



# A Systematic Review on Data Mining for Mathematics and Science Education

Dongjo Shin<sup>1</sup> · Jaekwoun Shim<sup>1</sup>

Received: 12 November 2019 / Accepted: 18 March 2020 / Published online: 14 May 2020  
© Ministry of Science and Technology, Taiwan 2020

## Abstract

Educational data mining is used to discover significant phenomena and resolve educational issues occurring in the context of teaching and learning. This study provides a systematic literature review of educational data mining in mathematics and science education. A total of 64 articles were reviewed in terms of the research topics and data mining techniques used. This review revealed that data mining in mathematics and science education has been commonly used to understand students' behavior and thinking process, identify factors affecting student achievements, and provide automated assessment of students' written work. Recently, researchers have tended to use such data mining techniques as text mining to develop learning systems for supporting teachers' instruction and students' learning. We also found that classification, text mining, and clustering are major data mining techniques researchers have used. Studies using data mining were more likely to be conducted in the field of science education than in the field of mathematics education. We discuss the main results of our review in comparison with the previous reviews of educational data mining (EDM) literature and with EDM studies conducted in the context of science and mathematics education. Finally, we provide implications for research and teaching and learning of science and mathematics and suggest potential research directions.

**Keywords** Educational data mining · Literature review · Mathematics education · Science education

---

**Electronic supplementary material** The online version of this article (<https://doi.org/10.1007/s10763-020-10085-7>) contains supplementary material, which is available to authorized users.

---

✉ Jaekwoun Shim  
[jaekwoun.shim@gmail.com](mailto:jaekwoun.shim@gmail.com)

<sup>1</sup> Gifted Education Center, Korea University, 315 Lyceum, 145 Anam-ro, Seongbuk-gu, Seoul, South Korea

## Introduction

Mathematics and science have long been perceived to be fundamental and essential subjects in the K-12 curriculum (McConney & Perry, 2010). Recently, since students' competencies in STEM (i.e. science, technology, engineering, and mathematics) have received considerable attention in many countries (English & King, 2019), mathematics and science are becoming much more important. Previous research, however, has shown that many students still have difficulties in learning mathematics and science and are likely to stop learning mathematics early in their schooling (e.g. Hodgen, Kuchemann, Brown, & Coe, 2009; Scarpello, 2007). In addition, although researchers largely agree that teachers should take into consideration students' in-the-moment thinking for effective and meaningful instruction (National Council of Teachers of Mathematics [NCTM], 2014; National Research Council [NRC], 2012), teachers have trouble building their instruction upon students' ideas and understanding (Jacobs, Lamb, & Philipp, 2010; Sánchez-Matamoros, Fernández, & Llinares, 2015). Therefore, there is a need for helping students and teachers deal with these issues in their learning and teaching.

Over the last few decades, the advancement of technological tools in education has contributed to the fields of mathematics and science education by addressing educational issues and assisting teachers and students in promoting their teaching and learning in effective and efficient ways. Educational data mining (EDM) can be considered an innovative tool or technique to provide new insights into teaching and learning. EDM is generally concerned with "developing, researching, and applying computerized methods to detect patterns in large collections of educational data that would otherwise be hard or impossible to analyze due to the enormous volume of data within which they exist" (Romero & Ventura, 2013, p. 12). Martin et al. (2015) argued that EDM techniques shed light on research in science education that has traditionally been difficult to perform (e.g. creating real-time assessment of students' science learning and uncovering the nature of student engagement and its long-term impact on students' choice of college major) as traditional data analysis methods were not equipped to cope with large amounts of unstructured data. Researchers have also used EDM approaches to develop an intelligent tutoring system that may support students' personalized learning of mathematics (Cai et al., 2019) and science and mathematics teachers' management of multiple groups (Schwarz et al., 2018; Tissenbaum & Slotta, 2019) by analyzing real-time data. In addition, a great volume of data collected by people (e.g. responses to the Third International Mathematics and Science Study [TIMSS] and the Program for International Student Assessment [PISA] questionnaires) and generated from online teaching and learning (e.g. massive open online courses [MOOCs]) have been analyzed to detect hidden but significant phenomena and information in mathematics and science education (e.g. Depren, 2018; Lee, 2019; She, Lin, & Huang, 2019).

With the recognition of the importance of EDM in the fields of education, researchers have conducted comprehensive review studies in various educational fields (e.g. Papamitsiou & Economides, 2014; Rodrigues, Isotani, & Zárate, 2018; Romero & Ventura, 2010; Saa, Al-Emran, & Shaalan, 2019; Shahiri, Husain, & Rashid, 2015). These studies, however, do not take into account the association between research topics and the EDM techniques to explore the topics. Rather, the previous studies

tended to categorize each article reviewed according to research topics and EDM techniques independently. In a review of EDM in the context of higher education, Aldowah, Al-Samarraie, and Fauzy (2019) considered the potential relationship between different educational problems and EDM techniques used to address the problems. However, Aldowah et al. did not include all journals on higher education to identify articles relevant to their review even though they used such different publication sources as Scopus, Web of Science, and Science Direct. Furthermore, none of the studies has investigated how EDM has been used to support teaching and learning of mathematics and science. Given the complexities of teaching and learning of science/mathematics and their corresponding educational issues, EDM techniques could supply meaningful support for teachers and students. Therefore, a systematic review of the EDM studies used in science and mathematics education might provide important insights into improving the quality of science/mathematics education.

## The Significance of the Study

Although science and mathematics are regarded as crucial subjects in school curricula around the world, not only do students struggle to learn fundamental concepts in mathematics and science, but teachers also have trouble teaching such concepts meaningfully. In recent years, a growing body of research has focused on EDM techniques to promote teaching and learning of science and mathematics (e.g. Cai et al., 2019; Martin et al., 2015). However, few literature reviews on this field have been systematically conducted to understand how such innovative techniques as EDM could be used to support science and mathematics education. To help researchers and teacher educators in science and mathematics education follow the trends in the EDM studies, integrate EDM techniques into their research, and design effective teaching and learning environments, it is important to review prior research studies on EDM comprehensively in the contexts of science and mathematics education.

## Previous Reviews on Educational Data Mining

Researchers have conducted comprehensive and systematic review studies on EDM from different perspectives. Cristobal Romero and Sebastian Ventura were two popular scholars who conducted comprehensive reviews on EDM techniques and their applications to educational systems. Reviewing EDM studies from 1995 to 2005, Romero and Ventura (2007) demonstrated that DM can be used in different educational systems (e.g. traditional classrooms, web-based courses, learning content management systems, and adaptive and intelligent web-based learning systems) and that DM techniques such as clustering, classification, outlier detection, association rule mining, sequential pattern mining, and text mining were usually used in educational research. In 2010, they extended their review by analyzing 300 EDM works up to the year 2009 (Romero & Ventura, 2010). They found that researchers frequently used EDM to provide feedback for supporting instructors, offer recommendations for students, predict student's performance, and analyze and visualize data. Developing a model of students' skills and knowledge, detecting undesirable student behaviors, and grouping students according to their personal characteristics and features

were other areas of interest that attracted many researchers' attention (Romero & Ventura, 2010). A review of empirical studies on EDM from 2008 to 2013 revealed that prior studies on EDM aimed to identify students' learning behaviors, predict students' performance, increase teachers' awareness of student learning, predict students' dropout and retention, improve feedback and assessment systems, and recommend educational resources from data pools (Papamitsiou & Economides, 2014).

Early reviews of EDM attempted to conduct a literature review covering EDM in general. More recent review studies on EDM have tended to focus on a particular educational context or a DM technique. Shahiri et al. (2015) and Saa et al. (2019) conducted systematic reviews on predicting students' academic performance and on identifying important factors (or attributes) affecting students' performance using EDM. Shahiri et al. found, perhaps not surprisingly, that students' cumulative grade point average is the most significant attribute to determine whether a student can complete a course. Internal assessments such as students' assignment grades, quiz and test scores, lab work, and attendance were also important attributes in predicting student performance. Two main EDM approaches to predict students' performance were classification and clustering (Saa et al., 2019).

The literature review of EDM by Rodrigues et al. (2018) focused on the process of evaluation in the context of e-learning. Rodrigues et al. identified main perspectives and trends in EDM: evaluation in traditional classrooms, evaluation of instructional actions of teachers, evaluation of administrative management, and evaluation of multimedia resources. In a comprehensive and systematic review of EDM in the context of students' learning in higher education, Aldowah et al. (2019) found that most of the EDM papers focused on computer-supported predictive analytics (e.g. predicting students' performance and drop/retention in a course) followed by computer-supported learning analytics (e.g. deriving information based on students' interactions). In terms of EDM techniques, classification and clustering were the two most commonly used approaches, which is in line with findings by Saa et al. (2019).

As observed, the previous literature reviews on EDM have provided insight into DM techniques/tools and their application to educational systems in general. Although some of them sought to highlight the application of EDM to specific educational contexts (e.g. predicting students' achievement and evaluating students' learning in e-learning), they did not consider how EDM has been applied to teaching and learning in such particular disciplines as mathematics and science. In addition, although the prior studies reviewed main topics of EDM and DM techniques systematically, they were unlikely to consider the association between the research topics and DM techniques. In other words, there seems to be a lack of evidence showing how EDM can shed light on solving educational problems in these disciplines. The aim of this study is to provide a review of the use of EDM in the context of mathematics and science education with the objective of helping researchers and educators address their educational issues using EDM.

## Methodology

The current study is a descriptive content analysis aimed at identifying and describing general issues and research trends in the integration of EDM into mathematics and science education (Çalık & Sözbilir, 2014). We used Kitchenham and Charters' (2007)

systematic literature review as a methodology to “summariz[e] all existing information about some phenomenon in a thorough and unbiased manner” (p. 7). Following their guidelines, our review consisted of three phases: a planning phase (specification of research questions and development of a review protocol), a conducting phase (research identification, study selection, study quality assessment, and data extraction), and a reporting phase.

## Research Questions

The research questions that guided this study were the following: (a) What are the topics of research studies that used EDM in mathematics and science education? (b) What types of DM techniques have been used to achieve the objectives in mathematics and science education?

## Identification of Research

The second step in this systematic literature review was to collect relevant research studies with regard to the research questions. With this objective, we determined the article pool and the key search strings. We used SCImago Journal Rank (SJR) as a starting point to identify appropriate and quality educational journals in mathematics and science. We selected “education” under the subject category, while other options were maintained as default (i.e. all subject areas, all regions/countries, and all types). This search resulted in 1222 journals. We then selected journals related to mathematics and science education. Journals related to computer education (e.g. *Computers & Education*) were also selected to identify research papers that used DM techniques in the teaching and learning of mathematics/science. Fifty-eight educational journals were chosen in total (17 mathematics education journals, 24 science education journals, 7 mathematics and science education journals, 10 computer education journals). A list of the journals considered is provided in the [Supplementary Material](#). Furthermore, we used Web of Science, a website which provides access to multiple databases in many different disciplines, to identify EDM articles related to mathematics and science education.

The search string we used in mathematics and/or science education journals was “data mining” OR “machine learning” OR “learning analytics,” while we used a more specific search string in computer education journals, [“data mining” OR “machine learning” OR “learning analytics”] AND [“mathematics” OR “science”], to find exclusively EDM papers related to mathematics and science. In the Web of Science, we used the search string [“data mining” OR “machine learning” OR “learning analytics”] AND [“mathematics” OR “science”] AND “education.” The period of the search was set from 2010 to 2019 (June). Any paper published after July 2019 was not included in this review. The initial stage of the search yielded 847 papers (630 from the SJR and 217 from the Web of Science).

## Study Selection

To select studies relevant to our research questions, we used the following inclusion and exclusion criteria, adapted from prior studies on EDM (e.g. Aldowah et al., 2019;

Papamitsiou & Economides, 2014). We did not restrict our study selection to journals published in a specific country (see Table 1).

Based on these criteria, the authors determined independently whether or not a paper would be included in the review. The inter-rater reliability between the authors was good (Cohen's kappa statistic 0.76). We discussed any discrepancies until consensus on paper selection was reached. Günel, Polat, and Kurt (2016), for example, used different DM techniques to extract learning concepts from calculus, abstract algebra, and computer science textbooks. Although Günel et al. used DM techniques to analyze mathematics textbooks at the college level, we decided to exclude this article from our analysis because it focused more on comparing different DM algorithms and their accuracy in detecting learning concepts than on providing implications for teaching and learning of mathematics. Filiz and Oz (2019) also compared different DM algorithms and their accuracy in predicting students' science performance, but we decided to include their article in our analysis because Filiz and Oz also used DM techniques to identify important factors influencing students' science performance. Finally, 64 articles were included in this review. Figure 1 shows the entire process of selecting the articles for this review. In the process of selecting studies for review, many articles were excluded. We randomly selected several articles that we excluded in the process of abstract review and found that these articles tended to include the search string (e.g. "data mining" and "education") somewhere in their text and references at least once but had nothing to do with our research questions.

## Data Extraction and Analysis

A data extraction form was designed to accurately record the information from the 64 articles we reviewed. We used an Excel spreadsheet to record the following information: authors, year, title, subject, data source, research objective, DM technique, results, and implications.

We began the analysis by categorizing each article according to research topic and DM technique using a coding scheme developed from Papamitsiou and Economides (2014). While coding the research topics and DM techniques, we noticed that some of the articles reviewed in the current study did not fit the coding scheme. For example, one of the codes used in Papamitsiou and Economides' review was prediction of dropout and retention. The current review, however, was not able to identify articles that used EDM to predict students' dropout and retention. Rather, many articles in mathematics and science education tended to use EDM to examine informative factors

**Table 1** Inclusion and exclusion criteria

Inclusion criteria	Exclusion criteria
<ul style="list-style-type: none"> <li>▪ DM or machine learning research</li> <li>▪ Empirical research</li> <li>▪ Research on teaching and learning mathematics or science</li> <li>▪ Research papers published in the last decade (i.e. from 2010 to 2019)</li> <li>▪ Research papers written in English</li> </ul>	<ul style="list-style-type: none"> <li>▪ Theoretical, conceptual, and review papers</li> <li>▪ Research papers focusing on data mining techniques without an educational perspective</li> <li>▪ Research papers that did not address any issues in mathematics or science education</li> <li>▪ Research papers whose full text was not accessible</li> </ul>

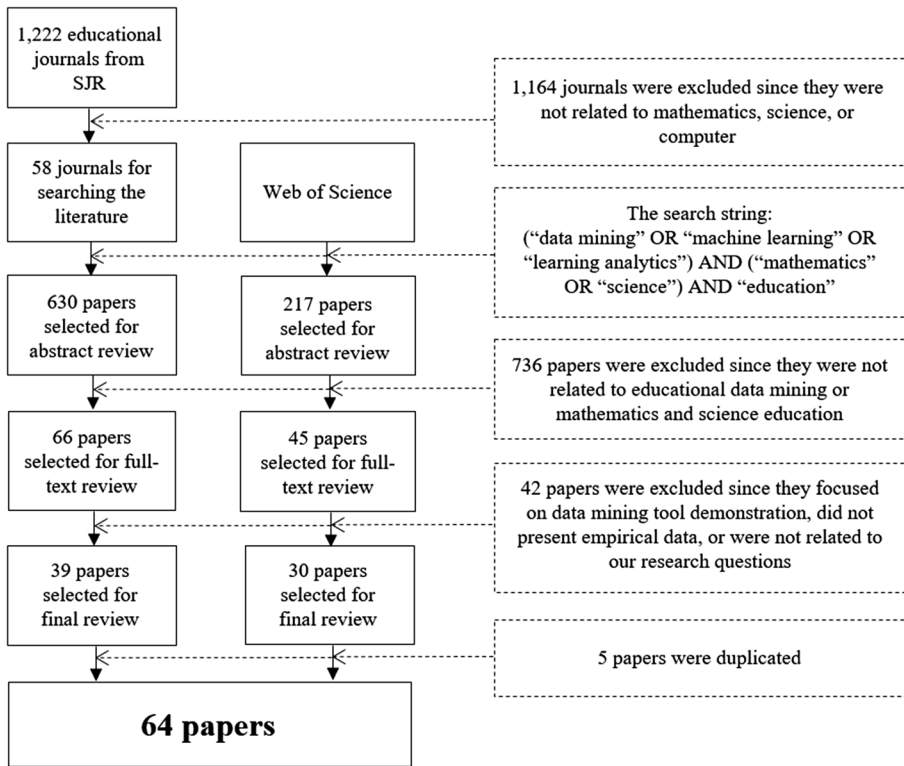


Fig. 1 Paper selection processes

in predicting students’ achievements in mathematics and science. Masci, Johnes, and Agasisti (2018), for instance, investigated student- and school-level factors affecting students’ PISA test scores in 9 different countries. Because we were unable to categorize Masci et al.’s article based on Papamitsiou and Economides’ coding scheme, we created a new code, *identification of factors affecting performance*, under the category of research topic.

Finally, we categorized the research topics for all articles into 7 major dimensions: *student modeling*, *automated assessment*, *identification of factors affecting performance*, *performance prediction*, *teacher support*, *student support*, and *document analysis*. We also classified the articles according to the DM techniques that the articles used. In this process, 5 different categories were found: *text mining*, *clustering*, *classification*, *social network analysis*, and *relationship mining*. We note that some articles were classified into more than one category due to some similarities between categories, especially when the articles addressed more than two research questions. As illustrated above, Filiz and Oz (2019) used DM methods not only to predict students’ science performance but also to identify factors affecting their performance; therefore, we assigned their article two codes: performance prediction and identification of factors affecting performance. Aiken, Henderson, and Caballero (2019) examined students’ course-taking patterns (coded as student modeling) as well as students’ academic and personal features affecting their academic success (coded as identification of factors



affecting performance). In these cases, we were unable to use a single code to categorize the articles.

## Results

This section summarizes and synthesizes literature on EDM conducted in the context of mathematics and science education. In general, previous studies were carried out much more frequently in the context of science (41 of 64 articles) than in that of mathematics (17 of 64 articles). Six studies were conducted in both mathematics and science (e.g. STEM). While 43 articles used DM algorithms to *analyze* educational data, 21 articles used DM techniques to *develop* intelligent tutoring systems.

### Research Topics

Table 2 shows that most of the reviewed papers focused on student modeling, followed by identification of factors affecting student performance and by automated assessment. A relatively small number of papers aimed to predict students' academic performance, support teachers' instruction, support students' learning, and analyze educational documents.

### Student Modeling

Using DM for modeling student behavior and thinking has drawn attention from many researchers. In particular, this topic was more likely to be conducted in science education than in mathematics education. The student modeling category was further divided into two sub-categories: modeling students' behavior and learning patterns and modeling students' thinking.

**Modeling Students' Behavior and Learning Patterns.** Researchers have sought to identify students' behaviors within online learning environments and to identify

**Table 2** Frequencies of articles reviewed according to research topic

Research topic	Science	Math	Both	Total
Student modeling	17	5	2	24
Identification of factors affecting performance	8	7	2	17
Automated assessment	14	1	0	15
Performance prediction	5	3	1	9
Teacher support	1	4	2	7
Student support	3	0	0	3
Document analysis	2	0	0	2
Total	50	20	7	77

Some articles were classified into more than one category. A list of references classified according to research topic is provided in the [Supplementary Material](#)



learning patterns, such as carelessness in scientific inquiry (Hershkovitz, de Baker, Gobert, Wixon, & Pedro, 2013), self-regulated learning (Kim, Yoon, Jo, & Branch, 2018), interaction among students in a MOOC (Tawfik et al., 2017), assignment submission (Lee, 2019), and interaction in online learning platforms (Hossain et al., 2018; Kinnebrew et al., 2016; Malmberg, Järvenoja, & Järvelä, 2013). Martin et al. (2015) engaged third- and fifth-grade students in an online fraction game in which students used repeated equipartitioning to create a particular combination of fractions (e.g. to create both  $1/6$  and  $1/9$ , a student may split the whole in thirds and then split one of the thirds into halves [i.e.  $1/6$ ] and another one of the thirds into thirds [i.e.  $1/9$ ]). Using a clustering algorithm, Martin et al. classified the students' splitting behaviors into several categories and examined how students in each category learned fractions over time. Other researchers have sought to identify undergraduate students' course-taking patterns (Aiken et al., 2019; Howard, Meehan, & Parnell, 2018; Wang, 2016).

**Modeling Students' Thinking.** DM has been used to investigate students' cognitive structures. Lamb and his colleagues used artificial neural networks to explore students' cognition rather than their content knowledge (Lamb, Annetta, Vallett, & Sadler, 2014; Lamb, Cavagnetto, & Akmal, 2016). Lamb et al. (2014) showed that students' cognitive processing of science-based tasks was non-linear and dynamic. The artificial neural networks were also useful to model the relationships between the cognitive attributes associated with game-based science tasks and develop a cognitive-attribute hierarchy (Lamb et al., 2016). Analyzing students' written responses, researchers also sought to model students' thinking about a specific science concept, such as the effects of a stop codon on DNA replication, transcription, and translation (Prevost, Smith, & Knight, 2016), global lunar patterns (Cheon, Lee, Smith, Song, & Kim, 2013), and the impact of a mutation in noncoding regions of a gene (Sieke, McIntosh, Steele, & Knight, 2019). Two studies focused on how students think about their teachers' teaching practices (Figueiredo, Esteves, Neves, & Vicente, 2016; Zhang, Qin, Jin, Deng, & Wu, 2017).

### Identification of Factors Affecting Performance

Seventeen articles explored the main factors (or attributes) influencing students' performance with the objective of providing high-quality education. This literature included similar numbers of mathematics and science education articles. Of the 17 articles, nine studies have used large volumes of data derived from the PISA or the TIMSS to identify important variables affecting student performance in mathematics and science (Chen, Zhang, Wei, & Hu, 2019; Depren, 2018; Depren, Aşkın, & Öz, 2017; Filiz & Oz, 2019; Gabriel, Signolet, & Westwell, 2018; Gorostiaga & Rojo-Álvarez, 2016; Liu & Whitford, 2011; Masci et al., 2018; She et al., 2019). There was only one study that examined informative factors in predicting prospective teachers' achievements. Using artificial neural networks, Akgün and Demir (2018) revealed that university placement scores were the most informative variable in predicting prospective elementary teachers' achievements in science and technology education courses. Other researchers sought to identify the factors that predicted students' attitudes towards mathematics

(Aksoy, Narli, & Idil, 2016), self-regulated learning (Kim et al., 2018), and their choice of undergraduate courses (Aiken et al., 2019; Suh, Upadhyaya, & Nadig, 2019).

### Automated Assessment

Although using open-ended questions may better reveal students' formal and informal ideas simultaneously than using multiple choice questions, resource and time limitations exist in assessing students' responses to open-ended questions (Sieke et al., 2019), especially when a teacher deals with a large number of students' written responses. To address these issues, researchers, particularly in science education, have used DM techniques to evaluate students' written explanations efficiently and accurately.

According to our review, Ha and Nehm were two scholars who contributed greatly to this field. Ha, Nehm, and their colleagues used computer-assisted scoring models such as SPSS Text Analysis and Summarization Integrated Development Environment (SIDE) to automatically evaluate students' written explanations of particular biology concepts (Beggrow, Ha, Nehm, Pearl, & Boone, 2014; Ha, Nehm, Urban-Lurain, & Merrill, 2011; Nehm, Ha, & Mayfield, 2012; Nehm & Haertig, 2012). These techniques require researchers (humans) to score a large number of student written answers first in order to train a machine (a computer). Ha et al. (2011) used the SIDE technique to develop an automated scoring model to evaluate students' written answers to questions about evolutionary change. Based on identification of key concepts, this model achieved satisfactory inter-rater agreement with human evaluators (kappa statistic larger than 0.8). In a similar study, Nehm et al. (2012) examined the impact of students' response length on assessment accuracy and noted that varying response lengths did not significantly influence scoring performance. DM was also used to detect students' cheating strategies in an e-learning environment (Northcutt, Ho, & Chuang, 2016). Although automated assessment is one of the interesting research topics in science education, it has not gotten much attention in mathematics education research.

### Performance Prediction

Academic performance in mathematics and science had a great influence on students' choice of major and career as well as educational policies and directions around the world. Some studies in this category focused on the accuracy of DM models to predict student performance (Abidi, Hussain, Xu, & Zhang, 2019; Depren et al., 2017; Filiz & Oz, 2019). Filiz and Oz (2019) demonstrated that students' science scores can be predicted with reasonable accuracy (more than 70%) based solely on students' responses to the TIMSS questionnaire which included questions about students' confidence in science and home educational resources. Other researchers focused on whether or not a DM technique can be used to make an early prediction of a student's failure in a course or graduation (Cooper & Pearson, 2012; Ismail & Abdulla, 2015).

### Teacher Support

Of the 7 research topics, teacher support was the only topic that received more attention in mathematics education. To help teachers develop their ability to respond to student thinking, Bywater, Chiu, Hong, and Sankaranarayanan (2019) developed an AI-based

tool, called the Teacher Responding Tool, to provide automatic recommendations for teachers to respond to students' written explanations. DM techniques also supported teachers in analyzing and reflecting on their own teaching practices (Duzhin & Gustafsson, 2018; Owens et al., 2017; Sergis et al., 2019). Two surveyed studies investigated how machine learning-based systems can help teachers orchestrate multiple groups in a collaborative setting (Schwarz et al., 2018; Tissenbaum & Slotta, 2019). Schwarz et al. (2018) used a system (System for Advancing Group Learning in Educational Technologies [SAGLET]) in which alerts are automatically sent to teachers as the system becomes aware of such critical moments as "idleness, off-topic talk, technical problems, explanation or challenge, confusion, correct solutions, and incorrect solutions" (p. 193) in online group learning. In this system, teachers observed multiple groups' learning and used the alerts about critical moments to guide the groups.

### Student Support

A small number of articles focused on the topic of student support, and the studies were likely to be conducted in the field of science education. Specifically, these studies involved the development of intelligent tutoring systems based on DM techniques to promote students' scientific argumentation (Huang et al., 2011; Lee et al., 2019). Lee et al. (2019) introduced a formative feedback system, HASbot, which uses the natural language processing (NLP) algorithm to assess students' written scientific arguments and provide immediate adaptive feedback. HASbot provides two types of feedback: (a) diagnostic (e.g. "You are scientific explanation level 1 because you did not use scientific evidence") and (b) suggestive (e.g. "Your argument will be stronger if you evaluate the strengths and weaknesses of the evidence from the model. What are you certain about from the groundwater model?") (Lee et al., 2019, p. 606). Another intelligent tutoring system was proposed by Rao and Saha (2019) to help students overcome learning difficulties in biology and facilitate students' independent learning. The system provided learning content in biology and assisted students by automatically identifying important biological concepts and questionable sentences, retrieving relevant images from websites for better understanding, conducting mock tests before actual examinations, and evaluating students' performance.

### Document Analysis

Two articles in our review used DM methods to analyze existing learning documents. Like the topic of student support, research on document analysis was limited to the area of science education. Wahlberg and Gericke (2018) conducted textbook analysis using DM techniques. Since protein synthesis tends to be included in both biology and chemistry upper secondary curricula, Wahlberg and Gericke used text summarization and text mining techniques in order to examine how the same concept is described in the two different curricula. Reitsma, Marshall, and Chart (2012) used intermediary-based standard cross-walking and the NLP technique to automatically align a substantial body of learning resources with varying educational standards in the USA.

**Table 3** Frequencies of articles reviewed according to data mining technique

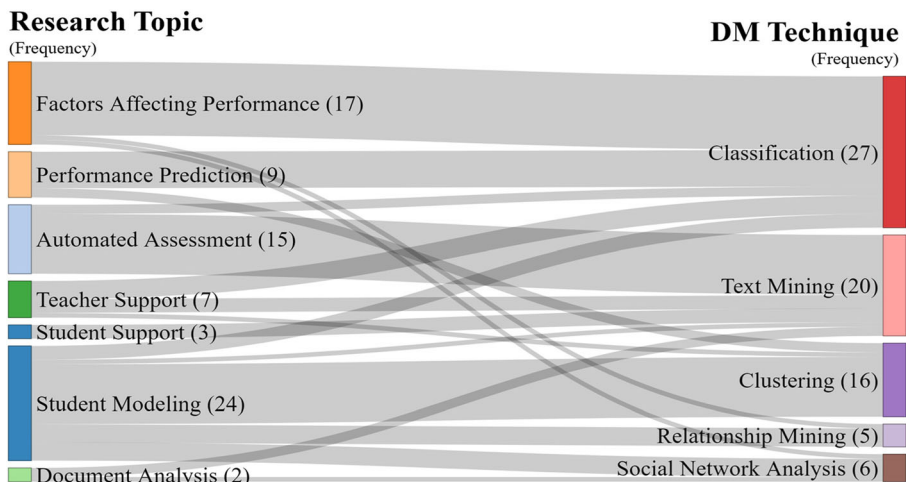
Research topic	Science	Math	Both	Total
Classification	12	11	4	27
Text mining	18	2	0	20
Clustering	9	5	2	16
Social network analysis	4	1	1	6
Relationship mining	3	0	2	5
Total	46	19	9	74

Some articles were classified into more than one category. A list of references classified according to DM technique used is provided in the [Supplementary Material](#)

### Data Mining Techniques

We classified the reviewed articles into 5 DM techniques (see Table 3). Classification was the most frequently used DM method, followed by text mining and by clustering. Whereas similar numbers of articles used classification algorithms to explore topics related to mathematics and science education, there was a difference in the number of articles that used text mining in each of the two disciplines. Furthermore, only a few articles used such DM methods as social network analysis and relationship mining to address educational issues in mathematics education.

We also found that different DM techniques were used to examine various research topics in mathematics and science education. In Fig. 2, the thickness of each link illustrates the strength of the relationship between the research topic and the DM technique used. Classification techniques were mostly used to identify factors affecting students’ performance and to predict students’ performance. Text mining was used to examine various research topics, especially focusing on automated assessment. Researchers tended to use clustering, social network analysis, and relationship mining methods to model students’ behavior and learning patterns. In the subsequent sections,



**Fig. 2** Relationships between research topics and data mining techniques

we briefly introduce the 5 DM techniques and describe how each technique was used to explore different research topics.

**Classification.** Classification algorithms aim to assign items in a data set to predefined classes. These algorithms can be used for “predicting student performance, achievement, knowledge, predicting/preventing student dropout, [and] detecting problematic student’s behavior in online courses/e-learning” (Aldowah et al., 2019, p. 23). In our review, some researchers have compared the classification accuracies of different algorithms to demonstrate the effectiveness of the algorithms (Abidi et al., 2019; Depren, 2018; Depren et al., 2017; Filiz & Oz, 2019; Gorostiaga & Rojo-Álvarez, 2016). Others have used a specific classification algorithm to identify important factors and predict student performance. White-box DM models such as tree-based algorithms (Aksoy et al., 2016; Figueiredo et al., 2016; Gabriel et al., 2018; Gobert, Kim, Sao Pedro, Kennedy, & Betts, 2015; Goggins, Xing, Chen, Chen, & Wadholm, 2015; Hershkovitz et al., 2013; Kirby & Dempster, 2015; Lavie Alon & Tal, 2015; Liu & Whitford, 2011; Masci et al., 2018; Owens et al., 2017; She et al., 2019; Wang, 2016) were preferable to such black-box models as random forest (e.g. Aiken et al., 2019; Kim et al., 2018) and artificial neural networks (e.g. Akgün & Demir, 2018; Cooper & Pearson, 2012; Roberts, Chung, & Parks, 2016), which were less comprehensible (Romero & Ventura, 2013).

**Text Mining.** Text mining is a method of identifying and extracting important information from a text (Aldowah et al., 2019; Romero & Ventura, 2013). In the current review, researchers have used text mining techniques to develop automated assessment systems (Begrow et al., 2014; Ha et al., 2011; Ha & Nehm, 2016; Liu, Rios, Heilman, Gerard, & Linn, 2016; Nehm & Haertig, 2012; Nehm et al., 2012; Northcutt et al., 2016; Prevost et al., 2016; Sieke et al., 2019; Wiley et al., 2017), assist teachers in noticing students’ mathematical thinking (Bywater et al., 2019) and in managing group learning (Schwarz et al., 2018; Tissenbaum & Slotta, 2019), help students improve their scientific arguments (Huang et al., 2011; Lee et al., 2019), and facilitate students’ self-regulated learning (Rao & Saha, 2019). Based on these observations, we noted that researchers in science education were much more likely to use text mining techniques in order to *develop* an intelligent system and evaluate its effectiveness than to use them to *analyze* large volumes of text data. There were few studies that used text mining to analyze text data and identify useful information. Wahlberg and Gericke (2018), for example, extracted keywords from science textbooks and compared them to examine how a science concept was described in the different textbooks.

**Clustering.** Clustering algorithms aim at “collecting similar objects together to form a group or cluster. Each cluster contains objects that are similar to each other but dissimilar to the objects of other groups” (Dutt, Ismail, & Herawan, 2017, p. 15991). In our review, K-means clustering (Figueiredo et al., 2016; Malmberg et al., 2013; Sergis et al., 2019; Zhang et al., 2017) and hierarchical clustering (Hossain et al., 2018; Lee, 2019; Liu & Lee, 2013; Martin et al., 2015) were the two most commonly used clustering techniques. In the context of mathematics and science education, clustering was used to identify patterns of undergraduate students’ resource choices (e.g. online videos and/or live lectures; Howard et al., 2018), identify the ways students solved homework and quiz problems in a physics

MOOC (Lee, 2019), classify students' responses to pretest questions before an instructional intervention (Magana et al., 2019), identify high- and low-achieving students' learning strategies in favorable or challenging learning situations (Malmberg et al., 2013), and detect students' behavior patterns in online learning environments (Araya et al., 2014; Kim et al., 2018; Martin et al., 2015). In addition, clustering was used to predict whether an undergraduate student was likely to graduate high school with a high or low cumulative GPA (Ismail & Abdulla, 2015).

**Social Network Analysis.** Social network analysis aims to “understand and measure the relationships between entities in networked information” (Romero & Ventura, 2013, p. 21). Based on network theory and graph theory, it characterizes the relationships in terms of nodes (e.g. concepts, keywords, individual actors) and links or edges (i.e. relationships between nodes). Visualization of networked structures helps researchers and instructors understand or discover hidden patterns or relationships in the context of teaching and learning (Aldowah et al., 2019). In the reviewed articles, researchers used social network analysis to model student thinking (Choi, Lim, & Son, 2017; Prevost et al., 2016; Sieke et al., 2019), identify important factors influencing students' attitudes towards mathematics (Aksoy et al., 2016), and analyze textbooks (Wahlberg & Gericke, 2018).

To examine students' recognition of mathematics and science, Choi et al. (2017) asked Korean middle graders to write down three keywords when they thought of mathematics and science. Results indicated that Korean middle graders associated science mainly with “experiment,” “chemistry,” “physics,” “scientist,” and “biology,” while mathematics was associated with “calculation,” “function,” “numbers,” “difficulty,” “equation,” “formula,” etc. Sieke et al. (2019) also used social network analysis to characterize students' thinking about the impact of mutation in a noncoding DNA region. Sieke et al. classified students' responses to question regarding noncoding mutation into 8 different categories and examined how the categories were related in a single answer.

**Relationship Mining.** Relationship mining is used to identify relationships between variables and discover rules to explain the relationships. Association rule mining is a commonly used method. Relationship mining also seeks to explore temporal associations, linear correlations, and causal relationships between variables (Romero & Ventura, 2013). Three of five articles classified as relationship mining in our review used association rule mining methods to model students' behavior and ideas. Chen and Chang (2017) used a fuzzy association rule mining technique to identify significant rules for explaining the relationships between grades in particular undergraduate courses and accumulated GPA. In a study of students' construction of concept maps in a biology context, Liu and Lee (2013) used an association rule mining algorithm, a priori, and identified 8 significant association rules that demonstrated the relationships between students' knowledge structures in biology.

## Conclusions and Discussion

The current review demonstrated that modeling students' behavior and thinking was the most preferred application of EDM in mathematics and science education, which is



in line with the findings of Papamitsiou and Economides (2014) and Peña-Ayala (2014). The higher frequency of articles dealing with the topic of student modeling may stem from the advancement of technology that captures a large amount of educational data in online learning environments (Lee, 2019; Levy & Wilensky, 2011). For example, log file data collected in online platforms (e.g. MOOCs and learning management systems) provide researchers with detailed information about what students have done and what they are doing (Lee, 2019). And, DM techniques have the potential to identify students' learning patterns based on the log file data. In the current review, we found that clustering was the most frequently used DM technique to model students' behavior and thinking, perhaps because of its effectiveness in identifying hidden patterns based on similarities and differences in the values of data.

We also found that relatively few studies in mathematics and science education focused on using DM techniques to predict student performance. This result may be surprising because previous reviews of EDM have consistently provided evidence that researchers have largely used DM methods to analyze students' performance and retention (e.g. Aldowah et al., 2019; Papamitsiou & Economides, 2014; Peña-Ayala, 2014; Romero & Ventura, 2010). The current review, however, revealed that more attention was paid to identifying important factors influencing students' mathematics or science performance than predicting their performance, perhaps because researchers in these areas focus more on the development of student performance and effective teaching and learning based on the important factors identified (Chen & Chang, 2017; Masci et al., 2018). Nonetheless, since mathematics and science have hierarchical structures, it is also important for educators to predict in advance whether a student will fail a mathematics or science class and provide an appropriate treatment as early as possible (Cooper & Pearson, 2012). DM techniques (e.g. classification algorithms) can be used to predict students' success, failure, or overall academic achievement in mathematics and science. Data from TIMSS and PISA can be particularly useful in that these programs have been collecting huge volumes of data on mathematics and science from many countries (e.g. Depren, 2018; Filiz & Oz, 2019). Students' log data in online learning systems can also be analyzed to predict students' performance and course completion as early in the course as possible (e.g. Gobert et al., 2015).

As seen in our review and the previous reviews of EDM, most of the studies were much more likely to concentrate on students' cognitive performance than on their affective features. Although Romero and Ventura (2013) argued that detecting and tracking students' emotions, motivations, and interests is a topic of interest to the EDM research community, we found that very little research in science and mathematics education has been conducted on how EDM can be used to analyze students' affective factors (e.g. Aksoy et al., 2016). Since affective domains significantly affect students' behavior and performance (Pantziara & Philippou, 2015), future research in science and mathematics education might examine how EDM can be used to detect useful information on students' affective factors, explore the relationships between affective constructs and cognitive domains, and assist students in developing a positive attitude (e.g. Aksoy et al., 2016).

Our review showed that applying DM techniques to automated assessment was the second most frequently addressed research topic, but only one article (Goggins et al., 2015) in mathematics education focused on automated assessment. A low number of



studies on automated assessment in mathematics education can be attributed to the complexity of developing such systems in mathematics. In our review, we noticed that automated assessment was more likely to involve text mining than other DM techniques (see Fig. 2). However, mining mathematical texts may not be straightforward because mathematics, in its nature, consists of many symbols and graphs. Automated assessment has many advantages in teaching and learning, such as rapid assessment of hundreds of students' written responses, immediate feedback and diagnosis of students' learning deficiencies, and consistent and unbiased assessment (Nehm & Haertig, 2012; Wiley et al., 2017). Future research in mathematics education might need to examine how students' written responses including mathematical symbols can be collected for analysis using DM and how automated assessment can be integrated into mathematics education to support teaching and learning of mathematics.

The low frequency of EDM studies under the categories of teacher support and student support may result from the fact that researchers paid more attention to analyzing enormous volumes of educational data using DM methods than developing DM-based systems to support teachers and students. In fact, there were 43 papers that used DM to analyze educational data and 21 papers that used DM to develop DM-based systems. Although a small number of studies focused on such topics as teacher support and student support, there was a difference in these topics between science and mathematics education. Whereas the topic of teacher support was more likely to be studied in the context of mathematics education, EDM research on student support was only performed in the area of science education. Based on the findings, two potential research directions in mathematics and science education can be suggested.

First, over the last decade, a growing body of research in mathematics education has investigated teachers' ability to notice and respond to in-the-moment student thinking (Jacobs et al., 2010; Kilic, 2018; Sánchez-Matamoros et al., 2015). Recent reforms in science education also support a vision of teaching and learning in which teachers pay attention to students' scientific ideas and build their teaching practices upon student thinking (NGSS Lead States, 2013; NRC, 2012). Barnhart and van Es (2015) found that preservice science teachers had difficulties in noticing and responding to students' thinking without appropriate support from teacher educators. Bywater et al. (2019) demonstrated that a DM-based tool assisted mathematics teachers in developing their expertise in noticing and responding to nuances in students' ideas and written explanations. Using the NLP technique, the tool provides teachers with three automatic recommendations, and the teachers can choose, edit, and/or combine any of the recommendations so that they can effectively respond to student ideas. Empirical results showed that the tool assisted teachers in noticing, interpreting, and responding to students' mathematical ideas by helping them notice nuances in students' explanations and make instructional decisions based on students' current understanding rather than evaluate the correctness of student solutions. Although Bywater et al. used DM techniques to examine mathematics teachers' noticing, future researchers may apply EDM in the context of science education to help novice science teachers develop their ability to notice student thinking as evidence of a successful classroom endorsed by the science education reforms.

Second, reviewing argumentation studies in science education thematically, Bağ and Çalık (2017) found that research on development of scientific argumentation skills was one of the topics that attracted many researchers' attention. In addition, the current

review showed that researchers in science education sought to develop DM-based tutoring systems and revealed that these systems helped students improve the quality of their scientific argumentation (Huang et al., 2011; Lee et al., 2019). NCTM (2014) stated that developing students' and novice teachers' mathematical arguments was one of the important practices in mathematics education. Although we did not find any articles that used DM techniques to develop students' mathematical argumentation skills, future researchers may integrate DM techniques into mathematics education to investigate students' and teachers' mathematical argumentation and to improve their levels of argumentation.

The findings in the current study have important implications for researchers and teacher educators. First, as discussed above, DM techniques can be used to analyze large volumes of educational data quantitatively. Although researchers have used statistical methods to analyze quantitative data, DM techniques (e.g. text mining) have the potential to analyze unstructured data, such as students' and teachers' discussion, written explanations, and online log data files, which are hard to analyze with traditional statistical methods. Therefore, researchers might use EDM to discover hidden phenomena and information in mathematics and science education. Second, although relatively small numbers of EDM papers in mathematics and science education focused on developing DM-based systems to support teachers and students, our review showed that DM techniques have the potential to deal with important issues in mathematics and science education (e.g. teachers' difficulties in noticing student thinking, managing different groups of students, and assessing many students' learning processes as well as personalizing students' learning and development of argumentation skills). Based on this review, we hope that EDM can be further used for improving the quality of mathematics and science education and addressing educational issues related to teaching and learning in mathematics and science education.

## References

- Abidi, S., Hussain, M., Xu, Y., & Zhang, W. (2019). Prediction of confusion attempting algebra homework in an intelligent tutoring system through machine learning techniques for educational sustainable development. *Sustainability*. Advance online publication. <https://doi.org/10.3390/su11010105>.
- Aiken, J. M., Henderson, R., & Caballero, M. D. (2019). Modeling student pathways in a physics bachelor's degree program. *Physical Review Physics Education Research*, Advance online publication. <https://doi.org/10.1103/PhysRevPhysEducRes.15.010128>.
- Akgün, E., & Demir, M. (2018). Modeling course achievements of elementary education teacher candidates with artificial neural networks. *International Journal of Assessment Tools in Education*, 5(3), 491–509.
- Aksoy, E., Narlı, S., & Idil, F. H. (2016). Using data mining techniques examination of the middle school students' attitude towards mathematics in the context of some variables. *International Journal of Education in Mathematics Science and Technology*, 4(3), 210–228.
- Aldowah, H., Al-Samarraie, H., & Fauzy, W. M. (2019). Educational data mining and learning analytics for 21st century higher education: A review and synthesis. *Telematics and Informatics*, 37, 13–46.
- Araya, R., Jiménez, A., Bahamondez, M., Calfucura, P., Dartnell, P., & Soto-Andrade, J. (2014). Teaching modeling skills using a massively multiplayer online mathematics game. *World Wide Web*, 17(2), 213–227.
- Bağ, H., & Çalık, M. (2017). A thematic review of argumentation studies at the K-8 level. *Education and Science*, 42(190), 281–303.

- Barnhart, T., & van Es, E. (2015). Studying teacher noticing: Examining the relationship among pre-service science teachers' ability to attend, analyze and respond to student thinking. *Teaching and Teacher Education*, *45*, 83–93.
- Beggrow, E. P., Ha, M., Nehm, R. H., Pearl, D., & Boone, W. J. (2014). Assessing scientific practices using machine-learning methods: How closely do they match clinical interview performance? *Journal of Science Education and Technology*, *23*(1), 160–182.
- Bywater, J. P., Chiu, J. L., Hong, J., & Sankaranarayanan, V. (2019). The teacher responding tool: Scaffolding the teacher practice of responding to student ideas in mathematics classrooms. *Computers & Education*, *139*, 16–30.
- Cai, W., Grossman, J., Lin, Z., Sheng, H., Wei, J. T. Z., Williams, J. J., & Goel, S. (2019). *MathBot: A personalized conversational agent for learning math*. Retrieved from <https://footprints.stanford.edu/papers/mathbot.pdf>. Accessed 16 Jan 2020.
- Çalık, M., & Sözbilir, M. (2014). Parameters of content analysis. *Education and Science*, *39*(174), 33–38.
- Chen, C. T., & Chang, K. Y. (2017). A study on the rare factors exploration of learning effectiveness by using fuzzy data mining. *EURASIA Journal of Mathematics, Science and Technology Education*, *13*(6), 2235–2253.
- Chen, J., Zhang, Y., Wei, Y., & Hu, J. (2019). Discrimination of the contextual features of top performers in scientific literacy using a machine learning approach. *Research in Science Education*. Advanced online publication. <https://doi.org/10.1007/s11165-019-9835-y>.
- Cheon, J., Lee, S., Smith, W., Song, J., & Kim, Y. (2013). The determination of children's knowledge of global lunar patterns from online essays using text mining analysis. *Research in Science Education*, *43*(2), 667–686.
- Choi, Y., Lim, Y., & Son, D. (2017). A semantic network analysis on the recognition of STEAM by middle school students in South Korea. *EURASIA Journal of Mathematics, Science and Technology Education*, *13*(10), 6457–6469.
- Cooper, C. I., & Pearson, P. T. (2012). A genetically optimized predictive system for success in general chemistry using a diagnostic algebra test. *Journal of Science Education and Technology*, *21*(1), 197–205.
- Depren, S. K. (2018). Prediction of students' science achievement: An application of multivariate adaptive regression splines and regression trees. *Journal of Baltic Science Education*, *17*(5), 887–903.
- Depren, S. K., Aşkın, Ö. E., & Öz, E. (2017). Identifying the classification performances of educational data mining methods: A case study for TIMSS. *Educational Sciences: Theory & Practice*, *17*(5), 1605–1623.
- Dutt, A., Ismail, M. A., & Herawan, T. (2017). A systematic review on educational data mining. *IEEE Access*, *5*, 15991–16005.
- Duzhin, F., & Gustafsson, A. (2018). Machine learning-based app for self-evaluation of teacher-specific instructional style and tools. *Education in Science*, *8*(1), 15. <https://doi.org/10.3390/educsci9040263>.
- English, L. D., & King, D. (2019). STEM integration in sixth grade: Designing and constructing paper bridges. *International Journal of Science and Mathematics Education*, *17*(5), 863–884.
- Figueiredo, M., Esteves, L., Neves, J., & Vicente, H. (2016). A data mining approach to study the impact of the methodology followed in chemistry lab classes on the weight attributed by the students to the lab work on learning and motivation. *Chemistry Education Research and Practice*, *17*(1), 156–171.
- Filiz, E., & Oz, E. (2019). Finding the best algorithms and effective factors in classification of Turkish science student success. *Journal of Baltic Science Education*, *18*(2), 239–253.
- Gabriel, F., Signolet, J., & Westwell, M. (2018). A machine learning approach to investigating the effects of mathematics dispositions on mathematical literacy. *International Journal of Research & Method in Education*, *41*(3), 306–327.
- 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.
- Goggins, S. P., Xing, W., Chen, X., Chen, B., & Wadholm, B. (2015). Learning analytics at “small” scale: Exploring a complexity-grounded model for assessment automation. *Journal of Universal Computer Science*, *21*(1), 66–92.
- 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.
- Günel, K., Polat, R., & Kurt, M. (2016). Analyzing learning concepts in intelligent tutoring systems. *International Arab Journal of Information Technology*, *13*(2), 281–286.
- Ha, M., & Nehm, R. H. (2016). The impact of misspelled words on automated computer scoring: A case study of scientific explanations. *Journal of Science Education and Technology*, *25*(3), 358–374.

- Ha, M., Nehm, R. H., Urban-Lurain, M., & Merrill, J. E. (2011). Applying computerized-scoring models of written biological explanations across courses and colleges: Prospects and limitations. *CBE Life Sciences Education*, 10(4), 379–393.
- Hershkovitz, A., de Baker, R. S. J., Gobert, J., Wixon, M., & Pedro, M. S. (2013). Discovery with models: A case study on carelessness in computer-based science inquiry. *American Behavioral Scientist*, 57(10), 1480–1499.
- Hodgen, J., Küchemann, D., Brown, M., & Coe, R. (2009). Children's understandings of algebra 30 years on. *Research in Mathematics Education*, 11(2), 193–194.
- Hossain, Z., Bumbacher, E., Brauneis, A., Diaz, M., Saltarelli, A., Blikstein, P., & Riedel-Kruse, I. H. (2018). Design guidelines and empirical case study for scaling authentic inquiry-based science learning via open online courses and interactive biology cloud labs. *International Journal of Artificial Intelligence in Education*, 28(4), 478–507.
- Howard, E., Meehan, M., & Pamell, A. (2018). Live lectures or online videos: Students' resource choices in a first-year university mathematics module. *International Journal of Mathematical Education in Science and Technology*, 49(4), 530–553.
- Huang, C. J., Wang, Y. W., Huang, T. H., Chen, Y. C., Chen, H. M., & Chang, S. C. (2011). Performance evaluation of an online argumentation learning assistance agent. *Computers & Education*, 57(1), 1270–1280.
- Ismail, S., & Abdulla, S. (2015). Design and implementation of an intelligent system to predict the student graduation AGPA. *Australian Educational Computing*, 30(2). Retrieved from <http://journal.acce.edu.au/index.php/AEC/article/view/53>. Accessed 16 Jan 2020.
- Jacobs, V. R., Lamb, L. L., & Philipp, R. A. (2010). Professional noticing of children's mathematical thinking. *Journal for Research in Mathematics Education*, 41(2), 169–202.
- Kilic, H. (2018). Pre-service mathematics teachers' noticing skills and scaffolding practices. *International Journal of Science and Mathematics Education*, 16(2), 377–400.
- Kim, D., Yoon, M., Jo, I. H., & Branch, R. M. (2018). Learning analytics to support self-regulated learning in asynchronous online courses: A case study at a women's university in South Korea. *Computers & Education*, 127, 233–251.
- Kinnebrew, J. S., Killingsworth, S. S., Clark, D. B., Biswas, G., Sengupta, P., Minstrell, J., . . . Krinks, K. (2016). Contextual markup and mining in digital games for science learning: Connecting player behaviors to learning goals. *IEEE Transactions on Learning Technologies*, 10(1), 93–103.
- Kirby, N., & Dempster, E. (2015). Not the norm: The potential of tree analysis of performance data from students in a foundation mathematics module. *African Journal of Research in Mathematics, Science and Technology Education*, 19(2), 131–142.
- Kitchenham, B., & Charters, S. (2007). *Guidelines for performing systematic literature reviews in software engineering (Version 2.3)*. Keele University and Durham University.
- Lamb, R., Annetta, L., Vallett, D., & Sadler, T. (2014). Cognitive diagnostic like approaches using neural-network analysis of serious educational videogames. *Computers & Education*, 70, 92–104.
- Lamb, R., Cavagnetto, A., & Akmal, T. (2016). Examination of the nonlinear dynamic systems associated with science student cognition while engaging in science information processing. *International Journal of Science and Mathematics Education*, 14(1), 187–205.
- Lavie Alon, N., & Tal, T. (2015). Student self-reported learning outcomes of field trips: The pedagogical impact. *International Journal of Science Education*, 37(8), 1279–1298.
- Lee, H. S., Pallant, A., Pryputniewicz, S., Lord, T., Mulholland, M., & Liu, O. L. (2019). Automated text scoring and real-time adjustable feedback: Supporting revision of scientific arguments involving uncertainty. *Science Education*, 103(3), 590–622.
- Lee, Y. (2019). Using self-organizing map and clustering to investigate problem-solving patterns in the massive open online course: An exploratory study. *Journal of Educational Computing Research*, 57(2), 471–490.
- Levy, S. T., & Wilensky, U. (2011). Mining students' inquiry actions for understanding of complex systems. *Computers & Education*, 56(3), 556–573.
- Liu, S. H., & Lee, G. G. (2013). Using a concept map knowledge management system to enhance the learning of biology. *Computers & Education*, 68, 105–116.
- Liu, O. L., Rios, J. A., Heilman, M., Gerard, L., & Linn, M. C. (2016). Validation of automated scoring of science assessments. *Journal of Research in Science Teaching*, 53(2), 215–233.
- Liu, X., & Whitford, M. (2011). Opportunities-to-learn at home: Profiles of students with and without reaching science proficiency. *Journal of Science Education and Technology*, 20(4), 375–387.

- Magana, A. J., Elluri, S., Dasgupta, C., Seah, Y. Y., Madamanchi, A., & Boutin, M. (2019). The role of simulation-enabled design learning experiences on middle school students' self-generated inheritance heuristics. *Journal of Science Education and Technology*, 28(4), 1–17.
- Malmberg, J., Järvenoja, H., & Järvelä, S. (2013). Patterns in elementary school students' strategic actions in varying learning situations. *Instructional Science*, 41(5), 933–954.
- Martin, T., Petrick Smith, C., Forsgren, N., Aghababayan, A., Janisiewicz, P., & Baker, S. (2015). Learning fractions by splitting: Using learning analytics to illuminate the development of mathematical understanding. *Journal of the Learning Sciences*, 24(4), 593–637.
- Masci, C., Johnes, G., & Agasisti, T. (2018). Student and school performance across countries: A machine learning approach. *European Journal of Operational Research*, 269(3), 1072–1085.
- McConney, A., & Perry, L. B. (2010). Science and mathematics achievement in Australia: The role of school socioeconomic composition in educational equity and effectiveness. *International Journal of Science and Mathematics Education*, 8(3), 429–452.
- National Council of Teachers of Mathematics. (2014). *Principles to actions: Ensuring mathematical success for all*. Reston, VA: Author.
- National Research Council. (2012). *A framework for K-12 science education: Practices, crosscutting concepts, and core ideas*. Washington, DC: National Academies Press.
- Nehm, R. H., Ha, M., & Mayfield, E. (2012). Transforming biology assessment with machine learning: Automated scoring of written evolutionary explanations. *Journal of Science Education and Technology*, 21(1), 183–196.
- Nehm, R. H., & Haertig, H. (2012). Human vs. computer diagnosis of students' natural selection knowledge: Testing the efficacy of text analytic software. *Journal of Science Education and Technology*, 21(1), 56–73.
- NGSS Lead States. (2013). *Next generation science standards: For states, by states*. Washington, DC: National Academies Press.
- 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.
- Owens, M. T., Seidel, S. B., Wong, M., Bejines, T. E., Lietz, S., Perez, J. R., . . . Balukjian, B. (2017). Classroom sound can be used to classify teaching practices in college science courses. *Proceedings of the National Academy of Sciences*, 114(12), 3085–3090.
- Pantziara, M., & Philippou, G. N. (2015). Students' motivation in the mathematics classroom. Revealing causes and consequences. *International Journal of Science and Mathematics Education*, 13(2), 385–411.
- Papamitsiou, Z., & Economides, A. A. (2014). Learning analytics and educational data mining in practice: A systematic literature review of empirical evidence. *Journal of Educational Technology & Society*, 17(4), 49–64.
- Peña-Ayala, A. (2014). Educational data mining: A survey and a data mining-based analysis of recent works. *Expert Systems with Applications*, 41(4), 1432–1462.
- Prevost, L. B., Smith, M. K., & Knight, J. K. (2016). Using student writing and lexical analysis to reveal student thinking about the role of stop codons in the central dogma. *CBE Life Sciences Education*, 15(4), ar65. <https://doi.org/10.1187/cbe.15-12-0267>.
- Rao, D. C., & Saha, S. K. (2019). An immersive learning platform for efficient biology learning of secondary school-level students. *Journal of Educational Computing Research*. Advanced online publication. <https://doi.org/10.1177/0735633119854031>.
- Reitsma, R., Marshall, B., & Chart, T. (2012). Can intermediary-based science standards crosswalking work? Some evidence from mining the standard alignment tool (SAT). *Journal of the American Society for Information Science and Technology*, 63(9), 1843–1858.
- Roberts, J. D., Chung, G. K., & Parks, C. B. (2016). Supporting children's progress through the PBS KIDS learning analytics platform. *Journal of Children and Media*, 10(2), 257–266.
- Rodrigues, M. W., Isotani, S., & Zárata, L. E. (2018). Educational data mining: A review of evaluation process in the e-learning. *Telematics and Informatics*, 35(6), 1701–1717.
- 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. (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.
- Romero, C., & Ventura, S. (2013). Data mining in education. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 3(1), 12–27.
- Saa, A. A., Al-Emran, M., & Shaalan, K. (2019). Factors affecting students' performance in higher education: A systematic review of predictive data mining techniques. *Technology, knowledge and learning*. Advanced online publication. doi:<https://doi.org/10.1007/s10758-019-09408-7>.

- Sánchez-Matamoros, G., Fernández, C., & Llinares, S. (2015). Developing pre-service teacher' noticing of students' understanding of the derivative concept. *International Journal of Science and Mathematics Education*, 13(6), 1305–1329.
- Scarpello, G. (2007). Helping students get past math anxiety. *Techniques: Connecting Education and Careers*, 82(6), 34–35.
- Schwarz, B. B., Prusak, N., Swidan, O., Livny, A., Gal, K., & Segal, A. (2018). Orchestrating the emergence of conceptual learning: A case study in a geometry class. *International Journal of Computer-Supported Collaborative Learning*, 13(2), 189–211.
- Sergis, S., Sampson, D. G., Rodriguez-Triana, M. J., Gillet, D., Pelliccione, L., & de Jong, T. (2019). Using educational data from teaching and learning to inform teachers' reflective educational design in inquiry-based STEM education. *Computers in Human Behavior*, 92, 724–738.
- Shahiri, A. M., Husain, W., & Rashid, N. A. (2015). A review on predicting student's performance using data mining techniques. *Procedia Computer Science*, 72, 414–422.
- She, H.-C., Lin, H.-s., & Huang, L.-Y. (2019). Reflections on and implications of the programme for international student assessment 2015 performance of students in Taiwan: The role of epistemic beliefs about science in scientific literacy. *Journal of Research in Science Teaching*. Advanced online publication. <https://doi.org/10.1002/tea.21553>.
- Sieke, S. A., McIntosh, B. B., Steele, M. M., & Knight, J. K. (2019). Characterizing students' ideas about the effects of a mutation in a noncoding region of DNA. *CBE Life Sciences Education*, 18(2), ar18. <https://doi.org/10.1187/cbe.18-09-0173>.
- Suh, S. C., Upadhyaya, A., & Nadig, A. (2019). Analyzing personality traits and external factors for stem education awareness using machine learning. *International Journal of Advanced Computer Science and Applications*, 10(5), 1–4.
- Tawfik, A. A., Reeves, T. D., Stich, A. E., Gill, A., Hong, C., McDade, J., . . . Giabbanelli, P. J. (2017). The nature and level of learner–learner interaction in a chemistry massive open online course (MOOC). *Journal of Computing in Higher Education*, 29(3), 411–431.
- Tissenbaum, M., & Slotta, J. D. (2019). Developing a smart classroom infrastructure to support real-time student collaboration and inquiry: A 4-year design study. *Instructional Science*. Advanced online publication, 47, 423–462. <https://doi.org/10.1007/s11251-019-09486-1>.
- Wahlberg, S. J., & Gericke, N. M. (2018). Conceptual demography in upper secondary chemistry and biology textbooks' descriptions of protein synthesis: A matter of context? *CBE Life Sciences Education*, 17(3), ar51. <https://doi.org/10.1187/cbe.17-12-0274>.
- Wang, X. (2016). Course-taking patterns of community college students beginning in STEM: Using data mining techniques to reveal viable STEM transfer pathways. *Research in Higher Education*, 57(5), 544–569.
- Wiley, J., Hastings, P., Blaum, D., Jaeger, A. J., Hughes, S., Wallace, P., ... & Britt, M. A. (2017). Different approaches to assessing the quality of explanations following a multiple-document inquiry activity in science. *International Journal of Artificial Intelligence in Education*, 27(4), 758–790.
- Zhang, W., Qin, S., Jin, H., Deng, J., & Wu, L. (2017). An empirical study on student evaluations of teaching based on data mining. *EURASIA Journal of Mathematics, Science and Technology Education*, 13(8), 5837–5845.