

## Exploring the Research Trend on Student Performance in Education Computing: A Bibliometric Analysis

<sup>1</sup>Noor Fadzilah Ab Rahman, <sup>1</sup>Siti Zaharah Mohid, & <sup>1</sup>Nor Musliza Mustafa,

<sup>1</sup>Faculty of Science and Information Technology, Kolej Universiti Islam Antarabangsa Selangor, Persiaran Putra, Bandar Seri Putra, 43000 Kajang, Selangor.

\*(Corresponding author) e-mail: [noorfadzilah@kuis.edu.my](mailto:noorfadzilah@kuis.edu.my)

### Article history:

Submission date: 12 Oct 2022

Received in revised form: 17 Nov 2022

Acceptance date: 29 December 2022

Available online: 31 December 2022

### Keywords:

Education Computing, student performance, bibliometric analysis.

### Funding:

This work is not supported through any grants.

### Competing interest:

The author(s) have declared that no competing interests exist.

### Cite as:

Ab Rahman, N. F., Mohid, S. Z., & Mustafa, N. M. (2022). Exploring the research trend on student performance in education computing: a bibliometric analysis. *Malaysian Journal of Information and Communication Technology (MyJICT)*, 11-21. DOI: <https://doi.org/10.53840/myjict7-2-41>



© The authors (2022). This is an Open Access article distributed under the terms of the Creative Commons Attribution (CC BY NC) (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact [myjict@uis.edu.my](mailto:myjict@uis.edu.my).

### ABSTRACT

The purpose of the study is to present a bibliometric analysis of the academic publication trends of student performance in education computing. These data were collected from the Scopus database that was published from 2012 to 2022. In the bibliometric analysis, important visual map and tables were produced to show publication trends from five different viewpoints including annual publications, countries, authors, publication sources, and keywords. The analysis results showed that the number of publications had increased rapidly by year during 2015 to 2019. The top countries include the United States, China, and India. Most of these publications were conference papers, and there are also article journals. The keyword co-occurrence analysis revealed that the research hotspots highlighted in this field were educational data mining, academic performance, classification and data mining mechanisms. Through this bibliometric research work, it can be expected that more researchers or scholars concentrating on utilizing two important algorithms; machine learning and deep learning, for analysing and forecasting student performance in higher educational computing institutions.

## Introduction

Education computing is a study and research area of computer-based education's applications, effects, and implications, including the design, development, and effectiveness of innovative computer hardware and software for use in the formal education system and educational environments (Rogers, 2000). As the discipline of education computing develops and matures, it has become essential to integrate the use of computer-based technologies with higher education (Tedre et al., 2018). In higher education, academic performance has been discovered to significantly affect students' academic success. In a study by Lorås et al., (2022), has shown that how students learn is critical and vital to their learning progress. In addition, student success in academic performance have become a significant research topic in educational institutions across the world, as it has become a primary strategic objective for most higher education institutions (Zeineddine et al., 2021).

In the past decade, predicting student academic performance has been a key point of debate among academics in various disciplines of education such as computing, engineering, computer science, medical, science, and management to improve student learning outcomes. Several studies have recognized that effective prediction techniques and approaches would help educators and institution administrator for design appropriate teaching and learning content to help students in their learning according to the predicted outcomes (Al-Barrak & Al-Razgan, 2015; Deng et al., 2019).

However, there are challenges to keep track on academic status through the prediction of student academic performance with massive educational data. Extraction of patterns and relationships from massive educational data becomes extremely challenging with the increasing availability of educational databases (Archana & Gandhi, 2016; Zafari et al., 2021). Various past studies have reviewed approaches, methods, algorithms and tools to improve students' performance prediction.

Several efforts and scientific research addressed the issues relying on used of student educational and behavioral records to classify students and predict their future performance using advanced statistics and Machine Learning. For example, a study by Khasanah & Harwati (2019) comprehensively reviewed student performance prediction using educational data mining techniques. These techniques also have been reviewed that are being applied on the educational data to solve the different problems faced by the students to improve the learning outcomes of the students (Gupta et al., 2020). Data mining techniques have been applied to massive educational databases to construct a classification model to predict the performance of students (Aziz & Awlla, 2019).

Other research has focused on using machine learning as one of the method to forecast students' academic achievement. Katarya et al., (2021) have performed a survey on which algorithm and techniques are best to use and which features should be considered while predicting the academic performance of students. Additionally, machine learning is becoming a significant tool for decision support in the education field today, with both educational institutions and students as the intended users. A systematic literature review by Balaji et al., (2021) have concluded that the machine learning techniques can predict the students' performance based on specified features as categorized and can be used by students as well as academic institutions.

Based on the previous literature, we noticed that particular issues and approaches in predicting student performance been reviewed and discussed. However, there is no prior studies have examined to provide a clear perspective and global development of the studies that have been done.

## Research Objective

Hence, this study is aimed to explore the trends of articles published on student performance in education computing from Scopus database between 2012 and 2022. The following are three research questions (RQ) to achieve the objective of the study:

RQ1: What is the current trend and productivity of publications on student performance in computing education computing?

RQ2: Which are the most productive and influential countries, institutions and authors on research of student performance in education computing?

RQ3: Which are the most prevalent themes studied among the authors?

## **Research Methodology**

This study explores the trend and productivity of research on student performance in education computing using bibliometric analysis. Bibliometric analysis is a widely used and rigorous method for exploring and analysing immense amounts of scientific (Donthu et al., 2021). Moreover, this bibliometric analysis is a content analysis method that reveals the connection of any publication, country, author or cited author with other publications and authors through visual maps (Su et al., 2021). A research by Zakaria et al. (2021) stated that a bibliometric study has been increasing in popularity as one of the approaches in demonstrating the trend of studies. The research methodology of this study consists of two phases: Data Collection and Data Analysis.

### ***Data Collection***

The bibliometric study was conducted on September 3, 2022 and we collected data from the Scopus database published from 2012 to 2022. The search query of the keyword “student performance” and “education computing” were used and applied to the article title, abstract and keyword within the Scopus database. This study was concentrated and limited only to the final article for the publication stage and no filtering on the language of documents was made. At the end of this phase, a total of 971 records were identified and retrieved from various fields of study.

### ***Data Analysis***

After the data collection phase was completely performed, the retrieved documents went through to analysis methods. Each collected document contains bibliometric data such as year, author name, subject area, source title, keyword, affiliation, country, and language. A total of 971 collected documents were analysed in three aspects such as performance analysis, frequency analysis and keyword analysis to reveal the research trends.

A few results in form of table, graph and visual map were produced to show the publication trends from five different perspectives including annual publications, countries, top authors, publication sources, and keywords. In this phase, we used Microsoft Excel 2016 and the open-source VOSviewer to conduct the bibliometric analysis.

## **Results and Discussion**

This section presents the bibliometric analysis results regarding the three research questions addressed in this study.

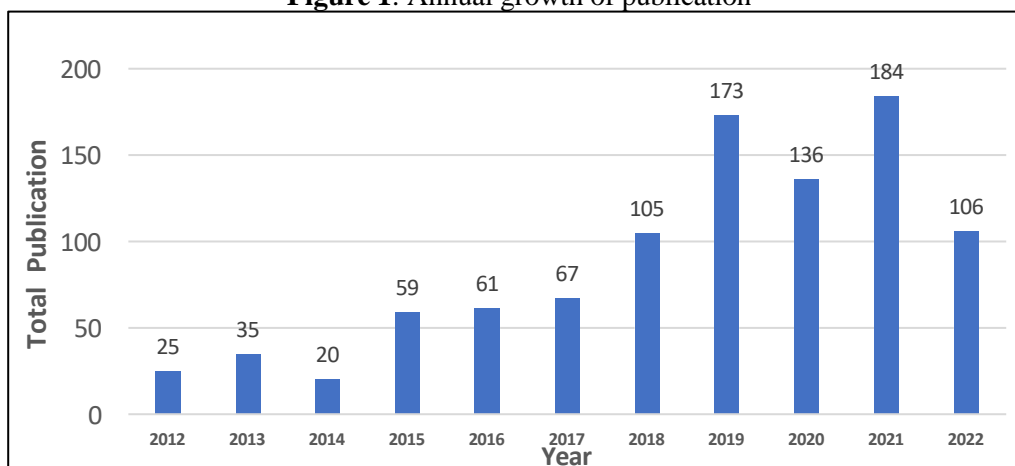
### ***What is the current trend and productivity of publications on student performance in computing education?***

In order to answer the first research question, the study analyzed of 971 retrieved documents to identify the research trends and productivity of publication based on five bibliometric indicators (a) annual growth of publication, (b) source types (c) document types (d) language of documents and (e) subject area of documents.

#### ***Annual growth of publication***

The annual growth of publication is presented in Figure 1, wherein the total number of publications is mapped against their respective year of publication. Figure 1 indicates that the research productivity and the number of publications had increased rapidly during 2015 to 2019 during in the past decade between 2012 and 2021. However, the publication activity was slightly dropped in 2014 and 2020. The greatest number of publications were in 2021 with 184 publications which indicates that research related on student performance is gaining attention among academics and researchers in the field of computing education.

**Figure 1: Annual growth of publication**



**Sources Types**

This study also explored to determine the distribution of published documents by analyzing them based on document source types. There are five source types of retrieved documents which are Conference Proceeding, Journal, Book Series, Book and Trade Journal. Table 1 shows that conference proceeding was the most common source, representing 622 (64.06%) of the total, followed by the journal (203 documents; 20.91%) and book series (141 documents; 14.52%).

**Table 1: Source Types of Documents**

Sources Type	No. of Publications	Percentage (%)
Conference Proceeding	622	64.06
Journal	203	20.91
Book Series	141	14.52
Book	4	0.41
Trade Journal	1	0.10

**Document Types**

The total of 971 collected documents were further analyzed based on document types. Document type refers to the originality of the document, such as conference papers, articles, reviews, and book chapters. The conference papers that appeared under document type were different from those appeared under source type in Table 1. Within document type, conference papers refer to the papers presented in conferences and were probably published as full journal articles. Some conference papers were also published either in conference proceedings or as a book chapter within source type although the document type was originated from a conference paper (Ahmi et al., 2019). The Scopus search retrieved four types of documents published as summarised in Table 2. As shown in the table, most of the total publications came from documents presented at a conference or symposium as conference papers (n=758; 78.06%). This was followed by articles (n= 198; 20.39%), book chapters (n= 10; 1.03%) and review papers (n=5; 0.51%).

**Table 2: Types of Documents**

Types of Documents	No. of Publications	Percentage (%)
Conference Paper	758	78.06
Article	198	20.39

Book Chapter	10	1.03
Review	5	0.51

### ***Language of documents***

The language of documents is another bibliometric indicator that is considered in this study. The language used for publications is presented in Table 3. Based on total publications, 98% of the retrieved documents were published in English (958 publications). Portuguese was used as a second language with 5 documents (0.5%), whereas only 3 documents (0.31%) used Chinese and Spanish languages for the publication. The rest documents were published in Japanese (1 publication), and Turkish (1 publication) collectively accounted for 0.1% of total publications.

**Table 3:** Language of documents

<b>Language</b>	<b>No. of Publications</b>	<b>Percentage (%)</b>
English	958	98.66
Portuguese	5	0.51
Chinese	3	0.31
Spanish	3	0.31
Japanese	1	0.10
Turkish	1	0.10

### ***Subject Area***

This study also examined and summarized the distribution of publications in the diversity of the subject area of documents. Table 4 shows the distribution of publications based on the subject area. Based on the total publication, it can be observed that most publications dominantly were in the field area of computer science (763 publications with 78.58%), then followed by social sciences (368 publications with 37.90%), engineering (315 publications with 32.44%), and mathematics (142 publications with 14.62%).

**Table 4:** Subject area

<b>Subject area</b>	<b>Total Publications</b>	<b>Percentage (%)</b>
Computer Science	763	78.58
Social Sciences	368	37.90
Engineering	315	32.44
Mathematics	142	14.62
Decision Sciences	94	9.68
Business, Management and Accounting	59	6.08
Physics and Astronomy	48	4.94
Materials Science	25	2.57
Medicine	25	2.57
Energy	18	1.85

### ***Which are the most productive and influential countries, institutions and authors on research of student performance in computing education?***

The most productive and influential countries and institutions, as well as prominent authors on student performance studies were examined by performing the authorship analysis in three bibliometric terms of (a) publication by country, (b) publication by the institution and (c) publication by the author.

**Publication by countries**

This study measures the most productive country by examining the documents collected based on the contribution of publication by country. There were 99 countries identified throughout the world that were collaborating on student performance in computing education research. Table 5 lists the top 10 most productive countries, each of which contributed more than 20 publications. Regarding the number of publications, the United States of America was the top-ranked country that contributed the most documents with a total of 260 documents (26.78%). China was second rank with 123 published publications (12.67%) and the third rank country was India with 90 papers (9.27%). Malaysia, Spain, Canada, Australia, Indonesia, United Kingdom, and Japan are the other countries that made publication contributions less than 50 articles.

**Table 5:** Top 10 most publication by country

Country	Total Publication	Percentage (%)
United States	260	26.78
China	123	12.67
India	90	9.27
Malaysia	42	4.33
Spain	37	3.81
Canada	35	3.60
Australia	34	3.50
Indonesia	31	3.19
United Kingdom	30	3.09
Japan	24	2.47

**Productive Authors**

This study also performed an authorship analysis to seek the most prominent author in contributed publication. Table 6 displays the top 10 most productive authors with more than three publications, in ascending order. Based on the Table 6, shows that Denny, P affiliated with University of Auckland had contributed six publications. In addition, similar authors from the University of California, San Diego include W.G. Griswold, S.N. Liao, and L. Porter; each contributed six articles. Following closely behind with 5 articles each were Leinonen, J. from the University of Helsinki in Finland and Becker, B.A. associated with University College Dublin in Ireland.

**Table 6:** Top 10 most productive authors

Author's name	Affiliation	Country	Total Publications
Denny, P.	University of Auckland	New Zealand	6
Griswold, W.G.	University of California, San Diego	United States	6
Liao, S.N.	University of California, San Diego	United States	6
Porter, L.	University of California, San Diego	United States	6
Becker, B.A.	University College Dublin	Ireland	5
Leinonen, J.	University of Helsinki	Finland	5
Ahmad, N.B.	Universiti Teknologi Malaysia	Malaysia	4
Hellas, A.	University of Helsinki	Finland	4

Mine, T.	Kyushu University, Fukuoka	Japan	4
Petersen, A.	University of Toronto Mississauga	Canada	4

### ***Productive Institutions***

Table 7 lists the top 10 most productive institutions with published more than 5 publications. As can be observed in Table 7, University of Toronto in Canada is the most productive institution in terms of number of publications with 15 documents. The North Carolina State University in United States was second most active institution with 33 published documents. In contrast, Carnegie Mellon University in United States is the third institution that actively contributed publications, with 10 articles. This analysis shows that top 10 most productive institutions are from country in Northern America and only Malaysia from Southeast Asia.

**Table 7:** Top 10 productive institutions

<b>Institutional</b>	<b>Country</b>	<b>Total publications</b>
University of Toronto	Canada	15
North Carolina State University	United States	13
Carnegie Mellon University	United States	10
Technologico de Monterrey	Mexico	9
Universiti Teknologi Malaysia	Malaysia	9
Arizona State University	United States	8
The University of Auckland	New Zealand	8
Worcester Polytechnic Institute	United States	8
University of California, San Diego	United States	8
University of Illinois Urbana-Champaign	United States	7

### ***Which are the most prevalent themes studied among authors?***

The study further performed a keyword analysis to identify the trend of frequent keywords are most widely used in the scientific literature on student performance studies in computing education. Two bibliometric indicators were used in this analysis based on term of (a) author keyword of document, and (b) author keyword co-occurrence.

### ***Keyword Analysis***

Table 8 summarizes the top 20 most frequently used author keywords which provided insights into the issues discussed in the educational community. After excluding the core keywords, student performance and education computing used in the search query, the keywords in Table 8 revealed the author keywords most frequently used are students, data mining, and forecasting. Moreover, from the result in Table 8, it presents that the author keywords most frequently used are seem likely related to educational technology, artificial intelligence (AI) technology, teaching, and learning systems.

**Table 8:** Top 20 frequency author keyword

<b>Author Keyword</b>	<b>Frequency</b>
Education Computing	971
Students	918

Student Performance	426
Data Mining	240
Forecasting	225
Learning Systems	210
E-learning	204
Teaching	200
Engineering Education	192
Computer Aided Instruction	173
Education	171
Curricula	149
Educational Data Mining	108
Academic Performance	107
Decision Trees	104
Machine Learning	98
Artificial Intelligence	71
Performance Prediction	66
Higher Education	63
Intelligent Tutoring System	59

#### ***Author keyword Co-occurrence analysis***

Then, this study further analysed the co-occurrence of author keywords to reveal the prevalent research topic. A software tool, Vosviewer was utilized to generate and visualize a bibliometric network map of the authors keywords co-occurrence. Figure 2 presents a network visualisation map of the author keywords with a minimum of 5 keyword occurrences using the fractional counting method. According to the default clustering method in VOSviewer, as can see in Figure 2 shows that 72 keywords were visualized and formed nine different color clusters to identify the research topic.

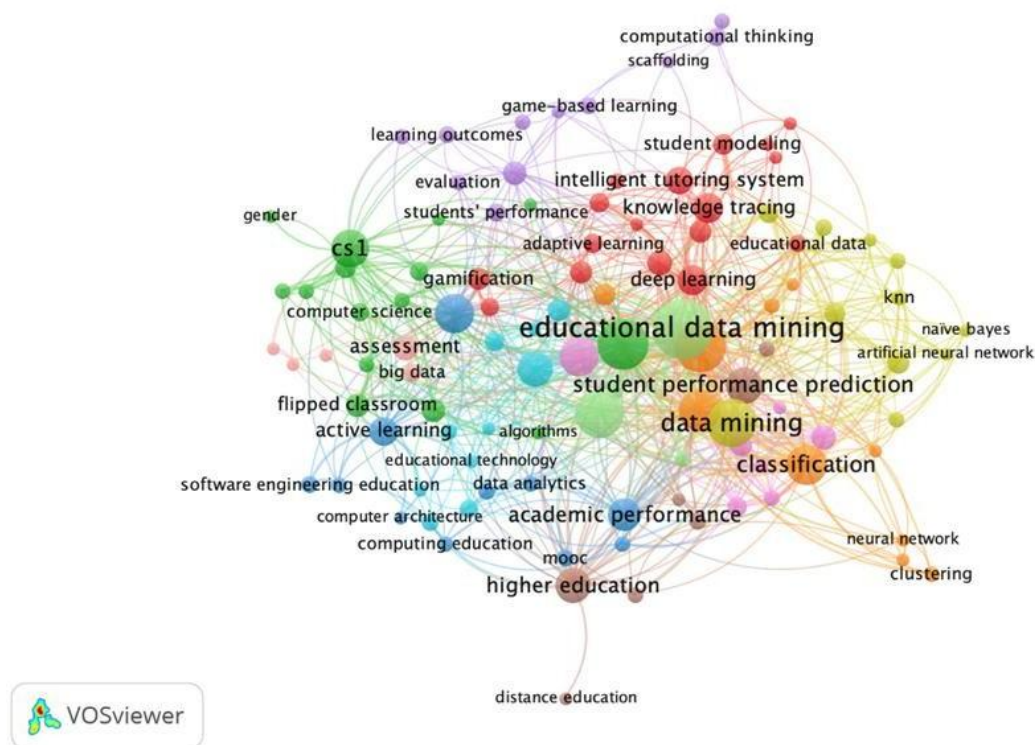
The color, node size, font size and thickness of connecting lines represent a relationship with other keywords. The bigger node size shows a larger number of keyword occurrences. In addition, the bigger occurrences of keywords also represent as the leading keyword to show the research topic. As shown in Figure 2, the highlighted leading keywords that can be seen were educational data mining, student performance prediction, data mining, higher education, classification, deep learning, machine learning algorithm and academic performance.

Based on the study's findings, we present the main findings in the context of the three research questions. Regarding the first study question, which concerned the trend and productivity of publication, our findings indicate that these factors are still present and continues as seen by the rising number of documents from year to year between 2012 and 2022. The majority of research publications were discovered in conference proceedings and journals as conference papers and articles. Additionally, the majority of published documents were released in English. According to our findings, research on student performance in education computing has largely been conducted in computer science and the social sciences, particularly as they relate to education.

Concerning the second research question, it appears that there is a fair amount of international collaboration in this field, according to country, institutions, and authors. Our present finding reveals that the top countries on more than 100 research publications published mostly in the United States are from various American academic institutions and followed by China. We have discovered that six of the top 10 most productive institutions in United States actively contributes more than five publications. Furthermore, three of the top 10 most productive authors with six publications are from the United States,

highlighting the United States as the country's prominence in this study when compared to China and other countries.

**Figure 2:** Network visualisation map of the author keywords



Regarding the third research question, the most common keywords are used to assess our observations of common themes or research topics. According to our findings, which were displayed by a generated visualisation map of the author keywords, revealed that the majority of studies on student performance in the field of education computing were focused on issues such as (a) educational data mining, (b) student performance prediction, (c) data mining, (d) classification, (e) deep learning, (f) machine learning algorithm, (g) academic performance, and (h) higher education.

## Conclusion

In this paper, we explored the research trend of student performance in education computing by performing a bibliometric analysis. A total of 971 publications were obtained from the Scopus database from 2012 until September of 2022. The findings indicate that publication trends in this field have increased rapidly from 2015 to 2019 and will continue to do so due to the use of digital learning and cutting-edge educational technology to enhance student performance in educational computing. More than half of the documents are sourced from conference proceedings. Almost all (98%) of the documents are in the English language. The findings also indicate that the publications are distributed in many subject areas, dominantly in Computer Science, followed by Social Sciences, Engineering, Mathematics, and Decision Sciences. The top ten most productive institutions and authors are all from the United States.

According to the finding of keyword analysis, researchers are increasingly interested in educational data mining, academic performance, data mining, and classification mechanisms. More researchers or scholars are expected to concentrate their efforts on utilizing two important algorithms; machine learning and deep learning for analysing and forecasting student performance in the higher educational institutions. Overall, we believe that the findings of this study can assist researchers in gaining insights into research trends,

country distributions, main contributors to this research area and the issues raised and discussed by the research communities in this field.

### Acknowledgment

We would like to thank the anonymous reviewer for their constructive comments and suggestions to improve this paper.

### References

- Ahmi, A., Elbardan, H., & Ali, R. H. R. M. (2019, January). Bibliometric analysis of published literature on industry 4.0. In *2019 International Conference on Electronics, Information, and Communication (ICEIC)* (pp. 1-6). IEEE.
- Al-Barrak, M. A., & Al-Razgan, M. S. (2015). PREDICTING STUDENTS' PERFORMANCE THROUGH CLASSIFICATION: A CASE STUDY. *Journal of Theoretical & Applied Information Technology*, 75(2).
- Archana, T., & Gandhi, U. D. (2016). Prediction of student performance in educational data mining-A survey. *International Journal of Pharmacy and Technology*, 8(3), 17757-17763.
- Aziz, S. M., & Awlla, A. H. (2019). Performance analysis and prediction student performance to build effective student using data mining techniques. *UHD Journal of Science and Technology*, 3(2), 10-15.
- Balaji, P., Alelyani, S., Qahmash, A., & Mohana, M. (2021). Contributions of machine learning models towards student academic performance prediction: a systematic review. *Applied Sciences*, 11(21), 10007.
- Deng, H., Wang, X., Guo, Z., Decker, A., Duan, X., Wang, C., ... & Abbott, K. (2019). PerformanceVis: Visual analytics of student performance data from an introductory chemistry course. *Visual Informatics*, 3(4), 166-176.
- Donthu, N., Kumar, S., Mukherjee, D., Pandey, N., & Lim, W. M. (2021). How to conduct a bibliometric analysis: An overview and guidelines. *Journal of business research*, 133, 285-296.
- Gupta, S. B., & Yadav, R. K. (2020). Analysis of Popular Techniques Used in Educational Data Mining. *International Journal of Next-Generation Computing*, 11(2).
- Katarya, R., Gaba, J., Garg, A., & Verma, V. (2021, March). A review on machine learning based student's academic performance prediction systems. In *2021 International Conference on Artificial Intelligence and Smart Systems (ICAIS)* (pp. 254-259). IEEE.
- Khasanah, A. U., & Harwati, H. (2019). Educational data mining techniques approach to predict student's performance. *International Journal of Information and Education Technology*, 9(2), 115-118.
- Lorås, M., Sindre, G., Trøtteberg, H., & Aalberg, T. (2021). Study behavior in computing education—a systematic literature review. *ACM Transactions on Computing Education (TOCE)*, 22(1), 1-40.
- Rogers, P. L. (2000). Barriers to adopting emerging technologies in education. *Journal of educational computing research*, 22(4), 455-472.
- Su, M., Peng, H., & Li, S. (2021). A visualized bibliometric analysis of mapping research trends of machine learning in engineering (MLE). *Expert Systems with Applications*, 186, 115728.
- Tedre, M., Simon, & Malmi, L. (2018). Changing aims of computing education: A historical survey. *Computer Science Education*, 28(2), 158-186.
- Zafari, M., Sadeghi-Niaraki, A., Choi, S. M., & Esmaily, A. (2021). A practical model for the evaluation of high school student performance based on machine learning. *Applied Sciences*, 11(23), 11534.
- Zakaria, R., Ahmi, A., Ahmad, A. H., & Othman, Z. (2021). Worldwide melatonin research: a bibliometric analysis of the published literature between 2015 and 2019. *Chronobiology International*, 38(1), 27-37.

Zeineddine, H., Braendle, U., & Farah, A. (2021). Enhancing prediction of student success: Automated machine learning approach. *Computers & Electrical Engineering*, 89, 106903.