

MACHINE LEARNING ALGORITHMS FOR INTELLIGENT DIAGNOSIS OF A PATIENT WITH HEART DISEASE: AN OVERVIEW

Ismailov Otabek Mirzalilovich¹

Khojamurotov Azizbek Shukhrat ugli²

¹Professor at Tashkent University of Information Technologies named after
Muhammad al-Khwarizmi, Department of Computer Systems

²Graduate student at Tashkent University of Information Technologies named after
Muhammad al-Khwarizmi, Department of Computer Systems

otabekismailov@gmail.com, azizbekxojamurotov2001@gmail.com

Abstract: This thesis is dedicated to the study of modern methods for the rapid detection and analysis of adverse changes in patients using machine learning algorithms. The successful results of this research will effectively aid doctors in improving the process of treating patients in the future. This work aims to bridge the gap between technological advancements and practical healthcare applications.

Keywords: machine learning, rapid detection, adverse changes, patient monitoring, medical algorithms, healthcare improvement, predictive analytics, clinical decision support.

Аннотация: Данная диссертация посвящена изучению современных методов быстрого выявления и анализа неблагоприятных изменений у пациентов с использованием алгоритмов машинного обучения. Успешные результаты этого исследования эффективно помогут врачам в улучшении процесса лечения пациентов в будущем. Эта работа направлена на преодоление разрыва между технологическими достижениями и практическим применением в здравоохранении.

Ключевые слова: машинное обучение, быстрое обнаружение, неблагоприятные изменения, мониторинг пациентов, медицинские алгоритмы, улучшение здравоохранения, прогнозная аналитика, поддержка принятия клинических решений.

Annotatsiya: Ushbu dissertatsiya mashinani o'rganish algoritmlaridan foydalangan holda bemorlardagi salbiy o'zgarishlarni tezkor aniqlash va tahlil qilishning zamonaviy usullarini o'rganishga bag'ishlangan. Ushbu tadqiqotning muvaffaqiyatli natijalari shifokorlarga kelajakda bemorlarni davolash jarayonini takomillashtirishda samarali yordam beradi. Ushbu ish texnologik yutuqlar va amaliy sog'liqni saqlash ilovalari o'rtasidagi tafovutni bartaraf etishga qaratilgan.

Kalit so'zlar: mashinani o'rganish, tezkor aniqlash, salbiy o'zgarishlar, bemorlarni kuzatish, tibbiy algoritmlar, sog'liqni saqlashni yaxshilash, bashoratli tahlillar, klinik qarorlarni qo'llab-quvvatlash.

INTRODUCTION

The World Health Organization has declared that cardiovascular diseases (CVDs) are the leading cause of death worldwide, responsible for 17.9 million deaths in 2019. These diseases encompass coronary heart disease, cerebrovascular disease, rheumatic heart disease, and others. Most CVDs could be prevented by avoiding risk behaviors such as excessive alcohol consumption, smoking, and unhealthy eating habits. The effects of CVDs may manifest in the human body through high blood pressure, high blood glucose levels, and high blood lipids. Timely detection of these pathologies is crucial to providing appropriate treatment to patients and preventing premature deaths. Artificial intelligence can assist in identifying patients at high risk of cardiovascular events by performing non-invasive tests, such as blood tests and electrocardiograms. Hospitals generate a substantial amount of medical data, necessitating modern approaches to utilize this information effectively. To achieve the goal of detecting high-risk patients, numerous public datasets containing patient health information can be utilized.

RELATED WORK

Given the significant impact of CVDs on society, extensive research has been conducted to develop techniques capable of accurately detecting and predicting these diseases. Many of these studies utilize machine learning models. A performance analysis summary of the models reviewed for this thesis is presented in Table 1.

Work	Algorithm	Accuracy
Ramesh et al. (2022) [1]	Naïve Bayes	86%
Miranda et al. (2021)[2]	Stochastic Gradient Descent	80%
Miranda et al. (2021)[2]	Logistic Regression	91.67%
Kavitha et al. (2021)[3]	Decision Tree and Random Forest	88%
Priscila et al. (2017)[4]	Neural Network	90%

Table 1: Performance of related works

The most commonly used approaches include Neural Networks, Support Vector Machines, Logistic Regression, Naive Bayes, and Decision Trees. The authors of the referenced studies all utilized the UCI dataset, which is also employed in this thesis. In this research, these methods have been implemented and analyzed using Python programming.

In the study referenced as [1], researchers proposed various methods for detecting CVDs using the UCI dataset. Specifically, they tested Tree Classifier, Random Forest, and K-Nearest Neighbors algorithms.

In work [3], a novel algorithm was proposed. Various approaches, including Decision Tree (DT), Random Forest (RF), and a hybrid model combining DT and RF, were analyzed. The hybrid model achieved superior accuracy compared to the individual models. Specifically, the posterior probabilities output from the Random Forest using the original dataset were used as input for the Decision Tree algorithm. Additionally, the researchers designed an application for heart disease prediction, which takes health examination results as input. The graphical user interface for this application was developed using Python's Tkinter library.

In [2], the researchers utilized Logistic Regression and Stochastic Gradient Descent to predict heart disease. Logistic Regression provided the best results in terms of accuracy, precision, and recall. The authors noted that a larger dataset could improve the outcomes, as the small dataset size limited their models' performance. They suggested that their proposed algorithms could serve as non-invasive diagnostic tools for detecting heart disease.

In research paper [4], the authors proposed an ensemble model of Neural Networks for heart disease prediction. They began by removing features based on their correlation coefficients. An ensemble of Neural Networks with randomly chosen parameters was initially considered, with only the most accurate networks being selected. Entropy was then used to choose the best components, ensuring the optimal selection of Neural Networks for the model.

In summary, the proposed systems can serve as decision support tools for medical professionals and as a preventative method for cardiovascular diseases (CVDs), with the goal of providing accessible assistance to all individuals and potentially reducing mortality rates worldwide. The works discussed in this section form the foundational basis for the algorithms selected to detect heart disease in this thesis.

CONCLUSION

In conclusion, the exploration of machine learning algorithms for the detection and prediction of cardiovascular diseases (CVDs) offers a promising pathway towards enhancing diagnostic accuracy and improving patient outcomes in cardiovascular healthcare. Through an in-depth analysis of existing research, it is evident that various machine learning techniques, including Logistic Regression, Support Vector Machines, Neural Networks, Decision Trees, and Random Forests, exhibit differential performances in CVD detection, with Logistic Regression and Support Vector Machines often demonstrating superior diagnostic accuracy. The significance of dataset size becomes apparent, as studies consistently indicate that larger datasets contribute to better algorithm performance, underscoring the importance of robust data collection strategies. Additionally, preprocessing techniques such as feature normalization have been proposed to enhance classifier performance, highlighting the potential for optimization through data preprocessing methods. Ensemble approaches, leveraging the strengths of multiple algorithms, have shown promise in enhancing predictive accuracy and represent a valuable avenue for future research. Moreover, the development of diagnostic systems for CVD detection requires careful consideration of fundamental steps, including data collection,

preprocessing, model development, patient recruitment, and medical evaluation, with machine learning models serving as invaluable tools in this process. Overall, the integration of machine learning into cardiovascular medicine holds immense potential for advancing diagnostic capabilities, optimizing treatment strategies, and ultimately improving patient outcomes in the fight against cardiovascular diseases. Continued research and innovation in this field are imperative for realizing the full potential of machine learning in cardiovascular healthcare and addressing the global burden of CVDs effectively.

References.

1. [8] Ramesh Tr et al. “PREDICTIVE ANALYSIS OF HEART DISEASES WITH MACHINE LEARNING APPROACHES”. In: Malaysian Journal of Computer Science 2022 (Apr. 2022), pp. 132–148. doi: 10.22452/mjcs.sp2022no1.10 (cit. on pp. 2, 33).
2. [7] Eka Miranda et al. “Intelligent Computational Model for Early Heart Disease Prediction using Logistic Regression and Stochastic Gradient Descent (A Preliminary Study)”. In: 2021 1st International Conference on Computer Science and Artificial Intelligence (ICCSAI). Vol. 1. 2021, pp. 11–16. doi: 10.1109/ICCSAI53272.2021.9609724 (cit. on pp. 2, 3).
3. M. Kavitha et al. “Heart Disease Prediction using Hybrid machine Learning Model”. In: 2021 6th International Conference on Inventive Computation Technologies (ICICT). 2021, pp. 1329–1333. doi: 10.1109/ICICT50816.2021.9358597 (cit. on p. 2).
4. S. Silvia Priscila and M. Hemalatha. “Improving the Performance of Entropy Ensembles of Neural Networks (EENNS) on Classification of Heart Disease Prediction”. In: 2017. url: <https://api.semanticscholar.org/CorpusID:201110435> (cit. on pp. 2, 3).