Educational Data Mining and Learning Analytics in the 21st Century

Georgios Lampropoulos
https://orcid.org/0000-0002-5719-2125
International Hellenic University, Greece

INTRODUCTION

The digitalization of everyday life and the rapid technological advancements have helped the new generation of students be familiar with and accustomed to digital technologies from a young age. As a result, their educational needs and requirements as well as their viewpoints regarding what effective learning is have drastically changed. Moreover, while the need for basic education is becoming urgent, the current educational system struggles to satisfy the new educational requirements (Prensky, 2001) and students' need for more personalized learning experiences.

Simultaneously, an exponentially increasing amount of heterogeneous data which is characterized by its volume, variety, veracity, velocity and value is being generated (McAfee et al., 2012). This vast volume of data is called Big Data and it paves the way for technological advances. Data mining is a scientific field which has experienced drastic advancements due to the rise of Big Data and its being more broadly applied in a wide variety of domains.

Data mining, also known as Knowledge Discovery in Databases (KDD), utilizes a variety of algorithms, techniques and methods in order to generate knowledge by discovering novel and useful information, patterns, relationships or structures from large data collections (Fayyad et al., 1996; Witten & Frank, 2002). Therefore, it can be used as an invaluable tool that supports and enhances decision-making (Peña-Ayala, 2014). As a result, data mining along with the necessary analysis tools are utilized in various sectors. The educational domain is no exception to that as data mining can offer several benefits.

The main aim of this chapter is to offer an overview of educational data mining and learning analytics and their essential role in improving 21st century education while capitalizing on the emergence of data science. For that reason, it presents and analyzes the concepts of learning analytics and educational data mining, their evolution as well as their role in modern education and highlights the impact of machine learning on them. Moreover, the chapter goes over the recent literature and extracts invaluable information according to the results and outcomes of related studies. Furthermore, it discusses their use as useful tools in educational settings and as a means to offer intelligent personalized learning in order to meet the new and upcoming educational needs and requirements. In addition, it goes over the merits that they can yield, and it suggests ways to address some of the current open challenges and limitations. Finally, it presents the summary of the main findings, drawn conclusions and provides directions for future research.

DOI: 10.4018/978-1-7998-9220-5.ch098

BACKGROUND

A lot of emphasis is put on the collection, processing and analysis of data so as to better comprehend and optimize the learning and teaching process and outcomes. Educational data mining and learning analytics are two fields which are becoming more popular due to this fact.

Educational data mining is a specialized form of data mining which focuses on educational environments. It aims at addressing educational issues by analyzing data and developing models that enhance the overall learning experience and outcomes and increase the institutional effectiveness (Baker and Yacef, 2009; Dutt et al., 2017). It is worth noting that due to its nature, it can be applied at all educational levels (Saa, 2016). According to the Educational Data Mining community, educational data mining can be defined as "an emerging discipline, concerned with developing methods for exploring the unique and increasingly large-scale data that come from educational settings and using those methods to better understand students, and the settings which they learn in" (Educational Data Mining Community, 2021). Based on this particular definition, it is apparent that educational data mining constitutes an interdisciplinary field. Therefore, it exploits machine learning, statistics, information retrieval, recommender systems and other innovative technologies and techniques (Romero & Ventura, 2010).

As a part of Technology-Enhanced Learning (TEL) research, learning analytics is another interdisciplinary scientific field which is gaining ground as it can enhance the existing education models (Ferguson, 2012; Siemens & Baker, 2012). It can be regarded as a powerful tool which explores how the large volume of data can be used to enhance the overall learning quality and to address a variety of educational challenges and issues (Bakharia et al., 2016; Pardo et al., 2019). Particularly, learning analytics involves the collection, analysis and visualization of data with a view to better comprehending and improving both learning experience and environment (Ferguson, 2012; Lang et al., 2017; Liñán & Pérez, 2015; Romero & Ventura, 2020; Siemens, 2013). The analysis and visualization of data can provide invaluable feedback for both students and teachers (Clow, 2013) and assist in fulfilling the new educational needs as well as positively affecting students' learning and progression (Slade & Prinsloo, 2013). Therefore, it can be regarded as an essential educational tool which can be further improved as the amount of data increases and machine learning algorithms and models become more advanced.

Even though these two fields seem similar at first sight, there are some distinct differences between them. Particularly, educational data mining follows a bottom-up approach as it aims at finding new patterns in data and developing new models while learning analytics follows a top-down approach to assess the learning theories and the educational process by utilizing existing models (Baker and Yacef, 2009; Bienkowski et al., 2012). Moreover, there are various distinctions between educational data mining and learning analytics regarding their origins, reduction and holism approaches, knowledge discovery aims and methods, adaptation and personalization strategies as well as the overall techniques and methods used (Liñán & Pérez, 2015; Siemens, 2012). It can be said that educational data mining focuses more on the technical challenges while learning analytics on the educational ones (Ferguson, 2012). Despite their main focus points, their successful implementation in educational settings can bring about numerous innovative and positive changes and yield several merits to satisfy the new educational needs and students' requirements.