Navigating Cardiac Health through Machine Learning – A Literature Survey

Sharmila Rathod[1], Simra Bhombal[2], Siddhesh S. Ghadi[3], Sakshi Jangir[4]
[1]*Professor, Computer Engineering, MCT's Rajiv Gandhi Institute of Technology, Mumbai University*
[2,3,4]*Student, Computer Engineering, MCT's Rajiv Gandhi Institute of Technology, Mumbai University*

***Abstract*- In this comprehensive literature survey, the current state of knowledge regarding Cardiac disease and Machine Learning algorithms used for this purpose is explored. Our investigation encompasses a wide range of scholarly sources, including peer-reviewed articles and reports, to provide a comprehensive overview of the existing research landscape. Cardiovascular diseases (CVDs) remain a leading cause of mortality worldwide, necessitating advanced tools and techniques for early detection, risk assessment, and personalized treatment. Throughout this exploration, we investigate various facets of ML in Cardiology, including the development of predictive models, risk assessment techniques, and the application of ML in medical imaging and Electronic Health Records (EHR) analysis. This study focuses on the application of machine learning algorithms in the analysis of Electrocardiogram (ECG) data for cardiac disease detection. Machine learning (ML) has emerged as a promising avenue to augment traditional medical practices, offering innovative solutions for navigating cardiac health. This review will also encourage researchers, scholars and curious minds to get guidance for future research and innovation.**

Keywords – Machine Learning, Cardiovascular Disease, Electrocardiogram (ECG), Electronic Health Record, Predictive Model, Cardiology, Risk Assessment, Medical Imaging

I. INTRODUCTION

A literature survey on heart disease is a comprehensive exploration of existing research and studies related to various aspects of cardiovascular health. This survey is here to dig deep and really understand what we know about a specific topic regarding heart disease, including its causes, risk factors, diagnostic methods, treatment options, and prevention strategies by synthesizing the findings, sharing insights on this research and complexities of heart disease, and contributing to the ongoing efforts to tackle this widespread and potentially deadly condition. This survey aims to provide insights into the applications of machine learning in understanding, diagnosing, treating, and preventing heart disease. By synthesizing findings from diverse sources, this review seeks to highlight the advancements, challenges, and opportunities in utilizing machine learning algorithms, such as predictive modeling, image analysis, and data-driven approaches, to enhance our understanding and management of heart disease in a data-driven and personalized manner.

II. LITERATURE SURVEY

This conference paper by Ch. Usha Kumari et al. showed methods for Heart Rhythm Abnormality Detection they used Machine Learning Technique. Electrocardiogram (ECG) analysis is pivotal in the diagnosis and classification of cardiac irregularities. Over the past few years, machine learning has taken center stage in boosting the precision and speed of systems that classify ECG data. This literature survey explores the state of the art in the field of heartbeat irregularity classification, with a specific focus on the approach outlined in the research under consideration. In this particular research endeavor, the classifier of choice is the Support Vector Machine (SVM). Impressively, the system achieves an exceptional overall accuracy rate of 97.14 percent. Feature extraction is used for early detection of cardiac disease.

In this Nutwara Chumrita et al. showed insights on detection of Arrhythmia with Support Vector Machines using Average Energy and Zero-crossing Features. In the field of analyzing ECG signals, this study utilizes a wide-ranging dataset compiled from three well-known databases: MIT-BIH arrhythmia, MIT-BIH normal sinus rhythm, and QT databases. The primary aim is to demarcate and discriminate

between normal and abnormal ECG signals, with a particular focus on the detection of arrhythmia instances. It is proved to be the fastest and effective technique for machine learning [1], [2], [3]. In the process of preparing data, feature extraction plays a vital role by converting raw data into a concise and meaningful set of characteristics. This process is instrumental in managing and condensing extensive datasets, effectively reducing redundancy without sacrificing essential information embedded within the original data [4]. These preparatory steps are vital in ensuring data consistency and facilitating subsequent analysis and classification tasks.

Ramya G. Franklin et al. surveyed upon machine learning approaches for identification of cardiac disease. This survey was done in India in 2020. This was the year people highly suffered from pandemic worldwide. The COVID-19 pandemic posed significant challenges for heart patients. In [5], the study delves into the prediction of heart diseases by examining several factors. It introduces the Heart Diseases Prediction System (HDPS), which relies on Data Mining (DM) techniques. The study employs a variety of algorithms, including SVM, Naïve Bayes (NB), Random Forest, Decision Tree (DT), K-Nearest Neighborhood (KNN), Naïve Bayes (NB) again, and Artificial Neural Network (ANN). This approach enhances the current decision-making system by integrating various feature selection methods and algorithms, ultimately aiming to improve and refine the existing system. In [6] it introduces the Heart Disease Prediction System (HDPS), which has the capability to assess an individual's heart condition. It leverages algorithms such as Decision Tree (DT) and Support Vector Machine (SVM) Classifier for this purpose.

In this paper authored by Erik Alonso et al. originating from Spain, the primary aim of their paper is to develop a dependable pulse detection algorithm. Their targeted population comprises heart patients, and their research centers on the critical task of creating a reliable algorithm for pulse detection. This algorithm depends on the examination of two crucial signals: the electrocardiogram (ECG) and thoracic impedance (TI), both typically accessible in defibrillators. The study utilizes a dataset comprising 1140 ECG and TI segments obtained from 187 patients who experienced out-of-hospital cardiac arrest (OHCA). In this dataset, 792 segments were categorized as pulse-generating rhythm (PR), while 348 were diagnosed as pulseless electrical activity (PEA) by a panel of OHCA experts. Interestingly, this study does not employ a conceptual or theoretical framework, opting instead to focus on the practical aspects of pulse detection using ECG and TI signals within the context of OHCA patients' data. To assess the classification performance of feature subsets, the study employed a two-step approach. Firstly, feature selection was conducted within an inner loop using the sequential floating forward selection (SFFS) algorithm [7]. Subsequently, a multivariate logistic regression algorithm was utilized to evaluate the chosen feature subsets.

In this paper from Saudi Arabia, Aqsa Rahim et al. focus on predicting cardiovascular disease accurately. They aim to help people with heart problems. They recognize the importance of improving prediction methods for such life-threatening conditions. They introduce a framework called MaLCaDD, which uses machine learning to make better predictions. First, it handles missing data by filling in with averages. Then, it balances the data to ensure fairness. Lastly, it picks the most important factors for better predictions. This approach aims to make cardiovascular disease prediction more precise and helpful for patients. The framework's validation was conducted using three benchmark datasets—Framingham, Heart Disease, and Cleveland—yielding impressive accuracy rates of 99.1%, 98.0%, and 95.5%, respectively. This comparative analysis demonstrates that MaLCaDD's predictions, achieved with a reduced set of features, outperform existing state-of-the-art methods. Consequently, MaLCaDD is a highly dependable tool suitable for real-world applications in the early diagnosis of cardiovascular diseases. Additionally, image-based diagnosis through cardiac imaging [8] and the utilization of various datasets such as heart disease [20], Cleveland [21], Framingham [22], and cardiovascular disease [23] with diverse attributes have been explored by researchers for predicting cardiovascular diseases.

The research conducted by Anjan Nikhil Repaka and Sai Deepak Reavikanti, originating in India, aims to address the critical issue of heart disease prediction in the context of patient care. Their study focuses on

employing the Naïve Bayes algorithm to enhance the accuracy of heart disease diagnosis. To achieve this, the researchers delve into an extensive literature survey, collecting and analyzing previous data and information related to heart disease. By synthesizing this wealth of knowledge, they strive to develop a robust predictive model that can assist healthcare professionals in identifying and mitigating heart disease risks among patients. This research holds significant promise in advancing the field of cardiovascular health and improving patient outcomes. The researchers are focused on finding better ways to spot heart disease. They use computer tricks to find important stuff hidden in big data [12]. These tricks help them come up with different ideas [13], like searching for hidden treasures in data, to understand and detect heart problems more effectively. Their goal is to make it easier for doctors to diagnose and treat heart issues, ultimately improving people's health.

M. Haider and Mohammad Haikal Satria, researchers from Malaysia, have developed a helpful tool for people with heart issues. The system relies on 13 medical factors such as gender, blood pressure, and cholesterol to estimate a patient's risk of getting heart disease [14]. They made something called a "parameter tuning framework" for artificial neural networks (ANNs). These ANNs are like computer programs that can help with heart problems. To see how well their tool works, they used a bunch of data. This data helps them figure out if their tool is good at diagnosing and treating heart conditions. This research has the potential to significantly impact the care of individuals with cardiac conditions. In the healthcare sector, there's a wealth of data that includes information about patients and the diagnosis of different illnesses [15].

III. CONCLUSION

In summary, the literature survey highlights the promising potential of ECG data and machine learning in predicting cardiac diseases. Researchers have made substantial strides in developing accurate predictive models that can assist in early diagnosis. However, the field still faces challenges such as ensuring the quality of ECG data, addressing model interpretability, and translating these advancements into real-world clinical practice. To fully harness the benefits of these technologies, ongoing collaboration between healthcare professionals and machine learning experts is crucial to refine these models and implement them effectively for better patient care.

REFERENCE

[1] S. Hunta, N. Aunsri, and T. Yooyativong, “Drug-drug interactions prediction from enzyme action crossing through machine learning approaches,” in Proceedings of Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology (ECTI-CON), 2015 12th International Conference, June 2015.

[2] L. Chen, Y. Feng, M. A. Maram, Y. Wang, M. Wu, K. Hirota, and W. Pedrycz, “Multi-SVM based dempster–shafer theory for gesture intention understanding using sparse coding feature,” Applied Soft Computing, vol. 85, p. 105787, 2019.

[3] L. Sun, B. Zou, S. Fu, J. Chen, and F. Wang, “Speech emotion recognition based on DNN-decision tree SVM model,” Speech Communication, vol. 115, pp. 29 – 37, 2019.

[4] S. Ghosh, P. R N, and R. Tripathy, Heart Sound Data Acquisition and Preprocessing Techniques: A Review, 02 2020, p. 300.

[5] Sarangam Kodati & Dr. R. Vivekanandam (2018), “Analysis of Heart Disease using in Data Mining Tools Orange and WEKA”, Global Journal of Computer Science and Technology, Vol. 18.

[6] Shalet K.S, V. Sabarinathan, V. Sugumaran, V. J. Sarath Kumar (2015), “Diagnosis of Heart Disease Using Decision Tree and SVM Classifier”, International Journal of Applied Engineering Research, Vol. 10, No.68, pp. 598-602.

[7] P. Pudil, F. Ferri, J. Novovicova, and J. Kittler, ‘‘Floating search methods for feature selection with nonmonotonic criterion functions,’’ in Proc. 12th Int. Conf. Pattern Recognit. (IAPR), Oct. 1994, pp. 279–283.

[8] C. Martin-Isla, V. M. Campello, C. Izquierdo, Z. Raisi-Estabragh, B. Baeßler, S. E. Petersen, and K. Lekadir, ‘‘Image-based cardiac diagnosis with machine learning: A review,’’ Frontiers Cardiovascular Med., vol. 7, p. 1, Jan. 2020.

[9] Ch. Usha Kumari, R. Ankita, T.Pavani, N. Arun Vignesh, N. Tarun Varma, Md. Aqeel Manzar, A. Reethika, “Heart Rhythm Abnormality Detection and Classification using Machine Learning Technique,” in

Proceedings of the Fourth International Conference on Trends in Electronics and Informatics (ICOEI 2020)
[10] Nutwara Chumrit , Chalita Weangwan, Nattapol Aunsria, “ECG-based Arrhythmia Detection Using Average Energy and Zero-crossing Features with Support Vector Machine” in proceedings of Fourth International Conference (2020 IEEE)
[11] Ramya G. Franklin, Dr. B. Muthukumar, “Survey of Heart Disease Prediction and Identification using Machine Learning Approaches” in Proceedings of the Third International Conference on Intelligent Sustainable Systems [ICISS 2020]
[12] Purushottama. C, Kanak Saxenab, Richa Sharma (2016), “Efficient Heart Disease Prediction System”, Elsevier, Procedia Computer Science, No. 85, pp. 962 – 969.
[13] Shalet K.S, V. Sabarinathan, V. Sugumaran, V. J. Sarath Kumar (2015), “Diagnosis of Heart Disease Using Decision Tree and SVM Classifier”, International Journal of Applied Engineering Research, Vol. 10, No.68, pp. 598-602.
[14] Abhishek Rairikar, Vedant Kulkarni, Vikas Sabale, Harshavardhan Kale, Anuradha Lamgunde “Heart Disease Prediction Using Data Mining Techniques” © IEEE, ICCC, 2017, p.p. 1-8
[15] Kipp W. Johnson, BS, Jessica Torres Soto, MS, Benjamin S. Glicksberg (2018), “Artificial Intelligence in Cardiology”, Elsevier, Journal of The American College of Cardiology, Vol. 71, No. 23, pp. 2668 – 2679
[16] M. Haider Abu Yazid, Muhammad Haikal Satria, Shukor Talib, Novi Azman, “Artificial Neural Network Parameter Tuning Framework for Heart Disease Classification”, in proceeding of EECSI 2018, Malang - Indonesia, 16-18 Oct 2018
[17] Anjan Nikhil Repaka, Sai Deepak Ravikanti, Ramya G Franklin, “Design and Implementing Heart Disease Prediction Using Naives Bayesian” in proceedings of the Third International Conference on Trends in Electronics and Informatics (ICOEI 2019)
[18] Erik Alonso, Unai Irusta, Elisabete Aramendi, Mohamud R. Daya, “A Machine Learning Framework for Pulse Detection During Out-of-Hospital Cardiac Arrest”, supported in part by the Spanish Ministry of Science, September 2 2020.
[19] Aqsa Rahim, Yawar Rasheed, Farooque Azam, Muhammad Waseem Anwar, Muhammad Abdul Rahim, Abdul Wahab Muzaffar, “An Integrated Machine Learning Framework for Effective Prediction of Cardiovascular Diseases”, College of Computing and Informatics, Saudi Electronic University, Riyadh 11673, Saudi Arabia, July 20 2021.
[20] Nilesh Rathod, Sunil Wankhade, “Optimizing neural network based on cuckoo search and invasive weed optimization using extreme learning machine approach”, Neuroscience Informatics, Volume 2, Issue 3, 2022, 100075, ISSN 2772-5286, https://doi.org/10.1016/j.neuri.2022.100075. (https://www.sciencedirect.com/science/article/pii/S2772528622000371)
[21] S. Gaikwad, S. Patel and A. Shetty, "Brain Tumor Detection: An Application Based on Machine Learning," 2021 2nd International Conference for Emerging Technology (INCET), Belagavi, India, 2021, pp. 1-4, doi: 10.1109/INCET51464.2021.9456347
[22] S. Gaikwad, "Study on Artificial Intelligence in Healthcare," 2021 7th International Conference on Advanced Computing and Communication Systems (ICACCS), Coimbatore, India, 2021, pp. 1165-1169, doi: 10.1109/ICACCS51430.2021.9441741.

WEBSITES
[23] Heart Disease Dataset by UCI. Accessed: Oct. 25, 2020. [Online]. Available: https://archive.ics.uci.edu/ml/datasets/Heart+Disease
[24] Cleveland Dataset by KEEL. Accessed: Nov. 15, 2020. [Online]. Available: https://sci2s.ugr.es/keel/dataset.php?cod=57
[25] Framingham Dataset by Kaggle. Accessed: Nov. 20, 2020. [Online]. Available: https://www.kaggle.com/amanajmera1/framingham-heart study-dataset
[26] cardiovascular disease by Kaggle. Accessed: Oct. 15, 2020. [Online]. Available: https://www.kaggle.com/sulianova/cardiovascular-disease dataset