MODESTUM EURASIA Journal of Mathematics, Science and Technology Education, 2025, 21(9), em2695
ISSN:1305-8223 (online)
OPEN ACCESS Review Article https://doi.org/10.29333/ejmste/16806

Use of machine learning in virtual learning environments: A bibliometric review

Alejandro Valencia-Arias [1*], Ezequiel Martínez Rojas [2], Sebastián Cardona-Acevedo [3], Leonel Alcides Castañeda-Peláez [4]

[1] Escuela de Ingeniería Industrial, Universidad Señor de Sipán, PERU
[2] Vicerrectoría de Investigación e Innovación, Universidad Arturo Prat, Iquique, CHILE
[3] Facultad de Ciencias Económicas y Administrativas, Instituto Tecnológico Metropolitano, Medellín, COLOMBIA
[4] Administración, Universidad Pontificia Bolivariana, Medellín, COLOMBIA

Received 25 April 2025 ▪ Accepted 05 June 2025

Abstract

The efficacy of machine learning (ML) in anticipating and enhancing the learning process of students within various educational contexts has been empirically demonstrated. The research focuses on the use of ML in virtual learning environments, with the objective of improving the effectiveness and personalization of online education. In order to achieve this, it is necessary to examine research trends in this field. Some areas demonstrate the need for a complete bibliography, which is the reason for this objective. The methodology employed a bibliometric analysis, conducted in accordance with the PRISMA 2020 guidelines, utilizing data sourced from Scopus and Web of Science. The results indicate an increase in the production of articles, with a particular focus on topics such as student participation and random forest.

Keywords: dynamic adaptability, PRISMA 2020, artificial intelligence, personalization of education, educational technology

INTRODUCTION

The application of machine learning (ML) in virtual learning environments (VLEs) has become a critical component of modern teaching methods. ML, a subfield of artificial intelligence, has proven effective in predicting and improving student learning outcomes in educational settings, offering significant benefits to educators and learners alike. Notably, research has used deep learning techniques to predict student performance based on VLE-generated data (Waheed et al., 2022). This field of study explores the potential of ML algorithms to analyze large amounts of educational data to enhance the learning experience, personalize teaching, and predict student performance. VLEs, which provide accessible and flexible educational platforms, stand to benefit greatly from the implementation of ML techniques. These techniques enable the continuous adaptation and optimization of educational resources (Chui et al., 2020). ML models have been shown to effectively identify students at risk of underperforming early on, allowing for timely interventions to improve their academic outcomes. Researchers have demonstrated that an ML algorithm can accurately predict which college students are at risk within a VLE, facilitating the implementation of targeted support strategies. This predictive capability is crucial in contemporary education, where personalized monitoring and assistance can substantially impact academic success (Chui et al., 2020).

Similarly, other authors emphasize the importance of assessing students' cognitive and emotional states automatically in virtual reality environments. They highlight the potential of ML to personalize educational training based on individual needs (Moon et al., 2020). ML models have also been used to identify at-risk students and implement early intervention strategies at various educational levels (Adnan et al., 2021).

The application of ML in VLEs is a significant area of contemporary educational research. Recent studies have demonstrated the value of ML techniques in analyzing and understanding student engagement and interaction in learning environments (Raj & VG, 2022). Additionally,

valenciajho@uss.edu.pe **(*Correspondence)** emartinezr@unap.cl sebastiancardona272247@correo.itm.edu.co
leonel_alcides@yahoo.com.mx

Contribution to the literature

- This article provides a comprehensive bibliometric analysis of the use of machine learning in virtual learning environments, highlighting key trends and thematic developments in the field.
- It identifies major conceptual clusters and emerging keywords that define current research priorities and gaps.
- The findings support the design of future research agendas and educational policies that integrate artificial intelligence to improve personalization, engagement, and early intervention in online learning.

the efficacy of predicting which students are at risk of failing autonomous learning was evaluated. It was demonstrated that neural network technology can identify students in need of additional support (Waheed et al., 2023). Deploying ML in VLEs is paramount, as it can transform and optimize education by offering innovative solutions to persistent educational problems. One of the most crucial aspects is these systems' ability to predict school dropouts by identifying students at risk early on. In this regard, a predictive model based on behavioral indicators has been developed to detect students at risk of dropping out. This model provides a valuable tool for implementing preventive interventions and improving retention rates (Hlioui et al., 2021).

Moreover, cloud-powered VLEs have been shown to play a pivotal role in mitigating the academic disparities between urban and rural regions. Research has demonstrated that the implementation of cloud-based VLEs can foster educational equity by ensuring equitable access to high-quality learning resources, irrespective of students' geographic location. This enhanced access has the potential to play a significant role in narrowing the academic proficiency gap between urban and rural students, thereby promoting a more inclusive and equitable education system. Consequently, integrating ML into VLEs not only enhances the personalization and effectiveness of learning but also plays a pivotal role in promoting equal educational opportunities (Safdar et al., 2022). The development of intelligent VLEs is confronted with a multitude of challenges which researchers are striving to surmount. For instance, they have identified several inherent issues associated with the creation of a smart VLE, including the integration of emerging technologies and the necessity for extensive customization at the system level. These challenges underscore the intricacy of developing systems that not only adapt to the individual needs of students but are also scalable and resource-efficient (Odrekhivskyy et al., 2019). A 2019 systematic review provides a foundation for further research in this area, as it demonstrates the need for additional studies to more effectively address the issue (Tamada et al., 2019).

Despite the mounting interest in leveraging ML algorithms to enhance educational outcomes, the extant literature reveals limitations concerning equity and fairness in learning analytics. The absence of comprehensive studies addressing the application of equity principles in the context of VLE signifies a necessity for more profound and systematic research in this area (Riazy et al., 2020). The objective of this research is to examine the current research trends on the use of ML in VLEs. In view of the aforementioned points, the subsequent research inquiries are hereby presented:

1. Which are the years in which the most interest has been presented in the use of ML in VLEs?
2. What type of growth has the number of scientific articles on the use of ML in VLEs shown?
3. What are the main research references on the use of ML in VLEs?
4. What is the thematic evolution derived from the scientific production on the use of ML in VLEs?
5. What are the main thematic clusters on the use of ML in VLEs?
6. What are the growing and emerging keywords in the research field on the use of ML in VLEs?
7. How are the keywords of the scientific literature on the use of ML in VLEs classified according to their function?
8. What themes are positioned as protagonists for the design of a research agenda on the use of ML in VLEs?

The present article is organized into five sections. In the introduction, the topic is contextualized and the relevance of the use of ML in VLEs is highlighted, as well as the objectives of the research. The methodology section delineates the bibliometric approach utilized, encompassing the selection criteria and the analysis of the sources. The results section presents the findings obtained from the analysis of the literature. The findings obtained from the analysis of the literature are then interpreted in the discussion, which relates them to the current state of knowledge and practical implications. In conclusion, the study offers a synopsis of its salient points, proffers suggestions for future research endeavors, and puts forward recommendations for the implementation of ML in the context of VLE.

METHODOLOGY

The present study is of an exploratory nature and is based on secondary sources. A bibliometric analysis is employed to examine the utilization of ML in VLEs. The study adheres to the standards set forth by the PRISMA-

2020 guidelines for conducting systematic reviews, thereby ensuring a rigorous and transparent methodology (Page et al., 2021).

Eligibility Criteria

The inclusion criteria for this research were based on the presence of specific terms in the titles and keywords. The primary metadata for the study were "machine learning" and "virtual learning environments." The search was conducted with consideration of the various forms of citation of these terms, with the aim of ensuring the comprehensiveness of the analysis. The combination of keywords facilitates the identification of relevant studies that address the application of ML techniques in virtual learning contexts, thereby ensuring the inclusion of a comprehensive range of research pertinent to the topic.

The exclusion process was conducted in three phases. Initially, all records exhibiting erroneous indexing were excluded, thereby ensuring that only correctly catalogued documents were analyzed. In the subsequent phase, documents for which the full text was inaccessible were excluded. However, it should be noted that this exclusion only applies to systematic literature reviews. This is due to the fact that, in bibliometrics, the analysis is focused exclusively on metadata. In the final exclusion phase, records with incomplete indexing and texts deemed to be of limited relevance to the study were excluded. This process ensured the quality and relevance of the data analyzed.

Information Sources

In this bibliometric study, the Scopus and Web of Science (WoS) databases were selected as the primary sources of information due to their esteemed reputation and comprehensive coverage of current scientific literature. This approach ensures that bibliometrics draws from a variety of reliable sources to analyze trends and advances in the use of ML in VLEs, as discussed in the study. The present study integrates data from Scopus and WoS for the purpose of conducting citation-based studies (Kumpulainen & Seppänen, 2022).

Search Strategy

To search for the two selected databases, two bespoke search equations were devised. These equations were adapted to the previously defined inclusion criteria and the specificities of each platform. Given the distinctive search parameters of Scopus and WoS, these equations were meticulously crafted to guarantee the integrity and precision of the data collection process.

For the Scopus database: ((TITLE ("Machine Learning" OR "ML" OR "Computational Learning" OR "Automated Learning") AND TITLE ("Virtual Learning Environments" OR "VLE" OR "Online Learning Platforms" OR "Digital Learning Environments" OR "Web-Based Learning Systems")) OR (KEY ("Machine Learning" OR "ML" OR "Computational Learning" OR "Automated Learning") AND KEY (" Virtual Learning Environments" OR "VLE" OR "Online Learning Platforms" OR "Digital Learning Environments" OR "Web-Based Learning Systems"))).

For the WoS database: ((TI = ("Machine Learning" OR "ML" OR "Computational Learning" OR "Automated Learning") AND TI = ("Virtual Learning Environments" OR "VLE" OR "Online Learning Platforms" OR "Digital Learning Environments" OR "Web-Based Learning Systems")) OR (AK = ("Machine Learning" OR "ML" OR "Computational Learning" OR "Automated Learning") AND AK = ("Virtual Learning Environments" OR "VLE" OR "Online Learning Platforms" OR "Digital Learning Environments" OR "Web-Based Learning Systems"))).

Data Management

In the bibliometric context of the use of ML in VLEs, the tool Microsoft Excel® was utilized. In addition, the free software VOSviewer® was used to carry out the bibliometric analysis and to visualize the results obtained. Specifically, VOSviewer® was employed to generate representative graphs of the various bibliometric indicators obtained during the study. This decision was informed by the efficacy and adaptability of VOSviewer® in the analysis of bibliometric data, as outlined in the Methods for analyzing texts using this tool (Bukar et al., 2023). Furthermore, it was noted that the subject was addressed in relation to students' perceptions of data analysis in Microsoft Excel, with the objective of providing a contextual framework for the utilization of this tool within educational settings (Nath, 2021).

Selection Process

According to the PRISMA 2020 statement, as described in the study by Page et al. (2021), it is imperative to indicate whether an internal automated classifier is employed to facilitate the selection process and whether this is done with internal or external validation teams. A critical step in the bibliometric review process is the evaluation of the risk of missing studies or incorrect classification. In this study, automated Microsoft Excel® tools were utilized as an internal classification and selection mechanism. Each researcher utilized the tool autonomously to implement the inclusion and exclusion criteria.

In the initial phase of a bibliometric study that incorporates ML in a VLE, information sources were identified by implementing a specific search strategy in the Scopus and WoS databases. Following a thorough selection and exclusion process, a total of 104 articles were obtained. These articles comprise the central corpus of this bibliometric study, providing a substantial

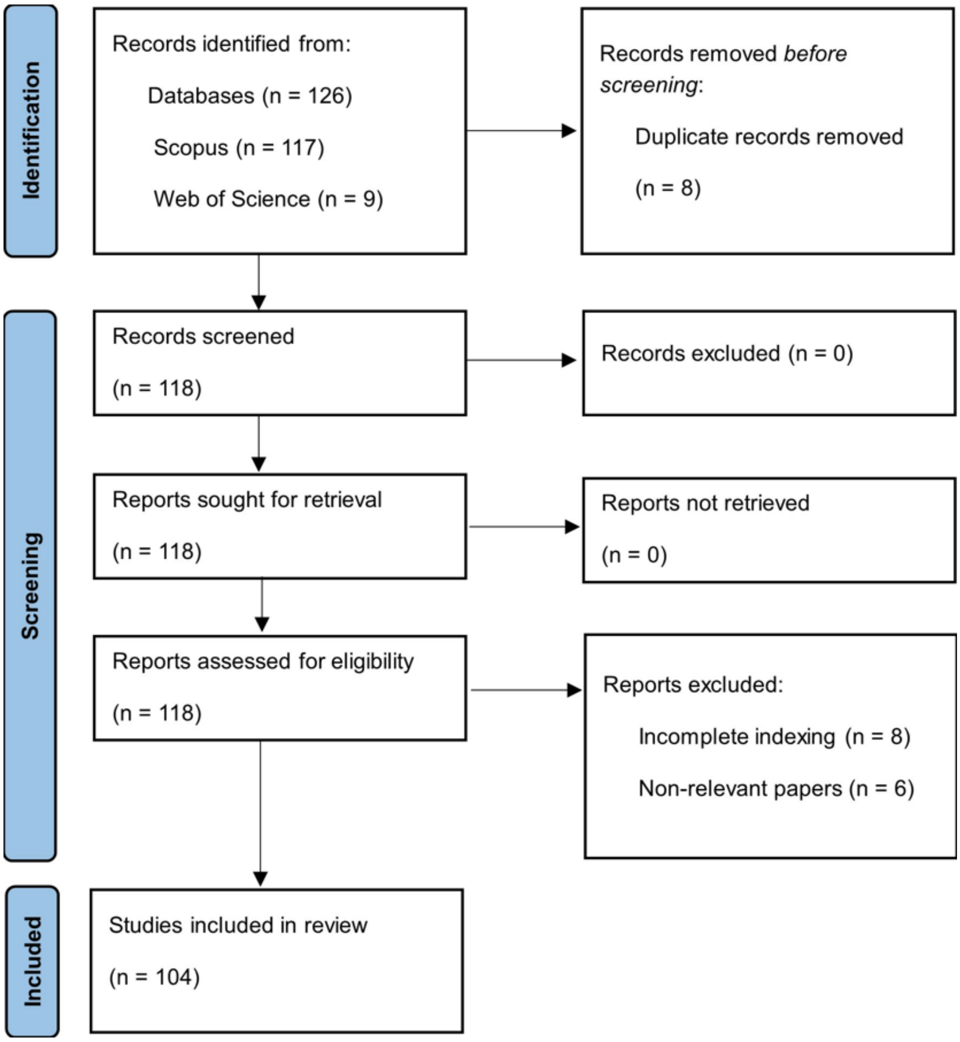

Figure 1. PRISMA-2020 flowchart (Source: Authors' own elaboration, based on Scopus and Web of Science)

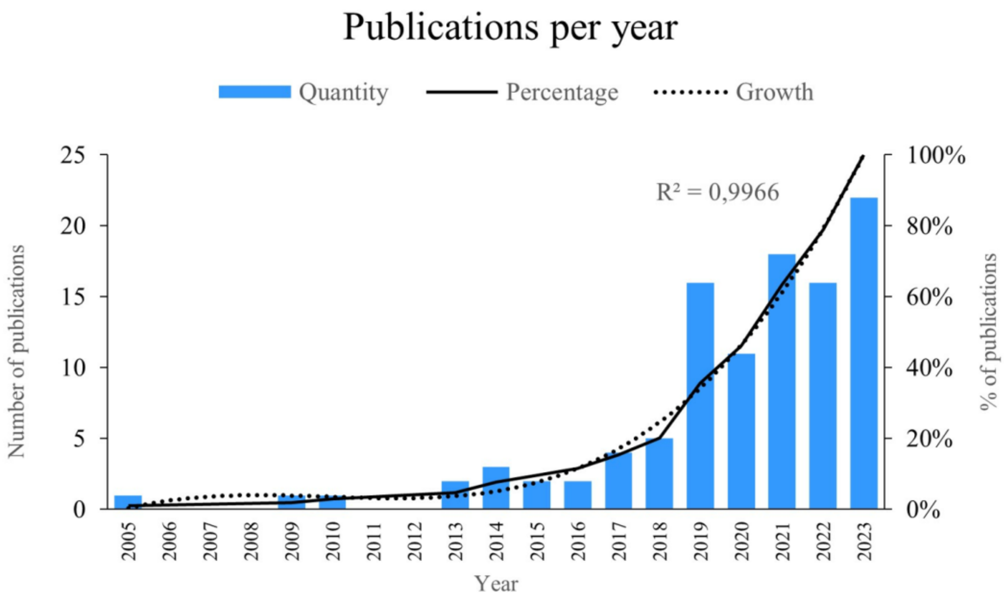

Figure 2. Publications per year (Source: Authors' own elaboration, based on Scopus and Web of Science)

foundation for analyzing trends and advancements in the utilization of ML in VLEs (**Figure 1**).

RESULTS

The current bibliometric analysis has enabled the identification of a substantial growth of 99.66% in academic production related to the utilization of ML in

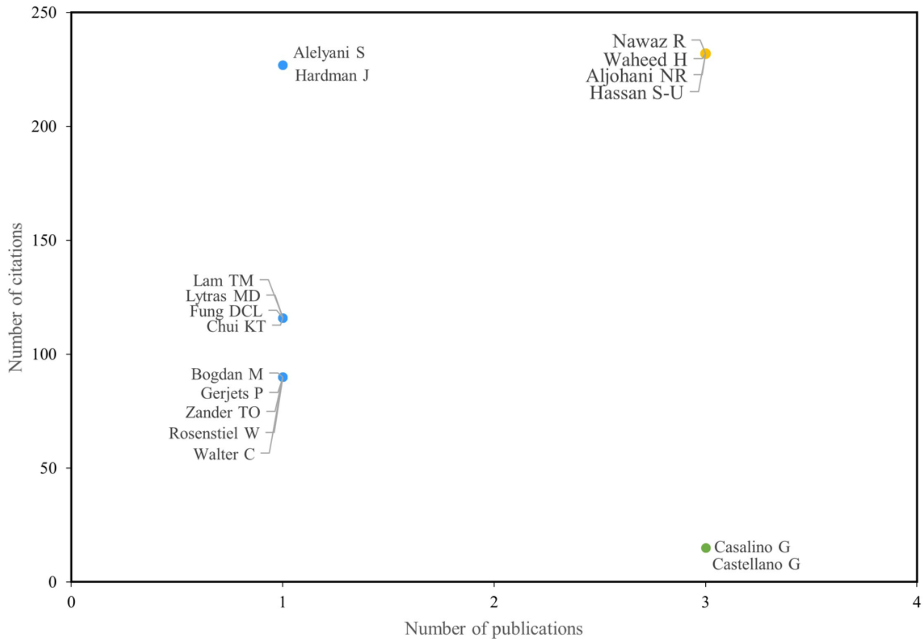

Figure 3. Main authors (Source: Authors' own elaboration, based on Scopus and Web of Science)

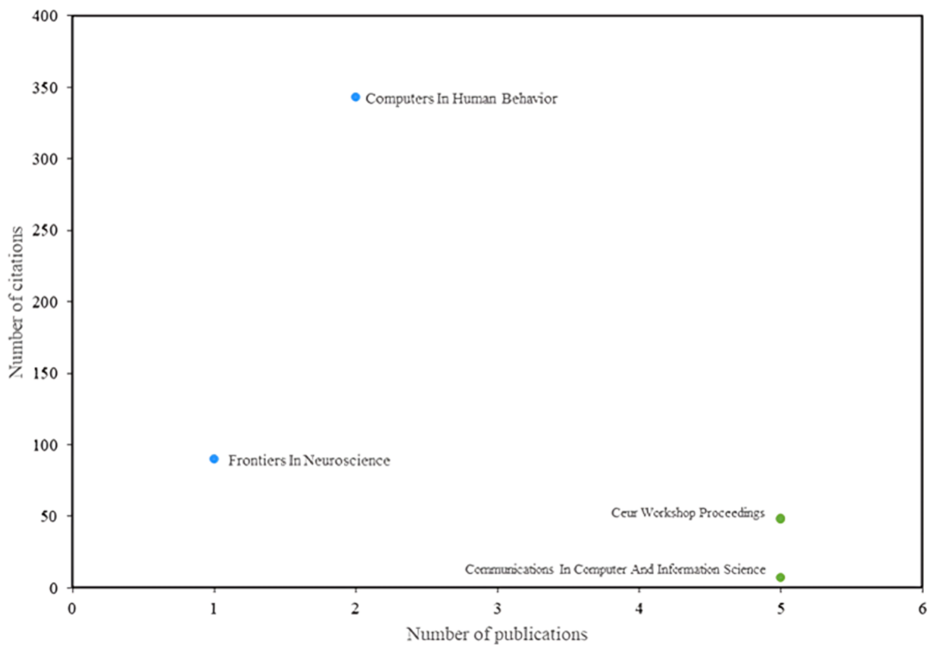

Figure 4. Main journals (Source: Authors' own elaboration, based on Scopus and Web of Science)

VLEs, as illustrated in **Figure 2**. It is noteworthy that the years 2023, 2021, 2019, and 2022 are particularly salient in terms of the quantity of publications on this subject that were documented. The substantial augmentation in scientific production in recent years is indicative of the mounting interest and significance of the implementation of ML in virtual educational environments. This phenomenon is concomitant with the necessity for meticulous analysis and comprehension of trends and developments in this domain.

This bibliometric analysis of the primary authors reveals three distinct groups of researchers who are making significant contributions to the field of the use of ML in VLEs, as illustrated in **Figure 3**. First, it is evident that the authors, spearheaded by Nawaz, R., Waheed, H., Aljohani, N. R., and Hassan, S.-U., are distinguished by their exceptional productivity and considerable research impact. This group has demonstrated its leadership in both the quantity of publications and the significant influence of their research within the scientific community. Conversely, a group of authors has been identified who have produced highly cited research results despite having a low productivity index. Among these authors, Aleryani, S. and Hardman, J. stand out. The third group of authors is distinguished by their exceptional scientific productivity. Aleryani, S., Hardman, J., Casalino, G., and Castellano, G. served as the primary representatives of the aforementioned group, which produced a substantial number of publications.

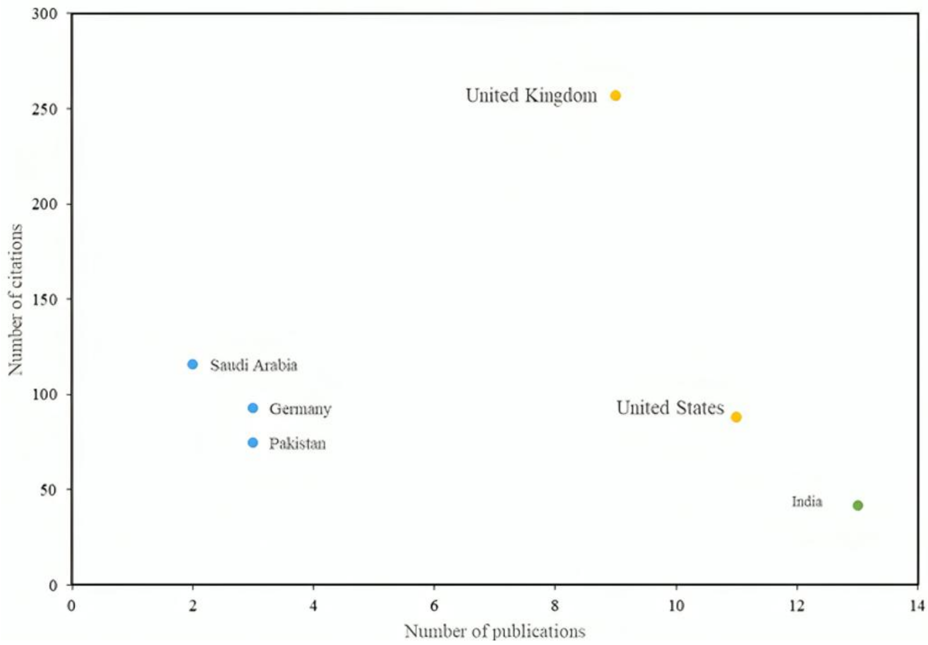

Figure 5. Main countries (Source: Authors' own elaboration, based on Scopus and Web of Science)

The analysis reveals the presence of two distinct groups of top journals, as illustrated in **Figure 4**. The initial group consists of journals that are distinguished by their high impact, despite their low index of scientific productivity. These include Computers In Human Behavior and Frontiers In Neuroscience, which are recognized for the quality and influence of their publications in the field. Conversely, the second group comprises journals that are distinguished by their notable scientific productivity, though not invariably by the number of citations they receive. Within this group, CEUR Workshop Proceedings merits particular attention. Despite receiving a comparatively modest number of citations, this publication makes a substantial contribution to the extant literature on the application of ML in VLEs.

With regard to the leading countries, three discrete clusters were identified, as illustrated in **Figure 5**. The initial cluster is constituted by countries that exhibit a distinctive level of scientific productivity and impact. This includes the United Kingdom (UK) and the United States (USA), which are unquestionable leaders in research on the utilization of ML in VLEs, contributing a considerable number of high-quality publications. In contrast, the second cluster consists of countries that have a notable impact on the scientific community, despite having relatively low productivity. This phenomenon is exemplified by Saudi Arabia and Germany, whose research output, despite its comparatively modest volume, frequently garners significant citations. The third cluster consists of countries that are distinguished primarily by their scientific productivity, although not necessarily by the number of citations they receive. India exemplifies this cluster, contributing significantly to the extant literature on the subject, though its research may receive less attention in terms of citations.

As demonstrated in **Figure 6**, the present study is concerned with the examination of thematic developments in the literature pertaining to the utilization of ML in VLEs, together with the identification of the most frequently addressed topics in each research year from 2005 to 2023. In the initial year of the research, 2005, the concept of a "virtual learning environment" began to emerge, thereby marking the commencement of research into virtual environments in the field of education. Conversely, subjects such as "student engagement," "random forests," "learning analytics," and "machine learning" have assumed prominence in recent years, reflecting heightened student engagement and a growing emphasis on data classification techniques, which are indicative of prevailing trends in research.

This bibliometric analysis offers a comprehensive overview of the keyword co-occurrence network within the field, which is represented through a total of eight thematic clusters, as illustrated in **Figure 7**. The blue cluster is the most prominent, comprising terms such as "learning analytics," "at-risk students," "clustering," "academic performance," and "higher education." This finding suggests the presence of a robust conceptual interrelationship among these fundamental concepts. The purple cluster, which includes terms such as "educational data mining," "oulad," and "big data," among others, suggests a significant thematic association. The remaining clusters, which are represented by light blue, yellow, green, orange, and red, account for other elements of conceptual affinity within the field of study. Consequently, they provide a comprehensive perspective of the main research areas and their interconnections.

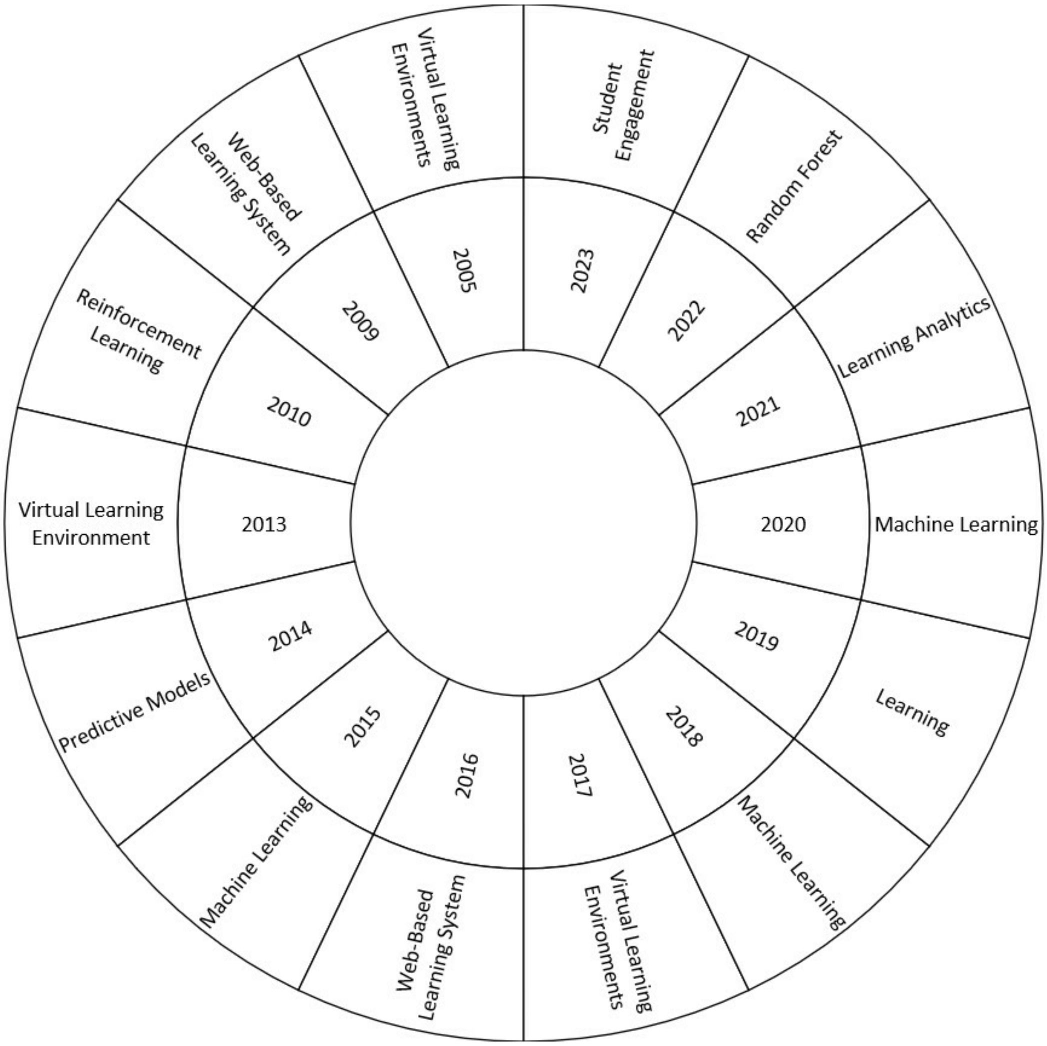

Figure 6. Thematic evolution (Source: Authors' own elaboration, based on Scopus and Web of Science)

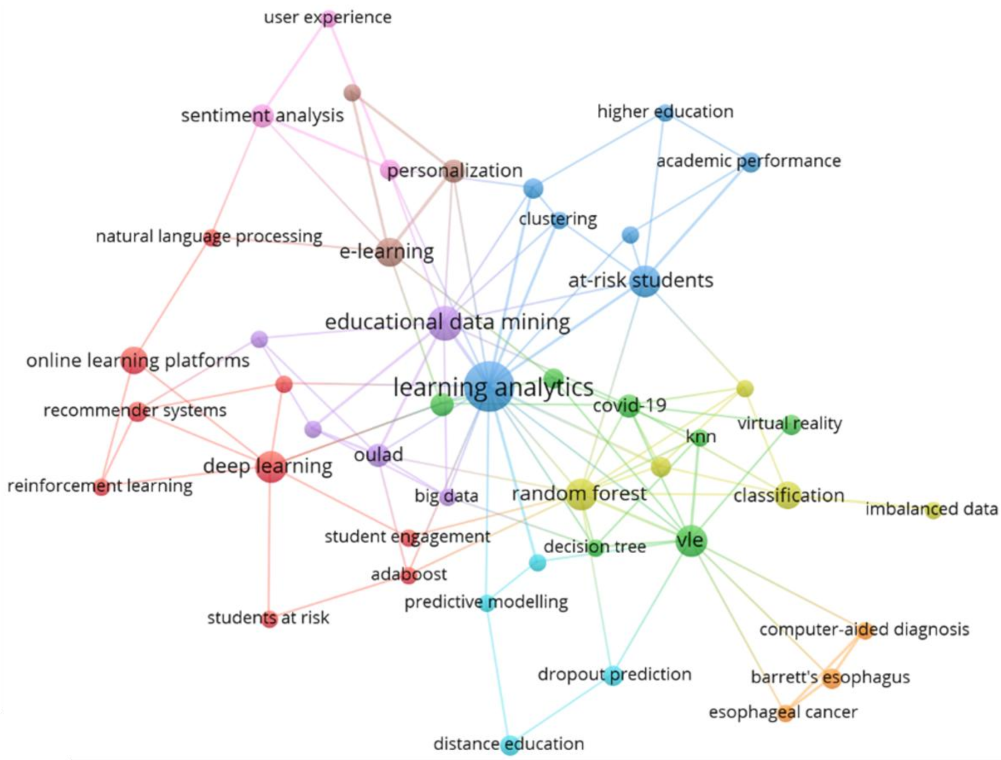

Figure 7. Keyword co-occurrence network (Source: Authors' own elaboration, based on Scopus and Web of Science)

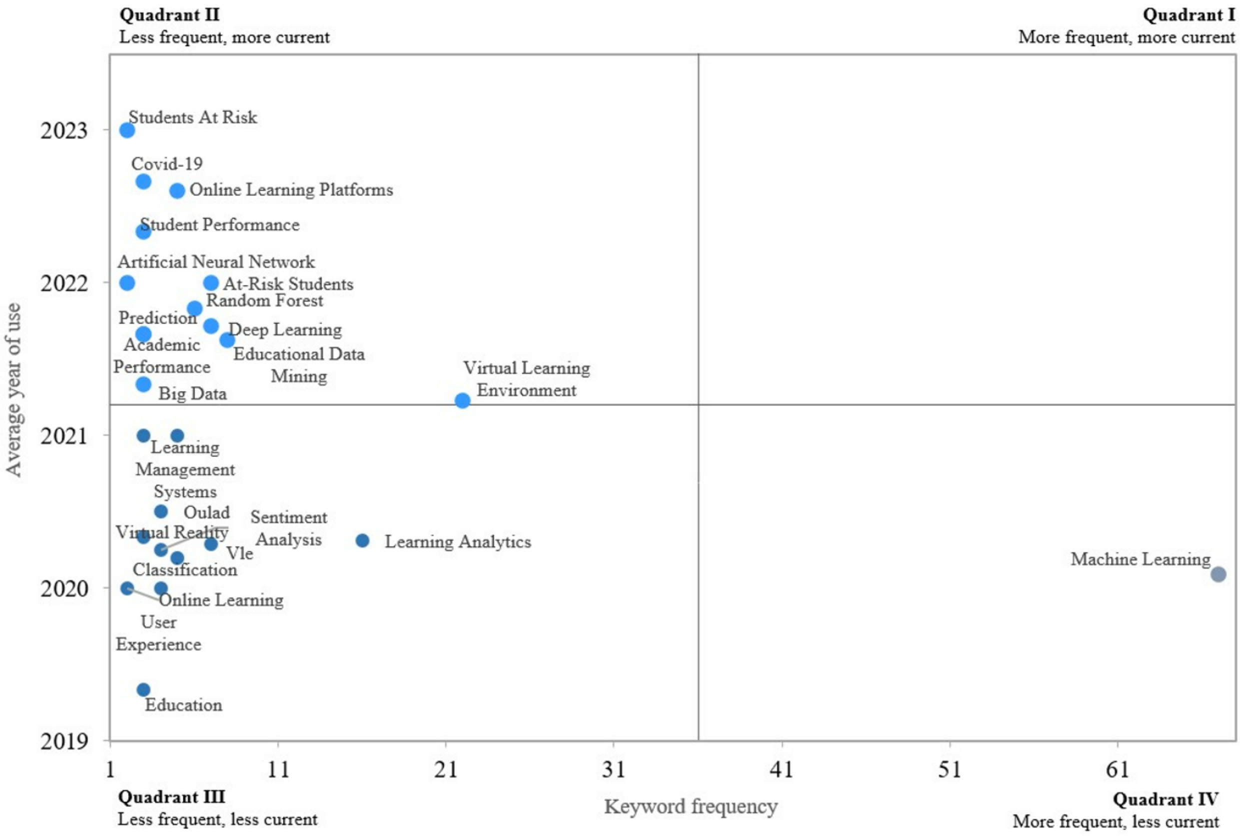

Figure 8. Validity and frequency of keywords (Source: Authors' own elaboration, based on Scopus and Web of Science)

As part of this study, **Figure 8** presents a Cartesian plane that combines the frequency of keyword use on the X-axis with the effectiveness of keyword use on the Y-axis, resulting in the formation of four distinct quadrants. The fourth quadrant is distinguished by a combination of declining frequency of use and freshness, as evidenced by the presence of terms such as "machine learning." Conversely, the second quadrant encompasses keywords that, while infrequent, maintain a high degree of validity. These terms are defined as emerging keywords. Examples of such terms include "at-risk students," "Covid-19," "online learning platforms," "student performance," "artificial neural networks," and "random forests."

DISCUSSION

This discussion section plays a pivotal role in this bibliometric study, offering a comprehensive analysis of the findings from research on the utilization of ML in VLEs. This section provides a detailed examination of trends, patterns, and new findings derived from an analysis of the scientific literature. Additionally, it delves into the practical ramifications of the findings and illustrates how they can contribute to the enhancement and evolution of online education, as well as the implementation of ML methodologies. Furthermore, the limitations of the study are addressed, and potential methodological limitations are identified, as well as avenues for future research.

Analysis of the Growth of Scientific Literature on the Use of Machine Learning in Virtual Learning Environments

In order to broaden the discourse and establish a link between the findings and relevant educational policies and pedagogical strategies, it is imperative to acknowledge the transformative capacity of ML in optimizing student outcomes and institutional decision-making. A growing body of research, particularly from 2023, 2021, 2019, and 2022, has demonstrated a paradigm shift toward data-driven education, where predictive analytics are increasingly used to personalize learning experiences and support at-risk learners.

The study by Badal and Sungkur (2023) demonstrates the integration of ML into VLEs to inform teaching strategies by identifying patterns in student engagement and performance. This finding carries substantial ramifications for educational policies that advocate evidence-based instruction and the integration of adaptive learning technologies. The incorporation of these models into national or institutional learning management systems empowers policymakers to mandate the early identification of struggling students and the timely deployment of targeted support services.

Adnan et al. (2021) also support the importance of continuous student monitoring through automated systems. From a pedagogical perspective, the implementation of ML-based alerts and dashboards empowers educators to adopt proactive and differentiated teaching approaches. For instance, instructors have the capacity to provide customized feedback, adjust course pacing, and design remedial activities tailored to the needs of students. From a policy perspective, this initiative supports the implementation of programs that prioritize the following:

(1) the retention of learners,

(2) the promotion of digital inclusion, and

(3) the incorporation of artificial intelligence into teacher training curricula.

Analysis of Research References on the Use of Machine Learning in Virtual Learning Environments

The lead authors, Nawaz, R. and Waheed, H., in collaboration with Aljohani, N. R., and Hassan, S.-U., have made a substantial contribution to the field of research on the utilization of ML in VLEs. The study conducted by Waheed, H. focused on the early prediction of at-risk students in a self-directed educational environment using neural networks. This finding is of considerable significance for the enhancement of student retention and the provision of personalized support (Waheed et al., 2023). In contrast, Alelyani, S. and Hardman, J., in collaboration with Nawaz, R., have made a notable contribution to the field through their study, which focuses on predicting students' academic performance from large data sets in VLEsusingdeep learning models (Waheed & Shafi, 2020).

The academic journals Computers in Human Behavior and Frontiers in Neuroscience have played a pivotal role in propelling the field forward by facilitating the dissemination of knowledge concerning the application of ML in VLEs. To illustrate, noteworthy contributions have been made to the field through the development of deep learning models designed to predict student academic performance based on extensive data sets derived from VLEs. This research postulates that the proliferation of accessible educational data, in conjunction with technology-enhanced learning platforms, provides avenues for investigating student learning behaviors (Waheed & Shafi, 2020). Similarly, Frontiers in Neuroscience has functioned as a substantial platform for studies investigating cognitive state monitoring and adaptive instructional design in digital environments.

This phenomenon is exemplified by research on assessing cognitive load using passive brain-computer interfaces (Gerjets et al., 2014). Conversely, CEUR Workshop Proceedings has distinguished itself in terms of productivity, providing a platform for the presentation of nascent research and preliminary findings in the domain of utilizing ML in VLEs. The challenges associated with the development of intelligent VLEs were elucidated, offering a critical perspective on the technical and practical obstacles encountered by researchers in this rapidly evolving field (Odrekhivskyy et al., 2019).

It is evident that countries with a high level of scientific production have played a pivotal role in advancing knowledge in this field. To illustrate, the UK and the USA have both made notable contributions to productivity and impact, providing pivotal research that has propelled the field forward. In Waheed and Shafi's (2020) study, a sophisticated artificial neural network is employed on a set of distinctive hand-crafted features extracted from clickstream data of VLEs to predict at-risk students and furnish measures for early intervention in such cases. Concurrently, other research underscores the significance of web-based intelligent tutoring systems, offering foundational insight into the utilization of artificial intelligence in pedagogical contexts (Piramuthu, 2005).

In contrast, other countries, including Saudi Arabia and Germany, have emerged as leaders in this field, contributing significantly to the advancement of research in ML in VLEs. An ML algorithm was presented to predict at-risk university students in VLEs, thereby emphasizing the significance of such tools in enhancing the educational experience (Chui et al., 2020). Furthermore, the exploration of cognitive state monitoring and adaptive instructional design in digital environments yielded invaluable insights into the optimization of learning effectiveness through the utilization of intelligent technologies (Gerjets et al., 2014).

Furthermore, India has exhibited noteworthy proficiency in productivity, particularly in the domain of VLEs. The advent of ML techniques has enabled the prediction of student engagement in such environments, thereby offering invaluable insights into the enhancement of student engagement and performance in digital contexts (Raj & VG, 2022).

Analysis of the Thematic Evolution on the Use of Machine Learning in Virtual Learning Environments

A thorough examination of the evolution of bibliometric themes concerning the implementation of ML in VLEs reveals a discernible progression in conceptual approaches through time. In 2005, the concept of the "VLE" was the subject of the greatest number of studies. These preliminary investigations were chiefly concerned with the creation and adaptation of VLEs (Piramuthu, 2005). However, in recent years, there has been a notable shift towards concepts such as "student engagement" and "random forests," as evidenced in the years 2023 and 2022. This increased focus is evidenced by a growing emphasis on student engagement and the integration of ML algorithms within educational settings. This development suggests a shift towards more sophisticated research methodologies and a holistic approach to enhancing the quality of teaching and learning in virtual environments (Badal & Sungkur, 2023).

Analysis of the Thematic Clusters on the Use of Machine Learning in Virtual Learning Environments

The analysis of the network of keyword co-occurrence in the field revealed the existence of various thematic clusters, which reflect areas of conceptual affinity. To illustrate, the primary cluster, represented by the blue color, comprises keywords such as "learning analytics," "at-risk students," "clustering," "academic performance," and "higher education." This cluster of articles demonstrates a robust interconnection between concepts pertaining to the identification of at-risk students and the analysis of academic performance in VLEs. Significant contributions have been made to this field by the development of ML models for the prediction and support of at-risk students in virtual environments (Chui et al., 2020; Shafiq et al., 2022).

In contrast, the second most significant cluster, represented by the purple color, encompasses keywords such as "educational data mining," "oulad," and "big data." This cluster underscores the pivotal role of educational data analysis in informing decision-making processes and enhancing teaching and learning in virtual environments. The diversity and complexity of the field of study was identified, as well as the necessity to address different aspects to advance the understanding and application of ML in virtual education. This objective may be realized through the exploration of how the utilization of time and educational data can enhance the adaptability of learning in digital environments (Casalino et al., 2020).

Analysis of the Frequency and Conceptual Validity Around the Use of Machine Learning in Virtual Learning Environments

In the fourth quadrant, the concept of ML was found to be in decline. A noteworthy study addresses the issue of early identification of students who are at risk. The research concentrated on the prediction of at-risk students at varying stages of the course, thus facilitating the implementation of efficacious early interventions. This dynamic and proactive approach exemplifies the potential of ML to adapt to various conditions and time periods, thereby offering a more complete and accurate view of educational scenarios. The findings of the study offer valuable insights into the optimization of early intervention strategies for the enhancement of student success, while also underscoring the versatility and utility of ML in the field of education (Adnan et al., 2021).

The second quadrant of the Cartesian plane of bibliometrics offers a promising avenue for exploring a novel concept of significant importance in the scientific domain of ML in VLEs. This approach is particularly salient in the present and imminent future, especially in the context of salient keywords such as "at-risk students," "Covid-19," and "online learning platform." In the context of the challenges inherent to online education amid the global pandemic caused by the SARS-CoV-2 virus, these key terms are of particular importance. A number of studies have examined these challenges, contributing to improvements in educational quality and offering valuable insights for the evolution of online education through the use of advanced ML techniques (Alsayat & Ahmadi, 2023; Colpo et al., 2024; Waheed et al., 2022).

Classification of Keywords on the Use of Machine Learning in Virtual Learning Environments According to Their Function

As indicated by the results of the bibliometric analysis, **Table 1** plays a pivotal role in the study. The categorization of emerging and growing keywords related to the utilization of ML in VLEs that is provided in **Table 1** is comprehensive. The categorization is further characterized by a clear delineation of their functionalities.

This classification enables the identification and analysis of the principal characteristics and applications of each classified function, thereby facilitating a comprehensive understanding of the diversity of approaches and research areas within this field.

Theoretical Implications

The analysis of the frequency of publications per year demonstrates a consistent upward trajectory in the level of attention devoted to this nexus between educational technology and automated learning. This increase in publications signifies a mounting interest among the academic community in comprehending and leveraging the capabilities of ML within the domain of virtual education.

By examining the principal theoretical references, the conceptual foundations underlying research in this domain can be identified. These references provide a robust framework for understanding and contextualizing empirical studies and practical applications of ML in online learning environments. Thematic evolution illustrates the heterogeneity of methodologies and domains of interest within the field, encompassing personalization of learning and the early identification of at-risk students.

The analysis of keyword co-occurrence offers a comprehensive understanding of the interconnections between topics and concepts within the scientific literature on this topic. This analysis elucidates the interconnections and interdependencies between disparate research domains, thereby facilitating a comprehensive understanding of the field. Furthermore, an analysis of emerging and growing keywords suggests the emergence of novel trends and the rapid evolution of

Table 1. Classification of keywords according to their function (prepared by the authors based on Scopus and WoS)

Keyword	Associated tools	Applications	Characteristics
Students at risk	Predictive analytics, ML models, learning analytics	Early identification of students at risk, intervention strategies, student support	Early identification of at-risk students, targeted intervention strategies, improved student outcomes
Covid-19	Natural language processing, data analysis, ML	Pandemic response, educational continuity, public health monitoring	Real-time monitoring of the impact of COVID-19 on education, adaptation of teaching methods and health knowledge
Online learning platforms	Learning management systems, learning analytics, artificial intelligence	Online learning, student engagement, data-based instruction	Seamless delivery of educational content, analytics to improve student engagement, and data-driven instruction
Student performance	Learning analytics, data analysis, predictive modeling	Academic assessment, adaptive learning, early intervention	Evaluation of individual and group academic performance, early intervention for students with difficulties and personalized learning paths
Artificial neural networks	Deep learning, deep neural networks, pattern recognition	Pattern recognition, image analysis, natural language processing	Advanced pattern recognition, image and text analysis, and decision-making systems that mimic human neural processing
Random forest	Decision trees, ensemble learning, ML	Data classification, predictive modeling, anomaly detection	Robust data classification, effective prediction modeling, and effective anomaly detection

research domains. This analysis functions as a compass, guiding the future trajectory of research in this domain.

A primary benefit of bibliometric analysis is its capacity to identify research gaps. These gaps signify domains within the field of ML in VLEs that have received inadequate exploration or study, thus indicating prospective avenues for future research endeavors. Addressing these gaps in knowledge is therefore essential for researchers seeking to contribute to a more comprehensive and nuanced understanding of this field, as well as to the continuous improvement of educational practices based on the use of ML.

Practical Implications

The thematic evolution towards a more profound examination of aspects related to student participation signifies a shift in research focus, moving from the mere exploration of VLEs to a more detailed analysis of the factors that influence the active participation of students in these environments. This shift is indicative of an increasing emphasis on understanding and enhancing the online learning experience, as well as on developing effective strategies to promote student participation and engagement.

The analysis of keywords and the identification of thematic clusters provide valuable insight into the predominant areas of interest and focus within the academic literature on this topic. The substantial presence of terms such as "learning analytics," "at-risk students," "clustering," "academic performance," and "higher education" in the primary thematic cluster underscores the importance of these concepts in research on the utilization of ML in VLEs. This conceptual affinity indicates the necessity for further investigation into the interrelationship between these areas and the potential for their effective integration into educational practice to enhance learning outcomes.

Conversely, an examination of the frequency and recency of keywords indicates emerging and declining trends within the field. The observed decline in the frequency of terms such as "machine learning" suggests a potential shift in research interest towards other areas or approaches. Conversely, the increasing frequency of keywords such as "at-risk students," "Covid-19," "online learning platforms," "student performance," "artificial neural networks," and "random forests" signifies a heightened focus on salient and emerging issues within the domain of online education.

By identifying research trends and emerging topics, this tool provides crucial information for policymakers and educational leaders regarding the allocation of resources and efforts to enhance the quality and effectiveness of online education. Moreover, it enables institutions to anticipate and adapt to changes in the educational landscape, thereby providing a solid basis for strategic planning and the implementation of policies that promote the effective integration of ML technologies in teaching and learning. A further notable practical consequence of bibliometrics in this context is its capacity to inform the training and professional development of teachers and practitioners in the field of online education. By identifying the most pertinent research domains and pivotal evolving concepts, this tool can inform the development of training programs that address the current and future needs of educators in the effective utilization of ML technologies. It is imperative to ensure that practitioners possess the necessary skills and knowledge to utilize these tools effectively in their pedagogical practices. This will enhance the quality and impact of online learning.

Moreover, the findings of this study may have practical ramifications with regard to the promotion of interdisciplinary collaboration and the establishment of research networks. By identifying areas of convergence and relationships between different disciplines and research topics, this tool can facilitate connections between researchers from disparate fields, thus fostering collaboration and knowledge sharing. Interdisciplinary collaboration is imperative for effectively addressing the intricate and multifaceted challenges associated with the integration of ML in VLEs. Such collaboration has the potential to yield substantial advancements in the comprehension and implementation of these technologies within the educational sector.

The results of this bibliometric analysis contribute to the identification of key research trends in the use of ML technologies in virtual environments. Moreover, they offer an empirical basis for the design and implementation of informed educational policies. In this regard, the results can guide policymakers in the development of evidence-based pedagogical strategies aimed at improving student engagement, personalizing learning, and enabling the early identification of at-risk students. For instance, the mounting emphasis on terms such as "at-risk students" and "learning analytics" underscores the necessity to institute institutional monitoring systems that integrate these technologies to avert attrition and foster educational inclusivity.

In terms of pedagogical strategies, the observed thematic evolution underscores the importance of designing student-centered teaching methodologies that foster active participation and autonomous learning. The incorporation of clustering algorithms or neural networks into learning platforms can facilitate the adaptation of content and learning pace to the individual characteristics of each student, aligning with the fundamental principles of adaptive instructional design. This practical orientation is reinforced by the prominence of keywords such as academic performance and online learning platforms, which reflect a growing concern with measuring and improving academic outcomes through the strategic use of data and intelligent technologies.

Furthermore, these findings can inform the development of curricula for teacher training programs, with an emphasis on cultivating advanced digital competencies. These competencies are designed to equip educators with the ability to understand and apply ML tools in both an effective and ethical manner. It is imperative that teacher training not be confined to the mere cultivation of technical aptitudes. Instead, it must encompass the cultivation of pedagogical, ethical, and collaborative competencies, in consideration of the intricacies inherent in digital learning environments. The identification of thematic clusters can function as a roadmap for the structuring of professional development content and the promotion of communities of practice focused on educational innovation.

Finally, the promotion of policies that encourage interdisciplinary research–in accordance with the findings–can strengthen institutional capacities to generate innovative solutions to the current challenges of online education. Initiatives that foster collaboration among experts in computer science, education, psychology, and sociology can contribute to the development of more inclusive and context-aware technologies. Bibliometrics, therefore, contribute not only to academic knowledge but also act as a strategic tool to align research, pedagogical practice, and public policy toward a more effective and equitable digital education.

Limitations

It is imperative to acknowledge the inherent limitations of bibliometric analysis. The validity and usefulness of this approach are contingent upon the availability and accuracy of data within the selected database. The quality of data can vary considerably between different sources, potentially introducing biases into the results and affecting the reliability of the conclusions drawn.

Furthermore, the selection of keywords and search criteria can have a significant impact on the results obtained. The employment of particular nomenclatures may impose limitations on the extent of the analysis, potentially resulting in the oversight of pertinent research that utilizes disparate terminology. This can result in a partial or biased view of the academic landscape on the topic in question.

It is also imperative to consider the possibility that the selective inclusion of certain publications may introduce biases into the results. The absence of uniformity in the establishment of inclusion/exclusion criteria has the potential to result in an incomplete or distorted representation of the extant body of literature. This phenomenon has the potential to influence the perception of the prevalence or significance of specific topics or methodologies within the field of study.

Additionally, the dynamics of scientific publication, including review and publication periods, can result in delays in the incorporation of new works into the database utilized for bibliometric analysis. This phenomenon can result in the propagation of an antiquated representation of research in the designated area of interest. This, in turn, engenders limitations in the comprehension of emergent trends and recent areas of interest.

Investigative Gaps

As illustrated in **Table 2**, a synopsis of the predominant research lacunae that have been ascertained is presented. These lacunae necessitate further consideration in subsequent research endeavors. **Table 2** offers a synopsis of the primary research lacunae that have been identified within the domain of the implementation of ML in VLEs.

Table 2. Research gaps (prepared by the authors based on Scopus and WoS)

Category	Investigative gaps	Justification	Questions for future researchers
Thematic gaps	1. Evaluation of the long-term impact of the use of ML on students' academic performance.	Despite growing interest in the use of ML in VLEs, there is a paucity of studies investigating its long-term impact on students' academic performance.	What are the long-term effects of using ML on the academic performance of students in virtual environments? What factors influence these effects?
	2. Research on the adaptability of ML models in different educational and cultural contexts.	The effectiveness of ML models may vary depending on the educational and cultural context in which they are implemented. However, there is a lack of research on how to adapt these models to meet the specific needs and characteristics of different environments.	How can ML models be adapted to be effective in different educational and cultural contexts? What are the specific challenges associated with this adaptation?
	3. Studies on the ethical and privacy implications of the use of ML algorithms in online education.	The use of ML algorithms in VLEs raises ethical and privacy concerns that require broader attention from the academic community. However, there is a lack of comprehensive research into these implications and how to mitigate their potential negative impacts.	What are the main ethical and privacy concerns associated with the use of ML algorithms in online education? How can these concerns be addressed effectively?
Geographic gaps	1. Lack of studies on the use of ML in VLEs in developing countries.	Most studies on the use of ML in online education have been carried out in developed countries, which limits our understanding of how these technologies can be implemented and be successful in resource-limited contexts.	What are the specific challenges and opportunities associated with implementing ML technologies in VLEs in developing countries? What strategies can be effective in overcoming these challenges?
	2. Lack of research on the differences in the implementation and effectiveness of ML models in VLEs in different geographic regions.	Socioeconomic, cultural, and infrastructure characteristics can influence the implementation and effectiveness of ML models in VLEs. However, there is a lack of comparative research examining these differences between geographic regions.	What regional factors influence the implementation and effectiveness of ML models in VLEs? How can these differences be addressed to improve equity and accessibility in digital education?
Interdisciplinary gaps	1. Need for interdisciplinary collaboration to address the complex ethical and privacy implications of using ML in online education.	Fully understanding the ethical and privacy implications of using ML in online education requires collaboration between experts in technology, ethics, law, and education. However, interdisciplinary research in this field is limited.	How can researchers from different disciplines collaborate effectively to address the complex ethical and privacy implications of using ML in online education? What strategies can be implemented to foster interdisciplinary collaboration in this field?
	2. Exploration of the interaction between the design of VLEs and the implementation of ML technologies.	The effective design of VLEs is essential to facilitate the successful implementation of ML technologies. However, research examining the interaction between these two aspects is limited, making it difficult to optimize the integration of these technologies in digital educational environments.	How can the design of VLEs be improved to facilitate the effective implementation of ML technologies? What design strategies can maximize the positive impact of these technologies on the student learning experience?
Temporal gaps	1. Need for longitudinal studies that evaluate the long-term impact of the use of ML on the learning experience and academic performance of students.	Most existing studies on the use of ML in online education are cross-sectional, which limits our understanding of the long-term Impact of these technologies on learning and academic performance. Longitudinal studies are required to address this gap.	What is the long-term impact of using ML on the learning experience and academic performance of students? How do these effects vary over time and what factors influence them?
	2. Research on the evolution of practices and policies related to the use of ML in online education.	The field of ML is constantly evolving, which poses challenges and opportunities in the educational field. However, research on the evolution of practices and policies related to the use of these technologies in online education is limited.	How have practices and policies related to the use of ML in online education evolved over time? What are the emerging trends and how can they influence the design and implementation of ML technologies in the future?

The aforementioned gaps are classified into four categories: thematic, geographic, interdisciplinary, and temporal. For each category, two or three significant gaps are identified, accompanied by a brief justification outlining the rationale for their designation as such. In addition, the inquiries presented herein are designed to serve as a framework for future researchers, guiding them in addressing these lacunae and contributing to the advancement of knowledge in this emerging field.

Research Agenda

The utilization of extensive data sets, commonly referred to as big data, is of paramount importance in the context of the deployment of ML in VLEs. The utilization of extensive data sets facilitates the implementation of ML algorithms, thereby enabling the discernment of patterns, trends, and latent relationships in student conduct and the efficacy of pedagogical strategies. Notwithstanding its significance, big data has become a relatively minor topic of recent research. It is recommended that future studies reinvigorate its importance by investigating new techniques and methodologies for the effective collection, management, and analysis of large-scale data in VLEs. In addition, it would be advantageous to investigate the potential of integrating big data techniques to enhance the personalization of learning and to furnish valuable information that could facilitate more informed educational decision-making.

Learning analytics has emerged as a potent instrument for the comprehension and enhancement of the learning experience in virtual environments. The application of data analysis techniques, such as ML, enables educators and course designers to obtain comprehensive insights into student progress, the efficacy of learning materials, and the factors that influence academic performance. Despite its continued relevance, attention towards learning analytics has decreased in recent research. Nevertheless, recent studies have the potential to reinvigorate this field by investigating the manner in which sophisticated ML methodologies can enhance the precision and practicality of learning analytics. In a similar vein, the potential of incorporating learning analytics to enhance feedback and personalization in VLEs could be examined. This examination could contribute to the re-emergence of learning analytics as a pivotal concept in the field.

Educational data mining is a field of study that employs data analysis techniques to identify and examine patterns in data generated by students and learning systems. This domain is pivotal to the implementation of ML in VLEs, as it facilitates the identification of relationships between key variables, such as academic performance and student behaviors. Notwithstanding its significance, the focus on educational data mining has diminished in recent research endeavors. However, the potential for future studies to reinvigorate its relevance lies in the exploration of novel data analysis techniques and ML algorithms meticulously engineered to tackle the distinctive challenges inherent to online education. In addition, the potential benefits of educational data mining in enhancing the early identification of at-risk students and personalizing learning in virtual environments should be investigated.

Deep learning is a subfield of ML that utilizes artificial neural networks, which are composed of multiple layers, to model and represent data in increasingly complex and abstract ways. In the context of ML in VLEs, deep learning has the potential to enhance the precision and efficacy of predictive and analytical models. Despite its acknowledged significance, there has been a notable decline in the focus on deep learning in recent research. Nevertheless, the potential exists for future studies to reinvigorate its relevance by investigating novel neural network architectures and deep learning algorithms specifically designed to address the challenges of online education. Furthermore, they could investigate the potential of deep learning to enhance the comprehension of student behavior and personalize educational content in virtual environments.

It is imperative to identify and address at-risk students in the context of ML in VLEs, as these students often face a multitude of academic and personal challenges that can significantly impact their educational success. Despite the growing recognition of the significance of this issue, recent research has indicated a decline in the attention devoted to at-risk students. However, the potential for future studies to reinvigorate their relevance exists. Such studies could investigate new ML techniques and models for the early detection of at-risk students in virtual environments. In addition, the potential benefits of data analysis and learning personalization in enhancing support and intervention targeting for these students should be investigated. This exploration could contribute to their reintegration by focusing on the most pertinent contemporary concepts.

VLEs are technological platforms designed with the specific purpose of facilitating online teaching and learning. These systems facilitate the dissemination of educational material, the facilitation of student-teacher interaction, and the evaluation of learning outcomes. Despite their significance in the context of ML applications in virtual education, there has been a notable decline in the focus on VLE in recent research. Nevertheless, the integration of ML techniques holds considerable potential for reinvigorating the relevance of VLE by enhancing its efficacy and usability. Furthermore, the adaptation and evolution of VLE to meet the evolving needs of students and educators in VLEs is a promising avenue for future research.

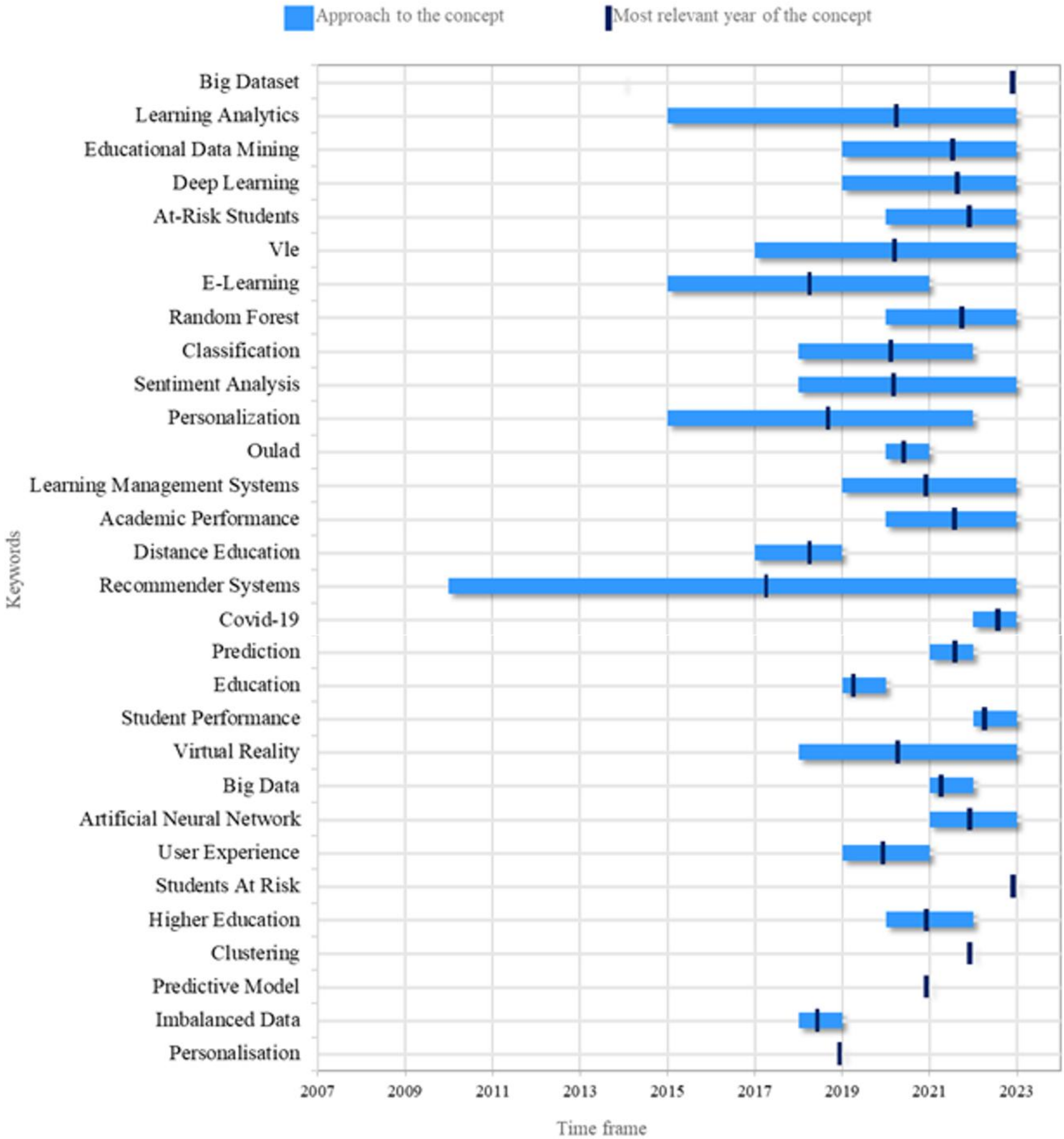

Figure 9. Research agenda (Source: Authors' own elaboration, based on Scopus and Web of Science)

Recommender systems are tools that utilize ML algorithms to suggest content or activities tailored to the user's expressed preferences and past behaviors. In the context of ML utilization in VLEs, recommender systems have the potential to enhance the learning experience by offering students suggestions of pertinent educational resources that are tailored to their individual requirements. Despite the acknowledged importance of recommender systems, there has been a decline in research attention toward them in recent times. Nevertheless, subsequent studies have the potential to reinvigorate their pertinence by investigating innovative techniques and models of personalized recommendation based on ML. These future investigations could explore the methods by which recommender systems can be effectively integrated into VLEs to enhance learning personalization and promote greater student engagement (**Figure 9**).

CONCLUSIONS

A review of the literature on the use of ML in VLEs provides valuable insights. A substantial surge in article production was evident in 2023 and 2021, attributable to the proliferation of cubic polynomials. Notable contributors include Nawaz, R. and Waheed, H. Among the most frequently cited journals are Computers in

Human Behavior and Frontiers in Neuroscience. Additionally, the UK and the USA are noteworthy as prominent countries, while the main thematic clusters offer insight into the current research landscape. These findings can inform researchers' decisions regarding publication venues and international collaborations.

The thematic development has undergone a shift in focus towards topics such as student engagement and Randolph forest. The emergence of novel keywords, such as "students at risk" and "Covid-19," signifies a responsiveness to the evolving demands of online education. The strategic allocation of keywords in research outputs underscores the imperative for future studies within this rapidly expanding field of inquiry to prioritize fundamental concepts.

Furthermore, the identification of research gaps provides the foundation for future studies that can address these knowledge gaps. Thematic, geographical, interdisciplinary, and temporal discrepancies indicate areas that necessitate further attention and investigation. These discrepancies present both challenges and opportunities for the academic community. These resources offer avenues for expanding theoretical understanding and for addressing practical problems in the implementation of ML technologies in VLEs.

In conclusion, the proposed research agenda, based on the results of the bibliometrics, serves as a guide for future studies in this field. This agenda identifies priority areas and emerging topics that require additional research to advance knowledge and practice in the use of ML in VLEs. The provision of a clear and specific direction is instrumental in enabling researchers to concentrate their efforts on areas that have the potential to make a significant contribution to the field. This direction also facilitates the identification of challenges and opportunities in bibliometrics.

Author contributions: EMR: contributed to the conception of the study, definition of objectives, and overall supervision of the research process; **SC-A:** was responsible for the methodological design and the collection of bibliometric data; **LAC-P:** conducted the analysis of results and the creation of visualizations; & **AV-A:** supported the writing of the manuscript and the critical review of its content. All authors agreed with the results and conclusions.

Funding: No funding source is reported for this study.

Ethical statement: The authors stated that this study involved no human or animal participants, and therefore did not require ethical approval.

Declaration of interest: No conflict of interest is declared by the authors.

Data sharing statement: This study's data availability statement has been duly registered and archived in the Zenodo open data repository, which is recognized for its commitment to the accessibility and preservation of scientific data. The data and materials presented in this study are publicly accessible and can be accessed via the following link: https://doi.org/10.5281/zenodo.14109407.

REFERENCES

Adnan, M., Habib, A., Ashraf, J., Mussadiq, S., Raza, A. A., & Abid, M. (2021). Predicting at-risk students at different percentages of course length for early intervention using machine learning models. *IEEE Access, 9*, 7519-7539. https://doi.org/10.1109/ACCESS.2021.3049446

Alsayat, A., & Ahmadi, H. (2023). A hybrid method using ensembles of neural network and text mining for learner satisfaction analysis from big datasets in online learning platform. *Neural Processing Letters, 55*(3), 3267-3303. https://doi.org/10.1007/s11063-022-11009-y

Badal, Y. T., & Sungkur, R. K. (2023). Predictive modelling and analytics of students' grades using machine learning algorithms. *Education and Information Technologies, 28*(3), 3027-3057. https://doi.org/10.1007/s10639-022-11299-8

Bukar, U. A., Sayeed, M. S., Razak, S. F. A., Yogarayan, S., Amodu, O. A., & Mahmood, R. A. R. (2023). A method for analyzing text using VOSviewer. *MethodsX, 11*, Article 102339. https://doi.org/10.1016/j.mex.2023.102339

Casalino, G., Castellano, G., & Vessio, G. (2020). Exploiting time in adaptive learning from educational data. In *Bridges and mediation in higher distance education. HELMeTO 2020. Communications in computer and information science, vol 1344* (pp. 3-16). Springer. https://doi.org/10.1007/978-3-030-67435-9_1

Chui, K. T., Fung, D. C. L., Lytras, M. D., & Lam, T. M. (2020). Predicting at-risk university students in a virtual learning environment via a machine learning algorithm. *Computers in Human Behavior, 107*, Article 105584. https://doi.org/10.1016/j.chb.2018.06.032

Colpo, M. P., Primo, T. T., & Aguiar, M. S. (2024). Lessons learned from the student dropout patterns on COVID-19 pandemic: An analysis supported by machine learning. *British Journal of Educational Technology, 55*(2), 560-585. https://doi.org/10.1111/bjet.13380

Gerjets, P., Walter, C., Rosenstiel, W., Bogdan, M., & Zander, T. O. (2014). Cognitive state monitoring and the design of adaptive instruction in digital environments: Lessons learned from cognitive workload assessment using a passive brain-computer interface approach. *Frontiers in Neuroscience, 8*. https://doi.org/10.3389/fnins.2014.00385

Hlioui, F., Aloui, N., & Gargouri, F. (2021). A withdrawal prediction model of at-risk learners based on behavioral indicators. *International Journal of Web-Based Learning and Teaching Technologies, 16*(2), 32-53. https://doi.org/10.4018/IJWLTT.2021030103

Kumpulainen, M., & Seppänen, M. (2022). Combining Web of Science and Scopus datasets in citation-based literature study. *Scientometrics, 127*(10), 5613-5631. https://doi.org/10.1007/s11192-022-04475-7

Moon, J., Ke, F., & Sokolikj, Z. (2020). Automatic assessment of cognitive and emotional states in virtual reality-based flexibility training for four adolescents with autism. *British Journal of Educational Technology, 51*(5), 1766-1784. https://doi.org/10.1111/bjet.13005

Nath, A. (2021). Analyzing learners perception about data analysis on Microsoft Excel. *Learning Community, 12*(02), 95-100. https://doi.org/10.30954/2231-458X.02.2021.2

Odrekhivskyy, M., Pasichnyk, V., Rzheuskyi, A., Andrunyk, V., Nazaruk, M., & Kunanets, O. (2019). Problems of the intelligent virtual learning environment development. In *Proceedings of the MoMLeT* (pp. 359-369).

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., & Mulrow, C. D. (2021). Updating guidance for reporting systematic reviews: Development of the PRISMA 2020 statement. *Journal of Clinical Epidemiology, 134*, 103-112. https://doi.org/10.1016/j.jclinepi.2021.02.003

Piramuthu, S. (2005). Knowledge-based web-enabled agents and intelligent tutoring systems. *IEEE Transactions on Education, 48*(4), 750-756. https://doi.org/10.1109/TE.2005.854574

Raj, N. S., & VG, R. (2022). Early prediction of student engagement in virtual learning environments using machine learning techniques. *E-Learning and Digital Media, 19*(6), 537-554. https://doi.org/10.1177/20427530221108027

Riazy, S., Simbeck, K., & Schreck, V. (2020). Systematic literature review of fairness in learning analytics and application of insights in a case study. In *Proceedings of the 12th International Conference* (pp. 430-449). https://doi.org/10.1007/978-3-030-86439-2_22

Safdar, S., Ren, M., Chudhery, M. A. Z., Huo, J., Rehman, H. U., & Rafique, R. (2022). Using cloud-based virtual learning environments to mitigate increasing disparity in urban-rural academic competence. *Technological Forecasting and Social Change, 176*, Article 121468. https://doi.org/10.1016/j.techfore.2021.121468

Shafiq, D. A., Marjani, M., Habeeb, R. A. A., & Asirvatham, D. (2022). A conceptual predictive analytics model for the identification of at-risk students in VLE using machine learning techniques. In *Proceedings of the 14th International Conference on Mathematics, Actuarial Science, Computer Science and Statistics* (pp. 1-8). IEEE. https://doi.org/10.1109/MACS56771.2022.10023143

Tamada, M. M., Magalhães Netto, J. F., & Lima, D. P. R. (2019). Predicting and reducing dropout in virtual learning using machine learning techniques: A systematic review. In *Proceedings of the IEEE Frontiers in Education Conference* (pp. 1-9). IEEE. https://doi.org/10.1109/FIE43999.2019.9028545

Waheed, A., & Shafi, J. (2020). Successful role of smart technology to combat COVID-19. In *Proceedings of the 4th International Conference on IoT in Social*. https://doi.org/10.1109/I-SMAC49090.2020.9243444

Waheed, H., Hassan, S. U., Nawaz, R., Aljohani, N. R., Chen, G., & Gasevic, D. (2023). Early prediction of learners at risk in self-paced education: A neural network approach. *Expert Systems with Applications, 213*, Article 118868. https://doi.org/10.1016/j.eswa.2022.118868

Waheed, H., Nisar, I., Khalid, M. U. N., Shahid, A., Aljohani, N. R., & Hassan, S. U. (2022). Predicting academic performance of students from the assessment submission in virtual learning environment. In A. Visvizi, O. Troisi, & M. Grimaldi (Eds.), *Research and innovation forum 2022. RIIFORUM 2022. Springer proceedings in complexity* (pp. 417-424). Springer. https://doi.org/10.1007/978-3-031-19560-0_33