Integrating Predictive Analytics and Computational Statistics for Cardiovascular Health Decision-Making

Siddhartha Nuthakki[1], Vinoth Kumar Kolluru[2], Yudhisthir Nuthakki[3], Sonika Koganti[4]

[1]Senior Data Scientist, First Object Inc, TX, USA,
[2]Researcher, Stevens Institute of Technology,
[3]Software Engineer, NJ,
[4]Software Engineer, Thought Circuits, TX, USA

Abstract
Cardiovascular diseases (CVDs) continue to be the primary cause of morbidity and death. Several machine learning techniques are used to build a model for predicting the occurrence of heart disease. This study set out to determine how predictive analytics and computational statistics may work together to transform cardiovascular health decision-making by providing new perspectives on risk assessment, individualized treatment, and resource management. Regardless of language restrictions, each work of published literature from its release in 2015 until April 2023 was carefully reviewed in various kinds of electronic databases. Out of the 32 publications that were found during the search, 12 were included in the systematic literature. The results of the study showed that the ratio of heart disease deaths has decreased when these computerized learning-based expert medical decision-making systems were implemented. For nations dealing with a physician shortage and an overburdened healthcare system, machine learning presents a great potential. It is crucial to medicine because it can identify patterns in large data sets and make it easier to identify diagnostic markers linked to risk or illness. Based on analysis, it has been shown that no one data mining approach or classifier can consistently yield the best results for all types of healthcare data. Anatomic models specific to each patient are created using 3D imaging methods like CT and MRI to study circulation in the human cardiovascular system. Unlike the anatomic model, the patient-specific physiologic model is an abstract model that is based on the equations regulating blood flow in arteries and unique to the patient's diagnostic data, particularly blood circulation and pressure measurements.

Keywords - Predictive analytics, Computational statistics, Cardiovascular, CVD, and health decision-making

1. Introduction

A vital aspect of everyone's life is their health. But due to a variety of factors, including poor lifestyle choices, psychological strain at work, stress at work, and environmental hazards like pollution and unsafe working conditions, millions of people globally suffer from chronic illnesses like cardiovascular diseases (CVD), which damage the heart and blood vessels and can be fatal or severely crippling [1]. Globally, cardiovascular diseases (CVDs) continue to be the primary cause of morbidity and death [2]. The Global Burden of Disease Report 2019 [3] states that the prevalence of CVD is rising gradually, with 523 million cases reported in 2019. Of those cases, 18.6 million fatalities occurred, or one-third of all deaths.

Experience both domestically and abroad demonstrates that early identification and efficient management of high-risk populations through intervention is a clear technical path and an affordable preventative and control program that can raise life expectancy, enhance health and quality of life, and lessen the burden of disease nationally [4]. It has been discovered that a range of variables, including as genetic information, symptoms, lifestyle, and risk factors, are necessary for the correct prediction of cardiovascular disease (CVD) [5]. Using data to find trends and forecast future occurrences with confidence is known as predictive analytics [6]. Clinical prediction models (CPMs) are tools for estimating the likelihood of certain clinical outcomes. They have the potential to enhance clinical decision-making and personalize treatment [7, 8]. In contrast, computational statistics use mathematical models and algorithms for the purpose of data analysis and interpretation. Combining these two fields of study might revolutionize the way decisions about cardiovascular health are made by offering a thorough and individualized approach [9, 10]. Thus, the purpose of this work is to demonstrate how the combination of computational statistics and predictive analytics may transform the way that cardiovascular health decisions are made.

2. Literature Review

Researchers have developed a model to predict cardiac illness using a variety of machine learning approaches. For the purpose of identifying cardiac problems, Singh and Kumar advise utilizing machine learning techniques [11]. The UCI machine learning repository provided the database that was used in this investigation. This study demonstrates that, with an accuracy rate of 90%, the Random Forest method produces the best outcomes when compared to other machine learning strategies. Different data mining strategies for evaluating heart disease prediction are described by Bhatla et al. [12]. Neural networks with fifteen characteristics perform better than any other data mining technique, according to the findings. Another study conclusion is that by using a GA and feature subset selection, the Decision Tree may achieve high accuracy. Using a data mining method, Dehkordi et al. created a prediction model on the basis of the prescription [13].

To improve the system's accuracy, they suggested using the skate algorithm. Boosting and bagging are similar to skating, an ensemble technique [14]. Four distinct categorization techniques—decision trees, Naive Bayes, K-NN, and skating—have been compared and contrasted. They've shown that skating is the most accurate classifier available. The accuracy rating of this classification method is 73.17% [15]. Jan et al. used two benchmark datasets collected from a UCI repository to create an ensemble data mining technique that makes use of five distinct classification algorithms [16].

Soni et al. have employed decision trees in combination with GA to enhance the effectiveness of classification, in contrast to the other two approaches—cluster-based classification and Naïve Bayes [17]. It is found that the proposed method has 99.2% accuracy. Mansoor et al.'s study [18] looked at how well the classification algorithms logistic regression and Random Forest performed in assessing the risk susceptibility of cardiovascular patients. They have shown that the random forest categorization method is not as effective as the logistic regression model. The random forest model has an accuracy of 88%, whereas the logistic regression model has an accuracy of 89%. Austin et al. [19] compared the performance of regression trees with that of classic classification trees. Conventional logistic regression is a useful tool for assessing cardiac disease risk [20].

2.1 Predictive Analytics in Cardiovascular Risk Assessment

According to estimates from the World Health Organization, roughly 17 million fatalities worldwide are attributed to CVDs, making up around 31% of all deaths. Many lives can be saved and patients may recover from CVD with early diagnosis [21, 22]. Cardiologists still have difficulties in making early diagnoses and treating patients. All conventional models of CVD risk assessment implicitly presume a linear relationship between each risk factor and the outcome of CVD. These models frequently oversimplify intricate connections, such as those involving several risk variables and non-linear interactions [21]. It is important to appropriately include a variety of risk variables and identify the subtle correlations between the risk factors and results. Routine clinical data and machine learning (ML) have not yet been used to prognostic CVD evaluation in a large-scale investigation [23]. Several ML-based models for the identification of CVD have been presented in recent years.A cardiologist can treat a patient more appropriately by using machine learning (ML) to anticipate ailments early on [24]. Numerous machine learning (ML) approaches exist, each having pros and cons, including decision trees, artificial neural networks, support vector machines, and K-Nearest Neighbor (K-NN) [25, 26]. These techniques have been used in larger domains such as skin illnesses [30], human heart (echocardiogram signals) [29], and liver [28]. Each approach produces different results because of various limitations. Further room exists for the development of automated CVD detection utilizing additional ML models with higher performance, according to observations from previous research [31].

Every common model for assessing CVD risk implicitly assumes that there is a linear relationship between every risk factor and the outcomes of CVD [32]. As a result, these models run the danger of oversimplifying intricate connections that involve several risk variables and non-linear interactions. It is necessary to investigate methods that more effectively take into account a variety of risk variables and ascertain more complex correlations between risk factors and results [33]. With its alternative approach to traditional prediction modeling, machine-learning (ML) has the potential to overcome existing drawbacks. By more effectively utilizing "big data" for algorithm development, it has the potential to revolutionize medicine [34]. Machine learning (ML) is an offshoot of the field of computational learning, or "artificial intelligence," and pattern recognition research [35]. By minimizing the error between expected and observed outcomes, this method depends on a computer to understand the intricate and non-linear relationships between variables [8]. ML may uncover latent variables, which are unlikely to be seen but may be deduced from other factors, in addition to possibly enhancing prediction [36].

This study set out to determine how predictive analytics and computational statistics may work together to transform cardiovascular health decision-making by providing new perspectives on risk assessment, individualized treatment, and resource management.

3. Methodology

The systematic review carefully followed a defined process and complied with PRISMA criteria [37]. Finding pertinent research by searching via the four databases was the first step.

3.1 Data Strategy

The ScienceDirect, Cochrane Library, IEEE Xplore, and PubMed databases were used in this investigation. The database of PubMed in 2023. These libraries were chosen because they are leaders in the area and consistently publish high-caliber, mostly peer-reviewed research. Following that, as indicated in Table 1, the following keywords were defined using various Boolean operators to search the query.

Table 1: Keywords

Topic	Keywords	Search
Integrating predictive analytics and computational statistics for cardiovascular health decision-making	Predictive analytics, Computational statistics, Cardiovascular, and health decision-making	("predict"[All Fields] OR "predictable"[All Fields] OR "predictably"[All Fields] OR "predicted"[All Fields] OR "predicting"[All Fields] AND ("analyte"[All Fields] OR "analytes"[All Fields] OR "analytes"[All Fields] OR "analytic"[All Fields] AND ("computational"[All Fields] AND "statistics"[All Fields]) OR "computational statistics"[All Fields]) AND ("cardiovascular system"[MeSH Terms] OR ("cardiovascular"[All Fields] AND "system"[All Fields]) OR "cardiovascular system"[All Fields] OR "cardiovascular"[All Fields] OR "cardiovasculars"[All Fields])

		AND ("health"[MeSH Terms] OR "health"[All Fields] OR "health s"[All Fields] OR "healthful"[All Fields] OR "healthfulness"[All Fields] OR "healths"[All Fields]) AND ("decision making"[MeSH Terms] OR ("decision"[All Fields] AND "making"[All Fields]) OR "decision making"[All Fields]))

It is ineffective to choose relevant items by doing keyword-specific library searches. In order to create efficient searches, logical operators like "AND" and "OR" were used to combine the terms. Next, each library is subjected to the search query to identify articles published within the previous ten years (2013–2023). Similar keywords are found for each keyword phrase because a single keyword is insufficient to find pertinent research studies.

3.2 Screening of Articles

Once the relevant articles have been obtained from all the databases, remove the duplicates. After that, the abstracts, titles, and readings of the entire texts were used to evaluate the articles. Ultimately, 12 papers were selected for additional review and quality assessment (Appendix 1).

3.3 Inclusion and Exclusion Criteria

All retrospective, cross-sectional, longitudinal, and interventional studies were considered. It seemed appropriate that the piece be published in English. The non-English title and abstract of a study written in a language other than English are assessed. An attempt is made to translate the complete text if it is thought to be pertinent to the research outcomes of interest. The publishing year must to fall between 2015 and 2023. a unique research; instead, it incorporates information from conference abstracts, reviews, and other sources. There is no usage of English in the language.

3.4 Quality Appraisal Tools

In a systematic review, factual analysis and an assessment of the evidence's dependability are required. The CASP approach must be applied in order to evaluate the data's dependability as well as each study's inherent biases. The appropriate risk of bias score was determined by accounting for the construction of timelines, measurement mistakes, blinding, partial assessments, selective the efficacy of, and other biases [38].

3.5 Data synthesis strategy

The findings were reported in accordance with PRISMA criteria. Using themes identified in the recovered data, the study's results were condensed into a narrative account. A preliminary synthesis of the research findings from the included studies, an examination of the linkages between the studies, and a grading of the synthesis's strength were all part of the technique used to perform the study.

4. Result and Discussion

The process of selecting studies resulted in the creation of 395 distinct publications. A total of thirty-two articles underwent full-text assessments, and twelve articles were rejected based only on their abstracts or titles. After applying the exclusion criteria, 12 studies were located, and their quality was assessed. One research was omitted as it lacked identifying information. Part of the already-highlighted point is that the comprehensive synthesis of the key concerns fully met the predetermined goals of the current systematic review. Predictive analytics and computational statistics may work together to transform cardiovascular health decision-making that are relevant to the goals of the current systematic review.

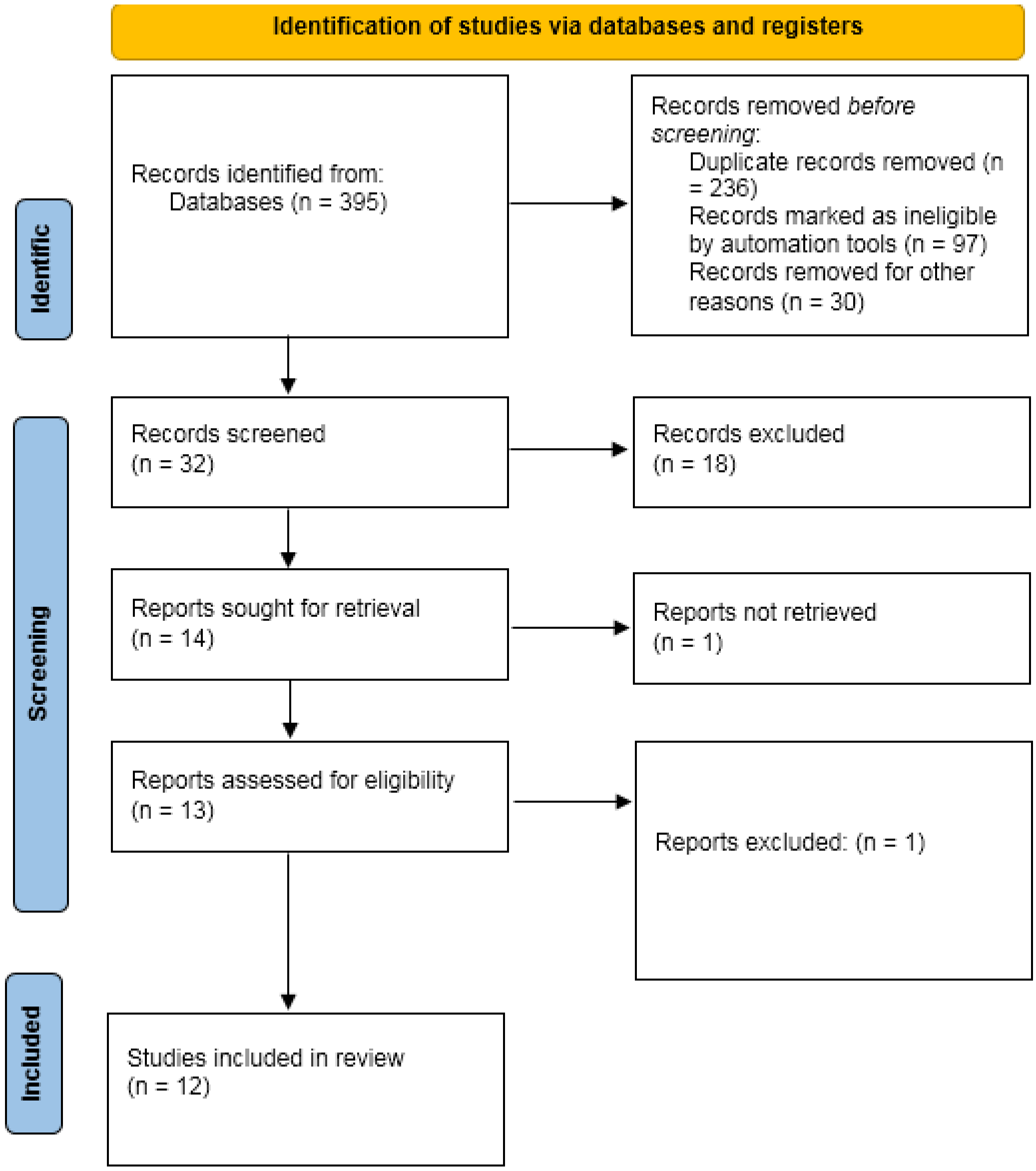

Figure 1: PRISMA Diagram

(The methodical review, elimination, and article selection processes are depicted in Figure 1.)

4.1 Theme 1: Predictive Analytics in Cardiovascular

ML techniques may support clinical management and experts investigate exceptional performance in several medical applications, including medical image analysis, language processing, and tumor or cancer cell identification. Algorithms for machine learning categorization have been introduced into clinical treatment [21]. Techniques for categorization can be used to extract knowledge. Practitioners can enhance patient outcomes and make better judgments by using accurate cardiac disease prediction [23]. Logistic regression (LR) and other machine learning-based techniques have been widely used in the early diagnosis of cardiac disease [11]. These expert medical judgment systems based on machine learning have reduced the ratio of death from heart disease. Machine learning holds considerable promise for nations grappling with a physician shortage and an overworked healthcare system. Because it can identify patterns in vast volumes of data and make it easier to identify diagnostic markers linked to disease or risk, it is crucial to medicine [17]. Created a cutting-edge deep learning system that uses face footage from an RGB camera to instantly predict heart rate [4]. An LSTM neural network model was presented by Singh and Kumar [11] for the early identification of heart failure (HF). Their suggested model was contrasted with many baseline models, including multilayer perceptrons (MLP), logistic regression, k-nearest neighbor (KNN), and support vector machines (SVM). The results show that the proposed model has the greatest accuracy when compared to other algorithms.

In a variety of health science applications, the subject of diabetes mellitus is expanding quickly [16]. Smart cardiac disease prediction systems and appropriate DM-based categorization approaches can result in low-cost, high-quality healthcare services with respect to accuracy. The decline in avoidable mistakes and the rising expense of healthcare are the primary drivers driving the digitalization of health data and the use of soft computing solutions. Figure 2[16] illustrates the basic operation of the fuzz computational-based DM framework for health informatics.

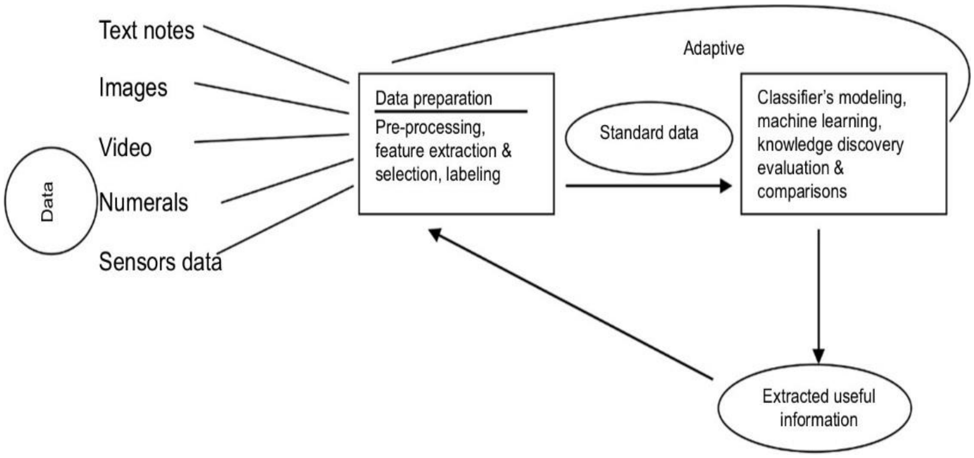

Figure 2: Adaptive DM model [16]

4.2 Theme 2: Computational Statistics for Personalized Medicine

The phrase "data mining" refers to a broad range of methods for obtaining valuable information from (big) data sets. Despite the fact that a sizable fraction of CVDs are avoidable, their prevalence is mostly due to insufficient preventative interventions. Heart attacks and strokes accounted for 7.3 million and 6.2 million, respectively, of the 17.3 million cardiovascular fatalities that occurred in 2008 [12].

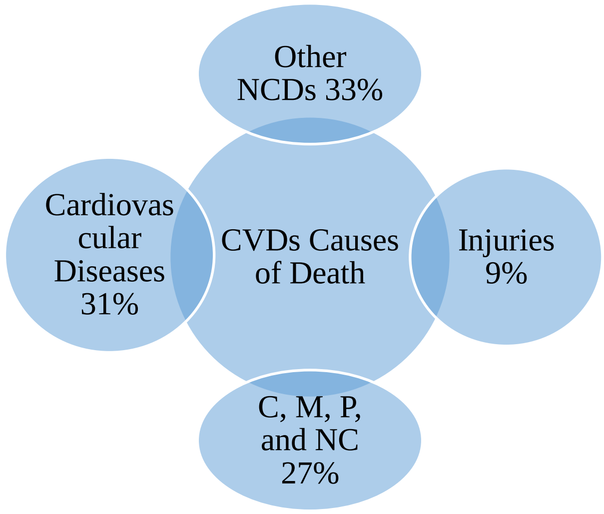

Figure 3: Distribution of major causes of deaths including CVDs.

Most researchers utilized the UCI machine learning library to diagnose heart disease in cases when only a few critical medical characteristics were present. According to observations, certain algorithms undoubtedly had the greatest accuracy but were unable to lower the number of heart disease detecting features [21]. A patient will naturally need to take fewer tests if there are fewer qualities employed in the identification of heart disease. By eliminating extraneous and irrelevant features from the dataset and selecting just those that are most useful for the diagnosis, algorithms can perform better [6].

Efficient diagnosis and/or prognosis of heart attacks by data mining techniques and approaches that significantly improve patient health and overall medical service quality. Based on analysis, it has been shown that no one data mining approach or classifier can consistently yield the best results for all types of healthcare data [23]. To get better results, hybrid or integrated data mining techniques can be applied, such as combining several classifiers, combining clustering with classification or association, etc. [17]. In conclusion, the UCI machine learning repository was utilized for testing, and the outcomes in risk prediction demonstrated a considerable improvement in accuracy, sensitivity, and specificity between the suggested clinical decision support system and the network-based system [12].

4.3 Theme 3: Optimization of Healthcare Resources

Clinical notes, including clinical, transfer, and discharge notes, can be transformed into structured text using natural language processing (NLP) in a way that is predefined and prepared for analysis [9]. For those working in health information management (HIM), this can be a very helpful tool because it can process text immediately. Additionally, it may be used by organizations to streamline communication between caregivers, offer a more cost-effective way to collect and analyze clinical information, and automate coding, which is a crucial administrative duty associated with keeping medical records, billing, and being paid. The drawback of most NLP systems is that they struggle to handle non-grammatical content that contains telegraphic phrases, bullet points, and incomplete sentences [9, 16].

The trend of automation in the coming decades will be sustained by the quick advancements in machine learning and artificial intelligence as well as the latest advancements in ubiquitous computing and neural network (neuromorphic) hardware technologies, which enable machines to match human-like cognitive and perceptual abilities [10]. Every day, diagnostic mistakes happen at a rate of around 3%–4% on average. The percentage of retrospective error for interpreting abnormal results is 32%. Despite decades of study, these large mistakes still occur. The availability of previous or comparable imaging, clinical history, interruptions during actual clinical justifications of imaging, diagnostic quality of imaging facilities involving artifacts, and processes factors like fatigue and turnaround time are a few examples of elements that may affect radiology performance [10, 21]. Expert radiologists' ability to identify chest X-ray pictures might be useful in supporting medical professionals in clinical situations. The models performed as predicted and required little effort to use. The developers were very motivated to present an elegant internal view of the workings of the deep learning model in the hidden layers and how it made the classification, given the advanced backend [14].

Due to its benefits and convenience of use, mobile health (mHealth) is becoming more and more popular. It has been demonstrated to increase service quality at a reasonable cost. The ECEB seeks to give information and skills pertaining to the majority of ENC components to women and families, as well as to educate all healthcare professionals. Based on the various infant situations, the ECEB action plan outlines care that must be provided from the moment of birth to 24 hours following delivery. It lessens the cognitive strain of caring for several infants and aids in spotting any performance gaps in medical professionals [20].

4.4 Theme 4: Computational Techniques in Decision-Making

Predicting the result of a therapeutic option selected for a specific patient based on information from "similar" medications previously administered to other persons with comparable characteristics is one aspect of decision-making, which is mostly based on experience and empirical data [20]. Although statistical tools for decision-support can help with this process, they still primarily rely on empirical data that might not be entirely correct in describing the specifics of that problem. In other words, this decision-making process lacks knowledge or projections about the future for each given patient, despite being loaded with information about past and present events for other subjects [10, 23].

In the predictive medicine paradigm presented here, a physician would build an anatomy and physiology model of a patient using diagnostic data. The physician may then predict how the patient would respond to different therapies given under different physiological conditions by utilizing simulation techniques incorporated into a simulation-based healthcare management computer program [9]. Anatomic models specific to each patient are created using 3D imaging methods such as CT and MRI to simulate blood flow in human blood vessels. Unlike the anatomic model, the patient-specific physiologic model is an abstract model based on the equations regulating blood flow in arteries and patient-specific diagnostic information including circulation stream and pressure measurements [12, 14].

Blood flow simulations in simulations representing the recommended therapies can be used to predict treatment results when the patient-specific model has been created [17]. Treatment planning has taken a revolutionary turn with the use of simulation-based, predictive medical planning techniques in clinical settings. In the event of a cardiovascular disease, these methods might assist medical professionals in developing medication regimens tailored to each patient that improve blood flow [11]. Understanding the effects of hemodynamic conditions on vascular adaptation and illness, as well as determining the optimal hemodynamic parameters for a given patient, should lead to the development of better medicines to improve patient care [4].

5. Conclusion

This study predicts cardiovascular disease using a variety of machine learning-based methods. The usefulness of these methods has been evaluated, and a comparative analysis has been finished. Various machine learning approaches to construct a heart disease prediction model. These expert medical judgment systems based on machine learning have reduced the ratio of death from heart disease. The study's result showed Machine learning holds considerable promise for nations grappling with a physician shortage and an overworked healthcare system. Because it can identify patterns in vast volumes of data and make it easier to identify diagnostic markers linked to disease or risk, it is crucial to medicine. Research has shown that no single data mining method or classifier can consistently yield the best results for all types of medical data. Anatomic models specific to each patient are created using 3D imaging methods like CT and MRI to study blood flow in the human body's cardiovascular system. By methodically examining the effectiveness of the various elements, the physicians will be assisted in efficiently preserving the data. Rather than capturing and archiving every characteristic, the data management team may store just the ones that are critical for heart disease prediction. Using a data set from the machine learning repository, the simulation was run, and the findings guaranteed that the accuracy, CPU use, and processing time of the suggested system would increase dramatically in the future. Future research may examine varied environments.

References

1. B. S. Wessler et al., “Clinical Prediction Models for Cardiovascular Disease,” Circulation: Cardiovascular Quality and Outcomes, vol. 8, no. 4, pp. 368–375, Jul. 2015, doi: https://doi.org/10.1161/circoutcomes.115.001693.
2. J. S. Rumsfeld, K. E. Joynt, and T. M. Maddox, “Big data analytics to improve cardiovascular care: promise and challenges,” Nature Reviews Cardiology, vol. 13, no. 6, pp. 350–359, Mar. 2016, doi: https://doi.org/10.1038/nrcardio.2016.42.
3. Global Burden of Disease Study, “Global Burden of Disease Study 2019 (GBD 2019) Healthcare Access and Quality Index 1990-2019 | GHDx,” ghdx.healthdata.org, 2019. https://ghdx.healthdata.org/record/ihme-data/gbd-2019-healthcare-access-and-quality-1990-2019
4. Kathiriya, S., Nuthakki, S., Mulukuntla, S., & Charllo, B. V. “AI and The Future of Medicine: Pioneering Drug Discovery with Language Models”, International Journal of Science and Research, vol. 12, no. 3,

pp. 1824-1829, Mar. 2023, doi: https://dx.doi.org/10.21275/SR24304173757
5. M. Bilal, M. Hussain, Iqra Basharat, and M. Fatima, "Cardiac data mining (CDM); organization and predictive analytics on biomedical (cardiac) data," Nucleation and Atmospheric Aerosols, Jan. 2013, doi: https://doi.org/10.1063/1.4825018.
6. Nuthakki, S., Kumar, S., Kulkarni, C. S., & Nuthakki, Y. "Role of AI Enabled Smart Meters to Enhance Customer Satisfaction". International Journal of Computer Science and Mobile Computing, Vol.11 Issue.12, December- 2022, pg. 99-107, doi: https://doi.org/10.47760/ijcsmc.2022.v11i12.010
7. R. TR, U. K. Lilhore, P. M, S. Simaiya, A. Kaur, and M. Hamdi, "Predictive Analysis of Heart Diseases With Machine Learning Approaches," Malaysian Journal of Computer Science, pp. 132–148, Mar. 2022, doi: https://doi.org/10.22452/mjcs.sp2022no1.10.
8. D. A. Jenkins et al., "Continual updating and monitoring of clinical prediction models: time for dynamic prediction systems?," Diagnostic and Prognostic Research, vol. 5, no. 1, Jan. 2021, doi: https://doi.org/10.1186/s41512-020-00090-3.
9. Singh, D., Bhogawar, S., Nuthakki, S., & Ranganathan, N. (2021). Enhancing Patient-Centered Care in Oncology through Telehealth: Advanced Data Analytics and Personalized Strategies in Breast Cancer Treatment. Breast cancer, 3, 12.
10. J. W. Gichoya, S. Nuthakki, P. G. Maity, and S. Purkayastha, "Phronesis of AI in radiology: Superhuman meets natural stupidity," arXiv.org, Mar. 27, 2018. https://arxiv.org/abs/1803.11244.
11. A. Singh and R. Kumar, "Heart Disease Prediction Using Machine Learning Algorithms," IEEE Xplore, 2020. https://ieeexplore.ieee.org/abstract/document/9122958
12. N. Bhatla and K. Jyoti, "An Analysis of Heart Disease Prediction using Different Data Mining Techniques," International journal of engineering research and technology, vol. 1, no. 8, Oct. 2012.
13. S. K. Dehkordi and H. Sajedi, "Prediction of disease based on prescription using data mining methods," Health and Technology, vol. 9, no. 1, pp. 37–44, Aug. 2018, doi: https://doi.org/10.1007/s12553-018-0246-2.
14. M. E. Hossain, S. Uddin, and A. Khan, "Network analytics and machine learning for predictive risk modelling of cardiovascular disease in patients with type 2 diabetes," Expert Systems with Applications, vol. 164, p. 113918, Feb. 2021, doi: https://doi.org/10.1016/j.eswa.2020.113918.
15. E. O. Olaniyi, O. K. Oyedotun, and K. Adnan, "Heart Diseases Diagnosis Using Neural Networks Arbitration," International Journal of Intelligent Systems and Applications, vol. 7, no. 12, pp. 75–82, Nov. 2015, doi: https://doi.org/10.5815/ijisa.2015.12.08.
16. Singh, D., Nuthakki, S., Naik, A., Mullankandy, S., Singh, P. K., & Nuthakki, Y. (2022). "Revolutionizing Remote Health: The Integral Role of Digital Health and Data Science in Modern Healthcare Delivery", Cognizance Journal of Multidisciplinary Studies, Vol.2, Issue.3, March 2022, pg. 20-30, doi: 10.47760/cognizance.2022.v02i03.002
17. J. Soni, U. Ansari, D. Sharma, and S. Soni, "Predictive Data Mining for Medical Diagnosis: An Overview of Heart Disease Prediction," International Journal of Computer Applications, vol. 17, no. 8, pp. 43–48, Mar. 2011, doi: https://doi.org/10.5120/2237-2860.
18. H. Mansoor, I. Y. Elgendy, R. Segal, A. A. Bavry, and J. Bian, "Risk prediction model for in-hospital mortality in women with ST-elevation myocardial infarction: A machine learning approach," Heart & Lung, vol. 46, no. 6, pp. 405–411, Nov. 2017, doi: https://doi.org/10.1016/j.hrtlng.2017.09.003.
19. P. C. Austin, J. V. Tu, J. E. Ho, D. Levy, and D. S. Lee, "Using methods from the data-mining and machine-learning literature for disease classification and prediction: a case study examining classification of heart failure subtypes," Journal of Clinical Epidemiology, vol. 66, no. 4, pp. 398–407, Apr. 2013, doi: https://doi.org/10.1016/j.jclinepi.2012.11.008.
20. Singh, D., Nuthakki, S., Naik, A., Mullankandy, S., Singh, P. K., & Nuthakki, Y. (2022). Revolutionizing

Remote Health: The Integral Role of Digital Health and Data Science in Modern Healthcare Delivery.

21. A. Muniasamy, V. Muniasamy, and R. Bhatnagar, “Predictive Analytics for Cardiovascular Disease Diagnosis Using Machine Learning Techniques,” Advances in intelligent systems and computing, pp. 493–502, May 2020, doi: https://doi.org/10.1007/978-981-15-3383-9_45.
22. World Health Organization, “Global recommendations on physical activity for health,” 2010, doi: https://doi.org/9789241599979.
23. V. Rajalakshmi, D. Sasikala, and A. Kala, “A Predictive Analysis for Heart Disease Using Machine Learning,” Advances in intelligent systems and computing, pp. 473–479, Sep. 2020, doi: https://doi.org/10.1007/978-981-15-5566-4_42.
24. D Singh, S Bhogawar, S Nuthakki, N Ranganathan, "Enhancing Patient-Centered Care in Oncology through Telehealth: Advanced Data Analytics and Personalized Strategies in Breast Cancer Treatment", International Journal of Science and Research (IJSR), Volume 10 Issue 9, September 2021, pp. 1707-1715, https://www.ijsr.net/getabstract.php?paperid=SR24108012724
25. T. A. Pearson et al., “Precision Health Analytics With Predictive Analytics and Implementation Research: JACC State-of-the-Art Review,” Journal of the American College of Cardiology, vol. 76, no. 3, pp. 306–320, Jul. 2020, doi: https://doi.org/10.1016/j.jacc.2020.05.043.
26. C. Krittanawong, H. Zhang, Z. Wang, M. Aydar, and T. Kitai, “Artificial Intelligence in Precision Cardiovascular Medicine,” Journal of the American College of Cardiology, vol. 69, no. 21, pp. 2657–2664, May 2017, doi: https://doi.org/10.1016/j.jacc.2017.03.571.
27. C. A. Taylor et al., “Predictive Medicine: Computational Techniques in Therapeutic Decision-Making,” Computer Aided Surgery, vol. 4, no. 5, pp. 231–247, Jan. 1999, doi: https://doi.org/10.3109/10929089909148176.
28. E. D. Peterson, “Machine Learning, Predictive Analytics, and Clinical Practice,” JAMA, vol. 322, no. 23, p. 2283, Dec. 2019, doi: https://doi.org/10.1001/jama.2019.17831.
29. M. Oyeleye, T. Chen, S. Titarenko, and G. Antoniou, “A Predictive Analysis of Heart Rates Using Machine Learning Techniques,” International Journal of Environmental Research and Public Health, vol. 19, no. 4, p. 2417, Feb. 2022, doi: https://doi.org/10.3390/ijerph19042417.
30. A. Dinh, S. Miertschin, A. Young, and S. D. Mohanty, “A data-driven approach to predicting diabetes and cardiovascular disease with machine learning,” BMC Medical Informatics and Decision Making, vol. 19, no. 1, Nov. 2019, doi: https://doi.org/10.1186/s12911-019-0918-5.
31. G. Chen, B. Ning, and T. Shi, “Single-Cell RNA-Seq Technologies and Related Computational Data Analysis,” Frontiers in Genetics, vol. 10, Apr. 2019, doi: https://doi.org/10.3389/fgene.2019.00317.
32. D. D. Miller, “Machine Intelligence in Cardiovascular Medicine,” Cardiology in Review, vol. 28, no. 2, pp. 53–64, Mar. 2020, doi: https://doi.org/10.1097/crd.0000000000000294.
33. Nuthakki, S., Neela, S., Gichoya, J. W., & Purkayastha, S. (2019). Natural language processing of MIMIC-III clinical notes for identifying diagnosis and procedures with neural networks. arXiv preprint arXiv:1912.12397.
34. D. Panda and S. R. Dash, “Predictive Analytics for Advance Healthcare Cardio Systems,” pp. 187–219, Jan. 2023, doi: https://doi.org/10.1007/978-981-99-1482-1_9.
35. S. Narejo, A. Shaikh, M. M. Memon, K. Mahar, Z. Aleem, and B. Zardari, “Big data analytics and classification of cardiovascular disease using machine learning,” Journal of Intelligent & Fuzzy Systems, pp. 1–9, Mar. 2022, doi: https://doi.org/10.3233/jifs-219302.
36. S. Molla et al., “A predictive analysis framework of heart disease using machine learning approaches,” Bulletin of Electrical Engineering and Informatics, vol. 11, no. 5, pp. 2705–2716, Oct. 2022, doi: https://doi.org/10.11591/eei.v11i5.3942.
37. P. S. Fleming, D. Koletsi, and N. Pandis, “Blinded by PRISMA: Are Systematic Reviewers Focusing on

PRISMA and Ignoring Other Guidelines?," PLoS ONE, vol. 9, no. 5, p. e96407, May 2014, doi: https://doi.org/10.1371/journal.pone.0096407.
38. H. A. Long, D. P. French, and J. M. Brooks, "Optimising the Value of the Critical Appraisal Skills Programme (CASP) Tool for Quality Appraisal in Qualitative Evidence Synthesis," Research Methods in Medicine & Health Sciences, vol. 1, no. 1, pp. 31–42, 2020, doi: https://doi.org/10.1177/2632084320947559.

Appendix 1

Table 2: Data extracted

Ref No.	**Aim**	**Model**	**Key Findings**
[4]	Assess the necessity of machine learning and look at the possible drawbacks and difficulties in applying it to cardiovascular.	➢ machine learning ➢ cognitive learning, ➢ deep learning ➢ and reinforcement learning-based methods	➢ The factors that influence the modeling challenge include data heterogeneity, ➢ data depth, data breadth, machine learning ➢ and feature selection techniques chosen, and orthogonal evidence.
[14]	An essential diagnostic tool for respiratory disorders is a chest X-ray.	Chest X-rays.	➢ X-rays of the chest and placing them into a GPU server with the algorithm for machine learning-based diagnosis of 14 lung diseases
[10]	how the performance of humans and AI might enhance for detecting pneumonia in chest X-rays.	Chest X-rays	➢ useful in actual clinical practice ➢ AI domain a more objective and replicable framework ➢ to reduce cognitive errors in imaging
[20]	Testing the Usability of a Mobile App for Decision Support	ECEB Program	➢ The ECEB action plan's usefulness, the complexities of tracking, monitoring, ➢ and caring for many newborns, ➢ and the key newborn care practices followed by healthcare providers
[9]	the performance of a NLP model	NLP and conventional machine learning models	➢ Deep learning models outperform other mapping techniques. ➢ 70.7% and 63.9% accuracy, respectively, are projected for the top-50 ICD-9 codes associated with diagnoses and procedures.
[11]	determine the machine learning	Machine Learning Algorithms	➢ Using the UCI repository dataset for training and testing,

	algorithms' accuracy in predicting cardiac disease.		➢ prestigious support is provided for the prediction algorithms
[17]	approaches for data mining used in the prediction of heart disease	KNN, Neural Networks	➢ the accuracy of Bayesian classification and decision trees is further improved. ➢ reduce the real data size to obtain the ideal subset of attributes required for heart disease prediction.
[23]	prediction model to be used in the diagnosis of medicine	Machine learning Cleveland heart disease dataset	➢ Treatment is provided to patients to reduce their chance of illness. ➢ Human tissues or blood are taken to forecast cardiovascular illnesses.
[21]	Making informed judgments in the presence of sickness analysis	machine learning and predictive analytics	➢ SVM and KNN have not outperformed ANN, RF, LDA, DT, and ➢ confusion matrix in terms of accuracy, recall, F1-score, and error rate.
[6]	evaluates the risk of CVD in individuals with type 2 diabetes using machine learning algorithms	prediction model utilising administrative data	➢ The network- and machine learning-based risk prediction model employing administrative data is demonstrated by the classifiers' accuracy, which varied from 79% to 88%.
[12]	recent introduction of data mining tools for the prediction of heart disease	Neural networks	➢ Using 15 characteristics, neural networks have surpassed all other data mining approaches. ➢ They have also demonstrated the greatest accuracy, or 100%, to date. ➢ Decision trees have also done well, using 15 attributes to achieve 99.62% accuracy.
[16]	An assessment of cardiovascular risk using predictive modeling	DM model	➢ Excellent diagnostic performance in terms of predicted accuracy and dependability; ➢ affordable and timely predictive option possessing an easy-to-use graphical user interface