WWW.IJRPR.COM

International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

Application of Machine Learning in Healthcare

Divyansh Paliya[1], Dr. Vishal Shrivastava[2], Dr. Akhil Pandey[3], Dr. Karuna Sharma[4], Dr. Ashok Kumar Kajla[5]

[1]B.Tech. Scholar, [2-5]Professor
Department of Artificial Intelligence & Data Science
Arya College of Engineering and I.T. Jaipur, India
divyanshpaliya720@gmail.com[1]
vishalshrivastava.cs@aryacollege.in[2], akhil@aryacollege.in[3],
skaruna.cs@aryacollege.in[4], ashokkajla@aryacollege.in[5]

ABSTRACT :

The rapid ascent of machine learning (ML) has spurred a radical change in the healthcare industry, allowing innovative strategies to be developed in order to tackle complex issues. This research study examines several implementations of ML in healthcare such as diagnostic precision enhancement, customization of treatment plans, optimization of resource allocation, and predictive analytics support. Machine learning has greatly contributed to the betterment of clinical decision-making, speeding up the discovery of drugs, and to real-time monitoring of patients with the use of wearable technologies. This incorporates methods such as profound learning, normal dialect handling, and fortification learning. The current research signifies the potential of machine learning in analysing vast medical data sets, which has tackled the problems of inefficiency and human error associated with traditional systems. In consideration of responsible implementation, this manuscript includes ethical factors including data privacy, bias with algorithms, and transparency in the framework. Focusing the research on the role played by machine learning in offering efficient, personalized, and accessible healthcare, this report includes practical applications and current progressions in the area.

Keywords: - Machine Learning, Healthcare Technology, Predictive Analytics, Diagnostic Accuracy, Personalized Medicine, Natural Language Processing (NLP), Medical Imaging, Health Monitoring Systems, Data Privacy in Healthcare, AI in Medicine, Deep Learning in Healthcare, Reinforcement Learning, Healthcare Decision Support

Introduction

The healthcare industry faces a wide range of different challenges, such as rising costs, low efficiency, and inequalities in patient care services. Recent developments in Machine Learning (ML) have produced very promising evidence in approaching these problems. As a special case of artificial intelligence, ML allows computers to adapt to large data sets, making way for predictions, pattern recognition, and maximization of processes.

This research paper addresses the diverse and multifaceted uses of machine learning, or ML, in the healthcare industry. The primary aim of this research is to provide a thorough and comprehensive overview of how ML is being used in a variety of fields, but not limited to, disease diagnosis, medical imaging, developing personalized treatment plans, and healthcare resource allocation. This paper also considers the ethical issues and critical concerns arising from the extensive use and integration of ML technologies in the healthcare sector. Some of the issues are, but not limited to, data privacy, the potential for bias inherent in algorithms, and the need for transparency in decision-making. Through critical examination of the benefits as well as the limitations of such technologies, this paper is trying to provide a comprehensive and balanced report on the revolutionizing impact being brought about by ML on the healthcare sector as a whole.

The ubiquitous and constantly increasing mass of medical information that we face today, from complete electronic health records (EHR) to advanced medical imaging and complete genomic data, has created a climate where it has become increasingly more difficult for older data analysis systems to keep up with this increase. In such a scenario, machine learning programs have shown a remarkable capacity for performing better at analysing large and complex data sets. These state-of-the-art algorithms give health care professionals beneficial and actionable insights that can eventually result in not only faster, but much more accurate diagnoses. In addition, these insights help create treatments that are more customized and tailored to patient needs, which in turn leads to overall improved outcomes in patients being treated.

Despite the potential to change the industry that it holds, the application of machine learning to the healthcare industry is far from challenge-free. There are some issues that are of an acute nature, e.g., bias in algorithms that can lead to unequal treatment for patients, the issue of privacy of data that raises serious issues of patient confidentiality, and the need for a transparent process for AI decision making. These matters need to be given close and careful scrutiny in order to allow these new tools to be utilised in a responsible and morally sound manner. In addition to this, there is still an urgent need for intersector collaboration between healthcare specialists, technologists, and policy makers. The reason for such collaboration is essential in order to provide an atmosphere in which the application of machine learning can easily be integrated in clinical practice as effectively as possible, ultimately supporting patient care and outcomes.

With such challenges and possibilities, the use of ML in healthcare has been a matter of greater importance and research. It is important to understand its existing applications, its limitations, and its future trends so that we can utilize its best potential to revolutionize healthcare.

The increasing incidence and availability of electronic health records (EHR), along with unprecedented computing horsepower, and ubiquitous availability of big data, have all contributed significantly to the universal uptake of machine learning (ML) technologies by the healthcare industry. Through the strategic application of a combination of ML algorithms, doctors and health workers are currently in a position to attain unprecedented diagnostic performance improvement, the quality of patient care, as well as hospital operational efficiency improvement beyond even what was dreamt of before. In addition, ML-driven applications incorporated into wearable health monitors and telemedicine platforms have changed the dynamics of remote healthcare provision, enabled real-time monitoring and tracked of patients, and the early diagnosis of diseases, which can result in improved health outcomes.

This paper discusses many applications that machine learning has applied in health care, such as disease diagnostics, treatment planning, enhancing operational efficiency, and predictive analytics. Considering its current use in health care as well as the ethical matters and direction for the future, it provides a rich overview regarding how machine learning is contributing to the course of events in health-care delivery. The goal is to focus on the many opportunities and challenges that exist in this highly dynamic field, highlighting ways in which these factors can be leveraged to improve patient care and overall healthcare system performance.

Overview of Machine Learning in Healthcare

Machine learning majorly divided into two types- supervised learning and unsupervised learning. All these algorithms use mathematical frameworks from both categories.

The algorithm is designed to enable computers to acquire the capability to execute specific tasks autonomously.

Machine Learning Techniques

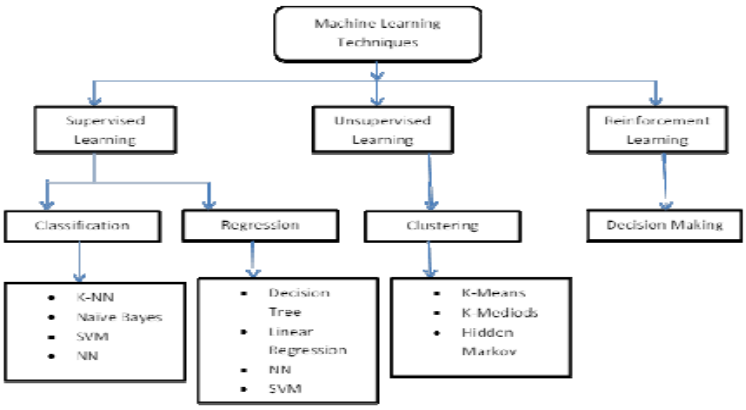


Figure1. Classification of ML Techniques

Supervised learning

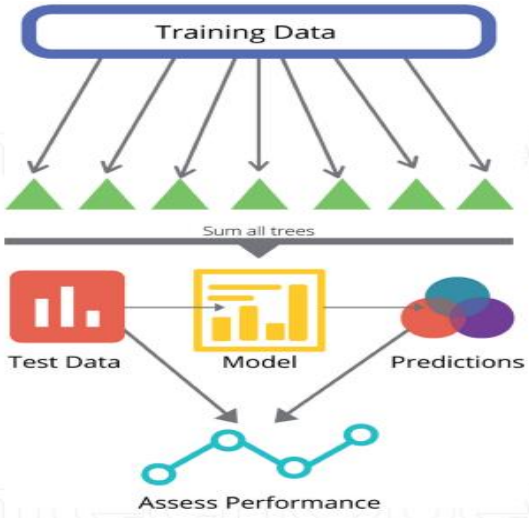
Supervised learning is perhaps the most common machine learning (ML) method applied in medicine, which allows computers to learn from labelled data to make predictions or classifications. Supervised learning algorithms are trained on input-output pairs, where medical data is input and known outcomes are output, such as disease diagnosis, patient risk classification, or the effectiveness of a treatment. Supervised learning models can generalize knowledge by looking at patterns in past data and apply it to new, unseen medical cases to help healthcare professionals make decisions, make early diagnoses, and implement individualized treatment plans.

Diagnosis and detection of diseases is perhaps one of the most important uses of supervised learning in the medical field. Machine learning systems with millions of clinical histories, laboratory findings, and imaging inputs can pinpoint patterns of diseases with high accuracy. For example, in the detection of cancer, supervised learning machines are given thousands of annotated mammograms to recognize tumours early. By separating out malignant from benign tumours, these models enable more precise diagnoses on the part of radiologists and fewer false positives and false negatives. In another area, the field of cardiology, supervised learning models look at electrocardiograms (ECGs) to identify diseased heart rhythm, enabling precocious treatment for conditions like atrial fibrillation.

Another important area in which supervised learning is making itself felt is that of personalized medicine. Conventional treatment protocols employ a generic one-size-fits-all strategy that fails to accommodate individual differences, but supervised learning allows for the creation of personalized therapy that is specifically tailored to individual patient specifications. In cancer, for example, machine learning models trained on large genomic and clinical data are capable of efficiently predicting how an individual patient will likely react to a specific chemotherapy regimen. Through close monitoring of genetic mutations and biomarker signatures, these sophisticated models allow oncologists to make the best current treatment choice, thereby reducing side effects while, at the same time, maximizing survival rates for patients. Personalized medication prescriptions based on supervised learning principles

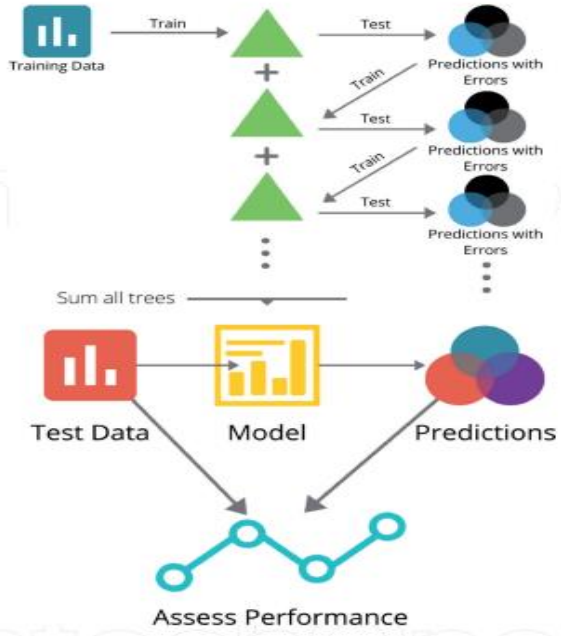
are also becoming commonly prevalent in the field of pharmacology, ensuring that medication administered is precisely developed to fit the unique genetic constitution of each patient.

Looking to the future, supervised learning is in general likely to go on contributing to the healthcare sector. This is done through its incredible ability of being able to make early disease detection possible, improve the precision of diagnosis, and give customized treatment protocols to individual patients. As quality medical data comes by in increasing measure, and machine learning algorithms continue to get stronger and more effective, supervised learning models are bound to be even more accurate. These models will become ever more efficient and capable of solving and solving complicated medical issues that occur. By integrating the cautious learning gained through machine learning with the invaluable experience and intuition of human healthcare experts in a highly efficient manner, the healthcare sector can embark on a journey towards a more data-oriented model of healthcare. This is bound to raise efficiency, as well as deliver a patient-centered model of medical care. Finally, these technologies have the ability of drastically improving health outcomes not just locally, but on the global level, for patients worldwide.



Unsupervised learning

The framework will take the choice by itself or maybe prepare on the premise of a few datasets. No labelling is given to the system that can be utilized for expectations. Unsupervised learning can be used to recover the covered-up design with the assistance of feature learning of the given data.



The clustering is an unsupervised learning approach that is utilized to partition the inputs into clusters. These clusters are not recognized prior. It builds bunches on the premise of likeness.

Semi-supervised learning

In Semi-supervised learning, the framework is expected to be partial preparing information. This sort of preparing is utilized with some trained information that can target a few lost comes about. This type algorithm is utilized on untagged information for preparing commitment.
The semi-supervised learning calculation prepared on both labelled and unlabelled information and this learning shows the highlights the features of both the unsupervised-learning and supervised learning calculations.

Historical Context and Evolution

The application of machine learning in the healthcare sector started way back in the 1980s, whose main applications were basic algorithms that were mainly for statistical analysis and as a decision support in clinical settings. Prime systems such as expert systems, based on established rules to assist doctors, were not sufficient in their abilities to evolve based on new information.
This advancement in ML was found when algorithms of high complexity joined with increased availability of high-volume health care data. Further development of robust computation powers, along with high utilization of electronic health records (EHR), facilitated high amounts of data collection relevant to analysis by ML. ML, therefore, accelerated pace in different applications starting from diagnostic imaging, surveillance of patients, and in prediction analytics in the beginning years of the 2000s.
Since then, there was a rapid evolution in ML methodology. Deep learning, NLP, and reinforcement learning showed exceptional promise. Deep learning form of neural network-based learning dramatically improved the accuracy rate in image recognition. Further, NLP transformed clinical unstructured text data-like doctors' notes-in ways that are processed and consequently used in decision-making.

Current Trends and Developments

Today, the field of healthcare is rapidly accelerating in terms of machine learning, pushed by progress in artificial algorithms, increased access to data, and greater computational capability. Among the trends apparent in that space are:
Integrating ML models with EHR systems is occurring to enable real-time decision support, prediction of patient outcomes, and pattern discovery in health data that may guide care strategies.
Precision Medicine: Machine learning is enabling personalized treatment strategies tailored to individual patients. By analysing genetic data, lifestyle factors, and medical history, ML algorithms can recommend specific treatments and predict patient responses to various therapies.
Medical Imaging and Diagnostics: ML, in deep learning particularly, has been drastically successful in the analysis of medical images including X-rays, MRIs, and CTs in identifying anomalies, such as tumours or fractures, usually better than radiologists. In this sense, these models guarantee that health service providers are able to produce much more timely and accurate diagnoses.
This acceleration has been done in machine learning algorithms about forecasting the effectiveness of compounds used as drugs, to know the side effects from them, and to get newer candidates for further development which result in the accelerated development of the therapy and treatments.
Real-time Patient Monitoring: This is being done using the combination of wearables, connected devices, and ML algorithms for continuous monitoring of patient vitals and health parameters. With these devices, one can predict health issues before they are critical, and therefore one can intervene preventively.

Machine Learning in Disease Diagnosis and Prediction

Machine learning (ML) has shown great promise in transforming disease diagnosis and prediction, enabling faster, more accurate, and personalized healthcare solutions. This section explores how ML is utilized in diagnosing various diseases, its role in predicting disease outcomes, and the technologies and methodologies that have been instrumental in this transformation.

Early Diagnosis and Screening

It's no secret that Machine Learning in supervised learning models has been helpful in the early stages of detecting certain illnesses. By analysing patient data and trying to look for a pattern, such models are able to recognize early signs of risks such as cancer or cardio or neuro degenerative diseases.
Cancer: Cancer on the other hand does not have such successful ML deep learning algorithms; This works for mammograms, cut scans and MRIs in trying to find tumours on the upper surface first. Reason being, It can identify such characteristics, fusion patterns that the average human eye cannot along with a decrease in false positives. For example, the efficiency with which ML techniques are able to detect breast cancer has also been reported to be much superior to that of the radiologists.
About Cardiovascular diseases: It can be said that ML algorithms have been working tremen-dously well since they were able to recognize arrhythmias and other coronary artery disease from electrocardiograms as well as echocardiograms. With decision tree and SVM methods, it has been possible to ascertain the chances of a stroke or heart attack in an individual depending on the patient’s history along with his imaging diagnostic methods.

Case Studies: Cancer, Cardiovascular Diseases, and Diabetes

ML has several potential real-world applications when it comes to diagnosing and predicting various kinds of diseases, and this is changing the way healthcare experts manage and diagnose these diseases.

In one case study, Google Health illustrated how screening for breast cancer could be performed faster and with more accuracy than ever before with the help of deep learning models. It was able to pick up more instances of cancerous lesions on mammograms than the radiologists and had fewer false alarms, and therefore a reduction in unnecessary biopsies, thus benefiting the patients.

Heart Disease Risk Prediction by SVM: In prevention of heart disease, one research that worked with Cleveland Heart Disease dataset applied Support Vector Machines to assess patient risk of heart disease from various demographics, clinical history and diagnostic tests. The ML model achieved quite a level of accuracy in predicting at risk patients and was actually quite helpful for suggesting suitable measures to such patients.

Diabetes Prediction Using Random Forests: A paper that was dedicated to Type 2 diabetes prediction onset time by random forest algorithms based on BMI, age and lifestyle factors was able to predict with accuracy. The model was used to suggest nutritional and physical interventions, demonstrating how ML can be integrated in the management of diseases before they progress to advanced stages.

NECESSITY OF MACHINE LEARNING IN HEALTHCARE

The use of machine learning (ML) in healthcare is not only a technological development but a necessity with the increased complexity of medical information, improved patient outcomes being required, and the necessity of efficiency in healthcare systems. Traditional healthcare procedures, while effective, are typically rooted in human instinct, expert judgment, and customary patterns of care, and may be inefficient, lead to misdiagnosis, and cause life-saving interventions to be delayed. With the exponentially increasing medical knowledge, clinical information, and diagnostic technology, healthcare professionals require intelligent systems that are able to handle large volumes of data, recognize patterns, and produce actionable information that is in excess of human intellectual capability.

Aside from its significant roles in disease prevention as well as accurate diagnosis, machine learning is significant in the optimization of hospital operations as well as optimal utilization of resources in hospitals. Health institutions are often confronted with the largest challenges such as human resource shortages, which have a negative effect on the treatment of patients, as well as far too long waiting times for patients being treated, and wasteful use of the very critical medical resources required for the smooth operation of these institutions. Machine learning-based predictive models are very crucial in the optimization of workflow management in hospitals; they achieve this by precisely predicting the rate of patient admissions, hence enabling hospitals to plan. The models also assist in the optimization of bed occupancy, with every available bed utilized efficiently, and in the optimization of the efficiency of scheduling operations and other medical procedures. This practice, therefore, ensures that healthcare providers are able to provide timely and effective treatment to patients while simultaneously reducing the operational expenses and the workload on the medical practitioners who work day and night within these institutions.

Another critical requirement of the discipline of machine learning in medicine is its core function in medical diagnostic and imaging techniques. Fields like radiology, pathology, and dermatology rely heavily on the analysis of images, where subtle differences seen in medical scans may indicate serious medical conditions that require treatment. Human analysis of such images is certainly valuable and useful, but it needs to be stated that it is prone to issues like fatigue, variability, and errors that can creep in while analysing. In contrast to this, image recognition software based on machine learning has been proven to have a much higher level of accuracy in identifying abnormalities visible in X-rays, MRIs, CT scans, and histopathological slides. Such sophisticated algorithms not only help radiologists diagnose faster and more accurately, but also help in automating mundane diagnostic routines. This automation helps medical professionals free up their attention for more complicated cases requiring their expertise and fine judgment.

In summary, it is worth appreciating that machine learning has evolved from a luxury to an absolute necessity in today's healthcare landscape. The technology's staggering ability to process huge volumes of medical data at an extremely high speed and efficiency greatly improves the accuracy of diagnosis, tailors' treatments to specific patients, automates various hospital operations, and provides remote healthcare possibilities. Such multiple uses of machine learning render it an absolute necessity in the continuously changing realm of medicine. As the healthcare demands continue to surge at an unprecedented level and the intricacies of medical issues escalate, it is clear that the use of machine learning will not only continue but also grow substantially. Such growth will continue to fuel further innovations that improve the quality of patient care, lower overall healthcare expenditure, and ultimately lead to saving more lives.

Healthcare Resource Optimization Using Machine Learning

One of the most particularly pressing issues for health care systems all over the world, especially today in the conditions of growing demands is healthcare resource optimization considering the non-growing or even shrinking resources such as medical personnel, technical equipment, or number of hospital beds. These emerging technologies open new horizons in health policies in terms of providing better management and optimization of health care resources ensuring the efficient delivery of services. Below, we detail how ML is applied in various aspects of healthcare resource optimization:

Predictive Analytics for Hospital Operations

Thanks to the integration of machine learning algorithms, predictive analytics has made it possible to gauge the demand for healthcare services. By analysing the history of patient admissions as well as seasonality patterns, it would be possible to forecast the referrals to the emergency room and the demand for hospitalization; therefore, healthcare providers plan beforehand by making sure that there is adequate staff as well as equipment to satisfy the required number of beds.

Example: ML-based models in ICU can make predictions about the necessity for critical interventions, such as intubation or support to organs, based on data from patients, vital signs, lab results, and historical patient trends.

More accurate triaging of patients, better management of critical care resources, faster intervention for high-priority cases, and improved patient outcomes.

Machine Learning in Remote Patient Monitoring and Telemedicine

Remote patient monitoring and telemedicine enable the provision and the follow up of the patients by the use of devices outside of the official health care centres.

These systems are all enhanced by machine learning in such a way that predictive analytic tools, anomaly detection systems and recommendation systems for health are provided.

In this category, real-time measurements of medical parameters including, heart rate, blood pressure, glucose pressure and oxygen saturation of Smart watches, fitness wrist reinforcers and more specialized medical devices are being gathered.

ML algorithms are utilized in this field for stream patterning, anticipating potential health complications and notifying the relevant health professionals in advance.

From Diabetes and hypertension to heart failure, ML predicts the progression of chronic conditions accurately.

The patient's records are utilized in predicting a future exacerbation and the risk of an emergency event is minimized by addressing the issue promptly.

The Natural Language Processing in telemedicine platforms enhances symptom assessment and aid those in their decision making by structuring conversations of patients and taking out useful health information.

The first point of triage is done by the use of machine learning chatbots and virtual assistants who also direct the patient to the appropriate care pathways. having noticed any change from the normal pattern, the systems are able to alert the doctor on giving any patient having deviated automatic

Ethical Considerations and Challenges in Machine Learning in Healthcare

The accepted perspectives and biases of ML models can be equated to the quality of data that is used to train themselves. If that data is not representative enough or is biased in any way, the ML systems would only serve to create further gaps in healthcare provision discrimination. For instance, diagnostic systems almost exclusively trained on some groups will fail to perform as intended for those groups who have been historically neglected and thus would mean discrimination in health services provision. To solve this problem one would have to utilize representative datasets across population groups and constantly test them for discrimination against groups that might be more vulnerable.

The health care data is prone to abuse, and even more so when ML systems are in play. Collection, storage, and processing of data is a recipe for disaster, it is a call for breach, tampering or unauthorized access to sensitive information. Laws such as HIPAA or GDPR law are tremendously important in its enforcement. There are such techniques that are now being developed and tried out that would allow the use of ML but would improve the level of security, hygiene of the data, such techniques include data anonymisation and federated learning among others.

The Future of Machine Learning in Healthcare

The future of machine learning (ML) in the healthcare sector is brimming with an incredible amount of promise that is just mind-boggling, promising to revolutionize not only patient care, but also medical research and the overall effectiveness of health systems. With technology still advancing at a rate never seen before, it is only forecasted that ML will become more and more integrated into a variety of aspects of the healthcare sector. This integration will drive significant innovations that will enhance the accuracy and speed of diagnoses, streamline many of the procedures involved in treatment, and ultimately lead to improvements in overall patient outcomes. With the growing availability of large medical datasets, combined with the refinement of AI-based models and the ongoing development of computational power, the healthcare industry is on the verge of a revolutionary transformation. In this new era, ML is to be at the centre of decision-making, automation, and the creation of personalized medicine, and therefore alter the delivery of healthcare.

Conclusions

The requirement of ML in medicine is its capability of addressing some of the most crucial medical problems such as the detection of diseases in their initial phases, continuous observation of the patient, and mechanization of the processes in a hospital. Its application to precision medicine has offered the hope of customized treatments which reduce the drug's side effects and increase the effectiveness of the treatment. Furthermore, predictive analytics with ML has introduced early intervention to manage chronic disease, with excellent patient results as well as reducing hospitalization. With constant advancements in deep learning, natural language processing, and computer vision, the application of ML to medicine is going to increase manifold in the coming years.

REFERENCES

[1] Copeland J. The Turing test. Minds and Machines. 2000;10(4):519-539

[2] French RM. The Turing test: The first 50 years. Trends in Cognitive Sciences. 2000;4(3):115-122

[3] Edwards S. World War II at sea: A global history. The Journal of American History. 2019;106(1):237

[4] Turing AM. Computing machinery and intelligence. Parsing the Turing Test: Philosophical and Methodological Issues in the Quest for the Thinking Computer. 2009. p. 23-65

[5] Samuel AL. Some studies in machine learning. IBM Journal of Research and Development. 1959;3(3):210-229

[6] Samuel AL. Programming computers to play games. Advances in Computers. 1960;1(C):165-192

[7] Fukushima K. Cognition: A self-organizing multilayered neural network. Biological Cybernetics. 1975;20(3-4):121-136

[8] Fukushima K. Noncognition: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position. Biological Cybernetics. 1980;36(4):193-202
[9] Huang K, Mashl RJ, Wu Y, Ritter DI, Wang J, Oh C, et al. Pathogenic germline variants in 10,389 adult cancers. Cell. 2018
[10] Silver D, Hassabis D. AlphaGo: Mastering the ancient game of go with machine learning. Google Research Blog. 2016
[11] Brewka G. Artificial intelligence—a modern approach by Stuart Russell and Peter Norvig, Prentice Hall. Series in Artificial Intelligence, Englewood Cliffs, NJ. The Knowledge Engineering Review. 1996;11(1): 78-79
[12] Alpaydin E. Introduction to Machine Learning. London: The MIT Press. 2014; 3:640
[13] Kotsiantis SB. Supervised machine learning: A review of classification techniques. Informatica (Ljubljana). 2007;31(3):249-268
[14] Hastie T, Tibshirani R, Friedman J. The Elements of Statistical Learning - Data Mining, Inference, and Prediction. Springer Series in Statistics. Switzerland: Springer. 2009;2(1):93-85
[15] Ho TK. A data complexity analysis of comparative advantages of decision forest constructors. Pattern Analysis & Applications. 2002;5(2):105-112
[16] Friedman JH. Stochastic gradient boosting. Computational Statistics and Data Analysis. 2002;38(4):367-378
[17] Friedman JH. Greedy function approximation: A gradient boosting machine. The Annals of Statistics. 2001;29(5):1189-1232
[18] Mason L, Baxter J, Bartlett P, Frean M. Boosting algorithms as gradient descent in function space. NIPS Conference Proceedings. 1999:512-518
[19] Libbrecht MW, Noble WS. Machine learning applications in genetics and genomics. Nature Reviews. Genetics. 2015;16(6):321-332
[20] Längkvist M, Karlsson L, Loutfi A. A review of unsupervised feature learning and deep learning for time-series modelling. Pattern Recognition Letters. 2014;42(1):11-24
[21] Lloyd SP. Least squares quantization in PCM. IEEE Transactions on Information Theory. 1982;28(2):129-137 References Smart Manufacturing - When Artificial Intelligence Meets the Internet of Things 22
[22] Schmid Huber J. Deep learning in neural networks: An overview. Neural Networks. 2015;61(1):85-117
[23] Hazan E, Klevan's A, Yuan Y. Hyperparameter optimization: A spectral approach. In: In: 6th International Conference on Learning Representations, ICLR 2018; Conference Track Proceedings. 2018
[24] Bardenet R, Brendel M, Kégl B, Sebag M. Collaborative hyperparameter tuning. In: 30th International Conference on Machine Learning; ICML. 2013. p. 2013
[25] Hutter F, Hoos H, Leyton Brown K. An efficient approach for assessing hyperparameter importance. In: 31st International Conference on Machine Learning; ICML. 2014. p. 2014
[26] Hawkins DM. The problem of Overfitting. Journal of Chemical Information and Computer Sciences. 2004;44(1):1-12
[27] Ng A. Regularization—Solving the Problem of Overfitting. Coursera; 2011. Available from: https://www.coursera. org/learn/machine-learning/lecture/ ACpTQ/the-problem-of-overfitting
[28] Srivastava N, Hinton G, Krizhevsky A, Sutskever I, Salakhutdinov R. Dropout: A simple way to prevent neural networks from overfitting. Journal of Machine Learning Research. 2014;15(56):1929-1958
[29] Picard RR, Cook RD. Cross-validation of regression models. Journal of the American Statistical Association. 1984;79(387):575-583
[30] Jothi N, Rashid NA, Husain W. Data mining in healthcare—A review. Procedia Computer Science. 2015;72(1):306-313
[31] Koh HC, Tan G. Data mining applications in healthcare. Journal of Healthcare Information Management. 2005;19(2):64-72
[32] Cottle M, Hoover W, Kanwal S, Kohn M, Strome T, Treister NW. Transforming Health Care through Big Data: Strategies for Leveraging Big Data in the Health Care Industry. Institute for Health Technology Transformation. New York, iHT2; 2013
[33] Wang Y, Kung LA, Byrd TA. Big data analytics: Understanding its capabilities and potential benefits for healthcare organizations. Technological Forecasting and Social Change. 2018;126(1):3-13
[34] Raghupathi W, Raghupathi V. Big data analytics in healthcare: Promise and potential. Health Information Science and Systems. 2014;2(1):3
[35] The Cancer Genome Atlas Program—National Cancer Institute [Internet]. 2019. Available from:
https://www.cancer.gov/aboutnci/organization/ccg/research/ structural-genomics/tcga
[36] tcga-infographic-enlarge. __ v100169753.png (1400×2580) [Internet]. 2019. Available from:
https://www.cancer.gov/PublishedContent/Images/images/nci/organization/tcga/tcgainfographicenlarge.__v100169753.png
[37] Peakman TC, Elliott P. The UK biobank sample handling and storage validation studies. International Journal of Epidemiology. 2008;37(2):234-244
[38] Sudlow C, Gallacher J, Allen N, Beral V, Burton P, Danesh J, et al. UK biobank: An open access resource for identifying the causes of a wide range of complex diseases of middle and old age. PLoS Medicine. 2015;12(3)
[39] Protocol for the UK Biobank— Wayback Machine [Internet]. 2019. Available from: https://web.archive.org/web/20060214144838/http://www.ukbiobank.ac.uk/docs/draft_protocol.pdf
[40] Elliott P, Peakman TC. The UK biobank sample handling and storage protocol for the collection, processing and archiving of human blood and urine. International Journal of Epidemiology. 2008;37(2):234-244
[41] Gao GF, Parker JS, Reynolds SM, Silva TC, Wang LB, Zhou W, et al. Before and after: Comparison of legacy and harmonized TCGA genomic data commons data. Cell Systems. 2019;9(1):24-34. e10
[42] Bycroft C, Freeman C, Petkova D, Band G, Elliott LT, Sharp K, et al. The UK biobank resource with deep phenotyping and genomic data. Nature. 2018;562(7726):203-209
[43] Background—EPIC [Internet]. 2019. Available from: http://epic.iarc.fr/ about/background.php
[44] Leitsalu L, Haller T, Esko T, Tammesoo ML, Alavere H, Snieder H, et al. Cohort profile: Estonian biobank of the Estonian genome centre, university of Tartu. International Journal of Epidemiology. 2015;44(4):1137-1147
[45] Master Decoder: A Profile of Kári Stefánsson | The Scientist Magazine® [Internet]. 2019. Available from: https://www.the-scientist.com/profile/master-decoder--a-profile-of-kristefnsson-65517