IJIRCCE

e-ISSN: 2320-9801 | p-ISSN: 2320-9798

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

Volume 12, Issue 1, January 2024

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

Impact Factor: 8.379

International Journal of Innovative Research in Computer and Communication Engineering

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 | A Monthly Peer Reviewed & Referred Journal |

|| Volume 12, Issue 1, January 2024 ||

| DOI: 10.15680/IJIRCCE.2024.1201058 |

Survey on Heart Disease Identification Method Using Machine Learning Classification in E-Healthcare

Mrs. S. A. Shete, Sahil Mazire, Venkatesh Saraf, Sarvesh Khade

Lecturer, Department of Information Technology, AISSMS's Polytechnic, Pune, Maharashtra, India[1]

Student, Department of Information Technology, AISSMS's Polytechnic, Pune, Maharashtra, India[2]

Student, Department of Information Technology, AISSMS's Polytechnic, Pune, Maharashtra, India[3]

Student, Department of Information Technology, AISSMS's Polytechnic, Pune, Maharashtra, India[4]

ABSTRACT: This survey paper investigates the integration of machine learning, specifically the Logistic Regression algorithm, in the realm of heart disease detection software. As cardiovascular health remains a critical concern globally, leveraging advanced technologies for accurate and timely diagnosis becomes imperative. The paper conducts a comprehensive review of existing literature, methodologies, and advancements related to the utilization of Logistic Regression in developing software solutions for heart disease detection.

The survey addresses the fundamental principles underlying Logistic Regression and its applicability in the context of medical software. It explores the various features and risk factors considered during model development, emphasizing the significance of data preprocessing and feature selection in enhancing the efficacy of the software.

Furthermore, the paper delves into the evolving landscape of machine learning applications in healthcare, emphasizing the role of Logistic Regression in creating interpretable and effective heart disease detection software. It discusses the challenges associated with imbalanced datasets, hyperparameter tuning, and the incorporation of diverse data sources to enhance the software's accuracy and reliability.

The survey concludes by highlighting potential areas for future research and development, envisioning advancements in algorithmic sophistication and real-world implementation of heart disease detection software. This paper aims to provide a valuable resource for researchers, practitioners, and developers in the field of machine learning-driven healthcare applications, with a specific focus on the Logistic Regression algorithm for heart disease detection software

KEYWORDS: 1. Machine Learning
2. Heart Disease
3. Cardiovascular Health
4. Data Mining
5. Feature Extraction
6. Classification Algorithms
7. Risk Prediction
8. Signal Processing
9. Healthcare Technology
10. Diagnostic Models

I. INTRODUCTION

Cardiovascular diseases (CVDs) represent a leading cause of global morbidity and mortality, underscoring the critical need for effective and efficient diagnostic tools. With the advent of machine learning in healthcare, there has been a paradigm shift towards leveraging advanced algorithms for early detection and management of heart diseases. This survey paper explores the integration of machine learning, specifically focusing on the Logistic Regression algorithm, in the development of heart disease detector software.

Machine learning algorithms offer a data-driven approach to medical diagnostics, enabling the extraction of patterns and insights from vast datasets that may elude traditional methods. Logistic Regression, a widely used classification algorithm, holds promise in providing interpretable and accurate predictions, making it a pertinent candidate for incorporation into heart disease detection software.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 | A Monthly Peer Reviewed & Referred Journal |

|| Volume 12, Issue 1, January 2024 ||

| DOI: 10.15680/IJIRCCE.2024.1201058 |

This survey aims to provide a comprehensive overview of the current state of research and development in this domain. By synthesizing existing literature, methodologies, and advancements, we seek to elucidate the strengths, limitations, and potential avenues for improvement in utilizing Logistic Regression for heart disease detection software.

As we delve into the intricacies of this interdisciplinary field, we will explore the fundamental principles of Logistic Regression, its application in medical diagnostics, and the specific considerations for developing software solutions targeting heart diseases. The survey will also navigate through challenges such as handling imbalanced datasets and optimizing hyperparameters, addressing the nuances of integrating machine learning into real-world healthcare scenarios.

By the conclusion of this survey, we aim to not only provide a comprehensive understanding of the current landscape but also to stimulate further research and innovation in the development of heart disease detection software. Ultimately, the integration of machine learning, particularly Logistic Regression, holds the potential to enhance the precision and efficiency of diagnostic tools, contributing to the advancement of cardiovascular healthcare.

II. STATEMENT ABOUT THE PROBLEM

. Research Problem

Real-World Problem: Cardiovascular diseases, especially heart diseases, present a pervasive global health challenge. The problem lies in the late diagnosis of these diseases, often after symptoms manifest or complications occur. Inadequate risk assessment and screening, coupled with resource limitations in healthcare systems, make it challenging to identify individuals at risk early in the disease progression. Challenges Faced by People: Many individuals are unaware of their susceptibility to heart diseases until it's too late, resulting in elevated mortality rates and increased healthcare expenditures. Traditional screening methods may lack the depth required for comprehensive risk evaluation, leading to underdiagnosis. Healthcare systems, burdened by resource shortages and time constraints, struggle to conduct thorough screenings for all.

Research Solution

The project, "Heart Disease Identification Method Using Machine Learning Classification in EHealthcare," was conceived to tackle these real-world issues. By harnessing the capabilities of machine learning and predictive modeling, this initiative seeks to: Facilitate Early Detection: Detect individuals at risk of heart diseases in their early stages, even before symptoms emerge, enabling timely medical intervention and potentially saving lives. Enhance Screening Accuracy: Provide a robust and accurate screening tool that goes beyond traditional clinical parameters, ensuring a more comprehensive assessment of risk factors. Address Resource Constraints: Offer an efficient and cost-effective solution that can assist healthcare professionals in prioritizing high-risk individuals within resource-constrained healthcare systems

Objective and scope of the project

- Develop a machine learning-based system for the early identification of heart diseases.
- Enhance the accuracy of heart disease prediction by leveraging advanced algorithms.
- Explore and optimize various machine learning models to identify the most effective one.
- Address class imbalance in the dataset to ensure reliable model performance.

Scope:

- Utilize the "Heart Disease Dataset" from Kaggle as the primary data source.
- Implement data preprocessing techniques, including scaling and balancing, to prepare the dataset for model training.
- Explore and evaluate machine learning models, including Logistic Regression, Support Vector
- The project's scope extends to contributing to improved cardiac healthcare outcomes through early detection and intervention

III. PROPOSED ALGORITHM

Logistic Regression is a statistical method used for binary classification, which means it is employed when the outcome variable has two possible classes. Despite its name, logistic regression is used for classification, not regression tasks. The algorithm models the relationship between a binary dependent variable and one or more independent variables by estimating probabilities using the logistic function.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 | A Monthly Peer Reviewed & Referred Journal |

|| Volume 12, Issue 1, January 2024 ||

| DOI: 10.15680/IJIRCCE.2024.1201058 |

In logistic regression, the logistic function (also known as the sigmoid function) is applied to a linear combination of the input features. The sigmoid function outputs values between 0 and 1, mapping any real-valued number to a

Logistic Regression is widely used in various fields for tasks such as medical diagnosis, spam detection, and credit scoring, among others.

IV. LITERATURE SURVEY

Sr. No	Paper/Author	OverView
1.	**Author: S. Sharma et al. (2020)** "Heart Disease Prediction Using Support Vector Machines with Optimized Feature Selection."	*In 2020, Sharma and colleagues conducted research on heart disease prediction using Support Vector Machines (SVM). Their study focused on feature selection and SVM optimization, achieving notable accuracy in diagnosing heart diseases, which aligns with the current project's objectives.*
2.	**Author: M. Johnson et al. (2019)** "Deep Learning-Based Approach for Heart Disease Identification and Risk Assessment."	*Johnson et al. published a study in 2019 that investigated the use of deep learning techniques for heart disease identification. Their work involved building neural network architectures similar to the one in this project. They highlighted the potential of deep learning in improving heart disease prediction accuracy.*
3.	**Author: F. K. Sun et al. (2018)** "Application of Machine Learning Algorithms in Early Detection of Heart Disease."	*In their study published in 2018, Sun et al. explored the use of machine learning techniques for heart disease prediction. They employed Logistic Regression and Random Forest models on a similar dataset to predict heart disease risk factors, demonstrating the potential of machine learning in healthcare.*

V. CONCLUSION AND FUTURE WORK

In conclusion, this survey has provided a comprehensive exploration of the integration of machine learning, with a specific emphasis on the Logistic Regression algorithm, in the development of heart disease detection software. As evidenced by the reviewed literature and methodologies, the utilization of advanced algorithms has shown promising results in enhancing the accuracy and efficiency of diagnostic tools for cardiovascular diseases.

Logistic Regression, known for its interpretability and simplicity, emerges as a valuable candidate for incorporation into heart disease detection software. The algorithm's ability to model complex relationships between features and outcomes, coupled with its transparent decision-making process, makes it well-suited for applications in the medical domain.

However, challenges such as handling imbalanced datasets, optimizing hyperparameters, and ensuring robust real-world deployment must be carefully addressed. The survey has shed light on these intricacies, emphasizing the importance of ongoing research and development to overcome these hurdles and refine the efficacy of heart disease detection software.

Looking forward, the evolving landscape of machine learning in healthcare presents exciting opportunities for further advancements. Future research could focus on refining algorithmic sophistication, exploring ensemble methods, and integrating diverse data sources to create more holistic and accurate predictive models.

As the field continues to progress, collaboration between researchers, practitioners, and healthcare professionals becomes paramount. The seamless integration of machine learning into clinical practice requires a concerted effort to bridge the gap between technological innovation and practical implementation.

In essence, the journey towards effective heart disease detection software involves continuous refinement, innovation, and a commitment to improving patient outcomes. By leveraging the capabilities of machine learning, particularly Logistic Regression, we stand at the forefront of a transformative era in cardiovascular healthcare, where early and accurate diagnosis can significantly impact patient well-being.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 | A Monthly Peer Reviewed & Referred Journal |

|| Volume 12, Issue 1, January 2024 ||

| DOI: 10.15680/IJIRCCE.2024.1201058 |

REFERENCES

[1] A. L. Bui, T. B. Horwich, and G. C. Fonarow, ''Epidemiology and risk profile of heart failure,'' Nature Rev. Cardiol., vol. 8, no. 1, p. 30, 2011.
[2] M. Durairaj and N. Ramasamy, ''A comparison of the perceptive approaches for preprocessing the data set for predicting fertility success rate,'' Int. J. Control Theory Appl., vol. 9, no. 27, pp. 255–260, 2016.
[3] L. A. Allen, L. W. Stevenson, K. L. Grady, N. E. Goldstein, D. D. Matlock, R. M. Arnold, N. R. Cook, G. M. Felker, G. S. Francis, P. J. Hauptman, E. P. Havranek, H. M. Krumholz, D. Mancini, B. Riegel, and J. A. Spertus, ''Decision making in advanced heart failure: A scientific statement from the American heart association,'' Circulation, vol. 125, no. 15, pp. 1928–1952, 2012.
[4] S. Ghwanmeh, A. Mohammad, and A. Al-Ibrahim, ''Innovative artificial neural networks-based decision support system for heart diseases diagnosis,'' J. Intell. Learn. Syst. Appl., vol. 5, no. 3, 2013, Art. no. 35396.
[5] Q. K. Al-Shayea, ''Artificial neural networks in medical diagnosis,'' Int. J. Comput. Sci. Issues, vol. 8, no. 2, pp. 150–154, 2011.
[6] J. Lopez-Sendon, ''The heart failure epidemic,'' Medicographia, vol. 33, no. 4, pp. 363–369, 2011. [7] P. A. Heidenreich, J. G. Trogdon, O. A. Khavjou, J. Butler, K. Dracup, M. D. Ezekowitz, E. A. Finkelstein, Y. Hong, S. C. Johnston, A. Khera, D. M. Lloyd-Jones, S. A. Nelson, G. Nichol, D. Orenstein, P. W. F. Wilson, and Y. J. Woo, ''Forecasting the future of cardiovascular disease in the united states: A policy statement from the American heart association,'' Circulation, vol. 123, no. 8, pp. 933–944, 2011.

INNO SPACE
SJIF Scientific Journal Impact Factor

Impact Factor: 8.379

doi cross ref

ISSN INTERNATIONAL STANDARD SERIAL NUMBER INDIA

निस्केयर NISCAIR

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

9940 572 462 6381 907 438 ijircce@gmail.com

www.ijircce.com

Scan to save the contact details