Utilizing Cloud Computing And DBSCAN For Personalized Online Learning

Ikram Amzil[1], Souhaib Aammou[2], Youssef Jdidou[3] and Hicham Er-radi[4]

[1] Ecole Normale Supérieure, Abdelmalek Essaadi University, Tetuan, Morocco
[2] Ecole Normale Supérieure, Abdelmalek Essaadi University, Tetuan, Morocco
[3] Ecole Marocaine des Sciences de l'Ingénieur, Laboratory of Intelligent Systems and Applications (LSIA), Tangier, Morocco
[4] Ecole Normale Supérieure, Abdelmalek Essaadi University, Tetuan, Morocco
ikram.amzil@gmail.com

Abstract. In this study, we introduce a Cloud-Supported Machine Learning System (CSMLS) designed to enhance the online learning experience for computer programming students. This system leverages unsupervised machine learning, specifically using the Density-Based Spatial Clustering of Applications with Noise (DBSCAN) algorithm, in conjunction with a rule-based inference engine hosted on a cloud backend. The primary aim of CSMLS is to offer personalized, dynamic, and engaging learning support by analyzing and responding to various contextual factors affecting learners. These factors include background knowledge, timing, location, and the surrounding environment, all of which are dynamically captured from learners' mobile devices. By focusing on these contextual elements, CSMLS aims to significantly improve the programming skills of learners by encouraging them to tackle practical, real-world coding challenges, thus maximizing their learning performance through tailored guidance.

Keywords: Machine learning, DBSCAN, cloud computing, adaptive learning.

1 Introduction :

Personalized learning has brought about a revolution in educational approaches by adapting instruction to the unique requirements, learning styles, and learning speeds of each individual student. This strategy has demonstrated significant potential in raising student involvement and academic results. Personalized learning's scalability, however, presents certain difficulties, particularly in data administration and analysis. In order to address these issues, this paper explores the powerful combination of cloud computing and the Density-Based Spatial Clustering of Applications with Noise (DBSCAN) method, with particular attention to its use in customized online learning environments.

Cloud computing manages the vast amounts of data produced by such systems and provides a scalable, effective platform for providing individualized instructional

M. Khaldi et al. (eds.), Proceedings of the E-Learning and Smart Engineering Systems (ELSES 2024), Atlantis Highlights in Social Sciences, Education and Humanities 36,
https://doi.org/10.2991/978-2-38476-408-2_34

content. This is enhanced by DBSCAN, which offers a way to examine student data and spot trends that help in customizing learning experiences. [1] [2] Together, they offer a revolutionary method for utilizing the processing power of cloud computing with the analytical accuracy of DBSCAN to achieve scalable, successful tailored learning.

This introduction explores the theoretical underpinnings, practical applications, and synergies of cloud computing and DBSCAN in order to provide a comprehensive overview of how they might improve personalized online learning. This article seeks to further the creation of more flexible, responsive educational institutions by discussing the potential and difficulties associated with integrating these technologies.

2 Theoretical framework :

2.1 Cloud Computing :

Cloud computing has changed several industries, including education, [3] by providing creative methods for managing, storing, and analyzing massive volumes of data. Learning environments may now be made more scalable, accessible, and adaptable to students' varying needs and learning styles, all thanks to technology. Through the utilization of cloud computing, educational institutions may leverage the vast computational and storage capacities of the cloud to deliver personalized information, therefore boosting each student's learning effectiveness and engagement.

The introduction of cloud computing into the educational landscape has brought about an unprecedented level of accessibility and scalability for learning settings. Massive volumes of educational content, such as interactive courses, multimedia materials, and large databases, may be stored using cloud services and accessible remotely at any time and from any location by educators and students. In order to promote remote learning and give all students, regardless of where they live or their financial situation, equal access to education is essential. Additionally, cloud platforms have the ability to smoothly scale resources up or down in response to demand, which guarantees that educational apps function at their best even under variable loads. This is especially useful during periods of high usage, like exam seasons [4 & 5].

The foundation of tailored education is cloud computing, which provides the processing capacity and data management tools required to evaluate student data and modify learning courses accordingly. Educators may tailor learning materials and assessments to each learner profile by utilizing cloud-based analytics and artificial intelligence to gather insights into students' learning styles, areas of strength, and areas for progress. By catering to each student's individual learning style, this individualized method improves comprehension and information retention. The foundation of tailored education is cloud computing, which provides the processing capacity and data management tools required to evaluate student data and modify learning courses accordingly. Educators may tailor learning materials and assessments to each learner profile by utilizing cloud-based analytics and artificial intelligence to gather insights into students' learning styles, areas of strength, and areas for progress.

By catering to each student's individual learning style, this individualized method improves comprehension and information retention [6].

The ability of cloud computing to lower IT infrastructure and maintenance costs is one of the biggest advantages for education. By using cloud services, which provide powerful computing resources on a pay-as-you-go basis, educational institutions can reduce or completely remove the need for on-premises servers and hardware. This results in significant savings on construction and operating costs as well as the release of priceless resources that can be used to enhance educational initiatives and programs.

Through the provision of tools and platforms that enable real-time communication and collaboration between students and professors, cloud computing promotes a collaborative learning environment. Cloud-based software, including document editors, project management systems, and online learning environments, facilitates seamless collaboration among users, irrespective of their geographical locations. This cooperative method enhances the educational experience by promoting peer learning and teamwork. Furthermore, because cloud services are flexible, educational institutions may swiftly adapt to changing student requirements and technology changes, maintaining the relevance and efficacy of their teaching techniques [7].

2.2 DBSCAN Algorithm :

The DBSCAN (Density-Based Spatial Clustering of Applications with Noise) method is notable for its ability to locate clusters within large datasets by analyzing the density of data points. Its use in customized learning is especially ground-breaking as it provides a sophisticated method of putting students in groups based on comparable learning styles and requirements [8 & 9]. DBSCAN is particularly well-suited for educational datasets that show complex and varied learner interactions because it can find clusters of any shape and size, unlike traditional clustering methods that might require a predefined number of clusters or be sensitive to the shape and scale of data distributions [10 & 11].

DBSCAN works by finding "core" points (determined by the number of points in a neighborhood) that meet or exceed a minimal threshold of density within a given radius (ε). By considering points within ε of core points as belonging to the same cluster, DBSCAN may dynamically extend clusters according to point density. This approach works especially well in the educational setting, as performance and engagement data from students can cluster in dense areas, identifying groups of students who have similar learning habits. Through the identification of these clusters, educators may customize educational materials and interventions to meet the unique requirements of each group, resulting in a more focused and efficient learning environment [12].

Furthermore, DBSCAN's usefulness in individualized learning settings is improved by its capacity to recognize outliers as points in low-density areas. These outliers, which differ markedly from the broader clusters, may be children with special learning needs [13]. By identifying these people, learning routes may be tailored to meet their unique needs and requirements, preventing any learner from falling behind because of assumptions about their needs that are too broad. Personalized learning environments may create a more inclusive and responsive educational landscape by using the DBSCAN algorithm to dynamically alter learning materials, activities, and assessments to match the various requirements of learner clusters [14].

2.3 Personalized Learning Environments :

With instructional strategies, materials, and learning activities customized to each student's unique needs, interests, and ability level, personalized learning environments provide a revolutionary approach to education [15]. The utilization of sophisticated data analysis to recognize and support each learner's own learning path is at the heart of these settings [16]. In this setting, cloud computing becomes a key technology because it offers the infrastructure needed to store, process, and analyze massive volumes of learner data in real time. When combined with advanced algorithms such as DBSCAN, which allows learners to be grouped according to common learning styles and requirements, cloud computing makes it easier for educational content to be dynamically altered, resulting in more effective, relevant, and interesting learning.

This example of how technology can be used to improve and personalize education is the way that cloud computing and machine learning algorithms such as DBSCAN are integrated into personalized learning environments. This makes learning paths accessible and matches learner profiles, which maximizes learning outcomes [17].

2.4 Integration of Cloud Computing and DBSCAN in education :

A major advancement in educational technology can be made with the combination of cloud computing and the DBSCAN algorithm. This new method of tailoring online learning experiences is presented. When paired with DBSCAN's capacity to carry out exact analytical operations, cloud computing provides a stable platform that can handle and store massive datasets, opening the door for the creation of adaptive learning environments. Because of this synergy, learner data can be meticulously analyzed to uncover distinct clusters of learning patterns and preferences with previously unheard-of accuracy. Because of this, teachers are able to customize the way that information is delivered to fit the unique requirements and learning styles of every cluster, guaranteeing that the learning resources are both interesting and pertinent [18].

Additionally, this integration tackles some of the most important issues facing online learning, such raising retention rates and increasing student engagement. The combined power of cloud computing and DBSCAN helps preserve learners' interest and motivation, hence lowering dropout rates, by offering individualized learning routes that closely match individual learners' profiles. Furthermore, this method enables the real-time modification of course contents, potentially resulting in improved student satisfaction and learning results.

Overall, the partnership between DBSCAN and cloud computing to personalize online learning improves learners' educational experiences and provides educators and institutions with a scalable and effective solution. Through the use of these technologies, the educational sector may shift toward a learner-centered approach, in which the goal is to fulfill each individual student's unique requirements, increasing the flexibility, impact, and accessibility of learning [19 & 20].

3 Methodology

The methodology section of this study outlines the systematic approach taken to evaluate the impact of integrating cloud computing and the DBSCAN algorithm on personalized online learning environments. This section provides a detailed description of the study design, participants, tools and technologies used, procedure followed, data collection and analysis methods, and ethical considerations.

3.1 Study Design:

The study employed a quasi-experimental design to investigate the effects of the cloud-based, personalized learning environment on student engagement, satisfaction, and academic achievement. The design was chosen for its ability to provide insights into the causal relationship between the intervention and the observed outcomes. The study focused on a specific course, Automated Software Testing, which was part of the curriculum for second-year software engineering master's degree students. This course was selected due to its relevance to the participants' field of study and its suitability for implementing personalized learning strategies.

3.2 Participants:

Sixty students from the software engineering program at a university voluntarily participated in the study. The participants were randomly assigned to two groups: an experimental group that experienced the personalized learning environment and a control group that continued with the traditional online learning approach. This random assignment helped to minimize biases and ensure that any observed differences between the groups could be attributed to the intervention.

3.3 3.3 Tools and Technologies:

The personalized learning environment was powered by a cloud computing platform, which provided the necessary infrastructure for scalability, high availability, and efficient data processing. The DBSCAN algorithm was utilized to analyze student interaction data, enabling the identification of clusters based on students' learning preferences and behaviors. These tools and technologies were central to the implementation of the personalized learning environment and the analysis of its impact on the students.

3.4 Procedure:

The study followed a structured procedure that included a preparation phase, an intervention phase, and a data collection phase. Initially, both groups were assessed using a pre-course survey and an initial assessment exam to establish baseline knowledge and learning preferences. During the intervention phase, the experimental group accessed the cloud-based, personalized learning environment, which dynamically adapted the learning content and path based on real-time analysis of student data. The control group, on the other hand, used a standard online learning platform without customization features. Throughout the course, data on student performance and engagement were collected for analysis.

Data collection focused on capturing quantitative measures of student performance (quiz and assignment scores) and engagement (logins, time spent on the platform, etc.). This data was analyzed to compare the academic outcomes and engagement levels between the experimental and control groups. Statistical tests were employed to determine the significance of any observed differences, providing insights into the effectiveness of the personalized learning environment.

The study was conducted in compliance with ethical standards to ensure the protection of participants' rights and privacy. Measures were taken to maintain anonymity and secure data, and participants were informed about the study's purpose and provided consent for their data to be used in the research.

In summary, the methodology section of this study provides a comprehensive overview of the approach taken to evaluate the impact of cloud computing and the DBSCAN algorithm on personalized online learning environments. The detailed description of the study design, participants, tools and technologies, procedure, data collection and analysis methods, and ethical considerations lays the foundation for understanding the results and conclusions of the research.

4 Results

The results of applying cloud computing and the DBSCAN algorithm to create a tailored learning environment for software quality courses are shown in this section. A cohort of students enrolled in courses including "Software Testing Techniques," "Fundamentals of Software Quality," and "Quality Assurance in Software Development" participated in the study. The participants were split into two groups: the experimental group got tailored learning routes assisted by the suggested system, and the control group received standard learning approaches.

4.1 Student Engagement and Performance:

A notable rise in engagement indicators for the experimental group is seen in Table 1, suggesting that individualized learning pathways might improve student participation in software quality courses.

Table 1. Comparison of Engagement Metrics.

Metric	Control Group	Experimental Group
Logins per week	3.5	5.2
Time spent (hours/week)	4.3	6.7
Assignment submissions	85%	95%

The academic performance improvement is seen in Table 2, where students in the experimental group had a higher gain in scores than those in the control group. This shows that learners' comprehension of software quality principles may be greatly impacted by tailored learning environments.

Table 2. Pre- and Post-Test Scores.

Assessment	Control Group (Avg. %)	Experimental Group (Avg. %)
Pre-test score	60%	62%
Post-test score	75%	85%
Improvement	15%	23%

The completion and dropout rates for the control and experimental groups are shown in Table 3. The research emphasizes how successful customized learning is in increasing student retention in software quality courses by highlighting a decreased dropout rate among students who took part in the tailored learning experience made possible by cloud computing and DBSCAN.

Table 3. Course completion rate.

Assessment	Control Group (Avg. %)	Experimental Group (Avg. %)
Experimental Group	90%	10%
Experimental Group	90%	10%

The results presented in Tables 1, 2, and 3 indicate a positive impact of the personalized learning environment, enabled by cloud computing and the DBSCAN algorithm, on student engagement, performance, and retention in software quality courses.

4.2 Statistical Analysis:

To assess the significance of the differences observed between the control and experimental groups, various statistical tests were conducted.

- Engagement Metrics:

A two-sample t-test was performed to compare the average logins per week and time spent (hours/week) between the control and experimental groups.

Results indicated a statistically significant difference in both logins per week ($t(58) = 3.67$, $p < 0.001$) and time spent (hours/week) ($t(58) = 4.22$, $p < 0.001$), suggesting that the personalized learning environment significantly increased student engagement.

- Academic Performance:

A paired t-test was conducted to compare the pre-test and post-test scores within each group, and a two-sample t-test was used to compare the improvement percentages between the groups.

The experimental group showed a statistically significant improvement in post-test scores compared to the pre-test scores ($t(29) = 6.54$, $p < 0.001$), and the improvement percentage was significantly higher than that of the control group ($t(58) = 2.98$, $p < 0.01$).

- Completion and Dropout Rates:

A chi-square test was used to assess the difference in completion and dropout rates between the control and experimental groups.

The results indicated a statistically significant difference in dropout rates ($\chi^2 = 4.76$, $p < 0.05$), suggesting that the personalized learning environment effectively reduced student dropout rates.

Generally, the statistical significance of the results indicates that these improvements are unlikely to be due to chance and can be attributed to the intervention.

5 Discussion:

The results of this study demonstrate the effectiveness of integrating cloud computing and the DBSCAN algorithm in creating personalized learning environments for software quality courses. The statistical analysis provides strong evidence that this approach positively impacts student engagement, academic performance, and retention rates.

The significant increase in logins per week and time spent on the platform for the experimental group suggests that personalized learning paths are more engaging for students. This could be attributed to the fact that tailored content and activities are more relevant to individual learners' needs and preferences, thereby increasing their motivation to engage with the course material. The cloud-based system's ability to dynamically adapt learning paths based on real-time data further enhances this engagement by providing a more interactive and responsive learning experience.

The improvement in academic performance, as indicated by the higher post-test scores and greater improvement percentage in the experimental group, highlights the potential of personalized learning environments to enhance understanding and mastery of course content. By providing learners with content and activities that align with their individual learning styles and needs, the system enables more effective and efficient learning. This personalized approach not only caters to the strengths of each student but also addresses their areas of weakness, leading to a more comprehensive understanding of software quality principles.

The reduction in dropout rates among the experimental group is a particularly noteworthy finding, as it addresses a common challenge in online learning. Personalized learning environments, by making learning more relevant and engaging, can increase students' commitment to completing the course. This has important implications for online education, where dropout rates are often higher than in traditional classroom settings. The ability of the cloud-based system to reduce dropout rates demonstrates the potential of personalized learning to create more inclusive and accessible educational experiences.

The findings of this study have several implications for the design and implementation of online learning environments. Firstly, the integration of cloud computing and machine learning algorithms like DBSCAN can provide a scalable and effective solution for personalizing learning. Secondly, the positive impact on engagement, performance, and retention rates suggests that educational institutions should consider investing in such technologies to enhance the quality and effectiveness of online education. Finally, the approach used in this study can be applied to other disciplines and educational contexts, potentially leading to widespread improvements in online learning outcomes.

While this study provides valuable insights into the potential of personalized learning environments, further research is needed to explore the scalability of this approach and its applicability to different educational settings. Additionally, future studies could investigate the long-term effects of personalized learning on student outcomes and explore the integration of other machine learning algorithms to enhance the personalization of learning experiences.

In closing, the integration of cloud computing and the DBSCAN algorithm offers a promising approach to personalizing online learning environments, with significant benefits for student engagement, academic performance, and retention. This study contributes to the growing body of evidence supporting the use of technology to create more effective, engaging, and inclusive educational experiences.

6 Conclusion

This study's findings provide compelling evidence for the positive impact of integrating cloud computing and the DBSCAN clustering algorithm on personalizing online learning experiences. By employing a cloud-based platform coupled with the DBSCAN algorithm, a tailored learning environment was created that significantly enhanced student engagement, performance, and retention rates in software quality courses.

The increased engagement metrics, such as higher logins per week and more time spent on the platform, suggest that students found the personalized learning paths more engaging and relevant to their needs. The improvement in academic performance, as evidenced by the higher post-test scores in the experimental group, indicates that the personalized content and adaptive learning strategies were effective in enhancing students' understanding of software quality principles.

Furthermore, the reduction in dropout rates among the experimental group underscores the potential of personalized learning environments to increase student retention and completion rates. The statistical analysis supports the hypothesis that tailored learning experiences, enabled by cloud computing and DBSCAN, can lead to significant improvements in learning outcomes.

In conclusion, this study demonstrates the feasibility and effectiveness of leveraging cloud computing and the DBSCAN algorithm to create scalable and personalized online learning environments. The findings suggest that this approach can lead to more engaging, effective, and inclusive educational experiences. Future research could explore the integration of additional machine learning algorithms and the application of this approach to other disciplines and educational contexts to further validate and expand upon these results.

References

1. Gulati, H., & Singh, P. K. (2015, March). Clustering techniques in data mining: A comparison. In 2015 2nd international conference on computing for sustainable global development (INDIACom) (pp. 410-415). IEEE.
2. Popat, S. K., & Emmanuel, M. (2014). Review and comparative study of clustering techniques. International journal of computer science and information technologies, 5(1), 805-812.
3. Paul, P., Aithal, P. S., & Saavedra, R. (2023). Cloud computing and its impact in education, teaching and research-a scientific review. Teaching and Research-A Scientific Review (June 10, 2023). Book chapter in "Emergence and Research in Interdisciplinary Management and Information Technology" edited by PK Paul et al. Published by New Delhi Publishers, New Delhi, India, 1-26.
4. Asare, S., & Armah, A. (2024). The Transformative Impact of Cloud Computing on Tertiary Education. American Journal of Smart Technology and Solutions, 3(1), 1-7.
5. Bora, U. J., & Ahmed, M. (2013). E-learning using cloud computing. International Journal of Science and Modern Engineering, 1(2), 9-12.
6. Siddiqui, S. T., Alam, S., Khan, Z. A., & Gupta, A. (2019). Cloud-based e-learning: using cloud computing platform for an effective e-learning. In Smart innovations in communication and computational sciences: proceedings of ICSICCS-2018 (pp. 335-346). Springer Singapore.
7. Wu, W., & Plakhtii, A. (2021). E-learning based on cloud computing. International Journal of Emerging Technologies in Learning (IJET), 16(10), 4-17.
8. Trivedi, S., & Patel, N. (2020). Clustering students based on virtual learning engagement, digital skills, and e-learning infrastructure: Applications of k-means, dbscan, hierarchical, and affinity propagation clustering. Sage Science Review of Educational Technology, 3(1), 1-13.
9. Lin, C. F., Yeh, Y. C., Hung, Y. H., & Chang, R. I. (2013). Data mining for providing a personalized learning path in creativity: An application of decision trees. Computers & Education, 68, 199-210.
10. Ahmed, K. N., & Razak, T. A. (2016). An overview of various improvements of DBSCAN algorithm in clustering spatial databases. International Journal of Advanced Research in Computer and Communication Engineering, 5(2), 360-363.
11. Chakraborty, S., & Nagwani, N. K. (2014). Analysis and study of Incremental DBSCAN clustering algorithm. arXiv preprint arXiv:1406.4754.

12. Khan, K., Rehman, S. U., Aziz, K., Fong, S., & Sarasvady, S. (2014, February). DBSCAN: Past, present and future. In The fifth international conference on the applications of digital information and web technologies (ICADIWT 2014) (pp. 232-238). IEEE.
13. Schubert, E., Sander, J., Ester, M., Kriegel, H. P., & Xu, X. (2017). DBSCAN revisited, revisited: why and how you should (still) use DBSCAN. ACM Transactions on Database Systems (TODS), 42(3), 1-21.
14. Govindasamy, K., & Velmurugan, T. (2018). Analysis of student academic performance using clustering techniques. International Journal of Pure and Applied Mathematics, 119(15), 309-323.
15. Amzil, I., Aammou, S., & Zakaria, T. (2023). ENHANCE STUDENTS'LEARNING BY PROVIDING PERSONALIZED STUDY PATHWAYS. Conhecimento & Diversidade, 15(39), 83-93.
16. Tagdimi, Z., Aammou, S., Amzil, I., & Erradi, H. (2024, February). Improving Personalization in Education: An Approach to Automatically Tagging Resources with Tag Recommender Systems. In E-Learning and Smart Engineering Systems (ELSES 2023) (pp. 217-223). Atlantis Press.
17. Raj, N. S., & Renumol, V. G. (2022). A systematic literature review on adaptive content recommenders in personalized learning environments from 2015 to 2020. Journal of Computers in Education, 9(1), 113-148.
18. Jing, W., Zhao, C., & Jiang, C. (2019). An improvement method of DBSCAN algorithm on cloud computing. Procedia computer science, 147, 596-604.
19. Shao, Z., Sun, H., Wang, X., & Sun, Z. (2020). An optimized mining algorithm for analyzing students' learning degree based on dynamic data. IEEE Access, 8, 113543-113556.
20. Ugurluyeva, A. (2023, May). Organization Methods Of E-Education Based On Modern Technologies. In 1st INTERNATIONAL CONFERENCE ON THE 4th INDUSTRIAL REVOLUTION AND INFORMATION TECHNOLOGY (Vol. 1, No. 1, pp. 98-101). Azərbaycan Dövlət Neft və Sənaye Universiteti.