applied sciences

MDPI

Editorial

Medical Big Data and Artificial Intelligence for Healthcare

Yudong Zhang [1,2,*], Jin Hong [3] and Shuwen Chen [4]

[1] School of Computing and Mathematic Sciences, University of Leicester, Leicester LE1 7RU, UK
[2] Department of Information Systems, Faculty of Computing and Information Technology, King Abdulaziz University, Jeddah 21589, Saudi Arabia
[3] Brain Information and Human Factors Engineering Laboratory, Zhongshan Institute of Changchun University of Science and Technology, Zhongshan 528403, China
[4] School of Physics and Information Engineering, Jiangsu Second Normal University, Nanjing 211200, China
* Correspondence: yudongzhang@ieee.org; Tel.: +44-754-870-0453

Big data have altered the way we manage, explore, evaluate, analyze, and leverage data across many different industries [1]. Medicine is one of the most remarkable fields in which big data are making big changes [2], in the form of medical big data (MBD).

A search for MBD-related publications on Web of Science shows that the number of publications has increased every year since 2000 (Figure 1). (Years 2021 and 2022 were excluded from the search since the related publications have not yet been fully indexed in the Web of Science.)

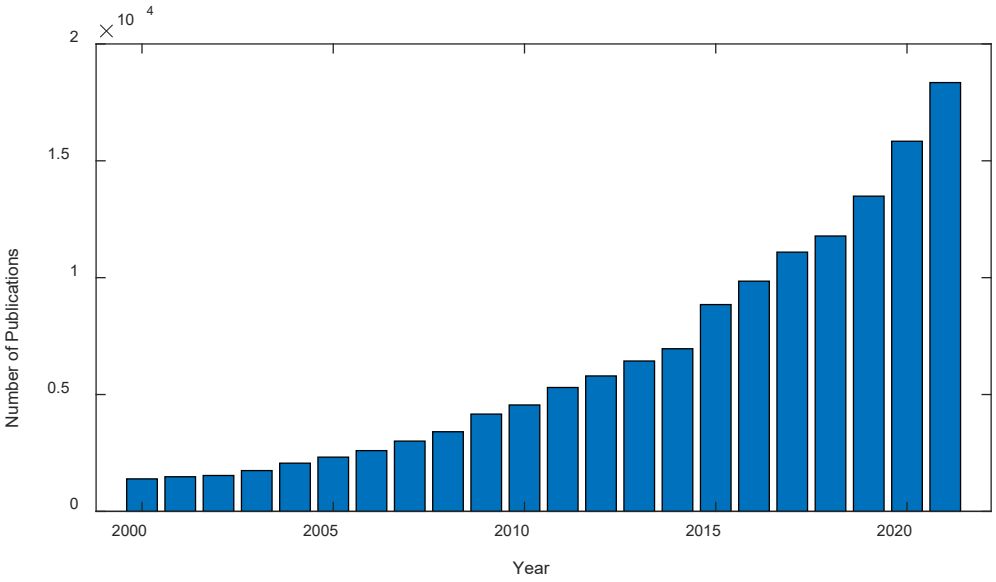

Figure 1. The number of MBD-related publications per year.

MBD [3] has the potential to benefit people in many ways. It can help healthcare professionals diagnose diseases more accurately, develop personalized treatment plans [4], and identify public health trends [5] and issues. Additionally, MBD can empower individuals to take control of their health by giving them access to their health data and allowing them to track their progress over time.

Specifically, MBD comprises patient records from various tests, X-rays [6], CT scans [7], MRI scans [8], ultrasound scans [9], the internet of medical things [10], and other diagnostic procedures, as well as electronic healthcare records [11,12] and medical and medication histories [13].

check for updates

Citation: Zhang, Y.; Hong, J.; Chen, S. Medical Big Data and Artificial Intelligence for Healthcare. *Appl. Sci.* **2023**, *13*, 3745. https://doi.org/10.3390/app13063745

Received: 16 February 2023
Revised: 9 March 2023
Accepted: 14 March 2023
Published: 15 March 2023

The growth of MBD, together with the expansion of computational models [14] in healthcare, has aided researchers and practitioners in data analysis [15]. The latest technologies will enable MBD to become a major part of data collection and processing, and a reliable form of diagnosis [16] and consultation in the future.

MBD-related research encourages exploratory medical and healthcare research [17], particularly when combined with artificial intelligence (AI) [18]. Such AI-based data-driven analysis [19] allows us to make faster progress than if we were to use a hypothesis-driven analysis [20].

There are roughly six main categories of common AI methods [21]: deep learning, machine learning, neural networks, computer vision, robotics, and natural language processing, as shown in Figure 2.

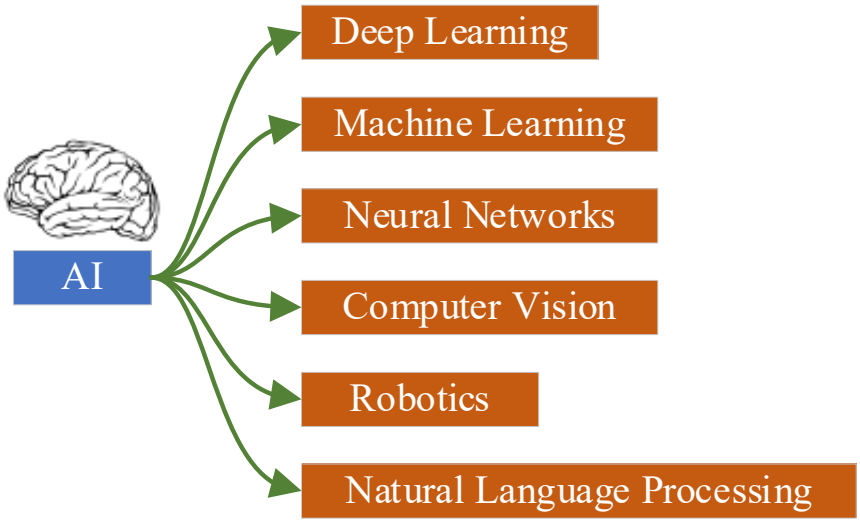

Figure 2. Six main categories of AI.

These six categories each contain many sub-categories and sub-sub-categories. For example, traditional machine learning can be further categorized into supervised learning, weakly supervised learning (WSL) [22], unsupervised learning [23], and reinforcement learning [24], as shown in Figure 3.

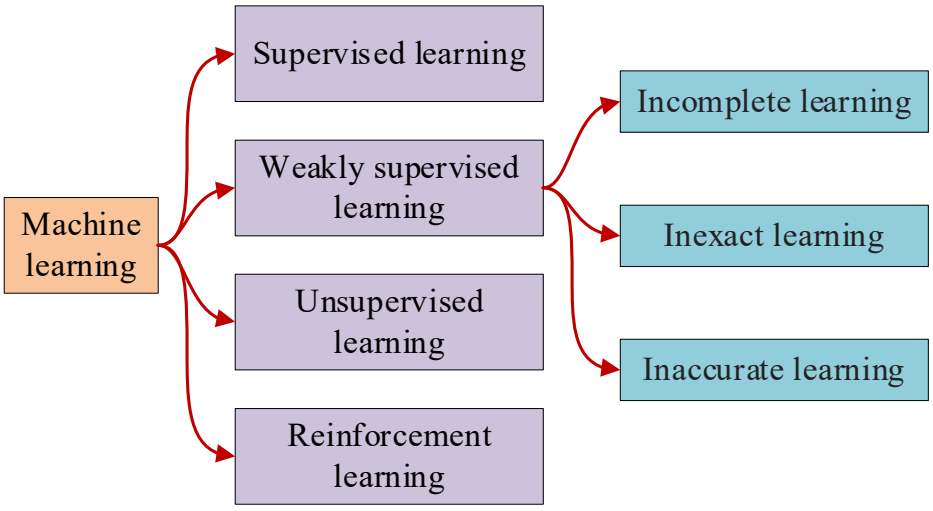

Figure 3. Sub-categories of machine learning.

If we continue to divide the sub-categories, WSL consists of three sub-sub-categories: incomplete learning, inexact learning, and inaccurate learning, as shown in Figure 3. The complexity of AI-related theories may impede the applications of AI to MBD healthcare

Deep learning (DL) is currently the most popular AI category. It uses artificial neural networks to analyze and learn from data. The "deep" aspect refers to the multiple layers of neurons used to build these neural networks.

There are many famous DL models, including convolutional neural networks, long short-term memory networks, recurrent neural networks [25], generative adversarial networks [26], deep belief networks, deep autoencoders, transfer learning models [27,28], graph neural networks [29], etc.

MBD analytics have already been used for many applications in medicine and healthcare, such as personalized medicine [30] and prescriptive analytics [31], waste reduction [32,33], clinical risk intervention [34], public response [35], and predictive analytics. MBD has also been applied in the automated reporting of external and internal patient data and the standardization of medical terms [36] and patient registries [37].

However, most MBD research thus far still focuses on either (i) traditional AI methods [38] due to the complexity of AI categories and sub-categories; or (ii) small-sized healthcare datasets involving the expensive collection of healthcare-related samples [39]. Therefore, further studies are needed to combine both MBD and recent deep learning methods [40].

Funding: This paper is partially supported by Medical Research Council, UK (MC_PC_17171); Royal Society International Exchanges Cost Share Award, UK (RP202G0230); BHF Accelerator Award, UK (AA/18/3/34220); Hope Foundation for Cancer Research, UK (RM60G0680); GCRF, UK (P202PF11); Sino-UK Industrial Fund, UK (RP202G0289); LIAS, UK (P202ED10, P202RE969); Data Science Enhancement Fund, UK (P202RE237); Fight for Sight, UK (24NN201); Sino-UK Education Fund, UK (OP202006); BBSRC, UK (RM32G0178B8).

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

References

1. Solatidehkordi, Z.; Ramesh, J.; Al-Ali, A.R.; Osman, A.; Shaaban, M. An IoT deep learning-based home appliances management and classification system. *Energy Rep.* **2023**, *9*, 503–509. [CrossRef]
2. Cerchione, R.; Centobelli, P.; Riccio, E.; Abbate, S.; Oropallo, E. Blockchain?s coming to hospital to digitalize healthcare services: Designing a distributed electronic health record ecosystem. *Technovation* **2023**, *120*, 102480. [CrossRef]
3. Shah, F.D.; Li, J.P.; Shah, Y.; Shah, F. Broad Big Data Domain via Medical Big Data. In Proceedings of the 4th International Conference on Systems and Informatics (ICSAI), Hangzhou, China, 11–13 November 2017; pp. 732–737.
4. Ajdari, A.; Liao, Z.X.; Mohan, R.; Wei, X.; Bortfeld, T. Personalized mid-course FDG-PET based adaptive treatment planning for non-small cell lung cancer using machine learning and optimization. *Phys. Med. Biol.* **2022**, *67*, 185015. [CrossRef] [PubMed]
5. Maro, A.; Asrat, H.; Qiu, W.Q.; Liang, R.A.; Sunny, S.; Aslam, S.; Abdallah, M.; Fornek, M.; Episcopia, B.; Quale, J. Trends in Clostridioides difficile infection across a public health hospital system in New York City 2019–2021: A cautionary note. *Am. J. Infect. Control* **2022**, *50*, 1389–1391. [CrossRef]
6. Awan, M.J.; Bilal, M.H.; Yasin, A.; Nobanee, H.; Khan, N.S.; Zain, A.M. Detection of COVID-19 in Chest X-ray Images: A Big Data Enabled Deep Learning Approach. *Int. J. Environ. Res. Public Health* **2021**, *18*, 10147. [CrossRef]
7. Phan, A.C.; Cao, H.P.; Le, T.N.H.; Trieu, T.N.; Phan, T.C. Improving liver lesions classification on CT/MRI images based on Hounsfield Units attenuation and deep learning. *Gene Expr. Patterns* **2023**, *47*, 119289. [CrossRef]
8. Palesi, F.; Nigri, A.; Gianeri, R.; Aquino, D.; Redolfi, A.; Biagi, L.; Carne, I.; De Francesco, S.; Ferraro, S.; Martucci, P.; et al. MRI data quality assessment for the RIN-Neuroimaging Network using the ACR phantoms. *Phys. Med.-Eur. J. Med. Phys.* **2022**, *104*, 93–100. [CrossRef] [PubMed]
9. Khizar, H.; Hu, Y.; Wu, Y.; Chen, W.; Bian, Y.; Le, C.; Huang, Z.; Ali, K.; Yang, J. Safety and efficacy of lumen-apposing metal stents and double-pigtail plastic stents for endoscopic ultrasound-guided drainage of walled-off necrosis; a systematic review and meta-analysis. *Ann. Med.* **2023**, *55*, 578–591. [CrossRef] [PubMed]
10. Wang, J.P.; Li, X.H. Secure Medical Data Collection in the Internet of Medical Things Based on Local Differential Privacy. *Electronics* **2023**, *12*, 307. [CrossRef]
11. Jin, S.J.; Chen, L.; Chen, K.; Hu, C.Z.; Hu, S.A.; Zhang, Z.H. Establishment of a Chinese critical care database from electronic healthcare records in a tertiary care medical center. *Sci. Data* **2023**, *10*, 49. [CrossRef]
12. Psarra, E.; Apostolou, D.; Verginadis, Y.; Patiniotakis, I.; Mentzas, G. Context-Based, Predictive Access Control to Electronic Health Records. *Electronics* **2022**, *11*, 3040. [CrossRef]

13. Tabassum, S.; Soomro, M.H.; Ali, A.; Ahmed, I.; Shafique, S.; Rahman, M.U.; Shah, S.H.; Memon, T.F.; Ali, M.M.; Memon, S. The Status Epilepticus in Adults: Results from a Tertiary Care Teaching Hospital in Karachi, Pakistan. *Bangladesh J. Med. Sci.* **2023**, *22*, 128–134. [CrossRef]
14. Khan, S.; Khan, H.U.; Nazir, S. Systematic analysis of healthcare big data analytics for efficient care and disease diagnosing. *Sci. Rep.* **2022**, *12*, 22377. [CrossRef] [PubMed]
15. Akhtar, M.M.; Ahamad, D.; Shatat, A.S.A.; Shatat, A.S.A. Big Data Classification in IOT Healthcare Application Using Optimal Deep Learning. *Int. J. Semant. Comput.* **2023**, *17*, 33–58. [CrossRef]
16. Bhachawat, S.; Shriram, E.; Srinivasan, K.; Hu, Y.C. Leveraging Computational Intelligence Techniques for Diagnosing Degenerative Nerve Diseases: A Comprehensive Review, Open Challenges, and Future Research Directions. *Diagnostics* **2023**, *13*, 288. [CrossRef]
17. Matharaarachchi, S.; Domaratzki, M.; Katz, A.; Muthukumarana, S. Discovering Long COVID Symptom Patterns: Association Rule Mining and Sentiment Analysis in Social Media Tweets. *JMIR Form. Res.* **2022**, *6*, e37984. [CrossRef]
18. Hulsen, T.; Petkovic, M.; Varga, O.E.; Jamuar, S.S. Editorial: AI in Healthcare: From Data to Intelligence. *Front. Artif. Intell.* **2022**, *5*, 909391. [CrossRef]
19. De Fazio, R.; Di Giovannantonio, R.; Bellini, E.; Marrone, S. Explainabilty Comparison between Random Forests and Neural Networks-Case Study of Amino Acid Volume Prediction. *Information* **2023**, *14*, 21. [CrossRef]
20. Sindermann, L.; Leehr, E.J.; Redlich, R.; Meinert, S.; Bohnlein, J.; Grotegerd, D.; Pollack, D.; Reepen, M.; Thiel, K.; Winter, A.; et al. Emotion processing in depression with and without comorbid anxiety disorder. *J. Affect. Disord.* **2022**, *314*, 133–142. [CrossRef]
21. Athanasopoulou, K.; Daneva, G.N.; Adamopoulos, P.G.; ScorilasORCID, A. Artificial Intelligence: The Milestone in Modern Biomedical Research. *BioMedInformatics* **2022**, *2*, 727–744. [CrossRef]
22. Tsuneki, M.; Kanavati, F. Weakly Supervised Learning for Poorly Differentiated Adenocarcinoma Classification in GastricEndoscopic Submucosal Dissection Whole Slide Images. *Technol. Cancer Res. Treat.* **2022**, *21*, 15330338221142674. [CrossRef]
23. Cohen, S.; Goldshlager, N.; Rokach, L.; Shapira, B. Boosting anomaly detection using unsupervised diverse test-time augmentation. *Inf. Sci.* **2023**, *626*, 821–836. [CrossRef]
24. Cui, H.; Peng, T.; Han, R.D.; Zhu, B.B.; Bi, H.J.; Liu, L. Reinforcement learning with dynamic completion for answering multi-hop questions over incomplete knowledge graph. *Inf. Process. Manag.* **2023**, *60*, 103283. [CrossRef]
25. Rezaei, M.; Mohammadifar, A.; Gholami, H.; Mina, M.; Riksen, M.; Ritsema, C. Mapping of the wind erodible fraction of soil by bidirectional gated recurrent unit (BiGRU) and bidirectional recurrent neural network (BiRNN) deep learning models. *Catena* **2023**, *223*, 106953. [CrossRef]
26. Liu, S.P.; Hu, W.L.; Xu, F.B.; Chen, W.J.; Liu, J.; Yu, X.C.; Wang, Z.F.; Li, Z.W.; Li, Z.W.; Yang, X.Y.; et al. Prediction of OCT images of short-term response to anti-VEGF treatment for diabetic macular edema using different generative adversarial networks. *Photodiagnosis Photodyn. Ther.* **2023**, *41*, 103272. [CrossRef]
27. Li, H.R.; He, F.Z.; Pan, Y.T. Multi-objective dynamic distribution adaptation with instance reweighting for transfer feature learning. *Knowl.-Based Syst.* **2023**, *263*, 110303. [CrossRef]
28. Yang, B.; Zhu, X.; Wei, B.; Liu, M.Z.; Li, Y.F.; Lv, Z.H.; Wang, F.M. Computer Vision and Machine Learning Methods for Heat Transfer and Fluid Flow in Complex Structural Microchannels: A Review. *Energies* **2023**, *16*, 1500. [CrossRef]
29. Singh, K.; Munchmeyer, J.; Weber, L.; Leser, U.; Bande, A. Graph Neural Networks for Learning Molecular Excitation Spectra. *J. Chem. Theory Comput.* **2022**, *18*, 4408–4417. [CrossRef]
30. Benucci, M.; Damiani, A.; Russo, E.; Guiducci, S.; Li Gobbi, F.; Fusi, P.; Grossi, V.; Amedei, A.; Manfredi, M.; Infantino, M. The Association of uPA, uPAR, and suPAR System with Inflammation and Joint Damage in Rheumatoid Arthritis: suPAR as a Biomarker in the Light of a Personalized Medicine Perspective. *J. Pers. Med.* **2022**, *12*, 1984. [CrossRef]
31. Rahman, M.S.; Reza, H. A Systematic Review Towards Big Data Analytics in Social Media. *Big Data Min. Anal.* **2022**, *5*, 228–244. [CrossRef]
32. Chang, C.T.; Hung, M.F. Mind the gap: Analyzing factors associated with consumers' single-use product reduction. *Sustain. Prod. Consum.* **2023**, *36*, 75–87. [CrossRef]
33. da Costa, T.P.; Gillespie, J.; Pelc, K.; Shenker, N.; Weaver, G.; Ramanathan, R.; Murphy, F. An Organisational-Life Cycle Assessment Approach for Internet of Things Technologies Implementation in a Human Milk Bank. *Sustainability* **2023**, *15*, 1137. [CrossRef]
34. Lieberman, B.; Kong, J.D.; Gusinow, R.; Asgary, A.; Bragazzi, N.L.; Choma, J.; Dahbi, S.E.; Hayashi, K.; Kar, D.; Kawonga, M.; et al. Big data- and artificial intelligence-based hot-spot analysis of COVID-19: Gauteng, South Africa, as a case study. *BMC Med. Inform. Decis. Mak.* **2023**, *23*, 19. [CrossRef]
35. Chen, F.Y.; Gao, C.C.; Gu, X.; Yue, T. Research on the Spatial and Temporal Differences in Public Response to Release-Type Communication to Stop Food Waste. *Appl. Sci.* **2023**, *13*, 736. [CrossRef]
36. Choi, S.; Joo, H.J.; Kim, Y.; Kim, J.H.; Seok, J. Conversion of Automated 12-Lead Electrocardiogram Interpretations to OMOP CDM Vocabulary. *Appl. Clin. Inform.* **2022**, *13*, 880–890. [CrossRef] [PubMed]
37. Elvas, L.B.; Serrao, C.; Ferreira, J.C. Sharing Health Information Using a Blockchain. *HealthCare* **2023**, *11*, 170. [CrossRef]
38. Gendy, M.E.G.; Yuce, M.R. Emerging Technologies Used in Health Management and Efficiency Improvement during Different Contact Tracing Phases against COVID-19 Pandemic. *IEEE Rev. Biomed. Eng.* **2023**, *16*, 38–52. [CrossRef] [PubMed]

39. Hamm, C.; Cavallo-Medved, D.; Moudgil, D.; McGrath, L.; Huang, J.; Li, Y.Y.; Stratton, T.W.; Robinson, T.; Naccarato, K.; Sundquist, S.; et al. Addressing the Barriers to Clinical Trials Accrual in Community Cancer Centres Using a National Clinical Trials Navigator: A Cross-Sectional Analysis. *Cancer Control* **2022**, *29*, 10732748221130164. [CrossRef]
40. Ramirez-Arias, F.J.; Garcia-Guerrero, E.E.; Tlelo-Cuautle, E.; Colores-Vargas, J.M.; Garcia-Canseco, E.; Lopez-Bonilla, O.R.; Galindo-Aldana, G.M.; Inzunza-Gonzalez, E. Evaluation of Machine Learning Algorithms for Classification of EEG Signals. *Technologies* **2022**, *10*, 79. [CrossRef]