boldsymbol{s(i) = \frac{4.0 - 2.5}{\max(2.5, 4.0)} = \frac{1.5}{4.0} = 0.375}$$

5. Visualization of Results

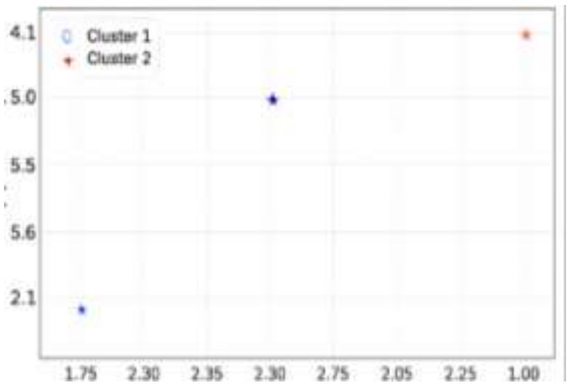

Figure 5. Clustering Result Visualization

Where

1. Title: "Clustering Results Visualization"
2. X-axis: "Student Satisfaction"
3. Y-axis: "Test Performance"
4. Legend: "Cluster 1" and "Cluster 2"

DOI: https://doi.org/10.31004/riggs.v4i4.5511

Visualization of clustering results based on the calculation example above. On this graph:

1. **Cluster 1** (in blue) contains points with a student satisfaction of 2.5 and an exam performance of 4.0.
2. **Cluster 2** (in red) contains points with student satisfaction of 3.5 and exam performance of 4.5.

The X-axis represents student satisfaction, while the Y-axis represents exam performance. Thus, this testing not only provides an understanding of patterns in student data, but also identifies groups of students who need a change in technology-based curriculum approaches.

4. Conclusion

This research shows that the application of data mining, especially through clustering algorithms, can provide important insights in analyzing and evaluating the curriculum used in digital education. By grouping students based on satisfaction and exam performance, three groups were found with different needs related to the curriculum they underwent. Cluster 1 calls for a more adaptive and technology-based curriculum, with a focus on more interactive learning and more support for students who are struggling. Cluster 2 shows that students with high satisfaction and performance require the maintenance of the quality of the existing curriculum, as well as the greater challenge of maintaining their results. Cluster 3 requires a more personalized approach to teaching, as well as the development of exam evaluation techniques that are more in-depth and appropriate to their learning style. By using data mining techniques, especially clustering, the results of this research provide a strong basis to improve the quality of the curriculum in the digital era, so that it is more relevant and in accordance with the increasingly diverse needs of students. Further evaluation and use of these techniques in future curriculum development can improve the effectiveness of the teaching and learning process and prepare students to face global challenges in the world of work.

References

[1] H. Saleh, R. Rachmat, and M. As'ad, "Application of Computer Vision for Customer Insights: Demographics and Visit Duration in Coffee Shops," *Res. Horiz.*, vol. 5, no. 5, pp. 2153–2164, 2025.

[2] S. Suherwin, S. N. Asia, and R. Rachmat, "Human-AI Interaction dengan Antarmuka Suara dalam Bahasa Lokal/Dialek Nusantara," *RIGGS J. Artif. Intell. Digit. Bus.*, vol. 4, no. 3, pp. 8618–8625, 2025.

[3] F. Faisal, R. Rachmat, S. N. Asia, S. Suherwin, and A. Ibrahim, "The Effectiveness of Interactive Learning Media Based on Augmented Reality in Enhancing Elementary School Students' Learning Motivation," *J. Tecnol. Quantica*, vol. 2, no. 5, pp. 135–143, 2025.

[4] H. Sani, R. Rasyid, S. N. Asia, S. Syamsuddin, S. Suherwin, and R. Șerban, "Real-Time IoT Integration for Coal Production And Distribution Management," *J. Inf. Syst. Technol. Res.*, vol. 4, no. 3, pp. 155–162, 2025.

[5] R. Rachmat, S. B. Iskandar, K. Kasmawaru, and S. Suherwin, "Comparison of coronary heart disease prediction using basic model and ensemble learning," *J. Intell. Decis. Support Syst.*, vol. 8, no. 2, pp. 74–84, 2025.

[6] R. Rachmat, "Optimalisasi Algoritma Ant Miner Berbasis Fuzzy Untuk Penemuan Informasi Tersembunyi Pada Dataset Hepatitis Dengan Menggunakan Teknik Feature Selection," *RIGGS J. Artif. Intell. Digit. Bus.*, vol. 4, no. 2, pp. 3417–3421, 2025.

[7] R. Rachmat and S. Suhartono, "Comparative analysis of single exponential smoothing and holt's method for quality of hospital services forecasting in general hospital," *Bull. Comput. Sci. Electr. Eng.*, vol. 1, no. 2, pp. 80–86, 2020.

[8] S. Suherwin, R. Rachmat, I. Said, and S. N. Asia, "Pembangunan Sistem AI Berdasarkan Analisis Aktivitas Digital Untuk Mengidentifikasi Gaya Belajar Siswa," *RIGGS J. Artif. Intell. Digit. Bus.*, vol. 4, no. 2, pp. 3117–3122, 2025.

[9] R. Rachmat, M. Yusuf, M. F. Basmar, and S. Suherwin, "Analysis of algorithms and data processing efficiency in movie recommendation systems based on machine learning," *J. Mandiri IT*, vol. 13, no. 2, pp. 273–279, 2024.

[10] H. Kasim, M. Yusuf, H. Haslinda, R. Rachmat, and M. F. Basmar, "Coal Spray Rate Prediction Based On Factor Analysis And Neural Network (Nn) Algorithm," *J. Soc. Res.*, vol. 2, no. 5, pp. 1489–1497, 2023.

[11] R. Rachmat and A. Ibrahim, "Decision Support System for Receiving Waste Retribution at the Housing and Sanitation Services Uses the Naïve Bayes Algorithm," *Bull. Comput. Sci. Electr. Eng.*, vol. 2, no. 2, pp. 84–92, 2021.

[12] M. F. Basmar, R. Racchmat, and M. Yusuf, "Decision Support System On The Influence Of Leadership Style And Work Motivation Of Upri Makassar Employees Using Naive Bayes Algorithm," *INFOKUM*, vol. 10, no. 5, pp. 544–549, 2022.

[13] R. Rachmat, M. Yusuf, I. Abbas, and M. F. Basmar, "Segmentasi Citra Menggunakan Algoritma K-Means Berbasis Particle Swarm Optimization," *METIK J. (AKREDITASI SINTA 3)*, vol. 6, no. 2, pp. 146–156, 2022.

[14] R. Rachmat and R. D. Liklikwati, "OPTIMASI k-Nearst-Neighbor (k-NN) DENGAN PARTICLE SWARM OPTIMIZATION PADA KLASIFIKASI NASABAH KREDIT," *JTRISTE*, vol. 6, no. 1, pp. 9–16, 2019.

[15] K. Tzafilkou, M. Perifanou, and A. A. Economides, "Development and validation of a students' remote learning attitude scale (RLAS) in higher education," *Educ. Inf. Technol.*, vol. 26, no. 6, pp. 7279–7305, 2021, doi: 10.1007/s10639-021-10586-0.