*International Journal on Recent and Innovation Trends in Computing and Communication*
*ISSN: 2321-8169 Volume: 12 Issue: 1*
***Article Received:** 10 July 2023 **Revised:** 12 August 2023 **Accepted:** 28 August 2023*

# The Importance of Educational Data Mining and Learning Analytics for Improving Teaching and Learning: An Issue Brief

**Lalhruaitluanga[1], C.Lalrinawma[2]**

[1]Asst. Professor, Dept of Computer Science, Govt.Zirtiri Residential Science College
[2]Asst. Professor, Dept of Computer Science, Govt.Zirtiri Residential Science College

**Abstract:**
**"**The words educational data mining and learning analytics are frequently used interchangeably, despite their being an increase in their investigation and implementation. This may be as a result of the fact that both areas have similar conceptual components. One way to ensure precision, homogeneity, and consistency It aims to pinpoint themes that are similar to and different from one other in the two domains as they develop. This a topic modelling study of papers on educational data mining and learning analytics was carried out in the elucidate the two areas' respective themes. In particular, we used structural topic modelling to find the two domains' subjects from the abstracts. For instructional purposes, we use structural topic modelling on N 1 4192 articles. For both educational data and survey data, we infer five-topic models analytics for mining and learning. While there may be disciplinary variations in research, our findings show that beyond their various lineages, there is no evidence to indicate a clear separation between the two disciplines. the area of educational research on the uses of advanced statistical methods is trending toward convergence for improving teaching and learning, discover how to mine massive data streams for insights that may be put to use. Over the past five years, both areas have converged on a growing emphasis on student behaviour. This study topic has advanced greatly, and a variety of related words, including Academic Analytics, Institutional Analytics, Teaching Analytics, Data-Driven Education, Data-Driven Decision-Making in Education, Big Data in Education, and Educational Data Science, are now used in the paper. The main publications, significant turning points, cycle of knowledge discovery, primary educational settings, specialised tools, freely accessible datasets, widely used methodologies, primary goals, and anticipated trends in this field of study are reviewed to provide the state of the art at this time.

**Keywords:** Educational Data Mining, Data Mining on Education, Data-Driven Decision-Making in Education, Big Data in Education, Educational Data Science, Learning analytics."

## 1. INTRODUCTION

"Education-related artificial intelligence (AI) research is still in its infancy. The two data-centric domains of educational data mining (EDM) and learning analytics (LA), which use machine learning in educational research, are the greatest examples of AI in education today (Baek & Doleck, 2020). Many evaluations have attempted to define these two complimentary domains, which emerged as a result of the open source and open data movements in educational research and the democratisation of access to computer power, in the future [1]. Learning analytics refers to the usage of patterns based on educational instructional data, whereas "educational data mining" is the name given to data mining based on educational data (Sahin & Yurdugul, 2019, p. 122). Data mining may be traced back to the early days of computerised record keeping, the introduction of object-relational databases, and the use of structured records for knowledge discovery through pattern analysis. With the proliferation of information, new methods for turning this data into practical insights have revolutionised practises [2]. Learning analytics is based on using educational data for informing its own objectives, i.e., teaching and learning, just like management analytics and business intelligence before it. If not as coextensive, then at least as two closely related undertakings committed to the application of sophisticated statistical learning approaches to the analysis of learning and educational data, respectively, LA and EDM are frequently treated as such in empirical research. Many people perceive the two fields as being largely interchangeable (Romero and Ventura, 2020), while some supporters insist that there should be a distinct distinction between the two. The main defence is based on an exclusionary definition of each field of study [3]. In essence, the distinction rests on the understanding that, despite any potential overlap, learning is not the same as education and vice versa. Thus, while EDM focuses on knowledge discovery from all created educational data sources, LA is focused on processes impacting learning at the individual and societal levels [4]. Two distinct groups have developed in the same region with a shared interest in how educational data may be used to advance both education and the science of learning (Baker & Inventado, 2014):"

- The goal of Educational Data Mining (EDM) is to provide techniques for examining the distinctive sorts of data that are generated by educational contexts (Bakhshinategh, Zaiane, ElAtia, & Ipperciel, 2018). It may also be described as the process of using data mining (DM) techniques on a particular type of dataset derived from educational contexts in order to answer crucial educational issues "(Romero & Ventura, 2013) [5].

- In order to understand and improve learning and the settings in which it takes place, learning analytics (LA) is the measurement, gathering, analysis, and reporting of data on learners and their surroundings (Lang, Siemens, Wise, & Gasevic, 2017). This concept depends on three essential components: data, analysis, and action (Siemens, 2013).
- The enhancement of educational practise is a shared objective of both groups, and both have an interest in data-intensive methods to educational research (Siemens & Baker, 2012; Lián & Pérez, 2015). While EDM is concentrated on the technology problem, LA is concentrated on the educational challenge. By utilising well-known prediction models, [6] LA is concentrated on data-driven decision-making and merging the technical and social/pedagogical elements of learning. On the other hand, EDM often searches for novel data patterns and creates fresh models and/or algorithms. In the end, rather than the methodologies employed, the distinctions between the two groups are more dependent on focus, research topics, and final application of models (Baker & Inventado, 2014). Despite the distinctions between the LA and EDM communities, there is a large amount of overlap between the two groups' goals for the research as well as its tactics and procedures [7].

Information retrieval, recommender systems, visual data analytics, domain-driven data mining, social network analysis, psychopedagogy, cognitive psychology, psychometrics, and other multidisciplinary fields are included in EDM and LA. In reality, they may be visualised as the fusion of three key fields: computer science, education, and statistics (Figure 1). Other subfields strongly connected to EDM and LA are formed at the junction of these three topics, including computer-based education (CBE), data mining and machine learning, and educational statistics [8].

### 1.1. Institutional analytics (IA) and academic analytics (AA)

For the purpose of generating institutional insight, academic programme activities such as courses, degree programmes, research, student fee revenue, course evaluation, resource allocation, and management are collected, analysed, and visualised (Campbell, DeBlois, & Oblinger, 2007; Siemens and Long, 2011). Thus, the political and economic problem is the main subject."

**a. Teaching analytics (TA)** is the term used to "describe the planning, development, and assessment of teaching activities as well as the analysis of teaching activities and performance data (Prieto, Sharma, Dillenbourg, & Jess, 2016). It focuses on the educational challenge as seen from the perspective of the educators [9].

**b. Data-Driven Education (DDE) and Data-Driven Decision-Making in Education (DDDM)** refer to the systematic collection and analysis of different types of educational data, which then informs a variety of decisions that may be made to enhance the performance of students and schools (Custer, King, Atinc, Read, & Sethi, 2018; Datnow & Hubbard, 2016).

**c. Big Data in Education (BDE)** refers to the use of big data techniques to data from educational environments, with the basic meaning summed up in volume, diversity, value, and velocity (Daniel, 2019). The use of data acquired from educational contexts or settings to address educational issues is known as Educational Data Science (EDS) (Romero & Ventura, 2017). Data science is a term used to describe the integration of statistics, data analysis, machine learning, and related techniques [10].

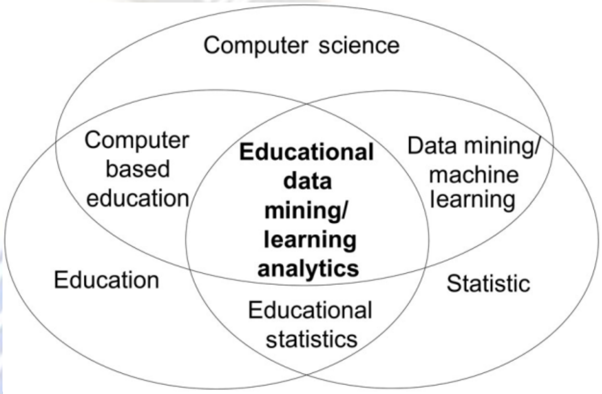


**FIGURE 1** Highlights of Educational Data Mining and Learning Analytics

A comprehensive assessment of the state of the art in EDM and LA has to be redone because six years have gone and a tonne of new studies have been released. The most significant changes since the last survey are: new terms related to education are used in the bibliography (AA, IA, TA, DDE, DDDM, BDE, and EDS); the number of books and papers published has increased exponentially (more in LA than EDM); the interest in data related to new types of educational environments has increased (Massive Open Online Courses, MOOCs, virtual and augmented reality learning, serious games, BL, etc.); [11] and more specialised tools and free datasets futuristic new trends.

## 2. BACKGROUND

LA and EDM are the results of two separate conferences and communities. The IEDM society hosted the inaugural Educational Data Mining Conference in Montreal, Canada in 2008, while the SOLAR society hosted the first Learning Analytics and Knowledge Conference in Banff, Canada in 2011. There are several further conferences that are connected (Table 1) [12]."

**TABLE 1 Listing of the most relevant conferences for learning analytics and educational data mining**

| Title | Acronym | Type | 1° year |
|---|---|---|---|
| International Conference on Artificial Intelligence in Education | AIED | Biannual | 1982 |
| International Conference on Intelligent Tutoring Systems | ITS | Biannual | 1988 |
| IEEE International Conference on Advanced Learning Technologies | ICALT | Annual | 2000 |
| European Conference on Technology-Enhanced Learning | EC-TEL | Annual | 2006 |
| International Conference on Educational Data Mining | EDM | Annual | 2008 |
| International Conference on User Modeling, Adaptation, and Personalization | UMAP | Annual | 2009 |
| International Conference on Learning Analytics and Knowledge | LAK | Annual | 2011 |
| Learning at Scale | L@S | Annual | 2014 |
| Learning and Students Analytics Conference | LSAC | Annual | 2017 |

"Data Mining in E-Learning, the first book addressing EDM/LA themes, was released in 2006. (Romero & Ventura, 2006). Since then, more books have been published than ever before (Table 2), particularly in recent years. We can observe that the phrases Data Mining in Education and Educational Data Mining were used in the titles throughout the initial years (from 2006 to 2014). The words Learning Analytics, Data Science, Big Data, and Data Mining were then used (from 2015 to 2017) [13]. Additionally, the phrases "Learning Analytics" have been more prevalent in title tags in recent years. The Handbook of Educational Data Mining (Romero et al., 2010) and the Encyclopaedia of Educational Data Mining are the two most significant works in the field. (CBLC)5 were both released in 2019. The number of publications or results returned by a publicly available web search engine like Google Schoolar when Googling the precise phrase is shown in Figure 2 to indicate the rising interest in EDM and LA over the past 20 years. From 2000 to 2018, the terms "Educational Data Mining" or "Learning Analytics" were searched for year. Both numbers, as can be seen, expand exponentially, demonstrating the considerable interest in both subjects. While LA has more allusions than EDM, LA didn't overtake EDM until 2011. The timing of the following significant events in can be used to explain this fact [14]. EDM and LA history includes, for example: the inaugural EDM workshop held by the Association for the Advancement of Montreal conference (Canada) 2009, the EDM guide (Romero et al., 2010), and the first Banff Learning Analytics & Knowledge (Canada) 2011, the first Palo Alto LA Summer Institute (USA) the first Learning at Scale Conference in 2013 2015 in Atlanta (USA), and the LA manual (Lang et al., 2017). This more significant increase in the number of A bibliometric analysis of works that use the phrase "Learning Analytics" rather than "Educational Data Mining" approach (Dormezil, Khoshgoftaar, & Robinson-Bryant, 2019) come to the conclusion that it is more accurate to characterise what seems to One domain (Learning Analytics) combining two (Educational Data Mining and Learning Analytics) notable subsets (i.e., Educational Data Mining) [15]."

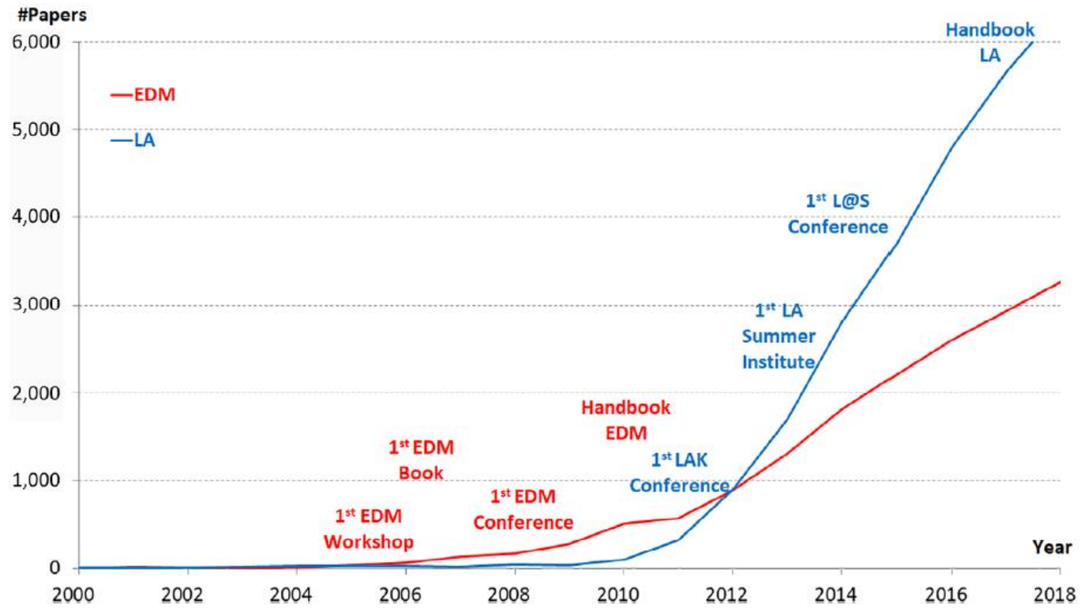


**FIGURE 2** By year, the number of articles and significant events related to educational data mining and learning analytics (January 1, 2019)

*International Journal on Recent and Innovation Trends in Computing and Communication*
*ISSN: 2321-8169 Volume: 12 Issue: 1*
***Article Received:** 10 July 2023 **Revised:** 12 August 2023 **Accepted:** 28 August 2023*

---

### 2.1. Knowledge discovery cycle for EDM/LA

Although there are some significant distinctions with particular traits in each phase as mentioned in the subsections below, the application of EDM/LA is a cycle application of the general knowledge discovery and data mining (KDD) process (Figure 3).

### 2.2. Educational environment

Different types of data can be gathered to address various educational issues depending on the type of educational environment (traditional classroom, computer-based, or blended learning) and the information system utilised, such as: LMS (Learning Management System), ITS (Intelligent Tutoring System), MOOC (Massive Open Online Course), etc (Romero & Ventura, 2013) [16].

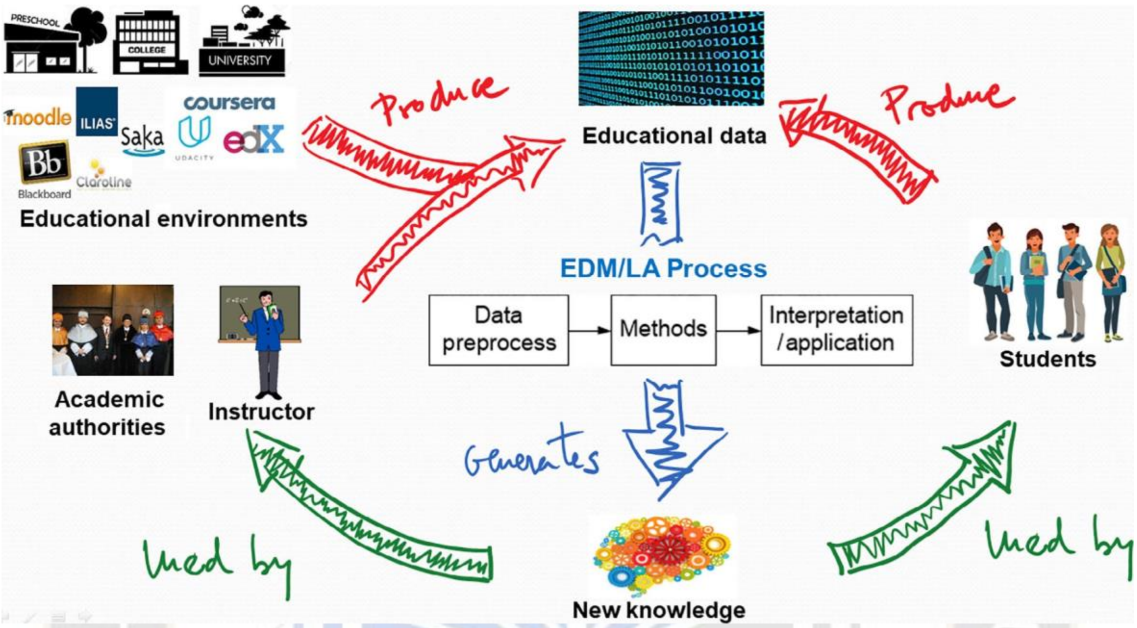


**FIGURE 3** Educational Data Mining/Learning Analytics knowledge discovery cycle process

### 2.3. Educational data

"(Romero, Romero, & Ventura, 2014) Educational data are gathered from a variety of sources, including interactions between teachers, students, and the educational system, administrative data (e.g., school and teacher information), demographic data (e.g., gender, age, etc.), student affectivity (e.g., motivation, emotional states), and so on. The amount of data that can be stored in educational environments is enormous, and it can come from a variety of sources, in a variety of formats, with varying degrees of granularity (from coarse to fine grain), or from a variety of levels of meaningful hierarchy (keystroke level, answer level [17], session level, student level, classroom level, and school level), which can provide more or less data (Figure 4). Since gathering and combining all of this raw data for mining are difficult undertakings in and of themselves, Pre-processing is essential.

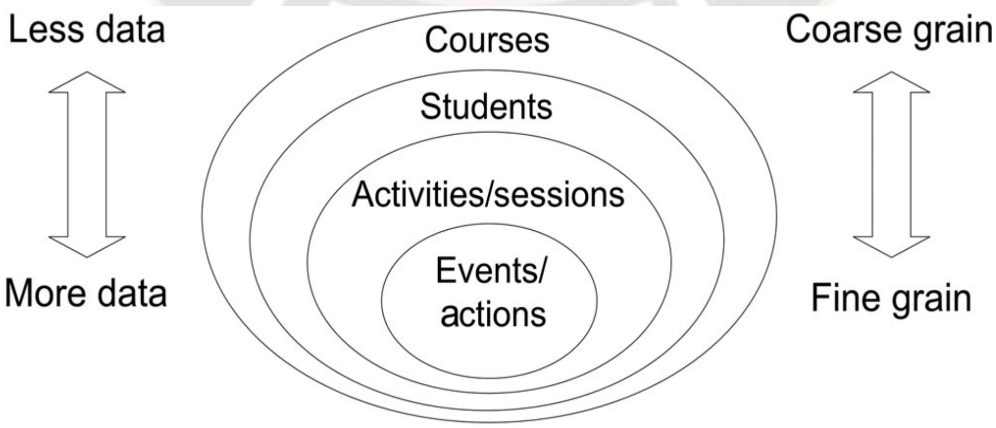


**FIGURE 4** Different levels of granularity and their relationship to the amount of data

*International Journal on Recent and Innovation Trends in Computing and Communication*
*ISSN: 2321-8169 Volume: 12 Issue: 1*
***Article Received:** 10 July 2023 **Revised:** 12 August 2023 **Accepted:** 28 August 2023*

### 2.4. Pre-processing

Data preparation is a challenging and complex activity, and it frequently consumes more than half of the time required to solve a data mining challenge (Bienkowski et al., 2012) [18][19]. The educational data (raw, original, or primary data) that are accessible to answer an issue are not in the right form (or abstraction). Consequently, it's essential to Transform the data into the proper form (updated data) in order to address each unique educational issue. This comprises selecting the data to be collected, concentrating on the questions to be addressed, and ensuring that the data match the questions with certain special difficulties, traditional preparation procedures are used with educational data (Romero et al., 2014) like the following ones. Feature engineering is the process of creating and choosing features or variables that provide data about the student success is crucial. Typically, we may summarise and translate all accessible information into a table for better understanding analysis [20]. In order to enhance, continuous attributes are often transformed/discretized into categorical attributes their understandability Finally, it's critical to uphold and safeguard the student's privacy by anonymizing removing any personally identifiable information from the data, including name, email, and phone number, that is not necessary for mining, so on. We can take into account using standards for ethical problems, data protection, and informed consent in this field of work etc. while utilising academic data (Pardo & Siemens, 2014) [21].

### 2.5. Methods and techniques

The bulk of conventional data mining methods, such as visualisation, classification, clustering, and association analysis methods, among others, have previously been effectively used in the educational sector (Baker, 2015). However, educational systems also have certain unique traits (hierarchical and longitudinal data) that call for a particular approach to the mining problem and data pre-processing. As outlined in the Methodologies section below, a variety of EDM and LA methods and strategies are employed to address various educational issues [22]."

### 2.6. Application of the new information and its interpretation

The final aim of any learning analytics process is action, and the outcomes of subsequent actions will indicate whether or not our analytical efforts were successful (Siemens, 2013). Therefore, in order to increase student learning performance, instructors and academic authorities must use the newly found knowledge obtained through the EDM/LA approaches to make interventions and decisions [23]. For the prior models produced by the EDM/LA method to be helpful for the decision-making process, it was crucial that they were understandable. Given that they are more accurate but less understandable than black-box DM models like neural networks, white-box DM models like decision trees are preferred in this context. Techniques for visualisation are also highly helpful for more easily understanding outcomes.

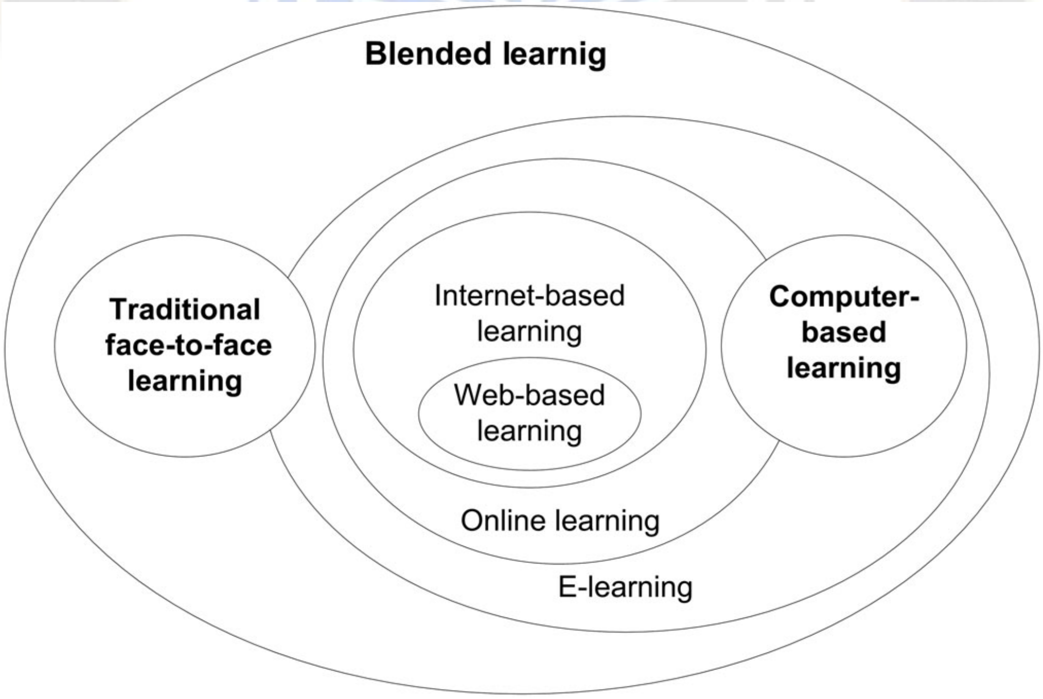


**FIGURE 5** Types of educational environments and systems

**TABLE 2** International levels of education

| Level | fundamental characteristics |
|---|---|
| Pre-primary education | Designed for children from age 3 to the start of primary education |
| Primary education or first stage of basic education | Normally starting between the ages of 5–7. |

***Article Received:** 10 July 2023 **Revised:** 12 August 2023 **Accepted:** 28 August 2023*

| | |
|---|---|
| Lower secondary education or second stage of basic education | Designed for children from ages 8–14 to complete basic education |
| Upper secondary education | More specialized education beginning at age 15 or 16 years. |
| Postsecondary nontertiary education | Programs that straddle the boundary between upper- and postsecondary education from an international point of view. |
| First stage of tertiary education | Tertiary programs having an educational content more advanced than those offered at previous le |
| Second stage of tertiary education | Tertiary programs leading to the award of an advanced research qualification, for example, Ph.D. |

## 3. EDUCATIONAL Area AND DATA

"There are several different types of educational settings, including traditional education, CBE, and BL (Figure 5). Each of them offers several data sources (Romero & Ventura, 2007) [24].

### 3.1. conventional face-to-face teaching

The most popular educational system is traditional education, or "back to the basics," which focuses mostly on face-to-face interaction between teachers and students through lectures, class discussions, small groups, individual seat work, and other activities. As shown in Table 3, traditional educational systems are categorised by UNESCO6 into several tiers. These systems collect data on things like student attendance, grades, curricular objectives, class information, and scheduling details[25]. Finally, it's crucial to remember that all of these conventional methods may employ computer-based educational technologies in addition to face-to-face instruction.

### 3.2. methods for teaching with computers

CBE refers to the use of computers in the classroom to manage or offer guidance for the student's instructions. Initially, CBE systems were straightforward standalone instructional programmes that operated on a local computer. There are many new web-based intelligent educational systems available today as a result of widespread Internet usage and the use of artificial intelligence (AI) technology. Table 3 lists and describes a few instances of modern computer-based educational systems [26]."

**TABLE 3Examples of computer-based educational systems**

| System | Description |
|---|---|
| Intelligent and Adaptive Hypermedia System (AIHS) | By creating a model of the objectives, preferences, and knowledge of every single student and applying this model during every encounter with that student, these systems make an effort to be more adaptable. Similar to ITS data, AIH data are recorded data. |
| System for Intelligent Tutoring (ITS) | By imitating student behaviour and altering its manner of contact with each student in accordance with that student's unique model, ITSs give students direct, tailored teaching or feedback. A domain model, student model, and pedagogical model are typically included. All student-tutor contact is recorded by ITSs (mouse clicks, typing, and speech). |
| Management of learning (LMS) | Administration, documenting, tracking, and reporting of training programmes, classroom and online events, e-learning programmes, and training content are all features provided by software suites that offer course delivery. They keep track of all student actions, including reading, writing, taking examinations, carrying out tasks in real life, and discussing happenings with classmates. |
| A very large online course (MOOC) | It alludes to a web-based course created to accommodate many of students. With no restrictions on enrolment, it may give educational materials online to anybody who wishes to enrol in a course. They keep the same data that LMS does. |
| Quiz and test system | These systems' principal objective is to assess students' understanding of one or more ideas or subjects by employing a sequence of questions, objects, and other prompts to elicit information from responders. They save a lot of data regarding the responses provided by pupils, computed grades, and statistics. |
| Other types | Virtual and augmented reality systems, concept maps, social networks, WIKIs, forums, wearable learning systems, learning object repositories, and so on. They save several kinds of data regarding their interactions with the kids. |

## 4. MATERIALS AND DATASETS

"Nowadays, a broad range of well-known general-purpose tools and frameworks, including RapidMiner, Weka, SPSS, Knime, Orange, Spark Lib, and others, may be utilised for the purposes of performing EDM and LA research (Slater, Joksimovic, Kovanovic, Baker, & Gasevic, 2017). However, these tools are difficult for educators to use since they need to choose a certain approach or algorithm to employ and must offer the right parameters beforehand in order to get effective findings or models. A certain level of competence is therefore required of the educators in order to identify the appropriate environments (Romero & Ventura, 2013) [27]. The usage of some of the particular EDM/LA software solutions available is a solution to this issue (Table 4). However, they can only use certain types of data. The majority of EDM/LA scholars often utilise their own data to address their own educational issues. Nevertheless, gathering and pre-processing educational data is a difficult and time-consuming undertaking (Romero et al., 2014) [28]. Using some of the public datasets that are now accessible for free online download is one more choice, we wish to draw attention to Data Shop (Koedinger et al., 2010), which is one of the first and largest datasets that also offers a tool for ITS research. Table 4 demonstrates that there are presently not many

*International Journal on Recent and Innovation Trends in Computing and Communication*
***ISSN: 2321-8169 Volume: 12 Issue: 1***
***Article Received:** 10 July 2023 **Revised:** 12 August 2023 **Accepted:** 28 August 2023*

publicly available datasets and that they do not represent all sorts of educational contexts (most of them are from e-learning systems). Therefore, we believe that the creation of a particular EDM/LA datasets repository, like to the generic UCI Machine Learning Repository, would be extremely helpful in the future. 7 It is crucial to keep in mind that these open data sets must be portable and adhere to the rules of data ethics, privacy, protection, and permission (Ferguson, Hoel, Scheffel, & Drachsler, 2016) [29]."

**Table 4.** data set description

| Name | URL | Description |
|---|---|---|
| DataShop | https://pslcdatashop.web.cmu.edu/ | It offers a collection of tools for analysis and reporting as well as a central repository for safeguarding and securing research material. |
| GISMO | http://gismo.sourceforge.net/ | Interactive graphic monitoring tool that gives educators a helpful visual representation of students' online course activity. |
| Inspire | https://moodle.org/plugins/tool_inspire | Machine learning backends are implemented in the Moodle Analytics API, which offers a descriptive and predictive analytics engine. |
| LOCO-Analyst | http://jelenajovanovic.net/LOCO-Analyst/ | a tool designed to give teachers feedback on the important elements of the learning process occurring in a virtual learning environment. |
| Meerkat-ED | http://www.reirab.com/MeerkatED | An instrument for monitoring student participation in a course assisted by computer-based collaborative learning resources |

### 4.1. Methods and applications

"With regard to EMD and LA (Baker & Inventado, 2014; Bakhshinategh et al., 2018; Romero & Ventura, 2013), there are several widely used techniques. The majority of these methods, including visualisation, prediction, clustering, outlier identification, connection mining, causal mining, social network analysis, process mining, and text mining, are widely accepted to be applicable to all forms of data mining. Others, such the reduction of data for human judgement, model-based discovery, knowledge tracing, and non-negative matrix factorization, are more prominent in the field of education [30]. However, there are a lot of potential goals or academic issues in EDM/LA, and this taxonomy does not include all of them. Actually, there are a lot more precise goals depending on who the end customer is. Although the learners and the instructors appear to be the only two primary categories of prospective users/stakeholders, there are really additional groups and purposes involved, as shown in Table 5 displays some of the most recent hot themes or more intriguing difficulties in the domain in order to provide more examples of the most potential uses of the EDM/LA."

**TABLE 5 lists some current uses or research areas of interest for the community of researchers studying Educational Data Mining and Learning Analytics (EDM/LA).**

| Topics of interest | Description | Source |
|---|---|---|
| Analyzing educational theories | To examine the potential integration of learning theories and learning analytics in educational research | (Wong et al., 2019) |
| Analyzing pedagogical strategies | To examine the use of EDM/LA approaches in the classroom and to assess their impact. | (Shen, Mostafavi, Barnes, & Chi, 2018) |
| Analyzing programming code | To employ EDM/LA methods for evaluating code from programming assignments, programming projects, and other sources. | (Li & Edwards, 2018) |
| Collaborative learning and teamwork group | Examine cooperative learning and forecast collaboration group performance | (Hernández-García, Acquila-Natale, Chaparro-Peláez, & Conde, 2018) |
| Curriculum mining/ analytics | In order to enhance curriculum development, programme quality, and other factors, it is necessary to assess programme structure, course grading, and administrative curricular data. | (Hilliger, Miranda, Celis, & Pérez-SanAgustín, 2019) |
| DashBoards and visual learning analytics | In order to investigate and comprehend pertinent user traces that have been gathered in (online) settings and to enhance (human) learning, a visualisation approach will be used. | (Millecamp, Broos, De Laet, & Verbert, 2019) |
| DashBoards and visual learning analytics | To implement multi-layered neural network architectures in the EDM/LA research field. | (Hernández-Blanco, Herrera-Flores, Tomás, & Navarro- Colorado, 2019) |
| Discovery causal relationships | To determine the causal connections between characteristics in a dataset for education. | (de Carvalho & Zarate, 2019) |
| Early warning systems | Predicting student performance and at-risk pupils as early as feasible will allow for early intervention to promote student achievement. | (Cano & Leonard, 2019) |
| Emotional learning analytics | To research how emotions impact learning and how important emotions are to learning. | (D'Mello, 2017) |
| Evaluating the efficacy of interventions | To assess the effectiveness of initiatives, data-driven student feedback, useful guidance, etc. | (Sonderlund, Hughes, & Smith, 2018) |

| | | |
|---|---|---|
| Feature engineering methods | Employing machine learning techniques, automatically create qualities or student traits | (Botelho, Baker, & Heffernan, 2019) |
| Game learning analytics | To incorporate data-mining and visualisation methods into serious gaming player interactions | (Alonso-Fernández, Calvo-Morata, Freire, Martínez-Ortiz,&Fernández-Manjón, 2019) |
| Interpretable and explanatory learnermodels | To create "white box" interpretable, informative, practical, and highly understandable leaner models. | (Rosé, McLaughlin, Liu, & Koedinger, 2019) |
| Learning foreign language | To improve foreign language acquisition by utilising EDM/LA approaches. | (Bravo-Agapito, Frances, & Seaone, 2019) |
| Measuring self-regulated learning | To assess students' self-regulated learning characteristics and behaviours using EDM/LA approaches. | (ElSayed, Caeiro- Rodríguez, MikicFonte, & Llamas-Nistal, 2019) |
| Multimodal learning analytics | Using machine learning with more cost-effective sensing technologies to deliver novel learning insights that occur across contexts | (Spikol et al., 2017) |
| Orchestrating learning analytics | To examine adoption, its practical applications, and other aspects of the current LA adoption activities in a classroom setting. | (Prieto, Rodríguez- Triana, Martínez-Maldonado, Dimitriadis, & Gaševi_c, 2019) |
| Providing personalized feedback | To automatically or partially produce customised feedback to aid in student learning | (Pardo, Jovanovic, Dawson, Gaševi_c, & Mirriahi, 2019) |
| Sentiment discovery | To automatically recognise the subjectivity, subjectivities, and underlying attitudes in learners and learning resources | (Han, Wu, Huang, Huang, & Zhao, 2019) |

## 4.2. Educational Data Mining

"To examine educational data primarily produced by students and teachers, EDM creates and implements statistical, machine-learning, and data-mining approaches. Their use could make it easier to examine how students learn by taking into account how they interact with their surroundings (Baker, Costa, Amorim, Magalhes, & Marinho, 2012). At first, sessions on intelligent tutoring systems and artificial intelligence in education were organised at conferences. Montreal hosted the inaugural International Conference on EDM in 2008 (Baker, Barnes & Beck, 2008). Since then, it has been held each year [31][32][33]. The literature in this area is wide and diverse. Bienkowski et al. (2012) give a widely referenced article in which they introduce EDM and LA as well as their foundations, difficulties in implementation, and areas of application. A special focus is placed on adaptive learning systems, which modify learning activities in response to model predictions. As far as we are concerned, three books—Romero & Ventura (2006), Romero, Ventura, Pechenizkiy, & Baker (2010), and Pea-Ayala—describe applications and techniques (2014). A survey with more than 300 references is presented by Romero and Ventura (2010). Applications of EDM techniques include a number of processes (Figure 6). The primary goal of the study and the necessary data are first determined as part of the design, which is then planned. The information is then retrieved from the relevant educational setting. Data often has to be pre-processed since it may come from several sources, have different formats, or be organised in different hierarchical levels. Applying EDM techniques results in models or patterns that must be understood [34]. The analysis is repeated after changing the teaching/learning process or the study design if the conclusions call for applying changes to the teaching/learning process or are not conclusive (because the problem has not been adequately addressed, the raw data are small or inappropriate, or the selected methods are not powerful enough)."

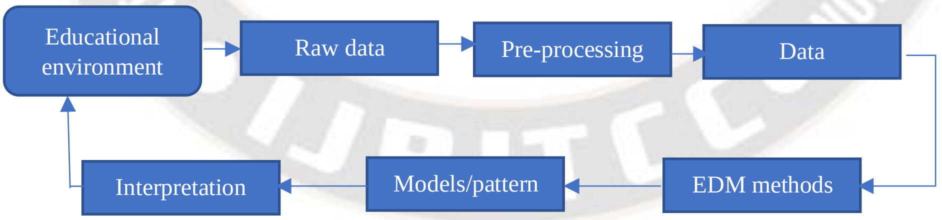


**Figure 6** shows an overview of how EDM techniques are used.

The number of EDM applications is rising. They may be divided into the following four groups, per Baker et al. (2012):[35]

- **Student modelling**: By modelling differences across students, a personalised learning process may be created using student data (such as knowledge, motivations, etc.) and EDM approaches.
- **Modelling of the domain's knowledge structure**: Techniques are developed for identifying data-based domain models that combine space-searching algorithms and psychometric modelling frameworks.
- **Pedagogical assistance**: Effective educational assistance may be found. Scientific research: apps can aid in the creation of new hypotheses as well as the development and testing of educational scientific ideas.
- Romero and Ventura (2013)[36] outline specific applications, including: student performance prediction, scientific inquiry, feedback for instructors, personalization and recommendation of learning materials to students, real-

*International Journal on Recent and Innovation Trends in Computing and Communication*
*ISSN: 2321-8169 Volume: 12 Issue: 1*
*Article Received: 10 July 2023 Revised: 12 August 2023 Accepted: 28 August 2023*

time alerting of stakeholders for disruptive student behaviour, student modelling (creating and fine-tuning cognitive models of students that represent their skills and declarative knowledge), domain modelling, student grouping and profiling, and more. Despite the fact that the majority of LMSs have their own tools that automatically provide customisable statistics reports on course development, they are frequently rather simple. For instance, Moodle (https://moodle.org/) enables the creation of several report kinds, including:

- logs for particular activities, students, items, and time periods;
- live logs, which include recent activity;
- activity reports, which show how many times each activity in a course has been viewed;
- course participation, which examines the behaviour of particular students for a particular time period and activity;
- completion data User activity in groups; user engagement in forums; and user activity. Google Analytics is a fascinating tool that is simple to use. It can offer data on demographics, average visit time, number of visits, pages visited, and so on. Lera-López, Faulin, Juan, and Cavaller (2009) discuss the capabilities offered by Sakai, WebCT/Blackboard, and Moodle for tracking student engagement and performance [37].

Blackboard (http://es.blackboard.com/sites/international/globalmaster/) also provides several types of reports, such as

a) user activity overview, which shows overall system and course activity for all students;
b) user statistics, which include the typical number of students and other users per month;
c) course activity overview, which displays overall system and course activity for each student.

## 4.3. Learning Analytics

"The First International Conference on Learning Analytics and Knowledge (LAK) calls for the measurement, collection, analysis, and reporting of data about learners and their contexts in order to comprehend and improve learning and the environments in which it takes place. (https://tekri.athabascau.ca/analytics/) In 2011, also in Canada, the inaugural International Conference on LAK (Long, Siemens, Conole, & Gaevi, 2011) took place [38]. Since then, it has been hosted every year. In the same year, the Association for Learning Analytics Research (SoLAR) (http://www.solaresearch.org) was established, making it the most active professional society[58].

One of the primary LA literary contributions is the 2014 book by Larusson & White. It focuses on the most recent ideas, results, tactics, tools, and case studies. How to:

(a) improve student and faculty performance;
(b) increase students' understanding of the course material;
(c) identify and assist struggling students;
(d) increase grading accuracy;
(e) permit instructors to identify and build upon their own strengths; and
(f) promote more effective use of resources at the institutional level [39].

The fundamental procedures for testing a learning/teaching process-related hypothesis are the same as those described for EDM: an iterative process where data is extracted from an educational environment and pre-processed before applying computational/quantitative methods in order to assist stakeholders (instructors, course managers, etc.) in decision-making [40].

## 4.3. Common methods in EDM and LA

Both EDM and LA use the majority of techniques that are suitable to educational data. Prediction, grouping, and connection mining are the most well-liked. There is other more, though, that cover a variety of uses. Table 5 displays the techniques, their explanations, and a few illustrations. Using data from Romero & Ventura (2007) and Baker et al. (2008) as well as Barnes, [41] Desmarais, Romero, & Ventura (2009), Baker & Yacef (2009) examine the percentage of works that used each category of approaches from 1995 to 2005 and from 2008 to 2009. First-period papers mostly used prediction algorithms (28%) or relationship mining approaches (43%). Clustering (15%) and human judgement or exploratory data analysis (17%) were also common. contrasted with connection Mining dropped to fifth rank (9%), while prediction methods rose to first (42%) (only publications from 2008) [42][43]. The percentages of applying human judgement and clustering approaches (12% and 15%, respectively) did not significantly alter. Since no publication from the first era employed this strategy, discovery using models gained representation (19%). Item response theory, Bayesian nets, and Markov decision procedures are also significant (28%), which is noteworthy [57]."

## 4.4. Similarities and differences between EDM and LA

There is undoubtedly a large amount of overlap between the two research areas. However, the research does point to certain distinctions. EDM and LA have the same objective of raising educational quality through data analysis and information extraction for stakeholders. Representative businesses from other industries, including banking, healthcare, and industry, have already implemented statistical, machine-learning, and data-mining approaches to improve performance through judgments based on past data [44]. Despite the fact that EDM research began a few years earlier, these fields of study have become more popular in recent years (Figure 7) [45]. According to Johnson, Adams, and Cummins (2012), [56]it is anticipated that these areas will continue to grow owing to the potential advantages (for students, instructors, administrators, researchers, and society at large) and the relevance of these subjects. Siemens & Baker (2012) claim that there are five essential differences between EDM and LA.

Which are:

- Automated discovery is a topic of interest for EDM researchers, and using human judgement is a tool for that;

in LA, the goal is to leverage human judgement instead [46].

- EDM breaks down systems into their component parts and looks at how those parts interact with one another, while LA seeks to comprehend whole systems.
- Beginnings: While LA origins are connected to the film industry, EDM origins are founded in educational software and student modelling. result prediction, the semantic web, "intelligent curriculum," and systemic interventions. The aforementioned writers claim that these disparities indicate widespread tendencies in each group, and as a result [47],
- As a result, they do not specify the appropriate scopes. Baker & Inventado express a similar viewpoint (2014),
- "The overlap and distinctions between the communities are essentially organic, developing from," where it is stated rather than a more fundamental philosophical divide, the interests and values of certain scholars[55].
- According to Bienkowski et al. (2012), LA encompasses more disciplines than EDM. Along with computer science,[48]
- LA has connections to sociology, information science, psychology, statistics, and the learning sciences. Consequently, even
- If the line between the two disciplines is hazy and their distinctions are in part dependent on their historical developments and current tendencies, they are nonetheless important to these authors[54].

### 4.5. Search and screening strategy

"We did our data collecting and analysis across multiple steps (Figure. 7). First, we used the following indexes to look for papers about EDM in the Web of Science database: SSCI, A&HCI, CICI-S, SCI-EXPANDED, EDCL, CCR-EXPENDED, CPCI-SSH, BKCI-S, BKCI-SSH, and IC. The chain of Prior literature evaluations has advised using scientific data bases as a database that contains a selection of superior publications (Xia and Zhong, 2018; Fu and Hwang, 2018; Xie et al., 2019) [53]. Our search keywords were, Data mining in education (TOPIC: "Learning Analytics") THEN DOCUMENT Types: [49] (Article) and 2015–2019 as the time range filters. We used the same exact procedure as with LA articles, however "Learning Analytics" search keywords (TOPIC: ("Learning Analytics") AND Types of documents: (Article). The search phrases we selected were "Educational Data Mining" and "Learning Analytics" to collect articles for the respective fields since we wish to examine the overall similarities and differences of the two fields. Inserting other search terms would have narrowed down the search results as well as including articles from other related fields. [50]."

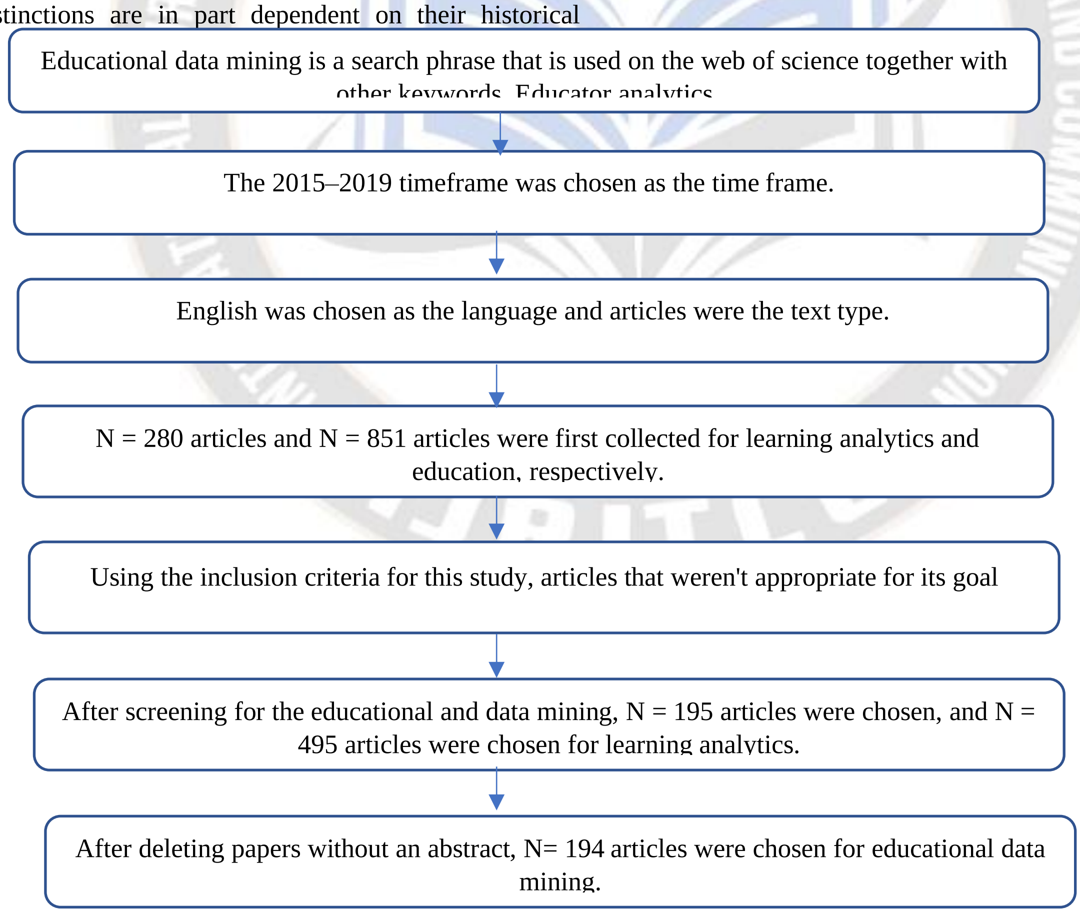


**Figure 7.** Flowchart for Study Selection for Educational Data Mining and Learning Analytics,

*International Journal on Recent and Innovation Trends in Computing and Communication*
*ISSN: 2321-8169 Volume: 12 Issue: 1*
***Article Received:** 10 July 2023 **Revised:** 12 August 2023 **Accepted:** 28 August 2023*

As some abstracts called for a detailed study of the full text to make sure the research met our inclusion criteria, we used the abstracts of the publications to weed out those that needed more examination and to eliminate those that did not. Second, we carried out a full-text screening to decide whether to include or omit the studies for those that needed more research (e.g., if a study evaluated a novel learning analytics model with empirical data or only provided a hypothetical model) [51][52]. 492 articles for LA and 194 for electronic dance music were gathered as a result of the filtering procedure."

## 5. ANALYSIS AND CONCLUSIONS

"We discovered that forecasting students' academic performance using educational data mining approaches such algorithms, prediction, and classification was the most frequent topic (Topic 1) for EDM publications.

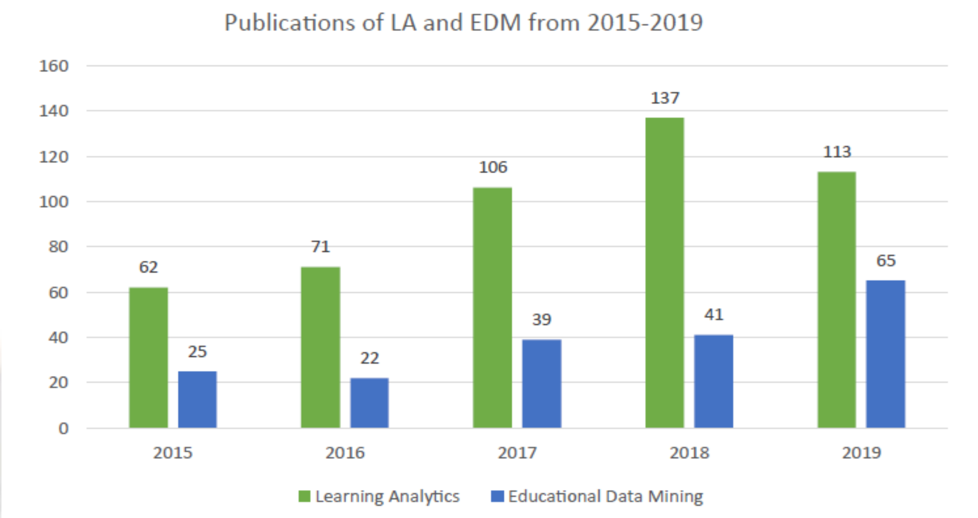


**Figure. 9.** Publications of LA and EDM from 2015 to 2019.

By doing independent topic modelling runs for the abstracts from each year, we also looked at how the study subjects in the LA and EDM literature evolve throughout the five-year period from 2015 to 2019. For the abstracts of the LA papers published in each year from 2015 to 2019 as well as the EDM articles, we specifically performed a topic modelling analysis. The topic modelling findings are included in Appendix A and indicate the research trends of each subject over a five-year period as well as how the two disciplines differ over time. Since these terms consistently rank among the top five subjects in the LA literature each year, it is clear that the field consistently places a high priority on student outcomes and performance in a course.

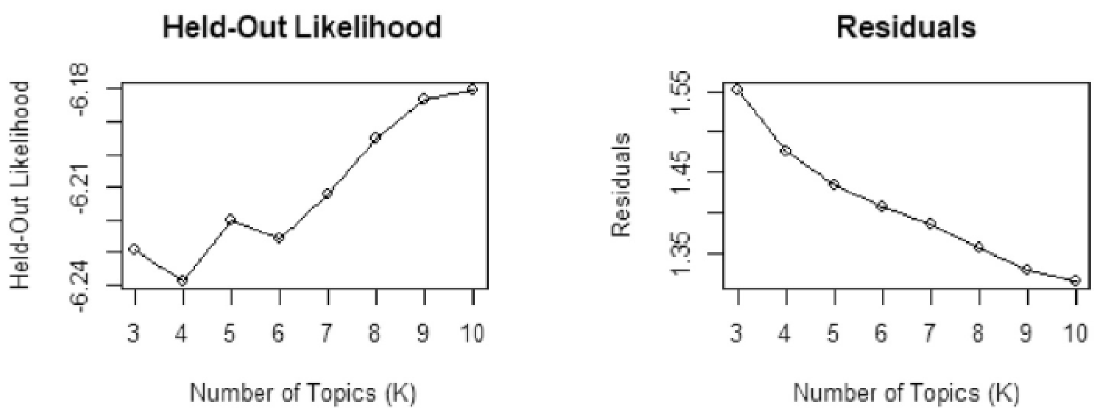


**Figure. 10.** Diagnostic values by number of topics for learning analytics.

Topics on the social and collaborative dimensions of learning were covered in some of the years. The social component of MOOCs was one of the most popular subjects in 2016, and social learning online was one of the most popular topics in 2019. Contrary to past LA literary subjects, which did not include issues concerning tools, the top topic of 2019 was about data analytics tools. Each year's top five EDM literary subjects clearly show how it differs from and relates to LA

*International Journal on Recent and Innovation Trends in Computing and Communication*
*ISSN: 2321-8169 Volume: 12 Issue: 1*
***Article Received:** 10 July 2023 **Revised:** 12 August 2023 **Accepted:** 28 August 2023*

literature. Themes on student performance and results in a course like the LA literature were frequently among the most popular topics in the EDM literature. The EDM literature did not emphasise social issues, in contrast to LA's most popular subjects.

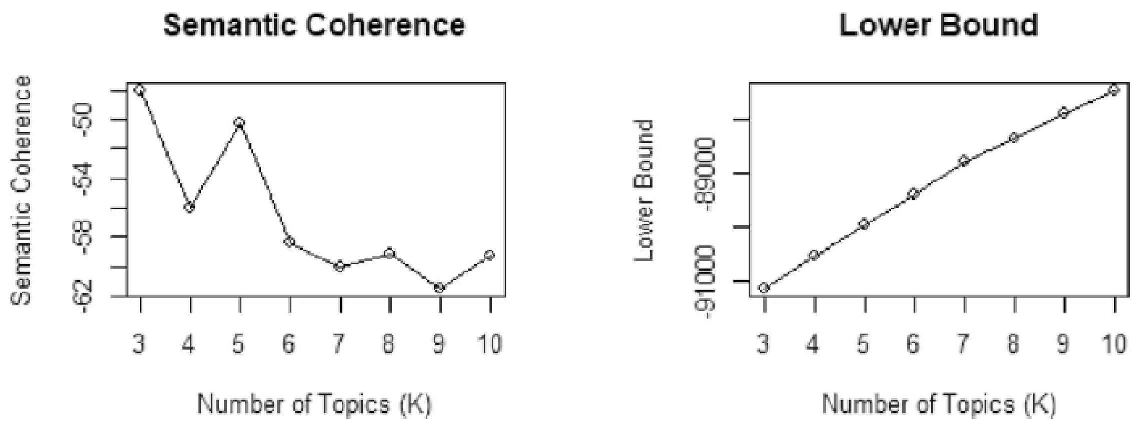


**Figure. 11.** Diagnostic values by number of topics for educational data mining."

**5.1. Discussion**

"The differences between the two disciplines' papers' themes are different to varying degrees rather than in fundamental ways. The emphasis in both sectors was on modelling student conduct as well as learning platforms and student achievement. LA newspapers emphasised social media, educational resources, and student interaction more Although EDM publications were primarily concerned with methodology, network analysis and data analysis techniques. This is consistent with the distinctions noted based on the material currently available, LA focuses more on the procedures of EDM and individual learning are more concerned with knowledge discovery. using several data sources (Papamitsiou and Economides, 2014). As with most. The evaluations note that the study environments and topics of interest are the two are becoming more similar (Aldowah et al., 2019; Papamitsiou and Economides, 2014; Ventura and Romero, 2009). This is demonstrated by the subjects that cross across. The top search terms for LA are "online," "social," whereas for EDM they are performance, prediction, and student collaboration machine, classification, machine analysis, and student Since this tends to the pragmatic distinction made by Sahin and Yurdugul (2019), according to which LA.

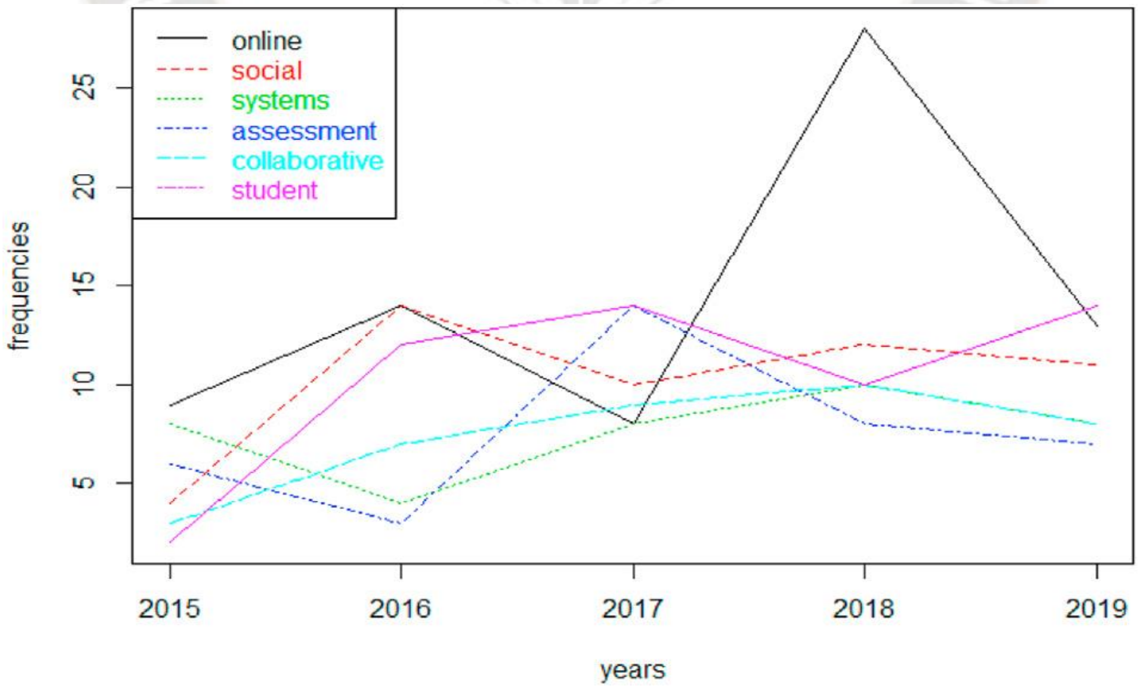


**Figure 12.** Trend of frequently used LA keywords over the 5-year period in the LA articles.

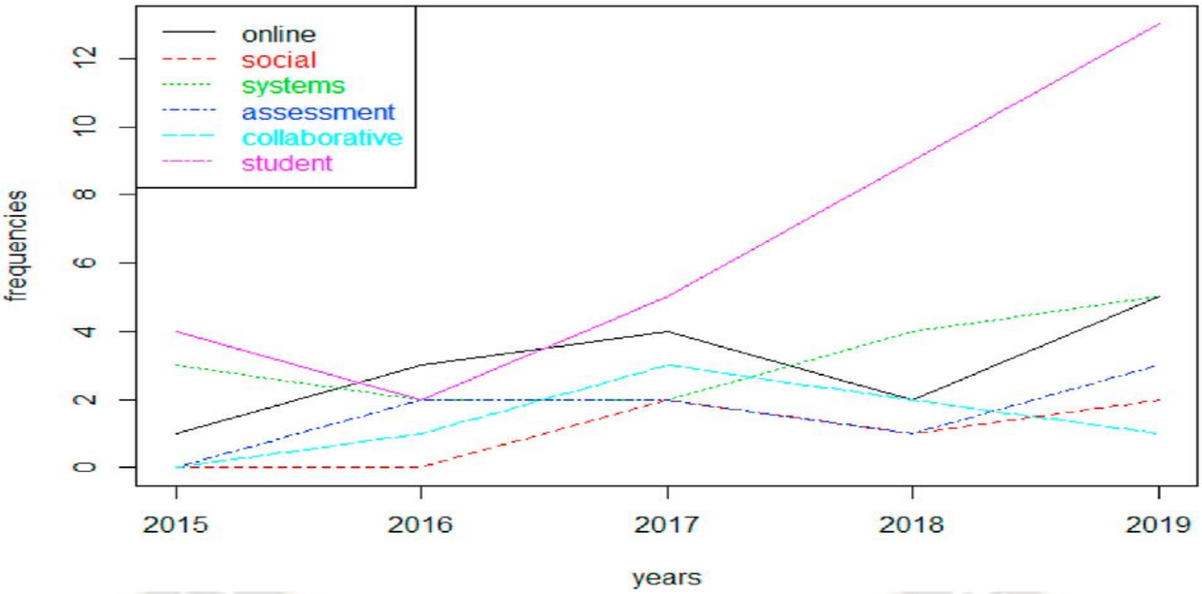


**Figure. 13.** Trend of frequently used LA keywords over the 5-year period in the EDM articles.

By examining feedback, evaluation, and results, LA and EDM both place a strong emphasis on students in a learning environment or during activities. It is essential to both areas to look at the outcomes and comments of students and teachers in various settings. A complex decision that takes time and effort to make is when and how to provide feedback (Leitner et al., 2019). Additionally, when student data is collected, instructors and other stakeholders may have ethical and privacy issues about LA and EDM and be reluctant to implement these technologies in learning environments (Leitner et al., 2019). Adopting LA and EDM models requires including instructors in the feedback and evaluation process."

## 6. CONCLUSION

"With the same goal of enhancing learning beginning with data, EDM and LA are two multidisciplinary groups comprising computer scientists, learning scientists, psychometricians, and researchers from other fields. In the past two decades, this field has rapidly expanded with the emergence of two distinct annual conferences (EDM and LAK), two specialised publications, and (JEDM and JLA), as well as an expanding body of books, papers, surveys, and reviews. Furthermore, they are an active for the purpose of adopting EDM/LA by educational institutions and schools globally, go from the lab to the general market and Analytics and data mining are anticipated to be used in all education research by the year 2020. (Baker and Inventado, 2014). All of this suggests that EDM/LA will soon develop into a mature field that will be extensively used by people other than scholars. but also, by lecturers, academic leaders, and associated businesses from throughout the globe. LA EDM has influenced how we think about learning and generated discoveries that have been applied to common practise or added to the hypothesis. The study emphasis and analytical complexity of this field of inquiry have advanced, although the influence on theory, frameworks, and practise has been less significant (Dawson, Joksimovic, Poquet, & Siemens, 2019). Therefore, encouragement from current research groups (such SoLAR and IEDM), funds, and resources are required. the active promotion of well-established works like the LACE initiative10, agencies, partnerships, and moving from exploratory models to more holistic, integrative systems-level approaches will boost the influence on practise research. Regarding potential trends in the field, our earlier study (Romero & Ventura, 2013) found two tendencies. One of them has not yet been fully accomplished, while the other still presents a barrier. The first was that in order for a larger and larger public to utilise EDM tools, more of them must be made freely available. There is presently a large selection of tools with specialised functions, as we can see in the section Tools and Datasets. However, in order to apply several tasks for solving various educational challenges from a single interface or tool, general purpose EDM/LA tools still need to be developed. Additionally, the mobility of the models produced by these tools has to be improved."

## 7. REFERENCES

1. Al-Emran, M., Malik, S. I., & Al-Kabi, M. N. (2020). A survey of internet of things (IoT) in education: Opportunities and challenges. In
2. Toward social internet of things (SIoT): Enabling technologies, architectures and applications (pp. 197–209). Cham, Switzerland: Springer.
3. Alonso-Fernández, C., Calvo-Morata, A., Freire, M., Martínez-Ortiz, I., & Fernández-Manjón, B. (2019). Applications of data science to game

4. learning analytics data: A systematic literature review. Computers & Education, 141, 1–14. Arnold, K. E., & Pistilli, M. D. (2012).
5. Course signals at Purdue: Using learning analytics to increase student success. In Proceedings of the 2nd International conference on learning analytics and knowledge, Vancouver, Canada (pp. 267–270).
6. Baker, R. S. (2015). Big data and education (2nd ed.). New York, NY: Teachers College, Columbia University.
7. Baker, R. S. (2019). Challenges for the future of educational data mining: The Baker learning analytics prizes. Journal of Educational Data Mining, 11(1), 1–17.
8. Baker, R. S. J.d., & Yacef, K. (2009). The state of educational data mining in 2009: A review and future visions. Journal of Educational Data Mining, 1(1), 3–17.
9. Baker, R. S. J.d., & Inventado, P. S. (2014). Educational data mining and learning analytics. In J. A. Larusson & B. White (Eds.), Learning analytics: From research to practice. Berlin, Germany: Springer.
10. Bakhshinategh, B., Zaiane, O. R., ElAtia, S., & Ipperciel, D. (2018). Educational data mining applications and tasks: A survey of the last 10 years. Education and Information Technologies, 23(1), 537–553.
11. Baradwaj, B. K., & Pal, S. (2012). Mining educational data to analyze students' performance. Computer Science, 2(6), 63–69.
12. Bienkowski, M., Feng, M., & Means, B. (2012). Enhancing teaching and learning through educational data mining and learning analytics: An issue brief (pp. 1–57). Washington, DC: U.S. Department of Education, Office of Educational Technology.
13. Bogarín, A., Cerezo, R., & Romero, C. (2018). A survey on educational process mining. WIREs: Data Mining and Knowledge Discovery, 8(1), e1230.
14. Botelho, A. F., Baker, R. S., & Heffernan, N. T. (2019). Machine-learned or expert-engineered features? Exploring feature engineering methods in detectors of student behavior and affect. In The twelfth international conference on educational data mining, Montréal, Canada.
15. Bravo-Agapito, J., Frances, C., & Seaone, I. (2019). Data mining in foreign language learning. WIREs: Data Mining and Knowledge Discovery, 10(1), e1287.
16. Campbell, J. P., DeBlois, P. B., & Oblinger, D. G. (2007). Academic analytics: A new tool for a new era. Educause Review, 42(4), 40. Cano, A., & Leonard, J. (2019). Interpretable multi-view early warning system adapted to underrepresented student populations. IEEE Transactions on Learning Technologies, 12, 198–211.
17. Cerezo, R., Calderón, V., & Romero, C. (2019). A holographic mobile-based application for practicing pronunciation of basic English vocabulary for Spanish speaking children. International Journal of Human-Computer Studies, 124, 13–25.
18. Custer, S., King, E. M., Atinc, T. M., Read, L., & Sethi, T. (2018). Toward data-driven education systems: Insights into using information to measure results and manage change. Washington, DC: Center for Universal Education at the Brookings Institution.
19. Daniel, B. K. (2019). Big data and data science: A critical review of issues for educational research. British Journal of Educational Technology, 50(1), 101–113.
20. Datnow, A., & Hubbard, L. (2016). Teacher capacity for and beliefs about data-driven decision making: A literature review of international research. Journal of Educational Change, 17(1), 7–28.
21. Dawson, S., Gaševi_c, D., Siemens, G., & Joksimovic, S. (2014). Current state and future trends: A citation network analysis of the learning analytics area. In Proceedings of the fourth international conference on learning analytics and knowledge, Indiana, USA (pp. 231–240).
22. Dawson, S., Joksimovic, S., Poquet, O., & Siemens, G. (2019). Increasing the impact of learning analytics. In Proceedings of the 9th international conference on learning analytics & knowledge, Tempe, Arizona (pp. 446–455).
23. Dawson, S., Poquet, O., Colvin, C., Rogers, T., Pardo, A., & Gasevic, D. (2018). Rethinking learning analytics adoption through complexity leadership theory. In Proceedings of the 8th international conference on learning analytics and knowledge, Sydney, Australia (pp. 236–244).
24. de Carvalho, W. F., & Zarate, L. E. (2019). Causality relationship among attributes applied in an educational data set. In Proceedings of the 34th ACM/SIGAPP symposium on applied computing (pp. 1271–1277). Limassol, Cyprus: ACM.
25. Ding, M., Wang, Y., Hemberg, E., & O'Reilly, U. M. (2019). Transfer learning using representation learning in massive open online courses. In Proceedings of the 9th international conference on learning analytics & knowledge, Tempe, Arizona (pp. 145–154).
26. Ding, W., Jing, X., Yan, Z., & Yang, L. T. (2019). A survey on data fusion in internet of things: Towards secure and privacy-preserving fusion. Information Fusion, 51, 129–144.
27. D'Mello, S. (2017). Emotional learning analytics. In Handbook of learning analytics (p. 115). New York, NY: SOLAR.
28. Dormezil, S., Khoshgoftaar, T., & Robinson-Bryant, F. (2019). Differentiating between educational data mining and learning analytics: A bibliometric approach. LABBEC Workshop (Learning analytics: Building bridges between the Education and the Computing communities, pp. 1–6),
29. Montreal, Canada. ElSayed, A. A., Caeiro-Rodríguez, M., MikicFonte, F. A., & Llamas-Nistal, M. (2019). Research in learning analytics and educational data mining to measure self-regulated learning: A systematic review. In World conference on mobile and contextual learning, Delft, Netherlands (pp. 46–53).

30. Ferguson, R. (2012). Learning analytics: Drivers, developments and challenges. International Journal of Technology Enhanced Learning, 4 (5/6), 304–317.
31. Ferguson, R., Hoel, T., Scheffel, M., & Drachsler, H. (2016). Guest editorial: Ethics and privacy in learning analytics. SoLAR, 3(1), 5–15. Ferreira-Mello, R., André, M., Pinheiro, A., Costa, E., & Romero, C. (2019). Text mining in education. WIREs: Data Mining and Knowledge Discovery, 9(6), e1332.
32. Giannakos, M. N., Sharma, K., Pappas, I. O., Kostakos, V., & Velloso, E. (2019). Multimodal data as a means to understand the learning experience.
33. International Journal of Information Management, 48, 108–119. Greller, W., & Drachsler, H. (2012). Translating learning into numbers: A generic framework for learning analytics. Educational Technology &
34. Society, 15(42), 42–57. Han, Z., Wu, J., Huang, C., Huang, Q., & Zhao, M. (2019). A review on sentiment discovery and analysis of educational big-data. WIREs:
35. Data Mining and Knowledge Discovery, 10(1), e1328. Hernández-Blanco, A., Herrera-Flores, B., Tomás, D., & Navarro-Colorado, B. (2019). A systematic review of deep learning approaches to educational data mining. Complexity, 2019, 1–22.
36. Hernández-García, _A., Acquila-Natale, E., Chaparro-Peláez, J., & Conde, M. _A. (2018). Predicting teamwork group assessment using log data-based learning analytics. Computers in Human Behavior, 89, 373–384.
37. Hilliger, I., Miranda, C., Celis, S., & Pérez-SanAgustín, M. (2019). Evaluating usage of an analytics tool to support continuous curriculum improvement. In EC-TEL practitioner proceedings, Delf, Netherland (pp. 1–14). Delft, Netherlands.
38. Joksimovi_c, S., Kovanovi_c, V., & Dawson, S. (2019). The journey of learning analytics. HERDSA Review of Higher Education, 6, 27–63. Kloos, C. D., Catálan, C., Muñoz-Merino,
39. P. J., & Alario-Hoyos, C. (2018). Design of a conversational agent as an educational tool. In Learning with MOOCS (LWMOOCS), Madrid, Spain (pp. 27–30).
40. Koedinger, K. R., Baker, R. S. J. D., Cunningham, K., Skogsholm, A., Leber, B., & Stamper, J. (2010). A data repository for the EDM community:
41. The PSLC datashop. In C. Romero, S. Ventura, M. Pechenizkiy, & R. S. J. D. Baker (Eds.), Handbook of educational data mining. Boca Raton, FL: CRC Press.
42. Lang, C., Siemens, G., Wise, A., & Gasevic, D. (2017). Handbook of learning analytics. SOLAR, Society for Learning Analytics and Research.
43. New York, NY: SOLAR. Leitner, P., Ebner, M., & Ebner, M. (2019). Learning analytics challenges to overcome in higher education institutions. In Utilizing learning analytics to support study success (pp. 91–104). Cham, Switzerland: Springer.
44. Li, Z., & Edwards, S. (2018). Applying recent-performance factors analysis to explore student effort invested in programming assignments.
45. In Proceedings of the international conference on frontiers in education: Computer science and computer engineering (FECS)
46. (pp. 3–10). Liñán, L. C., & Pérez, _A. A. J. (2015). Educational data mining and learning analytics: Differences, similarities, and time evolution. International
47. Journal of Educational Technology in Higher Education, 12(3), 98–112. Millecamp, M., Broos, T., De Laet, T., & Verbert, K. (2019). DIY: Learning analytics dashboards. In Companion proceeding of the 9th international conference on learning analytics & knowledge (LAK'19), Tempe, Arizona. (pp. 947–954).
48. Newton, P., & Newton, L. (2019). When robots teach: Towards a code of practice. Research in Education, 1–7. Pardo, A., Jovanovic, J., Dawson, S., Gaševi_c, D., & Mirriahi, N. (2019). Using learning analytics to scale the provision of personalised feedback.
49. British Journal of Educational Technology, 50(1), 128–138. Pardo, A., & Siemens, G. (2014). Ethical and privacy principles for learning analytics. British Journal of Educational Technology, 45(3), 438–450.
50. Romero, C., & Ventura, S. (2010). Educational data mining: A review of the state-of-the-art. IEEE Transaction on Systems, Man, and Cybernetics, Part C: Applications and Reviews, 40(6), 601–618.
51. Romero, C., & Ventura, S. (2013). Data mining in education. WIREs: Data Mining and Knowledge Discovery, 3(1), 12–27.
52. Romero, C., & Ventura, S. (2017). Educational data science in massive open online courses. WIREs: Data Mining and Knowledge Discovery, 7 (1), e1187.
53. Romero, C., Ventura, S., Pechenizky, M., & Baker, R. (2010). Handbook of educational data mining. Data Mining and Knowledge Discovery Series. Boca Raton, FL: Editorial Chapman and Hall/CRC Press, Taylor & Francis Group.
54. Romero, C., Ventura, S., & Salcines, E. (2008). Data mining in course management systems: Moodle case study and tutorial. Computers & Education, 51(1), 368–384.
55. Rosé, C. P., McLaughlin, E. A., Liu, R., & Koedinger, K. R. (2019). Explanatory learner models: Why machine learning (alone) is not the answer. British Journal of Educational Technology, 50(6), 2943–2958.
56. Shen, S., Mostafavi, B., Barnes, T., & Chi, M. (2018). Exploring induced pedagogical strategies through a Markov decision process framework: Lessons learned. JEDM|. Journal of Educational Data Mining, 10(3), 27–68.
57. Siemens, G. (2013). Learning analytics: The emergence of a discipline. American Behavioral Scientist, 57(10), 1380–1400.
58. Slater, S., Joksimovi_c, S., Kovanovic, V., Baker, R. S., & Gasevic, D. (2017). Tools for educational data mining: A review. Journal of Educational and Behavioral Statistics, 42(1), 85–106.