Machine Learning and Artificial Intelligence

Machine learning (ML) is a type of artificial intelligence (AI) consisting of algorithmic approaches that enable machines to solve problems deprived of explicit computer programming.

1. Introduction

Machine learning (ML) is a type of artificial intelligence (AI) consisting of algorithmic approaches that enable machines to solve problems deprived of explicit computer programming [1]. ML is becoming increasingly relevant in medicine as it can optimize the trajectory of clinical care of patients affected by chronic diseases and might inform precision medicine approaches and facilitate clinical trials. As shown in Figure 1, the number of articles applying ML to the medical field has been exponentially increasing, especially with regard to diagnostics and drug discovery. According to Accenture data, vital medical health AI applications can possibly create USD 150 billion in yearly savings for the United States healthcare sector by 2026 [2]. These data show that the healthcare industry can heavily leverage the possibilities provided by ML. This might also explain why AI companies are being increasingly involved in the area of medicine, from diagnosis to treatment and drug development. For instance, convolutional neural networks (used in image recognition and processing) have been able to effectively improve the diagnostic process of diabetic retinopathy [3][4]. Another example is rehabilitation, where learning agents can be trained to run by controlling the muscles attached to the virtual skeleton. Ideally, doctors might predict if a patient is able to walk, jump, or run properly after a specific treatment. Furthermore, data obtained during phases of rehabilitation might be later used to project new, AI designed, leg prostheses.

AI uses multiple layers of non-linear processing units to "teach" itself how to understand data, classify the records, or make predictions [5]. Thus, AI can produce electronic health records (EHRs) data and unstructured facts to make predictions about a patient's health. For instance, AI can rapidly read a retinal image or flag cases for follow up when several manual reviews would be too cumbersome [6].

When applied to big data, AI offers the promise of unlocking novel insights and accelerating breakthroughs. Paradoxically, although an unprecedented quantity of data is becoming available, only a fraction is being properly integrated, understood, and analyzed. The challenge lies in harnessing high volumes of data, integrating them from hundreds of sources, and understanding their various formats. AI offers potential for addressing these challenges since cognitive answers are explicitly intended to integrate and analyze big datasets. AI can understand diverse types of data such as lab calculations in a structured database or the script of a scientific publication. These software solutions are trained to understand technical, industry-specific content and use advanced reasoning, predictive modelling, and ML techniques to advance research.
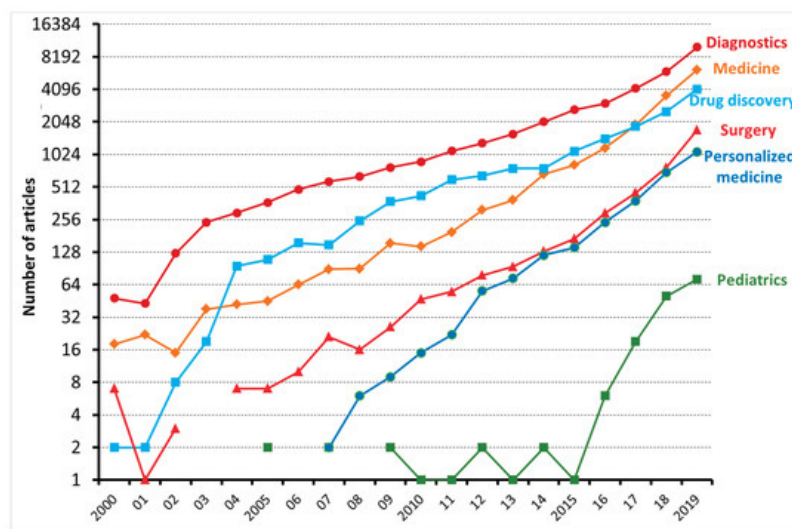
Figure 1. The number of articles, reviews, and editorials, dealing with machine learning and either diagnostics, medicine, drug discovery, surgery, personalized medicine, and pediatrics, published between 2000 and 2019 and indexed on the Web of Science.

2. Applications of ML in Medicine

2.1. Diagnostic Process

The ability of ML to detect diagnostic models reaching the level of clinical accuracy remains an objective not yet achieved, but seemingly feasible. This objective faces the challenge of finding ways to work with all the available data. This highlights the relevance of interdisciplinary collaborative work. In the area of brain diseases like depression, the Predicting Response to Depression Treatment (PReDicT) project has applied predictive analytics to help diagnose depression and identify the most effective treatment, with the overall goal of producing a commercially available emotional test battery for use in clinical settings [7]. In general, the use of ML to aggregate large datasets could significantly accelerate the diagnostic processes [8]. In Table 1, we have summarized information on ML in medicine.

Table 1. Applications of machine learning (ML) in medicine.

Application	Areas
Diagnostic testing	Personalized diagnostics Parkinson's disease progression prediction from mobile phone accelerometer data Predict viral failure in AIDS patients
Medical imaging	Clinical research: MRI and PET scans and deep learning Cellular image analysis: genotype, phenotype, classification, identification, cellular tracking
Oncology	Clinical research: Identify which genes are associated with breast cancer relapse. Prognosis: Predict probability of survival in 5 years
Remote patient monitoring	Real-time predictions using data from wearables Medication adherence monitoring

Of the numerous opportunities for the use of ML in clinical practice, medical imaging workflows are those that will be likely be most impacted in the near term. ML-driven algorithms that automatically process two- or three-dimensional image scans to recognize clinical signs (e.g., tumors or lesions) or articulate diagnoses are now available and some are progressing through regulatory steps toward the market [9]. Many of these use deep learning, a form of ML based on layered representations of variables, referred to as artificial neural networks. The latter can learn extremely complex relationships between features and labels and have been shown to exceed human abilities in performing tasks such as classification of images.

ML can improve diagnostic accuracy by analyzing not only medical images but also textual records. Indeed, ML allowed the identification of varicella cases in a pediatric Electronic Medical Record Database with a positive predictive value of 63.1% and a negative predictive value of 98.8% [10].

2.2. Predicting Prognosis

ML has been shown to achieve the same or better prognostic definition in several clinical conditions, as compared to conventional statistical methods. In particular, ML can better predict clinical deterioration in the ward [11], mortality in acute coronary syndrome [12], survival in patients with epithelial ovarian cancer [13], complications of bariatric surgery [14], and risk of metabolic syndrome [15]. On the other hand, other studies reported that ML and conventional statistical methods have similar prognostic usefulness in predicting mortality in intensive care units [16], readmission in patients hospitalized for heart failure [17], and all-cause mortality and cardiovascular events [18].

2.3. Drug Discovery

ML can facilitate various phases of the early stages of drug discovery, from initial screening of drug compounds to predicted success rates based on biological factors. This includes R&D technologies like next-generation sequencing. Precision medicine, which relies on the recognition of pathophysiological mechanisms and might serve the development of alternative therapeutic pathways, appears as the most innovative area. Much of this study encompasses unsupervised learning, which is in large part still limited to identifying patterns in data without predictions (the latter is still in the realm of supervised learning). Data from experimentation or manufacturing processes have the potential to aid pharmaceutical manufacturers to diminish the time required to produce drugs, leading to lowered costs and better replication. Adopting ML approaches could play a significant role in discovering new molecules or repurposing existing drugs for rare conditions or epidemics where urgency is key. With the increase in antibiotic resistance, exploiting ML techniques is already proving quite powerful in identifying new antibacterial agents in a faster and potentially inexpensive way [8]. For example, AI recently allowed the discovery of halicin, a compound structurally divergent from conventional antibiotics, acting against Clostridium difficile and pandrug-resistant Acinetobacter baumannii infections in murine models [19].

2.4. Personalized Treatment

Personalized medicine, which should lead to the identification of more effective treatment based on individual health data paired with predictive analytics, is closely related to better disease assessment. To meet the complexity of personalized medicine, new types of trials have been developed, such as basket, umbrella, or platform trials. The area is presently governed by supervised learning, which permits physicians, for instance, to select from further limited sets of diagnoses or estimate patient risk based on symptoms and genetic information.

Over the next decade, the increased use of micro biosensors and devices, as well as mobile apps with more sophisticated health measurement and remote monitoring capabilities, will provide an additional surge of data that can be used to help facilitate research and development, and treatment efficacy. This type of personalized treatment has significant consequences for the individual in terms of health optimization, but also for plummeting overall healthcare costs. If more patients adhere to following prescribed drug or treatment tactics, for instance, the reduction in health care charges will trickle up and back down.

Using ML in these settings depends on the collection and analysis of huge amounts of data, but with the emergence of big data comes the challenge of statistical inference from complex datasets to identify genuine patterns, while also restraining false classifications and making decisive judgments on diagnosis and treatment possibilities. Statistical bioinformatics has proven very useful in proteomic and genomic data analysis, and the adoption of ML to build predictors and classifiers has shown significant potential [8].

3. Conclusions

ML has the potential to transform the way medicine works [20]. However, increased enthusiasm has previously not been met by a corresponding interest from healthcare providers and operators.

There is no clear line between ML models and traditional statistical models, and a recent article summarizes the relationship between the two [21]. However, sophisticated new ML models (e.g., those used in "deep learning" [22][23]) are well suited to learn from the complex and heterogeneous kinds of data that are generated from current clinical care, such as medical notes entered by doctors, medical images, continuous monitoring data from sensors, and genomic data to aid make therapeutically significant predictions. Most ML classifiers perform uncertainly with risk prediction. Possibly much bigger sample sizes are required to gain reliable (calibrated) risk predictions [24]than reliable (diagnostic) classifications.

ML is creating a paradigm shift in medicine, from basic research to clinical applications, but it should be carefully implemented. Vulnerabilities such as security of data and adversarial attacks, where malicious manipulation in the input can affect a complete misdiagnosis, which could be employed for fraudulent interests, present a real threat to the

technology [8]. However, these vulnerabilities can be met with adequate efforts.

In the 1970s and 1980s, computerized tomography, based on the automatic elaboration of a huge bulk of X-rays images, revolutionized radio diagnostics, enabling radiologists to overcome the so-called "grey barrier". The use of CT allowed radiologists to improve their role in the healthcare system. However, the ML revolution seems to threaten one of physicians' most exclusive tasks, i.e., diagnostic activity. The new generation of practitioners should accept the challenge of ML, by learning how to comprehend, develop, and eventually, control it so as to improve patient care [9].

ML can analyze large amounts of data and turn that information into functional tools that can assist both doctors and patients. The increased integration of ML into everyday medical applications might improve the efficiency of treatments and lower costs in various ways. The challenge is to combine big data provided by genomics, transcriptomics, proteomics, and metabolomics with complex systems science, systems biology, and systems medicine of the body [25]. ML tools can be built for system-level interventions, comprising improving patient selection and enrolment for clinical trials, decreasing patient readmission, and automated follow-up of patients for scrutiny of complications.

References

1. Ngiam, K.Y.; Khor, I.W. Big data and machine learning algorithms for health-care delivery. Lancet Oncol. 2019, 20, e262–e273.

2. Stanfill, M.H.; Marc, D.T. Health Information Management: Implications of Artificial Intelligence on Healthcare Data and Information Management. Yearb. Med. Inform. 2019, 28, 56–64.

3. Arcadu, F.; Benmansour, F.; Maunz, A.; Willis, J.; Haskova, Z.; Prunotto, M. Deep learning algorithm predicts diabetic retinopathy progression in individual patients. NPJ Digit. Med. 2019, 2.

4. Lam, C.; Yi, D.; Guo, M.; Lindsey, T. Automated Detection of Diabetic Retinopathy using Deep Learning. AMIA Jt. Summits Transl. Sci. 2018, 2017, 147–155.

5. Hassabis, D.; Kumaran, D.; Summerfield, C.; Botvinick, M. Neuroscience-Inspired Artificial Intelligence. Neuron 2017, 95, 245–258.

6. Lee, C.S.; Tyring, A.J.; Wu, Y.; Xiao, S.; Rokem, A.S.; DeRuyter, N.P.; Zhang, Q.; Tufail, A.; Wang, R.K.; Lee, A.Y. Generating retinal flow maps from structural optical coherence tomography with artificial intelligence. Sci. Rep. 2019, 9, 1–11.

7. Kingslake, J.; Dias, R.; Dawson, G.R.; Simon, J.; Goodwin, G.M.; Harmer, C.J.; Morriss, R.; Brown, S.; Guo, B.; Dourish, C.T.; et al. The effects of using the PReDicT Test to guide the antidepressant treatment of depressed patients: Study protocol for a randomised controlled trial. Trials 2017, 18, 558.

8. Chen, C. Ascent of machine learning in medicine. Nat. Mater. 2019, 18, 407.

9. Saria, S.; Butte, A.; Sheikh, A. Better medicine through machine learning: What's real, and what's artificial? PLoS Med. 2018, 15.

10. Lanera, C.; Berchialla, P.; Baldi, I.; Lorenzoni, G.; Tramontan, L.; Scamarcia, A.; Cantarutti, L.; Giaquinto, C.; Gregori, D. Use of Machine Learning Techniques for Case-Detection of Varicella Zoster Using Routinely Collected Textual Ambulatory Records: Pilot Observational Study. JMIR Med. Inform. 2020, 8, e14330.

11. Churpek, M.M.; Yuen, T.C.; Winslow, C.; Meltzer, D.O.; Kattan, M.W.; Edelson, D.P. Multicenter comparison of machine learning methods and conventional regression for predicting clinical deterioration on the wards. Crit. Care Med. 2016, 44, 368.

12. Hernesniemi, J.A.; Mahdiani, S.; Tynkkynen, J.A.; Lyytikäinen, L.-P.; Mishra, P.P.; Lehtimäki, T.; Eskola, M.; Nikus, K.; Antila, K.; Oksala, N. Extensive phenotype data and machine learning in prediction of mortality in acute coronary syndrome–the MADDEC study. Ann. Med. 2019, 51, 156–163.

13. Paik, E.S.; Lee, J.-W.; Park, J.-Y.; Kim, J.-H.; Kim, M.; Kim, T.-J.; Choi, C.H.; Kim, B.-G.; Bae, D.-S.; Seo, S.W. Prediction of survival outcomes in patients with epithelial ovarian cancer using machine learning methods. J. Gynecol. Oncol. 2019, 30.

14. Nudel, J.; Bishara, A.M.; de Geus, S.W.L.; Patil, P.; Srinivasan, J.; Hess, D.T.; Woodson, J. Development and validation of machine learning models to predict gastrointestinal leak and venous thromboembolism after weight loss surgery: An analysis of the MBSAQIP database. Surg. Endosc. 2020.

15. Shimoda, A.; Ichikawa, D.; Oyama, H. Prediction models to identify individuals at risk of metabolic syndrome who are unlikely to participate in a health intervention program. Int. J. Med. Inform. 2018, 111, 90–99.

16. Kim, S.; Kim, W.; Park, R.W. A comparison of intensive care unit mortality prediction models through the use of data mining techniques. Healthc. Inform. Res. 2011, 17, 232–243.

17. Frizzell, J.D.; Liang, L.; Schulte, P.J.; Yancy, C.W.; Heidenreich, P.A.; Hernandez, A.F.; Bhatt, D.L.; Fonarow, G.C.; Laskey, W.K. Prediction of 30-day all-cause readmissions in patients hospitalized for heart failure: Comparison of machine learning and other statistical approaches. JAMA Cardiol. 2017, 2, 204–209.

18. Suzuki, S.; Yamashita, T.; Sakama, T.; Arita, T.; Yagi, N.; Otsuka, T.; Semba, H.; Kano, H.; Matsuno, S.; Kato, Y. Comparison of risk models for mortality and cardiovascular events between machine learning and conventional logistic regression analysis. PLoS ONE 2019, 14, e0221911.

19. Stokes, J.M.; Yang, K.; Swanson, K.; Jin, W.; Cubillos-Ruiz, A.; Donghia, N.M.; MacNair, C.R.; French, S.; Carfrae, L.A.; Bloom-Ackerman, Z. A deep learning approach to antibiotic discovery. Cell 2020, 180, 688–702.

20. Darcy, A.M.; Louie, A.K.; Roberts, L.W. Machine learning and the profession of medicine. JAMA J. Am. Med. Assoc. 2016, 315, 551–552.

21. Beam, A.L.; Kohane, I.S. Big data and machine learning in health care. JAMA J. Am. Med. Assoc. 2018, 319, 1317–1318.

22. Lecun, Y.; Bengio, Y.; Hinton, G. Deep learning. Nature 2015, 521, 436–444.

23. Hinton, G. Deep learning—A technology with the potential to transform health care. JAMA J. Am. Med. Assoc. 2018, 320, 1101–1102.

24. Van Calster, B.; McLernon, D.J.; Van Smeden, M.; Wynants, L.; Steyerberg, E.W.; Bossuyt, P.; Collins, G.S.; MacAskill, P.; McLernon, D.J.; Moons, K.G.M.; et al. Calibration: The Achilles heel of predictive analytics. BMC Med. 2019, 17.

25. Sekhar, H.; Rajula, R.; Mauri, M.; Fanos, V. Scale-free networks in metabolomics. Bioinformation 2018, 14, 140–144.