OPEN ACCESS
©NIJRMS ONLINE ISSN: 2992-5460
Publications 2024 PRINT ISSN: 2992-6041

NEWPORT INTERNATIONAL JOURNAL OF RESEARCH IN MEDICAL SCIENCES (NIJRMS)

Volume 5 Issue 3 2024

https://doi.org/10.59298/NIJRMS/2024/5.3.14883

The use of Machine Learning in Predicting Disease Progression

Gakire Munyaneza H.

Faculty of Engineering Kampala International University Uganda

ABSTRACT

Machine learning (ML) has emerged as a powerful tool in the healthcare sector, especially in predicting disease progression. This paper examines the application of various ML algorithms, including supervised, unsupervised, and reinforcement learning, in predicting disease trajectories, particularly in complex and chronic diseases like Alzheimer's, multiple sclerosis, cardiovascular disease, and diabetes. By leveraging clinical data, genetic information, and patient history, ML models like random forests and neural networks can accurately predict the time to disease progression. This has profound implications for early diagnosis, personalized treatment, and patient management. However, the integration of these models into clinical practice faces challenges related to data quality, interpretability, and deployment in real-world settings. Despite these limitations, the case studies reviewed demonstrate the transformative potential of machine learning in enhancing decision-making processes in healthcare.

Keywords: Machine learning, Disease progression prediction, Healthcare, Supervised learning, Chronic diseases.

INTRODUCTION

Machine learning is rapidly becoming one of the major methods in professional applications as well as experimental sciences, due to the increasing digitization of knowledge. In professional applications, the extension of the scope of machine learning from simple classification tasks to more sophisticated tasks like ranking and recording problems with a time horizon and high output costs is notable. These tasks open a lot of new application markets, especially in finance, such as trading, customer relationship management, information search and retrieval, user interfaces, and advisory systems. In experimental sciences and other fields, machine learning is used as a tool for data mining. In all these domains, the notion of confidence, which is computed using statistical properties of the data, provides a new level of augmented intelligence to users. However, the limited confidence concepts available in current black-box models are still far from describing the current state of the art in scientific knowledge [1, 2]. We feel that the combination of the sophistication of machine learning methods and the applicability of robust confidence measures opens not only new possible applications but also new complex research challenges in methodological and domain-specific topics. In this context, researchers are dealing with the overall problem of deciding if and when a presymptomatic subject will manifest the characteristic signs of a particular disease and if they can do this without having to rely upon a complex and potentially risky diagnostic test [3, 4].

FUNDAMENTALS OF MACHINE LEARNING

Machine learning algorithms include a multitude of classic statistical and logical tools. Machine learning algorithms fit into three general classes: supervised learning, in which the algorithm induces a model from training examples that capture the relation between the input attributes and the outcome or the goal attribute; unsupervised learning, in which the training set is unlabeled and the target is still to summarize and interpret an existing data set; and reinforcement learning, where an agent is trained to make a sequence of decisions while interacting with a malleable environment. Regardless of the methods used to develop the algorithm, the core steps involved in using a machine learning model include model development, in which the structure of the model is selected based on disparate training data; model evaluation to assess the extent of agreement between predictions and the outcomes for new data; and assessment of the model as an optimizing objective to measure how well the model is achieving the intended use [5, 6]. Prediction models are a specific class of models aiming to make as accurate predictions as possible for individual patients and are based on patients' prognostic factors. They belong to the application fields of supervised learning and have been developed using different statistical and machine learning techniques such as simple logistic regression, classification and regression trees, random forests, or neural networks. Prediction models are built using different techniques, and the selection of the technique used has a direct impact on the model's output. Furthermore, the selection and development of prediction factors are closely dependent on the intended use of the model, or the way the model is expected to generate recommendations for clinical action [7, 8].

PREDICTIVE MODELING IN HEALTHCARE

The Moonlighting Goal of Predictive Models: A Healthcare Perspective The purpose of most predictive models in healthcare is often purely for making more effective risk assessments and assisting in better targeting preventive actions. The predictions are mere means to the end of optimizing a specific decision-making objective. The predictive model itself is a supporting tool that fulfills goals that are not expressed in terms of predictive performance but in terms of making more informed and insightful decisions. This moonlighting nature of predictive models in healthcare often necessitates that predictive models are contrasted with decision tools with or without integrated predictive components. Considering types of models in this context helps interpret and propose new perspectives on their potential breakeven points. This analysis would be useful because there are diverse perspectives on what type of models are best suited for different healthcare use-case scenarios. Misunderstandings and misconceptions abound in both industry and academic worlds, and a deeper understanding may help with better selecting, using, or even designing predictive models for targeted healthcare applications [9].

APPLICATIONS OF MACHINE LEARNING IN DISEASE PROGRESSION PREDICTION

The objective of this work is to demonstrate the feasibility and utility of using machine learning techniques to predict the time from disease onset to a future clinical endpoint in two genetically sporadic diseases: multiple sclerosis and Alzheimer's disease. In these very different diseases, time to disease progression is critically important in efforts to prevent disease progression or to monitor the effects of potential agents. Here we have used standard clinical variables, genetic information, and gene sets derived from high-throughput experiments as input features to machine classifiers to predict time to disease progression. The primary finding of this work is that machine learning can be used to accurately predict time to progression in genetic diseases. Specifically, random forest classifiers are accurate and provide an easily interpretable view of predictive variables [10, 11]. The benefit provided by genetic information is relatively modest but significant in some cases. Final predictions are built upon previously known clinical predictions of time to disease progression. This indicates a general direction that might be useful in other medical situations where quantitative predictions are important. Accurate prediction depends upon enough patients having reached disease progression. In diseases where progression is rapidly fatal or where new treatments greatly skew the course of progression, the model will need to be retrained [12, 13].

CHALLENGES AND LIMITATIONS

In this paper, we have seen several novel machine-learning models proposed for disease progression prediction in the context of Alzheimer’s disease. In the next section, we will provide a focused review of other disease progression prediction models that have been proposed. These models are for other complex, severe, and chronic disease prognoses including Parkinson’s disease, multiple sclerosis, cataracts, breast cancer, cardiovascular disease, diabetes mellitus, chronic kidney failure, etc. An interesting finding of our review is that while many methods have been proposed for disease prognosis, only a handful of such studies have been investigated for disease progression prediction [14, 15]. It should be emphasized that, like most machine learning applications in clinical labs, building high-performance machine learning models for disease progression prediction and diagnosing as well as early diagnosis has proved to be a major public health challenge with a high degree of complexity. There are logistical, financial, human, and data-related reasons and obstacles. While much progress has been achieved in the development phase, large-scale successful deployment of these models is often a rather intricate problem because of various reasons. The challenges and limitations that make the application of machine learning primarily relevant to clinical settings are often discussed, and we will provide a synopsis that may be of utility to researchers from computer and information fields [16, 17].

CASE STUDIES AND SUCCESS STORIES

Predicting the future, and especially the future behavior of complex biological systems, is of major interest for medical and biological research. The emergence of data-rich, high-throughput technologies providing abundant information at various molecular levels, as well as phenotypic data at all scales, has provided the scientific community with a wide range of prediction tasks using machine learning techniques. There is an overwhelming literature full of success stories on diverse tasks such as the prediction of cardiovascular disease based on dietary patterns, identification of genes involved in complex disorders, prediction of specific functional interactions from biological data, and prediction protein function from sequence and structure. The overall picture, however, is that when using statistical predictions in medical and clinical settings, it makes a considerable difference to use information and predictions that are as accurate and reliable as possible [18, 19]. In this chapter, we provide a brief introduction to machine learning terminology, tasks, and techniques, and discuss how machine learning algorithms can be used in the more complex task of predicting disease progression based on biomolecular and genetic information. After reviewing several successful case studies demonstrating the validity of the approach, we discuss a widely used and freely available open-source package that enables users to combine a variety of machine learning algorithms and techniques into efficient multi-class predictors. Results of the application of the package to biological and clinical data on Huntington's disease demonstrate that machine learning algorithms constitute a powerful statistical tool in identifying and ranking disease-related phenotypic biomarkers [20, 21].

CONCLUSION

The application of machine learning in predicting disease progression offers a significant leap forward in healthcare, enabling more personalized, timely, and precise interventions. While machine learning models like random forests have shown great accuracy, the practical implementation of these tools in clinical environments remains complex due to factors such as data variability, algorithm interpretability, and integration challenges. Future work should focus on addressing these issues to enable the widespread adoption of predictive models in healthcare, ultimately enhancing patient outcomes. Success stories in chronic disease management highlight the promise of ML in revolutionizing healthcare practices, particularly in early diagnosis and treatment optimization.

REFERENCES

1. Sarker IH. Machine learning: Algorithms, real-world applications, and research directions. SN computer science. 2021 May;2(3):160.
2. Shehab M, Abualigah L, Shambour Q, Abu-Hashem MA, Shambour MK, Alsalibi AI, Gandomi AH. Machine learning in medical applications: A review of state-of-the-art methods. Computers in Biology and Medicine. 2022 Jun 1; 145:105458.
3. Wang Y, Zheng K, Gao W, Lv J, Yu C, Wang L, Wang Z, Wang B, Liao C, Li L. Asymptomatic and pre-symptomatic infection in Coronavirus Disease 2019 pandemic. Medical Review. 2022 Feb 23;2(1):66-88. degruyter.com
4. Benatar M, Wuu J, McHutchison C, Postuma RB, Boeve BF, Petersen R, Ross CA, Rosen H, Arias JJ, Fradette S, McDermott MP. Preventing amyotrophic lateral sclerosis: insights from pre-symptomatic neurodegenerative diseases. Brain. 2022 Jan 1;145(1):27-44. oup.com
5. Morales EF, Escalante HJ. A brief introduction to supervised, unsupervised, and reinforcement learning. InBiosignal processing and classification using computational learning and intelligence 2022 Jan 1 (pp. 111-129). Academic Press. researchgate.net
6. Shetty SH, Shetty S, Singh C, Rao A. Supervised machine learning: algorithms and applications. Fundamentals and methods of machine and deep learning: algorithms, tools, and applications. 2022 Feb 24:1-6. researchgate.net
7. Stiglic G, Kocbek P, Fijacko N, Zitnik M, Verbert K, Cilar L. Interpretability of machine learning-based prediction models in healthcare. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery. 2020 Sep;10(5): e1379. [PDF]
8. Yan L, Zhang HT, Goncalves J, Xiao Y, Wang M, Guo Y, Sun C, Tang X, Jing L, Zhang M, Huang X. An interpretable mortality prediction model for COVID-19 patients. Nature machine intelligence. 2020 May;2(5):283-8. nature.com
9. Ordu M, Demir E, Tofallis C, Gunal MM. A novel healthcare resource allocation decision support tool: A forecasting-simulation-optimization approach. Journal of the Operational Research Society. 2021 Mar 14;72(3):485-500. herts.ac.uk
10. Quazi S. Artificial intelligence and machine learning in precision and genomic medicine. Medical Oncology. 2022 Jun 15;39(8):120.
11. Ibrahim I, Abdulazeez A. The role of machine learning algorithms for diagnosing diseases. Journal of Applied Science and Technology Trends. 2021 Mar 19;2(01):10-9. jastt.org
12. Chittora P, Chaurasia S, Chakrabarti P, Kumawat G, Chakrabarti T, Leonowicz Z, Jasiński M, Jasiński Ł, Gono R, Jasińska E, Bolshev V. Prediction of chronic kidney disease-a machine learning perspective. IEEE Access. 2021 Jan 22; 9:17312-34. ieee.org

13. Kim WR, Mannalithara A, Heimbach JK, Kamath PS, Asrani SK, Biggins SW, Wood NL, Gentry SE, Kwong AJ. MELD 3.0: the model for end-stage liver disease updated for the modern era. Gastroenterology. 2021 Dec 1;161(6):1887-95. sciencedirect.com
14. Feng Z, Yu Q, Yao S, Luo L, Zhou W, Mao X, Li J, Duan J, Yan Z, Yang M, Tan H. Early prediction of disease progression in COVID-19 pneumonia patients with chest CT and clinical characteristics. Nature communications. 2020 Oct 2;11(1):4968. nature.com
15. Battineni G, Sagaro GG, Chinatalapudi N, Amenta F. Applications of machine learning predictive models in the chronic disease diagnosis. Journal of personalized medicine. 2020 Mar 31;10(2):21. mdpi.com
16. Nguyen H, Vu T, Vo TP, Thai HT. Efficient machine learning models for prediction of concrete strengths. Construction and Building Materials. 2021 Jan 10; 266:120950.
17. Banegas-Luna AJ, Peña-García J, Iftene A, Guadagni F, Ferroni P, Scarpato N, Zanzotto FM, Bueno-Crespo A, Pérez-Sánchez H. Towards the interpretability of machine learning predictions for medical applications targeting personalised therapies: A cancer case survey. International Journal of Molecular Sciences. 2021 Apr 22;22(9):4394. mdpi.com
18. Alanazi R. Identification and prediction of chronic diseases using machine learning approach. Journal of Healthcare Engineering. 2022;2022(1):2826127.
19. Shah D, Patel S, Bharti SK. Heart disease prediction using machine learning techniques. SN Computer Science. 2020 Nov;1(6):345.
20. Hudson IL. Data integration using advances in machine learning in drug discovery and molecular biology. Artificial Neural Networks. 2021:167-84.
21. Rickert CA, Lieleg O. Machine learning approaches for biomolecular, biophysical, and biomaterials research. Biophysics Reviews. 2022 Jun 1;3(2).

CITE AS: Gakire Munyaneza H. (2024). The use of Machine Learning in Predicting Disease Progression. NEWPORT INTERNATIONAL JOURNAL OF RESEARCH IN MEDICAL SCIENCES, 5(3):1-4.
https://doi.org/10.59298/NIJRMS/2024/5.3.14883