electronics

MDPI

Editorial

Knowledge Engineering and Data Mining

Agnieszka Konys [1] and Agnieszka Nowak-Brzezińska [2,*]

[1] Faculty of Computer Science and Information Technology, West Pomeranian University of Technology Szczecin, Zolnierska 49, 71-210 Szczecin, Poland
[2] Institute of Computer Science, Faculty of Science and Technology, University of Silesia, ul. Będzińska 39, 41-200 Sosnowiec, Poland
* Correspondence: agnieszka.nowak-brzezinska@us.edu.pl

Knowledge engineering and data mining are the two biggest pillars of modern intelligent systems. Knowledge induction from data is often based on using a wide range of machine learning algorithms and feature selection or extraction algorithms. When we collect various data types, we need solutions that will allow us to supervise these data correctly. Recently, machine-learning-based methods are increasingly employed to solve such problems; however, the selection of an appropriate feature selection technique, sampling mechanism, and/or classifiers for building decision support systems is very challenging. To address this challenging task, article [1] examines the effectiveness of various data science techniques concerning the issue of credit decision support. In particular, a processing pipeline was designed that consists of methods for data resampling, feature discretization, feature selection, and binary classification.

The capability of machine learning to discover hidden patterns in large datasets encourages researchers to invent data with high-dimensional features. In contrast, not all features are needed by machine learning, and, in many cases, high-dimensional features decrease the performance of machine learning. The research presented in paper [2] investigates and proposes methods to determine the best feature selection method in the domain of psychosocial education.

Recommendation systems are powerful tools that are integral parts of a great many websites. Most often, recommendations are presented in the form of a list that is generated by using various recommendation methods. Typically, however, these methods do not generate identical recommendations, and their effectiveness varies between users. In order to solve this problem, the application of aggregation techniques was suggested in article [3], the aim of which is to combine several lists into one, which, in theory, should improve the overall quality of generated recommendations.

Ontologies, and especially formal ones, have traditionally been investigated as a means with which to formalize an application domain, so as to carry out automated reasoning on it. The union of the terminological part of an ontology and the corresponding assertional part is known as a knowledge graph. On the other hand, database technology has often focused on the optimal organization of data, so as to boost efficiency in their storage, management, and retrieval. Graph databases are a recent technology that specifically focus on element-driven data browsing rather than on batch processing.

Paper [4] proposes an intermediate format that can be easily mapped onto a formal ontology on the one hand, so as to allow complex reasoning, and onto a graph database on the other, so as to benefit from efficient data handling. Selecting the right supplier is a critical decision in sustainable supply chain management. Paper [5] proposes and implements an ontology-based approach for knowledge acquisition from the text for a sustainable supplier selection domain. This approach is dedicated to acquiring complex relationships from texts and coding these in the form of rules.

Whenever we need to analyze big data we need to do it effectively, with the shortest possible time and the highest possible accuracy. If we deal with multidimensional data that

check for updates

Citation: Konys, A.; Nowak-Brzezińska, A. Knowledge Engineering and Data Mining. *Electronics* **2023**, *12*, 927. https://doi.org/10.3390/electronics12040927

Received: 6 February 2023
Accepted: 8 February 2023
Published: 13 February 2023

Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

are computing-intensive, applications should be parallelized and run on modern multicore machines to reduce the execution time. In paper [6] the authors demonstrate how to apply an affine transformation framework and generate parallel 2D tiled code computing GLREs (general linear recurrence equations).

The most popular classification techniques are decision trees, k-nearest neighbor classifiers, naive Bayes classifiers, or neural networks. A very interesting approach is presented in paper [7], where the study developed an autocorrect system for UAV smoke tracing. An AI model was used to calculate smoke tube angle corrections, such that smoke tube angles could be immediately corrected when smoke is sprayed.

Another interesting approach was presented in [8]. The exploration of oil and gas in offshore regions is increasing due to global energy demand. The weather in offshore areas is truly unpredictable due to the sparsity and unreliability of metocean data. Using metocean data, offshore wave height and period are predicted from the wind speed by three state-of-the-art machine learning algorithms (an artificial neural network, a support vector machine, and random forest).

Another interesting research is presented in [9], where the authors present an original concept of the classification of types of project tasks, which will allow for the more beneficial use of collected data in management support systems in the IT industry. The classification algorithms presented in the article are based on the manual recognition of task types. Rules based on keywords are created, which allow for the automatic recognition of task types at subsequent occurrences, which will allow for the fully automated operation of a task classification as well as subtask classification algorithm on a real-time basis and, finally, for the comprehensive support of the management of the development process.

A knowledge-mining- and graph-convolutional-network-based method is described in paper [10], where the authors propose a novel graph-convolutional-network-based method for the knowledge mining of interactions between drugs from the extensive literature. Thus, identifying possible drug–drug interactions (DDIs) has always been a crucial research topic in the field of clinical pharmacology.

A convolutional neural network is also used by the authors of paper [11], in which a neural network helps to discern the morphological information hidden in Chinese characters and a pretrained model obtains vectors with medical features. The different vectors are stitched together to form a multi-feature vector. Deep learning requires a large amount of annotated data to train the model, as does the proposed model, but large-scale annotated data in the Chinese electronic medical record domain require medical experts for annotation annotate, which can be time-consuming.

The healthcare sector is one of the most sensitive sectors in our society, and it is believed that the application of specific and detailed database creation and design techniques can improve the quality of patient care. In this sense, the better management of emergency resources should be achieved. Paper [12] presents an optimized database designed for emergency care. The general objective of the project was to create a database that was as complete as possible and with a great diversity of information, which would represent, in detail, all possible aspects of emergency health activity. A multi-model database allowed for the exploitation of information with predictive models.

Knowledge delivery is the topic which has recently been explored in an enormous way. The reason for this is the post-COVID-19 era in university education, where instructors around the world were at the forefront of implementing hybrid learning spaces for knowledge delivery. The purpose of the study presented in paper [13] is not only to divert the primary use of a YouTube channel into a tool to support asynchronous teaching, it also aims to provide feedback to instructors and suggest steps as well as actions to implement in their teaching modules to ensure students' access to new knowledge while promoting their engagement and satisfaction, regardless of the learning environment, i.e., face-to-face, distance, and hybrid. By analyzing and interpreting data directly from YouTube channel reports, six variables were identified and tested to quantify the lack of statistically significant changes in learners' viewing habits.

In facial aesthetics, soft-tissue landmark recognition and linear as well as angular measurements play critical roles in treatment planning. Visual identification and judgment by hand are time-consuming and prone to errors. As a result, user-friendly software solutions are required to assist healthcare practitioners in improving treatment planning. Paper [14] presents "A Computational Tool for Detection of Soft Tissue Landmarks and Cephalometric Analysis". The goal of the authors is to create a computational tool that may be used to identify and save critical landmarks from patient X-ray pictures. The second goal is to create automated software that can assess the soft-tissue facial profiles of patients in both linear and angular directions by using the landmarks that have been identified.

A variety of different techniques with which to support decisions requires deep knowledge about the advantages and disadvantages of these techniques, especially when we need to deal with multicriteria tasks. Multicriteria methods have gained traction in academia and industry practices for effective decision making. Paper [15] provides a complete overview of multicriteria methods through a bibliometric study, enabling scholars to comprehend the current state and future development patterns of multicriteria decision-making methods research.

We believe that this Special Issue covers the entire knowledge engineering pipeline: from data acquisition and data mining to knowledge extraction and exploitation. For this reason, we tried to gather the many researchers operating in the field to contribute to a collective effort in understanding the trends and future questions in the fields of knowledge engineering and data mining.

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

References

1. Ziemba, P.; Becker, J.; Becker, A.; Radomska-Zalas, A.; Pawluk, M.; Wierzba, D. Credit Decision Support Based on Real Set of Cash Loans Using Integrated Machine Learning Algorithms. *Electronics* **2021**, *10*, 2099. [CrossRef]
2. Muttakin, F.; Wang, J.-T.; Mulyanto, M.; Leu, J.-S. Evaluation of Feature Selection Methods on Psychosocial Education Data Using Additive Ratio Assessment. *Electronics* **2022**, *11*, 114. [CrossRef]
3. Bałchanowski, M.; Boryczka, U. Aggregation of Rankings Using Metaheuristics in Recommendation Systems. *Electronics* **2022**, *11*, 369. [CrossRef]
4. Ferilli, S. Integration Strategy and Tool between Formal Ontology and Graph Database Technology. *Electronics* **2021**, *10*, 2616. [CrossRef]
5. Konys, A. An Ontology-Based Approach for Knowledge Acquisition: An Example of Sustainable Supplier Selection Domain Corpus. *Electronics* **2022**, *11*, 4012. [CrossRef]
6. Bielecki, W.; Błaszyński, P. Parallel Tiled Code for Computing General Linear Recurrence Equations. *Electronics* **2021**, *10*, 2050. [CrossRef]
7. Chao, P.-Y.; Hsu, W.-C.; Chen, W.-Y. Design of Automatic Correction System for UAV's Smoke Trajectory Angle Based on KNN Algorithm. *Electronics* **2022**, *11*, 3587. [CrossRef]
8. Azad, M.; Uddin, M.A. Prediction of Offshore Wave at East Coast of Malaysia—A Comparative Study. *Electronics* **2022**, *11*, 2527. [CrossRef]
9. Wysocki, W.; Miciuła, I.; Mastalerz, M. Classification of Task Types in Software Development Projects. *Electronics* **2022**, *11*, 3827. [CrossRef]
10. Xu, X.; Meng, F.; Sun, L. Knowledge Mining of Interactions between Drugs from the Extensive Literature with a Novel Graph-Convolutional-Network-Based Method. *Electronics* **2023**, *12*, 311. [CrossRef]
11. Li, J.; Liu, R.; Chen, C.; Zhou, S.; Shang, X.; Wang, Y. An RG-FLAT-CRF Model for Named Entity Recognition of Chinese Electronic Clinical Records. *Electronics* **2022**, *11*, 1282. [CrossRef]
12. Arias, J.C.; Cubillas, J.J.; Ramos, M.I. Optimising Health Emergency Resource Management from Multi-Model Databases. *Electronics* **2022**, *11*, 3602. [CrossRef]
13. Kanetaki, Z.; Stergiou, C.; Bekas, G.; Jacques, S.; Troussas, C.; Sgouropoulou, C.; Ouahabi, A. Acquiring, Analyzing and Interpreting Knowledge Data for Sustainable Engineering Education: An Experimental Study Using YouTube. *Electronics* **2022**, *11*, 2210. [CrossRef]
14. Azad, M.; Elaiwat, S.; Alam, M.K. A Computational Tool for Detection of Soft Tissue Landmarks and Cephalometric Analysis. *Electronics* **2022**, *11*, 2408. [CrossRef]
15. Basílio, M.P.; Pereira, V.; Costa, H.G.; Santos, M.; Ghosh, A. A Systematic Review of the Applications of Multi-Criteria Decision Aid Methods (1977–2022). *Electronics* **2022**, *11*, 1720. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.