applied sciences

MDPI

Editorial

Integrated Artificial Intelligence in Data Science

Jerry Chun-Wei Lin [1], Stefania Tomasiello [2,3,*] and Gautam Srivastava [4]

[1] Faculty of Automatic Control, Electronics and Computer Science, Silesian University of Technology, 44-100 Gliwice, Poland
[2] Institute of Computer Science, Faculty of Science and Technology, University of Tartu, 51009 Tartu, Estonia
[3] Department of Industrial Engineering, University of Salerno, 84084 Fisciano, Italy
[4] Department of Mathematics and Computer Science, Brandon University, Brandon, MB R7A 6A9, Canada
* Correspondence: stefania.tomasiello@ut.ee or stomasiello@unisa.it

Artificial Intelligence (AI) is increasingly pervading everyday life since it can be used to solve high-complexity problems, as well as determine optimal solutions, in various domains and for numerous applications. From this perspective, it has the potential to create a better society. The benefits of AI in science, medicine, technology, and the social sciences have already been clearly proven. Data science, also referred to as pattern analytics and mining, can be used to retrieve useful and meaningful information from databases, which subsequently allows us to make decisions efficiently and build strategies in different domains. In particular, as a result of the exponential growth of data in recent years, the dual concept of big data and AI has given rise to many research topics, such as scale-up behaviour from classical algorithms.

In addition, a recent critical challenge that has emerged is the integration of multiple AI technologies from different fields (e.g., vision, security, control, bioinformatics) to develop efficient and robust systems that interact in the real world. Despite the tremendous progress in core AI technologies in recent years, the integration of such capabilities into larger systems that are reliable, transparent, and maintainable is still in its infancy. There are numerous open issues from both theoretical and practical perspectives.

This Special Issue aimed to collect and present all breakthrough research in this emerging field. A total of 22 papers have been accepted for publication in this Special Issue. In [1], neural language processing (NLP) and clustering techniques were used as the basis for a job advertisement analysis algorithm. In [2], a fuzzy logic-based dynamic ensemble model was discussed and employed for malware detection. The authors of [3] proposed an approach based on convolutional neural network (CNN) dual-attention unit (DAU) and selective kernel feature synthesis (SKFS) for the enhancement of dark images under low-light conditions. The topic of [4] is a new training algorithm for graph convolutional networks (GCNs) and the graph attention model (GAT). In [5], a fuzzy multicriteria decision-making (MCDM) system based on Z numbers was developed to support order-picking process management. Large-scale building information modeling was the application case studied in [6]. The authors used machine learning (ML) for the automated classification of elements and the detection of anomalies based on geometric inputs. In [7], a red fox optimizer for a deep learning-enabled microarray gene expression classification model was investigated. The article [8] deals with temporal fuzzy utility itemset mining, proposing a one-phase tree-structure method to find the high temporal fuzzy utility itemsets in a temporal database. In [9], several ML techniques were compared to assess the appropriateness of surgical prophylaxis. In [10], a deep learning (DL) model was developed for the early prediction of the learning performance of students, as well as for the analysis of student behaviour in a virtual learning environment. A comparative analysis of ensemble models for road traffic congestion was discussed in [11]. Advanced genetic Bollinger Bands and the trading strategy of correlation coefficient-based pairs have been proposed in [12]. In [13], the authors incorporate a crowdsourcing-based approach for annotated data acquisition in order to enhance the Active Learning procedure. In [14], an enhanced image captioning

check for updates

Citation: Lin, J.C.-W.; Tomasiello, S.; Srivastava, G. Integrated Artificial Intelligence in Data Science. *Appl. Sci.* **2023**, *13*, 11612. https://doi.org/10.3390/app132111612

Received: 29 September 2023
Accepted: 11 October 2023
Published: 24 October 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/).

model for the automatic textual description of images is discussed. A hybrid DL model was proposed in [15] for distributed denial of service (DDoS) attacks. The article [16] illustrates the means with which DL can be used to enhance an end-to-end table detection model. The article [17] likewise concerns a DL scheme, but is properly improved for pneumonia image classification. The authors of [18] presented an adaptive online task scheduling algorithm based on a variant of Q-learning for the efficient management of virtual machine resources in cloud computing. In [19], a variant of the data particle geometrical divide was proposed to tackle classification problems in the presence of imbalanced data. The article [20] deals with group multi-objective optimization, and the specific case study is related to portfolio optimization. In [21], a DL-based approach combined with natural language description was discussed and incorporated into a smart surveillance system. Finally, [22] presents a more general implementation of the Monte Carlo Tree Search using neural networks when there is an insufficient knowledge of the problem.

Funding: S. Tomasiello acknowledges funding from the European Social Fund via the IT Academy program and the Estonian Research Council, grant PRG1604.

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

References

1. Lukauskas, M.; Šarkauskaitė, V.; Pilinkienė, V.; Stundžienė, A.; Grybauskas, A.; Bruneckienė, J. Enhancing Skills Demand Understanding through Job Ad Segmentation Using NLP and Clustering Techniques. *Appl. Sci.* **2023**, *13*, 6119. [CrossRef]
2. Atacak, I. An Ensemble Approach Based on Fuzzy Logic Using Machine Learning Classifiers for Android Malware Detection. *Appl. Sci.* **2023**, *13*, 1484.
3. Das Mou, T.; Alam, S.B.; Rahman, M.H.; Srivastava, G.; Hasan, M.; Uddin, M.F. Multi-Range Sequential Learning Based Dark Image Enhancement with Color Upgradation. *Appl. Sci.* **2023**, *13*, 1034. [CrossRef]
4. Verma, A.K.; Saxena, R.; Jadeja, M.; Bhateja, V.; Lin, J.C.W. Bet-GAT: An Efficient Centrality-Based Graph Attention Model for Semi-Supervised Node Classification. *Appl. Sci.* **2023**, *13*, 847.
5. Stević, Ž.; Zavadskas, E.K.; Tawfiq, F.M.; Tchier, F.; Davidov, T. Fuzzy Multicriteria Decision-Making Model Based on Z Numbers for the Evaluation of Information Technology for Order Picking in Warehouses. *Appl. Sci.* **2022**, *12*, 12533.
6. Xiao, M.; Chao, Z.; Coelho, R.F.; Tian, S. Investigation of Classification and Anomalies Based on Machine Learning Methods Applied to Large Scale Building Information Modeling. *Appl. Sci.* **2022**, *12*, 6382. [CrossRef]
7. Vaiyapuri, T.; Liyakathunisa; Alaskar, H.; Aljohani, E.; Shridevi, S.; Hussain, A. Red Fox Optimizer with Data-Science-Enabled Microarray Gene Expression Classification Model. *Appl. Sci.* **2022**, *12*, 4172. [CrossRef]
8. Hong, T.P.; Lin, C.Y.; Huang, W.M.; Li, S.M.; Wang, S.L.; Lin, J.C.W. A One-Phase Tree-Structure Method to Mine High Temporal Fuzzy Utility Itemsets. *Appl. Sci.* **2022**, *12*, 2821. [CrossRef]
9. Shi, Z.Y.; Hon, J.S.; Cheng, C.Y.; Chiang, H.T.; Huang, H.M. Applying Machine Learning Techniques to the Audit of Antimicrobial Prophylaxis. *Appl. Sci.* **2022**, *12*, 2586. [CrossRef]
10. Chen, H.C.; Prasetyo, E.; Tseng, S.S.; Putra, K.T.; Kusumawardani, S.S.; Weng, C.E. Week-Wise Student Performance Early Prediction in Virtual Learning Environment Using a Deep Explainable Artificial Intelligence, Week-Wise Student Performance Early Prediction in Virtual Learning Environment Using a Deep Explainable Artificial Intelligence. *Appl. Sci.* **2022**, *12*, 1885. [CrossRef]
11. Bokaba, T.; Doorsamy, W.; Paul, B.S. A Comparative Study of Ensemble Models for Predicting Road Traffic Congestion. *Appl. Sci.* **2022**, *12*, 1337. [CrossRef]
12. Chen, C.H.; Lai, W.H.; Hung, S.T.; Hong, T.P. An Advanced Optimization Approach for Long-Short Pairs Trading Strategy Based on Correlation Coefficients and Bollinger Bands. *Appl. Sci.* **2022**, *12*, 1052. [CrossRef]
13. Boinski, T.M.; Szymanski, J.; Krauzewicz, A. Active Learning Based on Crowdsourced Data. *Appl. Sci.* **2022**, *12*, 409. [CrossRef]
14. Chang, Y.H.; Chen, Y.J.; Huang, R.H.; Yu, Y.T. Enhanced Image Captioning with Color Recognition Using Deep Learning Methods. *Appl. Sci.* **2022**, *12*, 209. [CrossRef]
15. Alghazzawi, D.; Bamasag, O.; Ullah, H.; Asghar, M.Z. Efficient Detection of DDoS Attacks Using a Hybrid Deep Learning Model with Improved Feature Selection. *Appl. Sci.* **2021**, *11*, 11634. [CrossRef]
16. Lee, S.-H.; Chen, H.-C. U-SSD: Improved SSD Based on U-Net Architecture for End-to-End Table Detection in Document Images. *Appl. Sci.* **2021**, *11*, 11446. [CrossRef]
17. Jiang, Z.P.; Liu, Y.Y.; Shao, Z.E.; Huang, K.W. An Improved VGG16 Model for Pneumonia Image Classification. *Appl. Sci.* **2021**, *11*, 11185. [CrossRef]
18. Li, K.; Peng, Z.; Cui, D.; Li, Q. SLA-DQTS: SLA Constrained Adaptive Online Task Scheduling Based on DDQN in Cloud Computing. *Appl. Sci.* **2021**, *11*, 9360. [CrossRef]

19. Rybak, L.; Dudczyk, J. Variant of Data Particle Geometrical Divide for Imbalanced Data Sets Classification by the Example of Occupancy Detection. *Appl. Sci.* **2021**, *11*, 4970. [CrossRef]
20. Fernández, E.; Rangel-Valdez, N.; Cruz-Reyes, L.; Gomez-Santillan, C. A New Approach to Group Multi-Objective Optimization under Imperfect Information and Its Application to Project Portfolio Optimization. *Appl. Sci.* **2021**, *11*, 4575. [CrossRef]
21. Dilawari, A.; Khan, M.U.G.; Al-Otaibi, Y.D.; Rehman, Z.U.; Rahman, A.U.; Nam, Y. Natural Language Description of Videos for Smart Surveillance. *Appl. Sci.* **2021**, *11*, 3730. [CrossRef]
22. Cotarelo, A.; García-Díaz, V.; Núñez-Valdez, E.R.; González García, C.; Gómez, A.; Chun-Wei Lin, J. Improving Monte Carlo Tree Search with Artificial Neural Networks without Heuristics. *Appl. Sci.* **2021**, *11*, 2056. [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.