DOI: 10.29345/futvis.160 Gelecek Vizyonlar Dergisi (fvj: Future Visions Journal) 4(4): 2020, 1-17

Science Mapping Research on Educational Data Mining: A Bibliometric Review of International Publications

Ergün Akgün[1] and Berran Patan Öztürk[2]

1 Assistant Professor; Bahcesehir University, Educational Sciences Faculty ergun.akgun@de.bau.edu.tr
2 Instructor; National Defence University, berranpatan@gmail.com
*Correspondence: ergun.akgun@de.bau.edu.tr

Abstract

The growing literature on Data Mining since the 2000s has recently expanded to include the sub-field "Educational Data Mining (EDM)". The present study seeks to measure the contribution to the literature of specific articles written about EDM, by studying the publication volume, authors, article types, impact of the citations, and common citation networks. We employed bibliometric research methods to analyze the 259 articles published on EDM selected for the study. In recent years, there has been a significant increase in the number of EDM articles published in international journals, and this study sought to analyze the most cited researchers, those with the highest contribution to the literature. Additionally, using common word analysis, the most used three term patterns in the articles have been identified. The study also provides suggestions to improve EDM literature in the international field.

Keywords: Educational data mining, bibliometric review, text mining, science mapping.

1 Introduction

The information that a researcher acquires or collects is called data. The word data is based on the word "datum", originating from the Latin "fact" and "reality". Even though data means "fact", because the working mechanism of the information or computer technology is numerical, this "fact" may not always be shown concretely (Zins, 2007). An unidentified sound, image, or object can be referred to as data. Based on these definitions, data can be defined as raw information that is not processed.

The word information originates from the Latin word "information" and is defined as "to provide knowledge or intelligence". The word knowledge can be defined as information that has gained meaning. The most important distinguishing factor between knowledge and information is that knowledge is a concept belonging to humans and, being a social concept, it is related to past and future (Zaim, 2005). Therefore, while knowledge is something kept in the brain, data and information are the forms added later to the knowledge in the brain (Akgün & Keskin, 2003). However, the definitions regarding these three concepts are not limited to this. For Çakırer (2013), information is the unorganized form of raw data. The organized data forms meaning and is shaped for a specific purpose. Güçlü and Sotirofski (2006) maintain that data are organized by others and only make sense for the person who organizes them. According to Abdullah et al. (2005), knowledge is defined clearly and expressed easily without any obscurity. In short, it can be said that to gain knowledge, information is required, and to gain information, data are needed. With this need, the concept of "data mining", which is very popular, gains importance.

Data mining is a concept that we encounter often and which is used in many areas is to analyze, and discover new relations within, big data. Generally, data mining aims to provide computer systems with features such as classification, clustering, information retrieval, and decision-making (Cabena et al., 1998; Diwani & Sam 2014). For Özbay (2015), data mining seeks to acquire meaningful information from a large amount of data. The subsequent aim is that this meaningful information is used to estimate and eliminate the defects in a system. As such, it prepares the data for the next step.

Data mining is divided into "Predictive" and "Descriptive" models. While predictive models consist of classification, regression, time analysis, and prediction, descriptive models include clustering, summarization, association rules, and sequence discovery (Dunham, 2003). In predictive models, with the available data, a prediction model is formed. With the prediction model, the factors whose results are unknown can be analyzed. On the other hand, descriptive models are used to examine the features of the data. The data mining used in education is called educational data mining (EDM). The aim of this new concept in education is to acquire meaningful information for the education industry by analyzing large sums of data from the educational field (Baker, 2010). Using EDM, the needs analysis made with the available data provides guidance for educators in terms of determining or changing the strategies they use. EDM is the research field that enables analysis of the data coming from the education system. With the increasing amount of data in education, EDM has become a new practice area for the larger data mining field (Güldal & Çakıcı, 2017). While Romero and Ventura (2010) hold that EDM is a practice which uses the methodology and techniques of data mining in the analysis of data from the education field to answer educational questions, Baker (2010) sees EDM as a discipline interested in researching data types in the educational environment and developing new methods.

EDM aims to increase the quality of education by obtaining information about learning and teaching processes (Romero et al., 2004). EDM uses the data types coming from education platforms to develop and analyze student models showing the current knowledge and motivation of the student (Mohamad & Tasir, 2013) and also can be used to increase student performance and level of success at the end of each school year, or to decrease failure rates (Fernandes et al., 2019). Even though most traditional data mining can be applied to EDM, there are some targets, techniques, and data sets that differentiate EDM from traditional data mining. In all three cases, there might be some adjustments that need to be made because of the characteristics of the education environment (Hanna, 2004). It is important to note that EDM is often confused with learning analytics. While EDM is the common ground for computer science, education, and statistics, learning analytics is only a sub-domain of statistics (Romero & Ventura, 2010). This is the fundamental point of differentiation between the two fields.

Previous EDM studies have provided benefits for those working in the education field. As described by Güldal and Çakıcı (2017), among others, these are: for students, to increase motivation; for teachers, to enhance learning by finding new methods in learning and teaching processes; for education executives, to gain data as a result of EDM studies; and for researchers, to answer important questions about the field with the new methods that will be used in education or the data that will be acquired to increase student performance. The main purpose of this study is to measure the influence of publication in the specific database about educational data mining and to show the academic impact of a piece of studies. To achieve this purpose, this research seeks to answers the research questions below:

1. How has the publication volume of the articles on Educational Data Mining and the number of articles published evolved over time?
2. What is the citation impact of the authors in Educational Data Mining?
3. What topics are studied most in Educational Data Mining research?

2 Method

This research employed a bibliometric analysis to study EDM publications. Bibliometric studies use quantitative methods to analyze data (White & MacCain, 1998). Bibliometrics is the application of mathematical and statistical methods to other communication tools (Pritchard, 1969), and the analysis of the specific features of published documents, such as the author, subject, resources cited by the author, and the relationship between the resources cited (Al & Tonta, 2004). The bibliometric analyses used in this study include co-authoring, citation, common citation, and common word analysis.

2.1 Identification of the Resources

The first step of this research was to develop a search strategy (Fig 1.), to determine how and where we would locate resources. To that end, we decided on a key word that would be used for the research. We searched (search made in the last quarter of 2019) in the Science Direct database for the word group "Educational Data Mining" and found in total 303 publications. The purpose of using a key word search is to identify the highest number of articles written on educational data mining. Among these publications, the studies on data mining used in engineering research, reports, books and book chapters, reviews and publications focusing on other disciplines such as data analysis were disregarded.

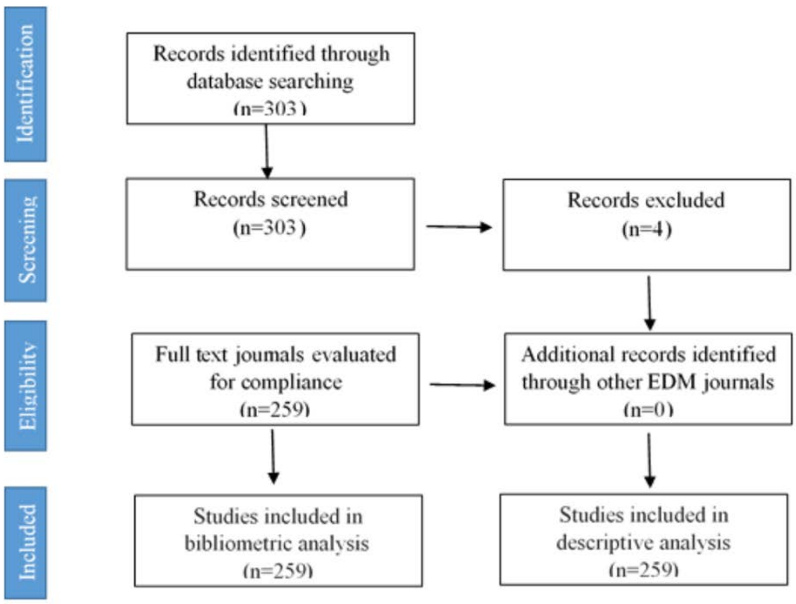

Figure 1. PRISMA Flow Diagram (A diagram showing the stages in determining and scanning the resources) (Moher, Liberati, Tetzlaff & Altman, 2009).

2.2 Data Collection and Analysis

Bibliometric analyses were made in VOSviewer, a program widely used to make bibliometric analysis and to form a visual map of the data. All articles were gathered in one RIS file. This file included title, publication, author, date, and summary information. We calculated the citation number (citation analysis) of an author or an article in Science Direct. Citation analysis tries to find the similarity between two articles by measuring the appearance frequency of articles in one data set in the bibliography (Small, 1974). Citation analysis helps find the resources used the most among the articles written on a subject, the authors with the highest contribution, and the subjects which these authors study (Al & Tonta, 2004).

The data, therefore, were first copied on a new, editable format and then united with the data gathered from 1 article identified in additional research. All these articles (n=259) have been analyzed via VOSviewer.

3 Findings

In this part of our research, we present information about each research question mentioned in the introduction.

3.1 Volume and Evolution of EDM Data Base

All 259 articles in our database were published in international journals. As shown in Figure 2, all the EDM articles in our data base were published after 2007. Based on this, it can be said that we didn't find any related studies before 2007 with our search criterias. The number of published articles on EDM increased until 2016, at which point it began to decline, reaching its lowest point within the last four years in 2017. The reason for this decrease might be the increase in research on Artificial Intelligence which emerged at about the same time as EDM.

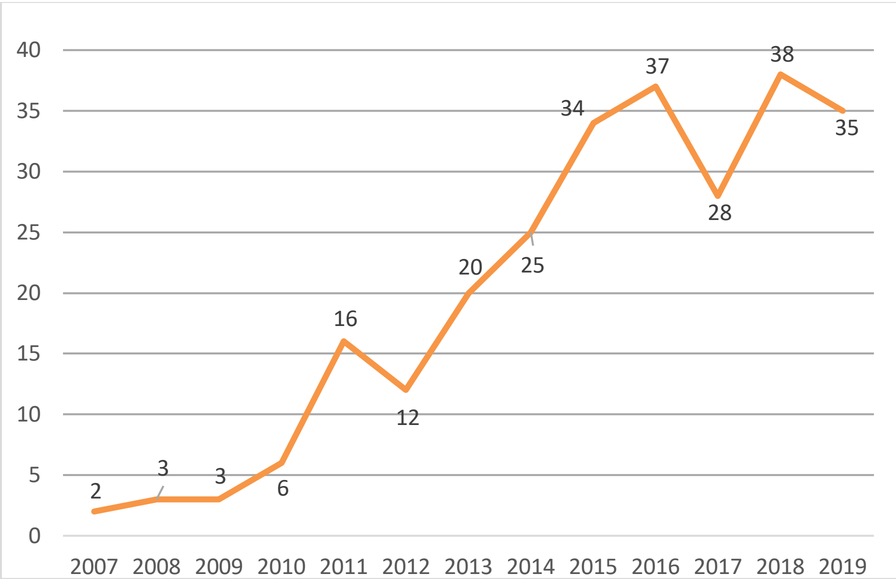

Figure 2: The Volume of International EDM Articles Published Between 2007 and 2019 (n=259) *Includes the Articles Which Have Been Published Until November 2019 in Science Direct.

Subsequently, we researched the international journals with the highest number of publications (Figure 3). From among the 9 international journals contributing to EDM literature, the journals that published more than 40 articles were Procedia Computer Science (PCS), Computers in Human Behavior (CHB), and Computers & Education (C&E). Procedia- Social and Behavioral Sciences (PSBS), Expert Systems with Applications (ESA), The Internet and Higher Education (TIHE), Knowledge-Based Systems (KBS), Neurocomputing (N), and Future Generation Computer Systems (FGCS) contributed each between 6 and 19 articles.

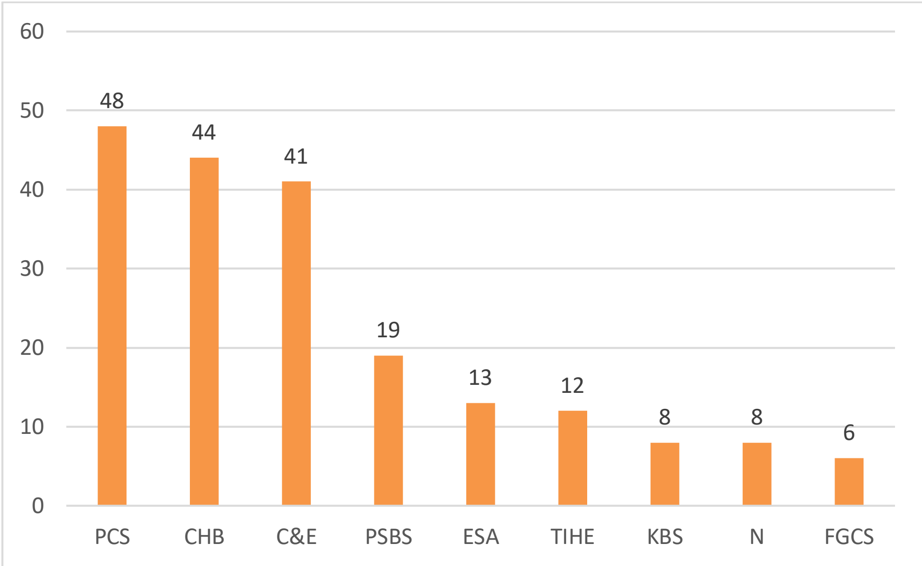

Figure 3: The Distribution of International EDM Articles According to Most Popular Journals (N=259). *This Image Includes Journals That Published At Least Six Relevant Articles. In Addition, There Are 42 Journals Which Published Four or Fewer Articles.

3.2 Influential Authors in EDM Literature

To find the most influential authors in international EDM literature, we researched the authors using VOSviewer. As seen in Figure 4, the size of the circles represents most influential authors effectiveness. Our author common citation analysis includes all international authors who have contributed to the EDM literature. The size of the circles in Figure 4. is indicative of how many citations the author has received. VOSviewer creates a citation map for the authors based on their citation frequency. The authors who share a common tradition tend to be close to each other and can be included in the same colored "cluster" in the citation analysis.

Based on our Science Direct database, we tried to find the common citation analysis of the authors (n=259). In the analysis of 766 different authors, we used a threshold value of at least 2 citations using the common citations analysis and set the VOSviewer to show 89 authors in the common citations analysis (Figure 4).

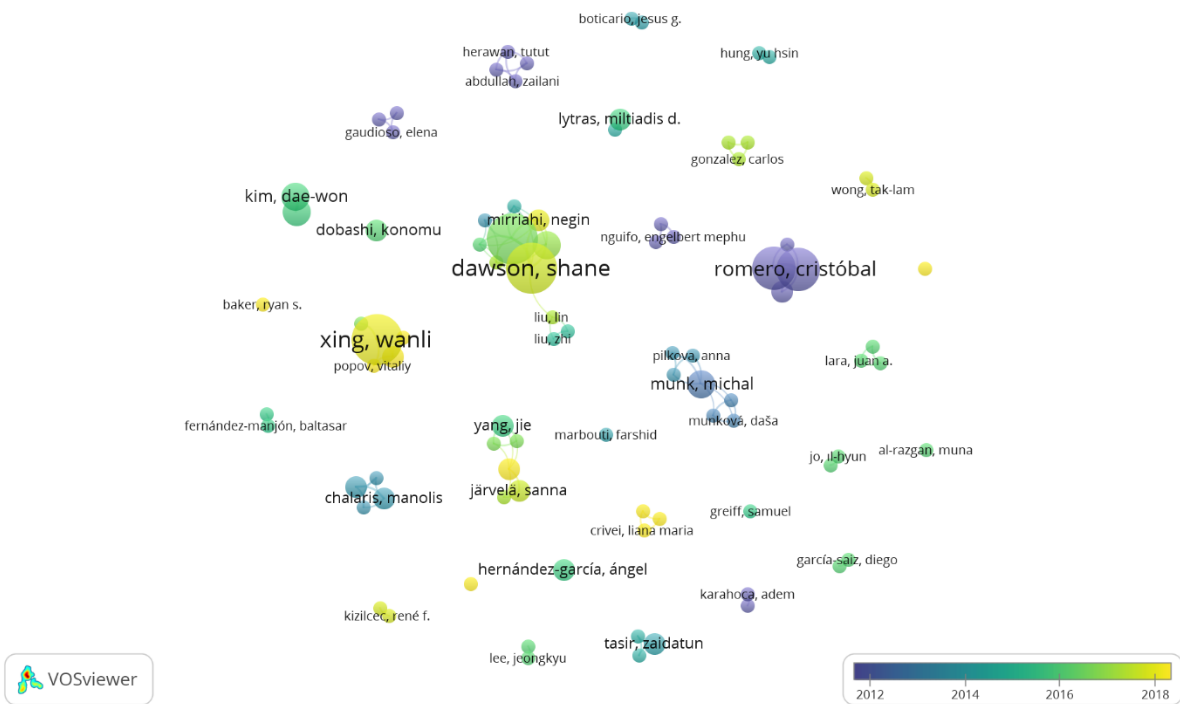

Figure 4: The Common Author Map of International EDM Articles (N=766 Authors; Threshold Value 2 Articles, 89 Authors).

The VOSviewer that we used for common author analysis shows 33 clusters. The biggest circle in the middle of the cluster shows the most influential authors. When the size of the circle gets smaller, the influence of the author decreases.

Subsequently, we checked the number of citations of the most influential authors publishing articles in EDM literature (Table 1).

Table 1: International EDM Article Authors With Total Link Strength of 6 And Higher (N=766 Authors, Threshold Value 2 Articles, 89 Authors).

Authors	Weight		
	Links*	Total Link Strength**	Documents***
gašević, dragan	7	19	7
dawson, shane	7	15	7
mirriahi, negin	5	10	3
pardo, abelardo	5	10	4
ventura, sebastián	3	10	6
munk, michal	6	9	4
romero, cristóbal	3	9	6
xing, wanli	4	8	7
chalaris, manolis	3	7	3
tsolakidis, anastasios	3	7	3
abdullah, zailani	3	6	2
ahmad, noraziah	3	6	2
chalaris, ıoannis	3	6	2
deris, mustafa mat	3	6	2
herawan, tutut	3	6	2
joksimović, srećko	5	6	2
jovanović, jelena	5	6	2
kapusta, jozef	3	6	2
kirschner, paul a.	5	6	3
kovanović, vitomir	5	6	2
pilkova, anna	3	6	2
skourlas, christos	3	6	2
svec, peter	3	6	2
zhu, gaoxia	3	6	3

*Links: The number of links with other authors.

**Total Link Strength: The total link strength with other authors.

***Documents: Out of 259 articles, the number of articles that belong to the author.)

Here we have included all the international article authors (8 different authors were disregarded who haven't been cited). Among these authors, 5 of them have total link strength of minimum ten and higher, 76 of them have total link strength with other authors minimum two and higher.

3.3 Most Studied Subjects in EDM Literature

We first tried to form a common word analysis by determining a key word. Here, we applied pair word to all articles in our database (n=259). We searched all the key words by adjusting our search in such a way that it would have a minimum 3 threshold value and 47 key words on the map (Figure 5).

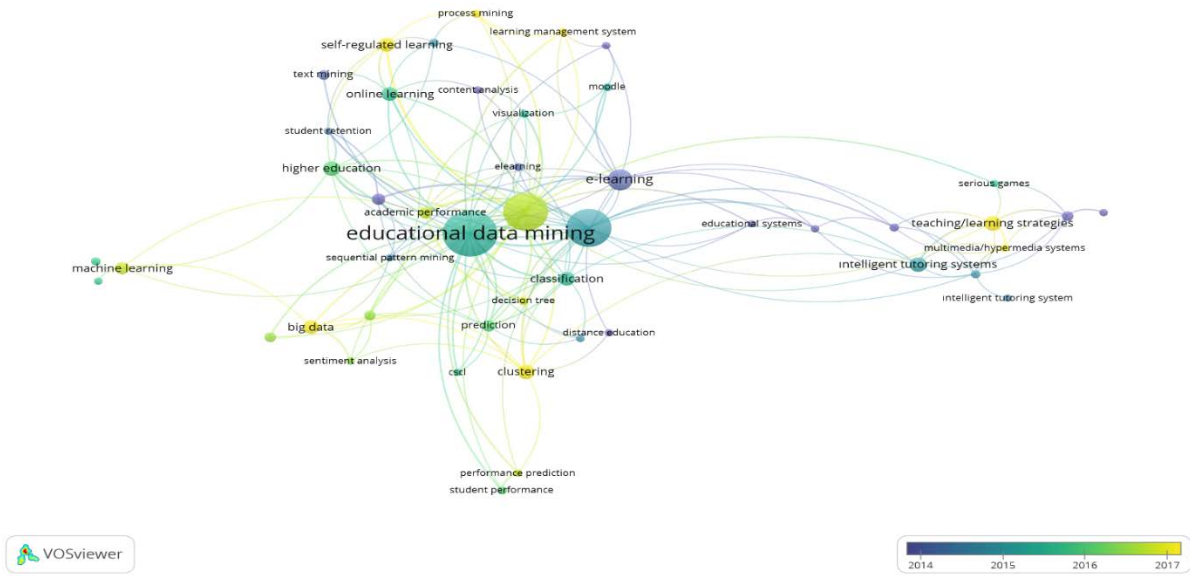

Figure 5: Co-occurrence Map of Key Words (n=859 key words; threshold value 3, screen 47 key words).

Interpretation of the common word analysis follows a similar approach to common citation analysis. The common word analysis here revealed nine clusters. The green cluster, the first cluster, is focused on subjects like educational data mining, e-learning, and text mining. The second cluster in density focuses on distant learning multimedia systems, and the third cluster focuses on prediction and clustering. The other six clusters concentrate on big data, e-learning, and mining subjects which are also included in the main three clusters. The clustered structure of the map shows the direct link between educational data mining and data mining. This is in line with the literature.

We also checked the articles within our sample for the usage number of the words used most frequently. Here we included all international articles. The most frequently encountered words in this literature are: Educational Data Mining, Data Mining, Learning Analytics, and E-learning. The position of Educational Data Mining, Data Mining, and Learning Analytics shows the centralities organizing the concepts in EDM literature (Table 2).

Table 2: International EDM Articles with Occurrence Value of 3 And Higher (threshold value 3, 47 key words).

Keyword	Weight		
	Links*	Total Link Strength**	Occurrences
educational data mining	25	54	53
data mining	25	47	40
learning analytics	25	52	40
e-learning	13	16	16
clustering	10	15	9
higher education	11	17	9
teaching/learning strategies	7	11	9
intelligent tutoring systems	8	11	9
big data	8	11	8
classification	11	16	8
online learning	8	12	8
self-regulated learning	6	11	8
machine learning	6	6	7
academic performance	9	13	6
association rules	7	13	6
prediction	9	13	6
collaborative learning	6	7	5
learning management systems	3	4	5
text mining	2	3	5
ınteractive learning environments	6	8	5
decision tree	6	9	4

distance education and telelearning	6	7	4
education	1	1	4
elearning	6	6	4
evaluation of cal systems	9	10	4
learning strategies	5	5	4
moodle	4	6	4
process mining	5	6	4
student performance	3	5	4
visualization	6	8	4
content analysis	4	4	3
Computer supported collaborative learning (cscl)	3	4	3
distance education	3	3	3
educational systems	4	5	3
genetic algorithm	5	5	3
learning management system	5	6	3
multi-label feature selection	1	1	3
multimedia/hypermedia systems	5	7	3
performance prediction	4	6	3
sentiment analysis	4	5	3
sequential pattern mining	5	5	3
serious games	3	4	3
simulations	1	2	3
student retention	5	8	3
web mining	5	7	3
weka	4	4	3
intelligent tutoring system	1	1	3

*Links: The number of links with other authors. **Total Link Strength: The total link strength with other authors.

4 Discussion, Conclusion and Implications

The fast growth in international EDM research has encouraged researchers in recent years to undertake systematic studies of EDM literature. To contribute to the improvement of the EDM database, this article analyzed articles published on educational data mining. In this part, we present our opinions on the limits and results of the study.

This research has revealed that EDM literature has been growing fast and has been mentioned in the journals more in the last four years. We believe that the EDM literature will grow more in the coming years. Our author analysis revealed that more than 500 different authors have contributed to the journals available via the Science Direct database. We analyzed the authors who have made the highest contribution to the field. We saw that 89 international researchers have published many articles on the field and other researchers have written two or more articles. It means that, EDM is fast growing but limited reasearchers are interested in this field.

Based on the results of the research, we deduce that international EDM research is quite influential and unique. The increasing amount of EDM research is the sign of a correct research culture in the scientific discipline. The most cited articles are from the last four years. As such, if the current trend continues, we expect that research in the EDM field will increase proportionally. Additionally, Figure 3 shows the impact of specific journals on the number of citations. It is seen that Procedia Computer Science (PCS), Computers in Human Behavior (CHB), and Computers & Education (C&E), all of which publish articles on EDM, are the most prolific.

Moreover, the common word analysis in Table 2 shows that the key words of the top three study subjects in EDM literature are educational data mining, data mining, and learning analytics. Apart from that, other frequently encountered key words are e-learning, higher education, teaching/learning strategies, and big data. As a result of this analysis, we are now able to say the likely direction of EDM research's future evolution and which authors and journals will contribute the most to the development of the field.

5 Limitedness

This research includes 259 research articles reached via Science Direct. It does not include post graduate theses, conference manifestations, or books on EDM. Because the research was designed and undertaken within one database, it does not consist of all journal articles. For this reason, we refrain from claiming that the results of this research speak for the entirety of EDM literature.

6 References

The studies with the (*) mark are included in the bibliometric-analysis.

Abdullah, r., Selamat, M. H., Sahibudin, S., & Alias, R. A. (2005). A framework forknowledge management system implementation in collaborative environment for higher learning institution. Journal of Knowledge Management Practice, 6(1), 1-8.

*Abdullah, Z., Herawan, T., Ahmad, N., & Deris, M. M. (2011a). Extracting highly positive association rules from students' enrollment data. *Procedia - Social and Behavioral Sciences, 28*, 107-111.

*Abdullah, Z., Herawan, T., Ahmad, N., & Deris, M. M. (2011b). Mining significant association rules from educational data using critical relative support approach. *Procedia - Social and Behavioral Sciences, 28*, 97-101.

*Adekitan, A. I., & Salau, O. (2019). The impact of engineering students' performance in the first three years on their graduation result using educational data mining. *Heliyon, 5*(2), e01250.

*Agudo-Peregrina, Á. F., Iglesias-Pradas, S., Conde-González, M. Á., & Hernández-García, Á. (2014). Can we predict success from log data in VLEs? Classification of interactions for learning analytics and their relation with performance in VLE-supported F2F and online learning. *Computers in Human Behavior, 31*, 542-550.

*Aher, S. B., & Lobo, L. M. R. J. (2013). Combination of machine learning algorithms for recommendation of courses in E-Learning System based on historical data. *Knowledge-Based Systems, 51*, 1-14.

*Ahmed, A. M., Rizaner, A., & Ulusoy, A. H. (2016). Using data Mining to Predict Instructor Performance. *Procedia Computer Science, 102*, 137-142.

*Akçapınar, G. (2015). How automated feedback through text mining changes plagiaristic behavior in online assignments. *Computers & Education, 87*, 123-130.

Akgün, Ali E., & Keskin, H. (2003). Sosyal bir etkileşim araci olarak bilgi yönetimi ve bilgi yönetimi süreci, *Gazi Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi*, 1:175–188.

Al, U., & Tonta, Y. (2004). Citation analysis: sources cited in dissertations completed at Hacettepe University department of librarianship. *Bilgi Dünyası,* 5(1):19-47.

*Al-Khalifa, S., & Al-Razgan, M. (2016). Ebsar: Indoor guidance for the visually impaired. *Computers & Electrical Engineering, 54*, 26-39.

*Al-Twijri, M. I., & Noaman, A. Y. (2015). A New Data Mining Model Adopted for Higher Institutions. *Procedia Computer Science, 65*, 836-844.

*Aldape-Pérez, M., Yáñez-Márquez, C., Camacho-Nieto, O., López-Yáñez, I., & Argüelles-Cruz, A.-J. (2015). Collaborative learning based on associative models: Application to pattern classification in medical datasets. *Computers in Human Behavior, 51*, 771-779.

*Alencar, D., Carvalho, D., Koenders, E., Mourão, F., & Rocha, L. (2017). Devising a computational model based on data mining techniques to predict concrete compressive strength. *Procedia Computer Science, 108*, 455-464.

*Alexandron, G., Ruipérez-Valiente, J. A., Chen, Z., Muñoz-Merino, P. J., & Pritchard, D. E. (2017). Copying@Scale: Using Harvesting Accounts for Collecting Correct Answers in a MOOC. *Computers & Education, 108*, 96-114.

*Ali, L., Hatala, M., Gašević, D., & Jovanović, J. (2012). A qualitative evaluation of evolution of a learning analytics tool. *Computers & Education, 58*(1), 470-489.

*Altujjar, Y., Altamimi, W., Al-Turaiki, I., & Al-Razgan, M. (2016). Predicting Critical Courses Affecting Students Performance: A Case Study. *Procedia Computer Science, 82*, 65-71.

*Angeli, C., Howard, S. K., Ma, J., Yang, J., & Kirschner, P. A. (2017). Data mining in educational technology classroom research: Can it make a contribution? *Computers & Education, 113*, 226-242.

*Aparicio, F., Morales-Botello, M. L., Rubio, M., Hernando, A., Muñoz, R., López-Fernández, H., . . . Buenaga, M. d. (2018). Perceptions of the use of intelligent information access systems in university level active learning activities among teachers of biomedical subjects. *International Journal of Medical Informatics, 112*, 21-33.

*Ashraf, M., Zaman, M., & Ahmed, M. (2018). Using Ensemble StackingC Method and Base Classifiers to Ameliorate Prediction Accuracy of Pedagogical Data. *Procedia Computer Science, 132*, 1021-1040.

*Asif, R., Merceron, A., Ali, S. A., & Haider, N. G. (2017). Analyzing undergraduate students' performance using educational data mining. *Computers & Education, 113*, 177-194.

*Baba, Y., Kinoshita, K., & Kashima, H. (2016). Participation recommendation system for crowdsourcing contests. *Expert Systems with Applications, 58*, 174-183.

*Bach, M. P., Čeljo, A., & Zoroja, J. (2016). Technology Acceptance Model for Business Intelligence Systems: Preliminary Research. *Procedia Computer Science, 100*, 995-1001.

*Badr, G., Algobail, A., Almutairi, H., & Almutery, M. (2016). Predicting Students' Performance in University Courses: A Case Study and Tool in KSU Mathematics Department. *Procedia Computer Science, 82*, 80-89.

Baker, R.S.J.D. (2010) *Data mining*, International Encyclopedia of Education 3:112-118.

*Belsis, P., Chalaris, I., Chalaris, M., Skourlas, C., & Tsolakidis, A. (2014). The Analysis of the Length of Studies in Higher Education based on Clustering and the Extraction of Association Rules. *Procedia - Social and Behavioral Sciences, 147*, 567-575.

*Blagojević, M., & Micić, Ž. (2013). A web-based intelligent report e-learning system using data mining techniques. *Computers & Electrical Engineering, 39*(2), 465-474.

*Burgos, C., Campanario, M. L., Peña, D. d. l., Lara, J. A., Lizcano, D., & Martínez, M. A. (2018). Data mining for modeling students' performance: A tutoring action plan to prevent academic dropout. *Computers & Electrical Engineering, 66*, 541-556.

*Campagni, R., Merlini, D., Sprugnoli, R., & Verri, M. C. (2015). Data mining models for student careers. *Expert Systems with Applications, 42*(13), 5508-5521.

*Canaleta, X., Vernet, D., Vicent, L., & Montero, J. A. (2014). Master in Teacher Training: A real implementation of Active Learning. *Computers in Human Behavior, 31*, 651-658.

*Cantabella, M., Martínez-España, R., Ayuso, B., Yáñez, J. A., & Muñoz, A. (2019). Analysis of student behavior in learning management systems through a Big Data framework. *Future Generation Computer Systems, 90*, 262-272.

*Castiello, C., Fanelli, A. M., Lucarelli, M., & Mencar, C. (2019). Interpretable fuzzy partitioning of classified data with variable granularity. *Applied Soft Computing, 74*, 567-582.

*Cerezo, R., Sánchez-Santillán, M., Paule-Ruiz, M. P., & Núñez, J. C. (2016). Students' LMS interaction patterns and their relationship with achievement: A case study in higher education. *Computers & Education, 96*, 42-54.

*Chalaris, M., Chalaris, I., Skourlas, C., & Tsolakidis, A. (2013). Extraction of Rules based on Students' Questionnaires. *Procedia - Social and Behavioral Sciences, 73*, 510-517.

*Chalaris, M., Gritzalis, S., Maragoudakis, M., Sgouropoulou, C., & Tsolakidis, A. (2014). Improving Quality of Educational Processes Providing New Knowledge Using Data Mining Techniques. *Procedia - Social and Behavioral Sciences, 147*, 390-397.

*Chassignol, M., Khoroshavin, A., Klimova, A., & Bilyatdinova, A. (2018). Artificial Intelligence trends in education: a narrative overview. *Procedia Computer Science, 136*, 16-24.

*Chen, C.-M., & Chen, M.-C. (2009). Mobile formative assessment tool based on data mining techniques for supporting web-based learning. *Computers & Education, 52*(1), 256-273.

*Cheng, M.-T., Rosenheck, L., Lin, C.-Y., & Klopfer, E. (2017). Analyzing gameplay data to inform feedback loops in The Radix Endeavor. *Computers & Education, 111*, 60-73.

*Chowanda, A., & Chowanda, A. D. (2018). Generative Indonesian Conversation Model using Recurrent Neural Network with Attention Mechanism. *Procedia Computer Science, 135*, 433-440.

*Chui, K. T., Fung, D. C. L., Lytras, M. D., & Lam, T. M. (2018). Predicting at-risk university students in a virtual learning environment via a machine learning algorithm. *Computers in Human Behavior*.

*Costa, E. B., Fonseca, B., Santana, M. A., de Araújo, F. F., & Rego, J. (2017). Evaluating the effectiveness of educational data mining techniques for early prediction of students' academic failure in introductory programming courses. *Computers in Human Behavior, 73*, 247-256.

*Cuadros, J., Artigas, C., Guitart, F., & Martori, F. (2015). Analyzing a Virtual-Lab Based Contextualized Activity from Action Logs. *Procedia - Social and Behavioral Sciences, 182*, 441-447.

*Czibula, G., Czibula, I. G., Miholca, D.-L., & Crivei, L. M. (2019). A novel concurrent relational association rule mining approach. *Expert Systems with Applications, 125*, 142-156.

Çakırer, M., A. (2013). *Bilgi yönetimi*, E-Ticaret Ekin Yayınevi, s.60-65.

*Davis, D., Chen, G., Hauff, C., & Houben, G.-J. (2018). Activating learning at scale: A review of innovations in online learning strategies. *Computers & Education, 125*, 327-344.

*de Klerk, S., Veldkamp, B. P., & Eggen, T. J. H. M. (2015). Psychometric analysis of the performance data of simulation-based assessment: A systematic review and a Bayesian network example. *Computers & Education, 85*, 23-34.

*de-Marcos, L., García-López, E., García-Cabot, A., Medina-Merodio, J.-A., Domínguez, A., Martínez-Herráiz, J.-J., & Diez-Folledo, T. (2016). Social network analysis of a gamified e-learning course: Small-world phenomenon and network metrics as predictors of academic performance. *Computers in Human Behavior, 60*, 312-321.

*Deater-Deckard, K., El Mallah, S., Chang, M., Evans, M. A., & Norton, A. (2014). Student behavioral engagement during mathematics educational video game instruction with 11–14 year olds. *International Journal of Child-Computer Interaction, 2*(3), 101-108.

*Dillenbourg, P. (2013). Design for classroom orchestration. *Computers & Education, 69*, 485-492.

Diwani, S. A. & Sam, A., (2014). Data mining awareness and readiness in healthcare sector: a case of tanzania, *Advances in Computer Science: An International Journal*, 3(1), No.7.

*Dobashi, K. (2015). Time Series Analysis of the in class Page View History of Digital Teaching Materials using Cross Table. *Procedia Computer Science, 60*, 1032-1040.

*Dobashi, K. (2016). Development and Trial of Excel Macros for Time Series Cross Section Monitoring of Student Engagement: Analyzing Students' Page Views of Course Materials. *Procedia Computer Science, 96*, 1086-1095.

*Dobashi, K. (2017). Automatic data integration from Moodle course logs to pivot tables for time series cross section analysis. *Procedia Computer Science, 112*, 1835-1844.

*Durahim, A. O., & Coşkun, M. (2015). #iamhappybecause: Gross National Happiness through Twitter analysis and big data. *Technological Forecasting and Social Change, 99*, 92-105.

Dunham, M., H. (2003). Data mining introductory and advanced topics, Prentice Hall, Pearson Education Inc., New Jersey, s.8.

*Durand, G., Belacel, N., & LaPlante, F. (2013). Graph theory based model for learning path recommendation. *Information Sciences, 251*, 10-21.

*Elia, G., Solazzo, G., Lorenzo, G., & Passiante, G. (2019). Assessing learners' satisfaction in collaborative online courses through a big data approach. *Computers in Human Behavior, 92*, 589-599.

*Engelmann, K., & Bannert, M. (2019). Analyzing temporal data for understanding the learning process induced by metacognitive prompts. *Learning and Instruction*.

*Engin, G., Aksoyer, B., Avdagic, M., Bozanlı, D., Hanay, U., Maden, D., & Ertek, G. (2014). Rule-based Expert Systems for Supporting University Students. *Procedia Computer Science, 31*, 22-31.

*Enochsson, A.-B. (2011). Who benefits from synchronous online communication?: A comparison of face-to-face and synchronous online interviews with children. *Procedia - Social and Behavioral Sciences, 28*, 15-22.

*Espinosa, R., García-Saiz, D., Zorrilla, M., Zubcoff, J. J., & Mazón, J.-N. (2019). S3Mining: A model-driven engineering approach for supporting novice data miners in selecting suitable classifiers. *Computer Standards & Interfaces, 65*, 143-158.

*Estévez-Ayres, I., Arias Fisteus, J., & Delgado-Kloos, C. (2017). Lostrego: A distributed stream-based infrastructure for the real-time gathering and analysis of heterogeneous educational data. *Journal of Network and Computer Applications, 100*, 56-68.

*Fahrudin, T., Buliali, J. L., & Fatichah, C. (2019). Ina-BWR: Indonesian bigram word rule for multi-label student complaints. *Egyptian Informatics Journal*.

*Fernandes, E., Holanda, M., Victorino, M., Borges, V., Carvalho, R., & Erven, G. V. (2019). Educational data mining: Predictive analysis of academic performance of public school students in the capital of Brazil. *Journal of Business Research, 94*, 335-343.

*Fernández-Gallego, B., Lama, M., Vidal, J. C., & Mucientes, M. (2013). Learning Analytics Framework for Educational Virtual Worlds. *Procedia Computer Science, 25*, 443-447.

*Filvà, D. A., Forment, M. A., García-Peñalvo, F. J., Escudero, D. F., & Casañ, M. J. (2019). Clickstream for learning analytics to assess students' behavior with Scratch. *Future Generation Computer Systems, 93*, 673-686.

*Fournier-Viger, P., Faghihi, U., Nkambou, R., & Nguifo, E. M. (2012). CMRules: Mining sequential rules common to several sequences. *Knowledge-Based Systems, 25*(1), 63-76.

*Fulantelli, G., Taibi, D., & Arrigo, M. (2015). A framework to support educational decision making in mobile learning. *Computers in Human Behavior, 47*, 50-59.

*Gaeta, M., Loia, V., Mangione, G. R., Orciuoli, F., Ritrovato, P., & Salerno, S. (2014). A methodology and an authoring tool for creating Complex Learning Objects to support interactive storytelling. *Computers in Human Behavior, 31*, 620-637.

*García, E., Romero, C., Ventura, S., & de Castro, C. (2011). A collaborative educational association rule mining tool. *The Internet and Higher Education, 14*(2), 77-88.

*Gašević, D., Dawson, S., Rogers, T., & Gasevic, D. (2016). Learning analytics should not promote one size fits all: The effects of instructional conditions in predicting academic success. *The Internet and Higher Education, 28*, 68-84.

*Gašević, D., Mirriahi, N., Dawson, S., & Joksimović, S. (2017). Effects of instructional conditions and experience on the adoption of a learning tool. *Computers in Human Behavior, 67*, 207-220.

*Gaudioso, E., Montero, M., & Hernandez-del-Olmo, F. (2012). Supporting teachers in adaptive educational systems through predictive models: A proof of concept. *Expert Systems with Applications, 39*(1), 621-625.

*Gaudioso, E., Montero, M., Talavera, L., & Hernandez-del-Olmo, F. (2009). Supporting teachers in collaborative student modeling: A framework and an implementation. *Expert Systems with Applications, 36*(2, Part 1), 2260-2265.

*Gauthier, A., Corrin, M., & Jenkinson, J. (2015). Exploring the influence of game design on learning and voluntary use in an online vascular anatomy study aid. *Computers & Education, 87*, 24-34.

*Gobert, J. D., Kim, Y. J., Sao Pedro, M. A., Kennedy, M., & Betts, C. G. (2015). Using educational data mining to assess students' skills at designing and conducting experiments within a complex systems microworld. *Thinking Skills and Creativity, 18*, 81-90.

*Goga, M., Kuyoro, S., & Goga, N. (2015). A Recommender for Improving the Student Academic Performance. *Procedia - Social and Behavioral Sciences, 180*, 1481-1488.

*Gómez-Aguilar, D. A., Hernández-García, Á., García-Peñalvo, F. J., & Therón, R. (2015). Tap into visual analysis of customization of grouping of activities in eLearning. *Computers in Human Behavior, 47*, 60-67.

*Gorostiaga, A., & Rojo-Álvarez, J. L. (2016). On the use of conventional and statistical-learning techniques for the analysis of PISA results in Spain. *Neurocomputing, 171*, 625-637.

*Greiff, S., Niepel, C., Scherer, R., & Martin, R. (2016). Understanding students' performance in a computer-based assessment of complex problem solving: An analysis of behavioral data from computer-generated log files. *Computers in Human Behavior, 61*, 36-46.

*Greiff, S., Wüstenberg, S., & Avvisati, F. (2015). Computer-generated log-file analyses as a window into students' minds? A showcase study based on the PISA 2012 assessment of problem solving. *Computers & Education, 91*, 92-105.

*Gutiérrez, F., Seipp, K., Ochoa, X., Chiluiza, K., De Laet, T., & Verbert, K. (2018). LADA: A learning analytics dashboard for academic advising. *Computers in Human Behavior*.

Güçlü, N. & Sotirofski, K. (2006). Information management, *The Journal of Turkish Educational Sciences*, 4(4): 351-371.

Güldal, H. & Çakıcı, Y. (2017). Educational data mining, *Balkan Educational Studies*, 1: 135-143.

Gümüş, S., Bellibaş M. Ş., Gümüş, E. & Hallinger, P. (2019). Science mapping research on educational leadership and management in Turkey: a bibliometric review of international publications, *School Leadership and Management*. 1-22.

*Hambach, J., Kümmel, K., & Metternich, J. (2017). Development of a Digital Continuous Improvement System for Production. *Procedia CIRP, 63*, 330-335.

*Hamsa, H., Indiradevi, S., & Kizhakkethottam, J. J. (2016). Student Academic Performance Prediction Model Using Decision Tree and Fuzzy Genetic Algorithm. *Procedia Technology, 25*, 326-332.

Hanna, M. (2004). Data mining in the e-learning domain, *Computers & Education Journal*, 42(3), 267–287.

*Hassan, S. M., & Al-Razgan, M. S. (2016). Pre-University Exams Effect on Students GPA: A Case Study in IT Department. *Procedia Computer Science, 82*, 127-131.

*He, W. (2013). Examining students' online interaction in a live video streaming environment using data mining and text mining. *Computers in Human Behavior, 29*(1), 90-102.

*He, Y., Xie, H., Wong, T.-L., & Wang, X. (2018). A novel binary artificial bee colony algorithm for the set-union knapsack problem. *Future Generation Computer Systems, 78*, 77-86.

*Helal, S., Li, J., Liu, L., Ebrahimie, E., Dawson, S., Murray, D. J., & Long, Q. (2018). Predicting academic performance by considering student heterogeneity. *Knowledge-Based Systems, 161*, 134-146.

*Henrie, C. R., Halverson, L. R., & Graham, C. R. (2015). Measuring student engagement in technology-mediated learning: A review. *Computers & Education, 90*, 36-53.

*Henriques, R., Antunes, C., & Madeira, S. C. (2015). A structured view on pattern mining-based biclustering. *Pattern Recognition, 48*(12), 3941-3958.

*Hernández-García, Á., Acquila-Natale, E., Chaparro-Peláez, J., & Conde, M. Á. (2018). Predicting teamwork group assessment using log data-based learning analytics. *Computers in Human Behavior, 89*, 373-384.

*Hernández-Lara, A. B., Perera-Lluna, A., & Serradell-López, E. (2019). Applying learning analytics to students' interaction in business simulation games. The usefulness of learning analytics to know what students really learn. *Computers in Human Behavior, 92*, 600-612.

*Hershkovitz, A., & Nachmias, R. (2011). Online persistence in higher education web-supported courses. *The Internet and Higher Education, 14*(2), 98-106.

*Hesketh, B., Griffin, B., & Loh, V. (2011). A future-oriented Retirement Transition Adjustment Framework. *Journal of Vocational Behavior, 79*(2), 303-314.

*Hlosta, M., Zdrahal, Z., & Zendulka, J. (2018). Are we meeting a deadline? classification goal achievement in time in the presence of imbalanced data. *Knowledge-Based Systems, 160*, 278-295.

*Hodges, G. W., Wang, L., Lee, J., Cohen, A., & Jang, Y. (2018). An exploratory study of blending the virtual world and the laboratory experience in secondary chemistry classrooms. *Computers & Education, 122*, 179-193.

*Hogo, M. A. (2010). Evaluation of e-learning systems based on fuzzy clustering models and statistical tools. *Expert Systems with Applications, 37*(10), 6891-6903.

*Holanda, O., Isotani, S., Bittencourt, I. I., Dermeval, D., & Alcantara, W. (2017). An object triple mapping system supporting detached objects: A performance and memory usage empirical comparison. *Engineering Applications of Artificial Intelligence, 62*, 234-251.

*Hooshyar, D., Ahmad, R. B., Yousefi, M., Fathi, M., Horng, S.-J., & Lim, H. (2016). Applying an online game-based formative assessment in a flowchart-based intelligent tutoring system for improving problem-solving skills. *Computers & Education, 94*, 18-36.

*Howard, S. K., Ma, J., & Yang, J. (2016). Student rules: Exploring patterns of students' computer-efficacy and engagement with digital technologies in learning. *Computers & Education, 101*, 29-42.

*Hu, Y.-H., Lo, C.-L., & Shih, S.-P. (2014). Developing early warning systems to predict students' online learning performance. *Computers in Human Behavior, 36*, 469-478.

*Huang, T. C.-K., Huang, C.-H., & Chuang, Y.-T. (2016). Change discovery of learning performance in dynamic educational environments. *Telematics and Informatics, 33*(3), 773-792.

*Hung, Y. H., Chang, R. I., & Lin, C. F. (2016). Hybrid learning style identification and developing adaptive problem-solving learning activities. *Computers in Human Behavior, 55*, 552-561.

*Ivanović, M., Milićević, A. K., Aleksić, V., Bratić, B., & Mandić, M. (2018). Experiences and perspectives of Technology-enhanced learning and teaching in higher education – Serbian case. *Procedia Computer Science, 126*, 1351-1359.

*Järvelä, S., Malmberg, J., Haataja, E., Sobocinski, M., & Kirschner, P. A. (2019). What multimodal data can tell us about the students' regulation of their learning process? *Learning and Instruction*.

*Jena, R. K. (2019). Understanding academic achievement emotions towards business analytics course: A case study among business management students from India. *Computers in Human Behavior, 92*, 716-723.

*Jovanović, J., Gašević, D., Dawson, S., Pardo, A., & Mirriahi, N. (2017). Learning analytics to unveil learning strategies in a flipped classroom. *The Internet and Higher Education, 33*, 74-85.

*Jovanovic, J., Mirriahi, N., Gašević, D., Dawson, S., & Pardo, A. (2019). Predictive power of regularity of pre-class activities in a flipped classroom. *Computers & Education, 134*, 156-168.

*Juhaňák, L., Zounek, J., & Rohlíková, L. (2019). Using process mining to analyze students' quiz-taking behavior patterns in a learning management system. *Computers in Human Behavior, 92*, 496-506.

*Jurado, F., & Rodriguez, P. (2015). Sentiment Analysis in monitoring software development processes: An exploratory case study on GitHub's project issues. *Journal of Systems and Software, 104*, 82-89.

*Kaoropthai, C., Natakuatoong, O., & Cooharojananone, N. (2019). An intelligent diagnostic framework: A scaffolding tool to resolve academic reading problems of Thai first-year university students. *Computers & Education, 128*, 132-144.

*Kapusta, J., Munk, M., Svec, P., & Pilkova, A. (2014). Determining the Time Window Threshold to Identify User Sessions of Stakeholders of a Commercial Bank Portal. *Procedia Computer Science, 29*, 1779-1790.

*Kaur, P., Singh, M., & Josan, G. S. (2015). Classification and Prediction Based Data Mining Algorithms to Predict Slow Learners in Education Sector. *Procedia Computer Science, 57*, 500-508.

*Kim, C., Choe, S., Choi, C., & Park, Y. (2008). A systematic approach to new mobile service creation. *Expert Systems with Applications, 35*(3), 762-771.

*Kim, D., Park, Y., Yoon, M., & Jo, I.-H. (2016). Toward evidence-based learning analytics: Using proxy variables to improve asynchronous online discussion environments. *The Internet and Higher Education, 30*, 30-43.

*Kizilcec, R. F., Pérez-Sanagustín, M., & Maldonado, J. J. (2017). Self-regulated learning strategies predict learner behavior and goal attainment in Massive Open Online Courses. *Computers & Education, 104*, 18-33.

*Kotsiantis, S., Patriarcheas, K., & Xenos, M. (2010). A combinational incremental ensemble of classifiers as a technique for predicting students' performance in distance education. *Knowledge-Based Systems, 23*(6), 529-535.

*Kovanović, V., Gašević, D., Joksimović, S., Hatala, M., & Adesope, O. (2015). Analytics of communities of inquiry: Effects of learning technology use on cognitive presence in asynchronous online discussions. *The Internet and Higher Education, 27*, 74-89.

*Kurahashi, S., & Kuniyoshi, K. (2014). Effects of Collaborative Learning on a Complex Doubly Structured Network. *Procedia Computer Science, 35*, 910-917.

*Kurilovas, E. (2018). On data-driven decision-making for quality education. *Computers in Human Behavior*.

*Lara, J. A., Lizcano, D., Martínez, M. A., Pazos, J., & Riera, T. (2014). A system for knowledge discovery in e-learning environments within the European Higher Education Area – Application to student data from Open University of Madrid, UDIMA. *Computers & Education, 72*, 23-36.

*Lee, J., & Kim, D.-W. (2015a). Fast multi-label feature selection based on information-theoretic feature ranking. *Pattern Recognition, 48*(9), 2761-2771.

*Lee, J., & Kim, D.-W. (2015b). Memetic feature selection algorithm for multi-label classification. *Information Sciences, 293*, 80-96.

*Lee, J., & Kim, D.-W. (2016). An effective initialization method for genetic algorithm-based robot path planning using a directed acyclic graph. *Information Sciences, 332*, 1-18.

*Lee, J., & Kim, D.-W. (2017). SCLS: Multi-label feature selection based on scalable criterion for large label set. *Pattern Recognition, 66*, 342-352.

*Lee, Y. (2018). Effect of uninterrupted time-on-task on students' success in Massive Open Online Courses (MOOCs). *Computers in Human Behavior, 86*, 174-180.

*Lee, Y.-J. (2012). Developing an efficient computational method that estimates the ability of students in a Web-based learning environment. *Computers & Education, 58*(1), 579-589.

*Legaspi, R., Sison, R., Fukui, K.-i., & Numao, M. (2008). Cluster-based predictive modeling to improve pedagogic reasoning. *Computers in Human Behavior, 24*(2), 153-172.

*Leong, C. K., Lee, Y. H., & Mak, W. K. (2012). Mining sentiments in SMS texts for teaching evaluation. *Expert Systems with Applications, 39*(3), 2584-2589.

*Levy, S. T., & Wilensky, U. (2011). Mining students' inquiry actions for understanding of complex systems. *Computers & Education, 56*(3), 556-573.

*Li, A., Wu, J., & Liu, Z. (2017). Market Manipulation Detection Based on Classification Methods. *Procedia Computer Science, 122*, 788-795.

*Li, S., Zheng, J., & Zheng, Y. (2019). Towards a new approach to managing teacher online learning: Learning communities as activity systems. *The Social Science Journal*.

*Li, X., Rao, Y., Xie, H., Liu, X., Wong, T.-L., & Wang, F. L. (2017). Social emotion classification based on noise-aware training. *Data & Knowledge Engineering*.

*Liao, Z., & Zhu, Y. (2014). When a Classifier Meets More Data. *Procedia Computer Science, 30*, 50-59.

*Lim, L.-A., Gentili, S., Pardo, A., Kovanović, V., Whitelock-Wainwright, A., Gašević, D., & Dawson, S. (2019). What changes, and for whom? A study of the impact of learning analytics-based process feedback in a large course. *Learning and Instruction*.

*Lin, C. F., Yeh, Y.-c., Hung, Y. H., & Chang, R. I. (2013). Data mining for providing a personalized learning path in creativity: An application of decision trees. *Computers & Education, 68*, 199-210.

*Lin, H.-Y., & Yang, S.-Y. (2019). A cloud-based energy data mining information agent system based on big data analysis technology. *Microelectronics Reliability, 97*, 66-78.

*Lino, A., Rocha, Á., Macedo, L., & Sizo, A. (2019). Application of clustering-based decision tree approach in SQL query error database. *Future Generation Computer Systems, 93*, 392-406.

*Liu, W., Wang, J., Sangaiah, A. K., & Yin, J. (2018). Dynamic metric embedding model for point-of-interest prediction. *Future Generation Computer Systems, 83*, 183-192.

*Liu, Z., Liu, S., Liu, L., Sun, J., Peng, X., & Wang, T. (2016). Sentiment recognition of online course reviews using multi-swarm optimization-based selected features. *Neurocomputing, 185*, 11-20.

*Loderer, K., Pekrun, R., & Lester, J. C. (2018). Beyond cold technology: A systematic review and meta-analysis on emotions in technology-based learning environments. *Learning and Instruction*.

*Loh, C. S., & Sheng, Y. (2014). Maximum Similarity Index (MSI): A metric to differentiate the performance of novices vs. multiple-experts in serious games. *Computers in Human Behavior, 39*, 322-330.

*Lucke, U., & Rensing, C. (2014). A survey on pervasive education. *Pervasive and Mobile Computing, 14*, 3-16.

*Maitra, S., Madan, S., Kandwal, R., & Mahajan, P. (2018). Mining authentic student feedback for faculty using Naïve Bayes classifier. *Procedia Computer Science, 132*, 1171-1183.

*Maldonado-Mahauad, J., Pérez-Sanagustín, M., Kizilcec, R. F., Morales, N., & Munoz-Gama, J. (2018). Mining theory-based patterns from Big data: Identifying self-regulated learning strategies in Massive Open Online Courses. *Computers in Human Behavior, 80*, 179-196.

*Malmberg, J., Järvelä, S., Järvenoja, H., & Panadero, E. (2015). Promoting socially shared regulation of learning in CSCL: Progress of socially shared regulation among high- and low-performing groups. *Computers in Human Behavior, 52*, 562-572.

*Marbouti, F., Diefes-Dux, H. A., & Madhavan, K. (2016). Models for early prediction of at-risk students in a course using standards-based grading. *Computers & Education, 103*, 1-15.

*Markéta, D., & Kateřina, K. (2012). Complex Model of e-Learning Evaluation Focusing on Adaptive Instruction. *Procedia - Social and Behavioral Sciences, 47*, 1068-1076.

*Martinez-Maldonado, R., Yacef, K., & Kay, J. (2015). TSCL: A conceptual model to inform understanding of collaborative learning processes at interactive tabletops. *International Journal of Human-Computer Studies, 83*, 62-82.

*Mayilvaganan, M., & Kalpanadevi, D. (2015). Cognitive Skill Analysis for Students through Problem Solving Based on Data Mining Techniques. *Procedia Computer Science, 47*, 62-75.

*Michlík, P., & Bieliková, M. (2010). Exercises recommending for limited time learning. *Procedia Computer Science, 1*(2), 2821-2828.

*Miguéis, V. L., Freitas, A., Garcia, P. J. V., & Silva, A. (2018). Early segmentation of students according to their academic performance: A predictive modelling approach. *Decision Support Systems, 115*, 36-51.

*Miholca, D.-L., Czibula, G., & Crivei, L. M. (2018). A new incremental relational association rules mining approach. *Procedia Computer Science, 126*, 126-135.

*Minović, M., Milovanović, M., Šošević, U., & Conde González, M. Á. (2015). Visualisation of student learning model in serious games. *Computers in Human Behavior, 47*, 98-107.

*Modesto, D. M., & Ferreira, S. B. L. (2014). Guidelines for Search Features Development – A Comparison between General Users and Users with Low Reading Skills. *Procedia Computer Science, 27*, 334-342.

*Mohamad, S. K., & Tasir, Z. (2013). Educational Data Mining: A Review. *Procedia - Social and Behavioral Sciences, 97*, 320-324.

*Mohammad, M. N., Sulaiman, N., & Muhsin, O. A. (2011). A novel intrusion detection system by using intelligent data mining in weka environment. *Procedia Computer Science, 3*, 1237-1242.

Moher D, Liberati A, Tetzlaff J, Altman DG. (2009) The PRISMA group, preferred reporting items for systematic reviews and metaanalyses: The PRISMA statement, *PLoS Med* 6(7): e1000097.

*Montenegro, M., Clasing, P., Kelly, N., Gonzalez, C., Jara, M., Alarcón, R., . . . Saurina, E. (2016). Library Resources and Students' Learning Outcomes: Do All the Resources Have the Same Impact on Learning? *The Journal of Academic Librarianship, 42*(5), 551-556.

*Mödritscher, F. (2010). Towards a recommender strategy for personal learning environments. *Procedia Computer Science, 1*(2), 2775-2782.

*Munk, M., & Drlík, M. (2011). Impact of Different Pre-Processing Tasks on Effective Identification of Users' Behavioral Patterns in Web-based Educational System. *Procedia Computer Science, 4*, 1640-1649.

*Munk, M., Pilkova, A., Kapusta, J., Svec, P., & Drlik, M. (2013). Pillar 3 and Modelling of Stakeholders' Behaviour at the Commercial Bank Website during the Recent Financial Crisis. *Procedia Computer Science, 18*, 1747-1756.

*Munková, D., Stranovská, E., Fráterová, Z., & Ďuračková, B. (2013). Identity vs. Foreign Language and Culture in Production of Speech Acts. *Procedia - Social and Behavioral Sciences, 84*, 361-36.

*Munková, D., Stranovská, E., Munk, M., & Ďuračková, B. (2013). Impact of Cognitive Style "Category Width" on the use of Social and Expressive Factors in Politeness Speech Acts: Text Mining Application. *Procedia - Social and Behavioral Sciences, 82*, 775-779.

*Muñoz, P., R-Moreno, M. D., & Barrero, D. F. (2016). Unified framework for path-planning and task-planning for autonomous robots. *Robotics and Autonomous Systems, 82*, 1-14.

*N, N., & Pujari, N. V. (2016). A Tutor Assisting Novel Electronic Framework for Qualitative Analysis of a Question Bank. *Computers in Human Behavior, 65*, 9-13.

*Narciss, S., Sosnovsky, S., Schnaubert, L., Andrès, E., Eichelmann, A., Goguadze, G., & Melis, E. (2014). Exploring feedback and student characteristics relevant for personalizing feedback strategies. *Computers & Education, 71*, 56-76.

*Natek, S., & Zwilling, M. (2014). Student data mining solution–knowledge management system related to higher education institutions. *Expert Systems with Applications, 41*(14), 6400-6407.

*Nistor, N., & Hernández-Garcíac, Á. (2018). What types of data are used in learning analytics? An overview of six cases. *Computers in Human Behavior, 89*, 335-338.

*Nkambou, R., Fournier-Viger, P., & Nguifo, E. M. (2011). Learning task models in ill-defined domain using an hybrid knowledge discovery framework. *Knowledge-Based Systems, 24*(1), 176-185.

*Noroozi, O., Alikhani, I., Järvelä, S., Kirschner, P. A., Juuso, I., & Seppänen, T. (2018). Multimodal data to design visual learning analytics for understanding regulation of learning. *Computers in Human Behavior*.

*Northcutt, C. G., Ho, A. D., & Chuang, I. L. (2016). Detecting and preventing “multiple-account” cheating in massive open online courses. *Computers & Education, 100*, 71-80.

*Nouira, A., Cheniti-Belcadhi, L., & Braham, R. (2018). An Enhanced xAPI Data Model Supporting Assessment Analytics. *Procedia Computer Science, 126*, 566-575.

*Oeda, S., & Hashimoto, G. (2017). Log-Data Clustering Analysis for Dropout Prediction in Beginner Programming Classes. *Procedia Computer Science, 112*, 614-621.

*Ognjanovic, I., Gasevic, D., & Dawson, S. (2016). Using institutional data to predict student course selections in higher education. *The Internet and Higher Education, 29*, 49-62.

Özbay, Ö. (2015). Data mining concept and data mining applications in education, *The Journal of International Education Science*, 5: 262-272.

*Paassen, B., Mokbel, B., & Hammer, B. (2016). Adaptive structure metrics for automated feedback provision in intelligent tutoring systems. *Neurocomputing, 192*, 3-13.

*Palombi, O., Jouanot, F., Nziengam, N., Omidvar-Tehrani, B., Rousset, M.-C., & Sanchez, A. (2019). OntoSIDES: Ontology-based student progress monitoring on the national evaluation system of French Medical Schools. *Artificial Intelligence in Medicine, 96*, 59-67.

*Papamitsiou, Z., & Economides, A. A. (2017). Exhibiting achievement behavior during computer-based testing: What temporal trace data and personality traits tell us? *Computers in Human Behavior, 75*, 423-438.

*Park, Y., Yu, J. H., & Jo, I.-H. (2016). Clustering blended learning courses by online behavior data: A case study in a Korean higher education institute. *The Internet and Higher Education, 29*, 1-11.

*Prates, J. C., Fritzen, E., Siqueira, S. W. M., Braz, M. H. L. B., & de Andrade, L. C. V. (2013). Contextual web searches in Facebook using learning materials and discussion messages. *Computers in Human Behavior, 29*(2), 386-394.

Pritchard, A. (1969) Statistical bibliography or bibliometrics, *Journal of Documentation* 25: 348-349.

*Priyambada, S. A., Mahendrawathi, E. R., & Yahya, B. N. (2017). Curriculum Assessment of Higher Educational Institution Using Aggregate Profile Clustering. *Procedia Computer Science, 124*, 264-273.*Reamer, A. C., Ivy, J. S., Vila-Parrish, A. R., & Young, R. E. (2015). Understanding the evolution of mathematics performance in primary education and the implications for STEM learning: A Markovian approach. *Computers in Human Behavior, 47*, 4-17.

*Richey, J. E., Andres-Bray, J. M. L., Mogessie, M., Scruggs, R., Andres, J. M. A. L., Star, J. R.,. McLaren, B. M. (2019). More confusion and frustration, better learning: The impact of erroneous examples. *Computers & Education*.

*Rizvi, S., Rienties, B., & Khoja, S. A. (2019). The role of demographics in online learning; A decision tree based approach. *Computers & Education, 137*, 32-47.

*Robinson, K. (2013). The interrelationship of emotion and cognition when students undertake collaborative group work online: An interdisciplinary approach. *Computers & Education, 62*, 298-307.

*Rodrigues, F., & Oliveira, P. (2014). A system for formative assessment and monitoring of students' progress. *Computers & Education, 76*, 30-41.

*Romero, C., López, M.-I., Luna, J.-M., & Ventura, S. (2013). Predicting students' final performance from participation in on-line discussion forums. *Computers & Education, 68*, 458-472.

*Romero, C., & Ventura, S. (2007). Educational data mining: A survey from 1995 to 2005. *Expert Systems with Applications, 33*(1), 135-146.

*Romero, C., Ventura, S., & García, E. (2008). Data mining in course management systems: Moodle case study and tutorial. *Computers & Education, 51*(1), 368-384.

Romero, C., Ventura, S. (2010). Educational data mining: a review of the state of the art. IEEE Transactions On Systems, Man, And Cybernetics—Part C: *Applications and Reviews.* 40. 601 - 618.

Romero, C., Ventura, S., & De Bra, P. (2004). Knowledge discovery with genetic programming for providing feedback to courseware author, user modeling and user- adapted interaction, *The Journal of Personalization Research*, 14(5), 425–464.

*Romero, C., Ventura, S., Zafra, A., & Bra, P. d. (2009). Applying Web usage mining for personalizing hyperlinks in Web-based adaptive educational systems. *Computers & Education, 53*(3), 828-840.

*Romero-Zaldivar, V.-A., Pardo, A., Burgos, D., & Delgado Kloos, C. (2012). Monitoring student progress using virtual appliances: A case study. *Computers & Education, 58*(4), 1058-1067.

*Rowe, E., Asbell-Clarke, J., Baker, R. S., Eagle, M., Hicks, A. G., Barnes, T. M., . . . Edwards, T. (2017). Assessing implicit science learning in digital games. *Computers in Human Behavior, 76*, 617-630.

*Saggi, M. K., & Jain, S. (2018). A survey towards an integration of big data analytics to big insights for value-creation. *Information Processing & Management, 54*(5), 758-790.

*Salehi, M., & Nakhai Kamalabadi, I. (2013). Hybrid recommendation approach for learning material based on sequential pattern of the accessed material and the learner’s preference tree. *Knowledge-Based Systems, 48*, 57-69.

*Saleti, S., & R.B.V, S. (2019). A MapReduce solution for incremental mining of sequential patterns from big data. *Expert Systems with Applications, 133*, 109-125.

*Salmeron-Majadas, S., Santos, O. C., & Boticario, J. G. (2014). An Evaluation of Mouse and Keyboard Interaction Indicators towards Non-intrusive and Low Cost Affective Modeling in an Educational Context. *Procedia Computer Science, 35*, 691-700.

*Sandoval, A., Gonzalez, C., Alarcon, R., Pichara, K., & Montenegro, M. (2018). Centralized student performance prediction in large courses based on low-cost variables in an institutional context. *The Internet and Higher Education, 37*, 76-89.

*Santos, O. C., Boticario, J. G., & Pérez-Marín, D. (2014). Extending web-based educational systems with personalised support through User Centred Designed recommendations along the e-learning life cycle. *Science of Computer Programming, 88*, 92-109.

*Sarraipa, J., Baldiris, S., Fabregat, R., & Jardim-Goncalves, R. (2012). Knowledge Representation in Support of Adaptable eLearning Services for All. *Procedia Computer Science, 14*, 391-402.

*Saxena, A., Prasad, M., Gupta, A., Bharill, N., Patel, O. P., Tiwari, A., . . . Lin, C.-T. (2017). A review of clustering techniques and developments. *Neurocomputing, 267*, 664-681.

*Scheiter, K., Schubert, C., Schüler, A., Schmidt, H., Zimmermann, G., Wassermann, B., . . . Eder, T. (2019). Adaptive multimedia: Using gaze-contingent instructional guidance to provide personalized processing support. *Computers & Education, 139*, 31-47.

*Serrano-Laguna, Á., Martínez-Ortiz, I., Haag, J., Regan, D., Johnson, A., & Fernández-Manjón, B. (2017). Applying standards to systematize learning analytics in serious games. *Computer Standards & Interfaces, 50*, 116-123.

*Serrano-Laguna, Á., Torrente, J., Moreno-Ger, P., & Fernández-Manjón, B. (2014). Application of Learning Analytics in educational videogames. *Entertainment Computing, 5*(4), 313-322.

*Shahiri, A. M., Husain, W., & Rashid, N. a. A. (2015). A Review on Predicting Student's Performance Using Data Mining Techniques. *Procedia Computer Science, 72*, 414-422.

*Shivaprasad, G., Reddy, N. V. S., Acharya, U. D., & Aithal, P. K. (2015). Neuro-Fuzzy Based Hybrid Model for Web Usage Mining. *Procedia Computer Science, 54*, 327-334.

*Shoufan, A. (2019). Estimating the cognitive value of YouTube's educational videos: A learning analytics approach. *Computers in Human Behavior, 92*, 450-458.

*Shukor, N. A., Tasir, Z., & Van der Meijden, H. (2015). An Examination of Online Learning Effectiveness Using Data Mining. *Procedia - Social and Behavioral Sciences, 172*, 555-562.

*Shukor, N. A., Tasir, Z., Van der Meijden, H., & Harun, J. (2014). A Predictive Model to Evaluate Students’ Cognitive Engagement in Online Learning. *Procedia - Social and Behavioral Sciences, 116*, 4844-4853.

Small, H. (1974). Co-citation in the scientific literature: a new measure of the relationship between two documents, *Essays of an Information Scientist* 2:28-31.

*Soffer, T., Kahan, T., & Livne, E. (2017). E-assessment of online academic courses via students' activities and perceptions. *Studies in Educational Evaluation, 54*, 83-93.

*Sonnenberg, C., & Bannert, M. (2019). Using Process Mining to examine the sustainability of instructional support: How stable are the effects of metacognitive prompting on self-regulatory behavior? *Computers in Human Behavior, 96*, 259-272.

*Stahovich, T. F., & Lin, H. (2016). Enabling data mining of handwritten coursework. *Computers & Graphics, 57*, 31-45.

*Strohmeier, S., & Piazza, F. (2013). Domain driven data mining in human resource management: A review of current research. *Expert Systems with Applications, 40*(7), 2410-2420.

*Sunar, A. S., Abbasi, R. A., Davis, H. C., White, S., & Aljohani, N. R. (2018). Modelling MOOC learners' social behaviours. *Computers in Human Behavior*.

*Svetsky, S., & Moravcik, O. (2013). The Supportive System for the Processing of Human and Non-human Knowledge Sources and Combining Mining Techniques. *AASRI Procedia, 4*, 182-189.

*Tanes, Z., Arnold, K. E., King, A. S., & Remnet, M. A. (2011). Using Signals for appropriate feedback: Perceptions and practices. *Computers & Education, 57*(4), 2414-2422.

*Tezer, M., & Ertarkan, Z. (2010). Teachers' proficiency and infrastructural problems of using technology during the process of technology integration in pre-school education institutions of North Cyprus. *Procedia - Social and Behavioral Sciences, 9*, 1960-1968.

*Thai-Nghe, N., Drumond, L., Krohn-Grimberghe, A., & Schmidt-Thieme, L. (2010). Recommender system for predicting student performance. *Procedia Computer Science, 1*(2), 2811-2819.

*Theobald, M., Bellhäuser, H., & Imhof, M. (2018). Identifying individual differences using log-file analysis: Distributed learning as mediator between conscientiousness and exam grades. *Learning and Individual Differences, 65*, 112-122.

*Topîrceanu, A., & Grosseck, G. (2017). Decision tree learning used for the classification of student archetypes in online courses. *Procedia Computer Science, 112*, 51-60.

*Torabi, R., Moradi, P., & Khantaimoori, A. R. (2012). Predict Student Scores Using Bayesian Networks. *Procedia - Social and Behavioral Sciences, 46*, 4476-4480.

*Tsai, C.-J., & Shih, K.-J. (2019). Mining a new biometrics to improve the accuracy of keystroke dynamics-based authentication system on free-text. *Applied Soft Computing, 80*, 125-137.

*Tseng, H., Yi, X., & Yeh, H.-T. (2019). Learning-related soft skills among online business students in higher education: Grade level and managerial role differences in self-regulation, motivation, and social skill. *Computers in Human Behavior, 95*, 179-186.

*Tu, E., Cao, L., Yang, J., & Kasabov, N. (2014). A novel graph-based k-means for nonlinear manifold clustering and representative selection. *Neurocomputing, 143*, 109-122.

*Uddin, M. F., & Lee, J. (2016a). Utilizing Relevant Academic and Personality Features from Big Unstructured Data to Identify Good and Bad Fit Students. *Procedia Computer Science, 95*, 383-391.

*Uddin, M. F., & Lee, J. (2016b). We Are What We Generate - Understanding Ourselves Through Our Data. *Procedia Computer Science, 95*, 335-344.

*Ueda, H., & Nakamura, M. (2017). Data Analysis for Evaluation on Course Design and Improvement of "Cyberethics" Moodle Online Courses. *Procedia Computer Science, 112*, 2345-2353.

*Upendran, D., Chatterjee, S., Sindhumol, S., & Bijlani, K. (2016). Application of Predictive Analytics in Intelligent Course Recommendation. *Procedia Computer Science, 93*, 917-923.

*Ünlü, A., & Schrepp, M. (2015). Untangling comparison bias in inductive item tree analysis based on representative random quasi-orders. *Mathematical Social Sciences, 76*, 31-43.

*Vahdat, M., Oneto, L., Anguita, D., Funk, M., & Rauterberg, M. (2016). Can machine learning explain human learning? *Neurocomputing, 192*, 14-28.

*Valls, F., Redondo, E., Fonseca, D., Torres-Kompen, R., Villagrasa, S., & Martí, N. (2018). Urban data and urban design: A data mining approach to architecture education. *Telematics and Informatics, 35*(4), 1039-1052.

*Valsamidis, S., Kazanidis, I., Petasakis, I., Kontogiannis, S., & Kolokitha, E. (2014). E-Learning Activity Analysis. *Procedia Economics and Finance, 9*, 511-518.

Van Eck, N., & Waltman, L. (2010). Software survey: VOSviewer, a computer program for bibliometric mapping. *Scientometrics*, 84(2), 523-538.

*Verba, N., Chao, K.-M., Lewandowski, J., Shah, N., James, A., & Tian, F. (2019). Modeling industry 4.0 based fog computing environments for application analysis and deployment. *Future Generation Computer Systems, 91*, 48-60.

*Vieira, C., Parsons, P., & Byrd, V. (2018). Visual learning analytics of educational data: A systematic literature review and research agenda. *Computers & Education, 122*, 119-135.

*Wang, H.-C., Chang, C.-Y., & Li, T.-Y. (2007). The comparative efficacy of 2D- versus 3D-based media design for influencing spatial visualization skills. *Computers in Human Behavior, 23*(4), 1943-1957.

*Wang, X., Zhou, C., & Xu, X. (2019). Application of C4.5 decision tree for scholarship evaluations. *Procedia Computer Science, 151*, 179-184.

*Wang, Y., Zhao, Y., Zhou, Q., & Lin, Z. (2018). Image encryption using partitioned cellular automata. *Neurocomputing, 275*, 1318-1332.

*Wassan, J. T. (2015). Discovering Big Data Modelling for Educational World. *Procedia - Social and Behavioral Sciences, 176*, 642-649.

*Weng, C.-H. (2011). Mining fuzzy specific rare itemsets for education data. *Knowledge-Based Systems, 24*(5), 697-708.

White, H. D., & K. W. Mccain. (1998). Visualizing a discipline: an author co citation analysis of information science, 1972–1995., *Journal of the American Society for Information Science*, 49 (4): 327–355.

*Widyahastuti, F., & Tjhin, V. U. (2018). Performance Prediction in Online Discussion Forum: state-of-the-art and comparative analysis. *Procedia Computer Science, 135*, 302-314.

*Wise, A. F., Perera, N., Hsiao, Y.-T., Speer, J., & Marbouti, F. (2012). Microanalytic case studies of individual participation patterns in an asynchronous online discussion in an undergraduate blended course. *The Internet and Higher Education, 15*(2), 108-117.

*Xiao, X., Skitmore, M., & Hu, X. (2017). Case-based Reasoning and Text Mining for Green Building Decision Making. *Energy Procedia, 111*, 417-425.

*Xie, T., Zheng, Q., & Zhang, W. (2018). Mining temporal characteristics of behaviors from interval events in e-learning. *Information Sciences, 447*, 169-185.

*Xing, W., Chen, X., Stein, J., & Marcinkowski, M. (2016). Temporal predication of dropouts in MOOCs: Reaching the low hanging fruit through stacking generalization. *Computers in Human Behavior, 58*, 119-129.

*Xing, W., & Gao, F. (2018). Exploring the relationship between online discourse and commitment in Twitter professional learning communities. *Computers & Education, 126*, 388-398.

*Xing, W., Goggins, S., & Introne, J. (2018). Quantifying the Effect of Informational Support on Membership Retention in Online Communities through Large-Scale Data Analytics. *Computers in Human Behavior, 86*, 227-234.

*Xing, W., Guo, R., Petakovic, E., & Goggins, S. (2015). Participation-based student final performance prediction model through interpretable Genetic Programming: Integrating learning analytics, educational data mining and theory. *Computers in Human Behavior, 47*, 168-181.

*Xing, W., Popov, V., Zhu, G., Horwitz, P., & McIntyre, C. (2019). The effects of transformative and non-transformative discourse on individual performance in collaborative-inquiry learning. *Computers in Human Behavior, 98*, 267-276.

*Yahya, A. A. (2019). Swarm intelligence-based approach for educational data classification. *Journal of King Saud University - Computer and Information Sciences, 31*(1), 35-51.

*Yahya, A. A., Osman, A., & El-Bashir, M. S. (2017). Rocchio algorithm-based particle initialization mechanism for effective PSO classification of high dimensional data. *Swarm and Evolutionary Computation, 34*, 18-32.

*Yang, J., Lee, Y., & Chang, K. H. (2018). Evaluations of JaguarCode: A web-based object-oriented programming environment with static and dynamic visualization. *Journal of Systems and Software, 145*, 147-163.

*Yang, Z., Liu, Z., Liu, S., Min, L., & Meng, W. (2014). Adaptive multi-view selection for semi-supervised emotion recognition of posts in online student community. *Neurocomputing, 144*, 138-150.

*Yengin, I., Karahoca, A., & Karahoca, D. (2011). E-learning success model for instructors' satisfactions in perspective of interaction and usability outcomes. *Procedia Computer Science, 3*, 1396-1403.

*Yıldız, O., Bal, A., Gülseçen, S., & Kentli, F. D. (2012). A Genetic-Fuzzy Based Mathematical Model to Evaluate The Distance Education Students' Academic Performance. *Procedia - Social and Behavioral Sciences, 55*, 409-418.

*Yin, M. S., Haddawy, P., Suebnukarn, S., & Rhienmora, P. (2018). Automated outcome scoring in a virtual reality simulator for endodontic surgery. *Computer Methods and Programs in Biomedicine, 153*, 53-59.

*You, J. W. (2016). Identifying significant indicators using LMS data to predict course achievement in online learning. *The Internet and Higher Education, 29*, 23-30.

*Yüce, B., Karahoca, A., & Karahoca, D. (2011). The use of electronic curriculums in occupational education to evaluate and improve the cognitive capacity of candidate software engineers. *Procedia Computer Science, 3*, 1418-1424.

*Zafra, A., Romero, C., & Ventura, S. (2011). Multiple instance learning for classifying students in learning management systems. *Expert Systems with Applications, 38*(12), 15020-15031.

*Zafra, A., & Ventura, S. (2012). Multi-instance genetic programming for predicting student performance in web based educational environments. *Applied Soft Computing, 12*(8), 2693-2706.

Zaim, H. (2005) *Bilginin artan önemi ve bilgi yönetimi*, İşaret Yayınları: İstanbul.

*Zapata, A., Menéndez, V. H., Prieto, M. E., & Romero, C. (2015). Evaluation and selection of group recommendation strategies for collaborative searching of learning objects. *International Journal of Human-Computer Studies, 76*, 22-39.

*Zengin, K., Esgi, N., Erginer, E., & Aksoy, M. E. (2011). A sample study on applying data mining research techniques in educational science: Developing a more meaning of data. *Procedia - Social and Behavioral Sciences, 15*, 4028-4032.

*Zhang, M. (2015). Internet use that reproduces educational inequalities: Evidence from big data. *Computers & Education, 86*, 212-223.

*Zheng, J., Xing, W., & Zhu, G. (2019). Examining sequential patterns of self- and socially shared regulation of STEM learning in a CSCL environment. *Computers & Education, 136*, 34-48.

*Zhong, X., Sun, Z., Xiong, H., Heffernan, N., & Islam, M. M. (2017). Learning Curve Analysis Using Intensive Longitudinal and Cluster-Correlated Data. *Procedia Computer Science, 114*, 250-257.

*Zhu, G., Xing, W., & Popov, V. (2019). Uncovering the sequential patterns in transformative and non-transformative discourse during collaborative inquiry learning. *The Internet and Higher Education, 41*, 51-61.

*Zhuhadar, L., Marklin, S., Thrasher, E., & Lytras, M. D. (2016). Is there a gender difference in interacting with intelligent tutoring system? Can Bayesian Knowledge Tracing and Learning Curve Analysis Models answer this question? *Computers in Human Behavior, 61*, 198-204.

*Zhuhadar, L., Yang, R., & Lytras, M. D. (2013). The impact of Social Multimedia Systems on cyberlearners. *Computers in Human Behavior, 29*(2), 378-385.

Zins, Chaim. (2007). Conceptual approaches for defining data, information, and knowledge. *JASIST*. 58. 479-493. 10.1002/asi.20508.

*Zorrilla, M., & García-Saiz, D. (2013). A service oriented architecture to provide data mining services for non-expert data miners. *Decision Support Systems, 55*(1), 399-411.