applied sciences

MDPI

Editorial

Data Analytics and Machine Learning in Education

Juan A. Gómez-Pulido [1,*], Young Park [2], Ricardo Soto [3] and José M. Lanza-Gutiérrez [4]

[1] Department of Technology of Computers and Communications, Universidad de Extremadura, 10003 Cáceres, Spain
[2] Department of Computer Science and Information Systems, Bradley University, Peoria, IL 61625, USA
[3] School of Computer Engineering, Pontificia Universidad Católica de Valparaíso, Valparaíso 4059, Chile
[4] Department of Computer Sciences, Universidad de Alcalá, 28805 Alcalá de Henares, Spain
* Correspondence: jangomez@unex.es

The widespread application of information and communication technologies in education, especially in the context of learning management platforms, is generating a large amount of data related to the academic activities in which students and teachers participate. These data stand out not only for their quantity and heterogeneity, but also for their relationship with the behavior and performance of the educational actors. For this reason, these data must be properly stored, processed and analyzed, with the aim of extracting knowledge that can be highly useful for improving educational processes. For this purpose, this Special Issue aims to present cutting-edge research on the application of advanced data analysis and machine learning techniques in education.

Among the main data sources in the academic environment, technology-enhanced learning (TEL) platforms stand out. For example, virtual campus management platforms, which are very widespread today in education, generate data from all kinds of educational management tasks, the virtualization of teaching, monitoring of student academic progress, storage of teaching materials and student interaction with them, etc.

These data and the tools that generate them provide a great opportunity to undertake novel research lines that deserve to be explored, not only for the benefit of the educational community, but also as an opportunity for the advancement of knowledge. In this sense, many lines of research can be found, such as predicting students' behaviour, developing new tools for supporting learning stages, recommending resources, preventing dropout, enhancing activities, etc. To this end, computer and data sciences provide advanced methods for data processing and analysis for knowledge extraction. Data mining, big data, machine learning, deep learning, collaborative filtering, and recommender systems, among other fields related to data science and artificial intelligence, allow for the development of advanced techniques that provide significant potential for the above purposes, leading to new applications and more effective approaches in academic analysis and prediction.

In summary, this Special Issue provides a collection of papers of original advances in the analysis, prediction, and recommendation of applications propelled by artificial intelligence, data science, data analytics, big data, and machine learning, especially in the TEL context. Although all the papers included in this Special Issue cover their specific topics, we could group them under the following three broad approaches: (1) data analytics and machine learning methods for studying students' behaviour; (2) tools for the improvement of learning environments; and (3) optimization of the management in these environments. However, some of these papers could be classified into more than one of these groups.

With regard to data analytics and machine learning methods for studying students' behaviour, we find six contributions. An algorithm based on K-Means and clustering is proposed to analyze the living habits and learning performance of the students from four universities in "Analysis of University Students' Behavior Based on a Fusion K-Means Clustering Algorithm" [1], by Wenbing Chang et al. The article "Analysis and Prediction of Engineering Student Behavior and Their Relation to Academic Performance Using Data

Citation: Gómez-Pulido, J.A.; Park, Y.; Soto, R.; Lanza-Gutiérrez, J.M. Data Analytics and Machine Learning in Education. *Appl. Sci.* **2023**, *13*, 1418. https://doi.org/10.3390/app13031418

Received: 9 January 2023
Accepted: 18 January 2023
Published: 20 January 2023

Analytics Techniques" [2], by Hanns de la Fuente-Mella et al., focuses on identifying personality traits in computer science students and determining whether they are related to academic performance. The article "Towards Predicting Student's Dropout in University Courses Using Different Machine Learning Techniques" [3], by Janka Kabathova and Martin Drlik, analyzes the importance of the dataset's features when applying machine learning classifiers, in order to predict student's dropout. The paper "Quantifying the Impact of Student Enrollment Patterns on Academic Success Using a Hidden Markov Model" [4], by Shahab Boumi and Adan Vela, applies a Hidden Markov Model to distinguish and cluster students' enrollment strategies into the following three categories: full-time, part-time, and mixed. The paper "Predicting GPA of University Students with Supervised Regression Machine Learning Models" [5], by Lukáš Falát and Terézia Piscová, predicts the grade point average by applying machine learning methods and identifies the factors influencing this average. Finally, the paper "Learning Analytics to Determine Profile Dimensions of Students Associated with Their Academic Performance" [6], by Andrés Gonzalez-Nucamendi et al., determines the most important factors that lead to good academic performance by considering student's self-regulation learning and affective strategies.

With regard to the development of tools for the improvement of learning environments, we find three contributions. The paper "Automated Transformation from Competency List to Tree: Way to Competency-Based Adaptive Knowledge E-Evaluation" [7], by Asta Margienė et al., proposes a tool to convert the competency portfolio in list form to a tree-based competency portfolio, allowing the integration of different e-learning systems. The paper "Design and Implementation of an IoT-Based Smart Classroom Incubator" [8], by Mustafa Burunkaya and Kazim Duraklar, develops an IoT system and a smart classroom incubator algorithm to reduce the adverse impacts of environmental factors on learning. Finally, the paper "Visualizing Collaboration in Teamwork: A Multimodal Learning Analytics Platform for Non-Verbal Communication" [9], by René Noël et al. presents a Multimodal Learning Analytics platform to support a collaboration assessment based on the capture and classification of non-verbal communication interactions.

With regard to the management of processes in learning environments, we identified five contributions. The paper "Table Organization Optimization in Schools for Preserving the Social Distance during the COVID-19 Pandemic" [10], by Rubén Ferrero-Guillén et al., applies a Genetic Algorithm for optimization of the disposition of the tables at schools during the coronavirus pandemic. The paper "Online Blended Learning in Small Private Online Course" [11], by Yong Han et al., explores the applicability of small private online course advanced teaching concepts to computer network online experiment teaching. The paper "Quality Assurance for Performing Arts Education: A Multi-Dimensional Analysis Approach" [12], by Qingyun Li et al., utilizes a multi-dimensional analysis approach for senior student evaluation, developing an analytical framework to analyze the course evaluation data to make evidence-based recommendations. The paper "Data Analysis as a Tool for the Application of Adaptive Learning in a University Environment" [13], by William Villegas-Ch et al., analyzes educational data through a big data architecture to generate learning based on meeting the needs of students. Finally, in the paper "Toward a Better Understanding of Academic Programs Educational Objectives: A Data Analytics-Based Approach" [14], by Anwar Ali Yahya et al., a dataset of Program Education Objects in outcome-based engineering academic programs was built and analyzed to develop a better understanding of their correlations in order to provide useful actionable insights for empowering decision-making toward the systemization and optimization of academic programs processes.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Chang, W.; Ji, X.; Liu, Y.; Xiao, Y.; Chen, B.; Liu, H.; Zhou, S. Analysis of University Students' Behavior Based on a Fusion K-Means Clustering Algorithm. *Appl. Sci.* **2020**, *10*, 6566. [CrossRef]
2. de la Fuente-Mella, H.; Guzmán Gutiérrez, C.; Crawford, K.; Foschino, G.; Crawford, B.; Soto, R.; León de la Barra, C.; Cisternas Caneo, F.; Monfroy, E.; Becerra-Rozas, M.; et al. Analysis and Prediction of Engineering Student Behavior and Their Relation to Academic Performance Using Data Analytics Techniques. *Appl. Sci.* **2020**, *10*, 7114. [CrossRef]
3. Kabathova, J.; Drlik, M. Towards Predicting Student's Dropout in University Courses Using Different Machine Learning Techniques. *Appl. Sci.* **2021**, *11*, 3130. [CrossRef]
4. Boumi, S.; Vela, A.E. Quantifying the Impact of Student Enrollment Patterns on Academic Success Using a Hidden Markov Model. *Appl. Sci.* **2021**, *11*, 6453. [CrossRef]
5. Falát, L.; Piscová, T. Predicting GPA of University Students with Supervised Regression Machine Learning Models. *Appl. Sci.* **2022**, *12*, 8403. [CrossRef]
6. Gonzalez-Nucamendi, A.; Noguez, J.; Neri, L.; Robledo-Rella, V.; García-Castelán, R.M.G.; Escobar-Castillejos, D. Learning Analytics to Determine Profile Dimensions of Students Associated with Their Academic Performance. *Appl. Sci.* **2022**, *12*, 560. [CrossRef]
7. Margienė, A.; Ramanauskaitė, S.; Nugaras, J.; Stefanovič, P. Automated Transformation from Competency List to Tree: Way to Competency-Based Adaptive Knowledge E-Evaluation. *Appl. Sci.* **2022**, *12*, 1582. [CrossRef]
8. Burunkaya, M.; Duraklar, K. Design and Implementation of an IoT-Based Smart Classroom Incubator. *Appl. Sci.* **2022**, *12*, 2233. [CrossRef]
9. Noël, R.; Miranda, D.; Cechinel, C.; Riquelme, F.; Primo, T.T.; Munoz, R. Visualizing Collaboration in Teamwork: A Multimodal Learning Analytics Platform for Non-Verbal Communication. *Appl. Sci.* **2022**, *12*, 7499. [CrossRef]
10. Ferrero-Guillén, R.; Díez-González, J.; Verde, P.; Álvarez, R.; Perez, H. Table Organization Optimization in Schools for Preserving the Social Distance during the COVID-19 Pandemic. *Appl. Sci.* **2020**, *10*, 8392. [CrossRef]
11. Han, Y.; Wu, W.; Zhang, L.; Liang, Y. Online Blended Learning in Small Private Online Course. *Appl. Sci.* **2021**, *11*, 7100. [CrossRef]
12. Li, Q.; Li, Z.M.; Han, J.; Ma, H. Quality Assurance for Performing Arts Education: A Multi-Dimensional Analysis Approach. *Appl. Sci.* **2022**, *12*, 4813. [CrossRef]
13. Villegas-Ch, W.; Roman-Cañizares, M.; Jaramillo-Alcázar, A.; Palacios-Pacheco, X. Data Analysis as a Tool for the Application of Adaptive Learning in a University Environment. *Appl. Sci.* **2020**, *10*, 7016. [CrossRef]
14. Yahya, A.A.; Sulaiman, A.A.; Mashraqi, A.M.; Zaidan, Z.M.; Halawani, H.T. Toward a Better Understanding of Academic Programs Educational Objectives: A Data Analytics-Based Approach. *Appl. Sci.* **2021**, *11*, 9623. [CrossRef]