ISSN: 2395-7852

International Journal of Advanced Research in Arts, Science, Engineering & Management (IJARASEM)

Volume 11, Issue 2, March 2024

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

IMPACT FACTOR: 7.583

| www.ijarasem.com | ijarasem@gmail.com | +91-9940572462 |

International Journal of Advanced Research in Arts, Science, Engineering & Management (IJARASEM)

| ISSN: 2395-7852 | www.ijarasem.com | Impact Factor: 7.583 | Bimonthly, Peer Reviewed & Referred Journal|

| Volume 11, Issue 2, March 2024 |

Predicting Stroke Risk with Machine Learning and a Web-Based Early Intervention Tool

Dr.T.ARAVIND, VENKATESAN J, VIKRAM C, SURYA A

Assistant Professor, Department of CSE, Muthayammal Engineering College (Autonomous), Rasipuram, Tamil Nadu, India

Department of CSE, Muthayammal Engineering College (Autonomous), Rasipuram, Tamil Nadu, India

Department of CSE, Muthayammal Engineering College (Autonomous), Rasipuram, Tamil Nadu, India

Department of CSE, Muthayammal Engineering College (Autonomous), Rasipuram, Tamil Nadu, India

ABSTRACT: This study presents an exploratory data analysis (EDA) concentrated on recognizing risk factors connected with heart attacks by leveraging machine learning techniques. Given the complexity and multifactorial nature of heart disease, identifying and recognizing the most significant predictors are actually for early intervention and prevention strategies. Through feature importance ranking, correlation analysis, and predictive modeling, we highlight key factors contributing to increased heart attack risk. These findings underscore the importance of combining EDA with machine learning to enhance our understanding of heart disease dynamics, offering potential pathways for more personalized and effective preventive healthcare measures. The study targets not only to support the academic discourse on heart disease prediction but also to serve as a foundation for developing more sophisticated predictive models and health policy planning.

KEYWORDS: Exploratory Data Analysis, Machine Learning, Heart Attack Risk Factors, Predictive Modeling, Feature Importance, Health Data Analytics

I. INTRODUCTION

This study explores the intricate connection between various health indicators and their potential contribution to heart disease. It utilizes a dataset enhanced along with both categorical and numerical data. This dataset involves a wide scale of factors, including demographic details like age and sex, clinical parameters such as chest pain type, resting blood pressure, serum cholesterol, fasting blood sugar, and resting electrocardiographic results, among others. Each of these features provides a unique perspective on an individual's health profile page, providing essential insights into the multifaceted attributes of heart disease advancement. Incorporating an output variable showing the medical diagnosis of cardiovascular disease enhances the dataset's utility, paving the way for the development of predictive models targeted at very early detection and prevention.

Cardiovascular diseases, particularly cardiovascular disease, stand as the leading cause of mortality worldwide, providing a substantial public health challenge. The potential to accurately predict and alleviate the risk of heart attacks with the identification of crucial risk factors is of significant relevance in decreasing these statistics. Traditional strategies for recognizing cardiovascular disease have primarily depended on statistical analyses of risk factors such as age, sex, cholesterol levels, high blood pressure, diabetes mellitus, and also smoking practices. This research specifies the dataset preprocessing steps to prepare the data for analysis and correlation analysis.

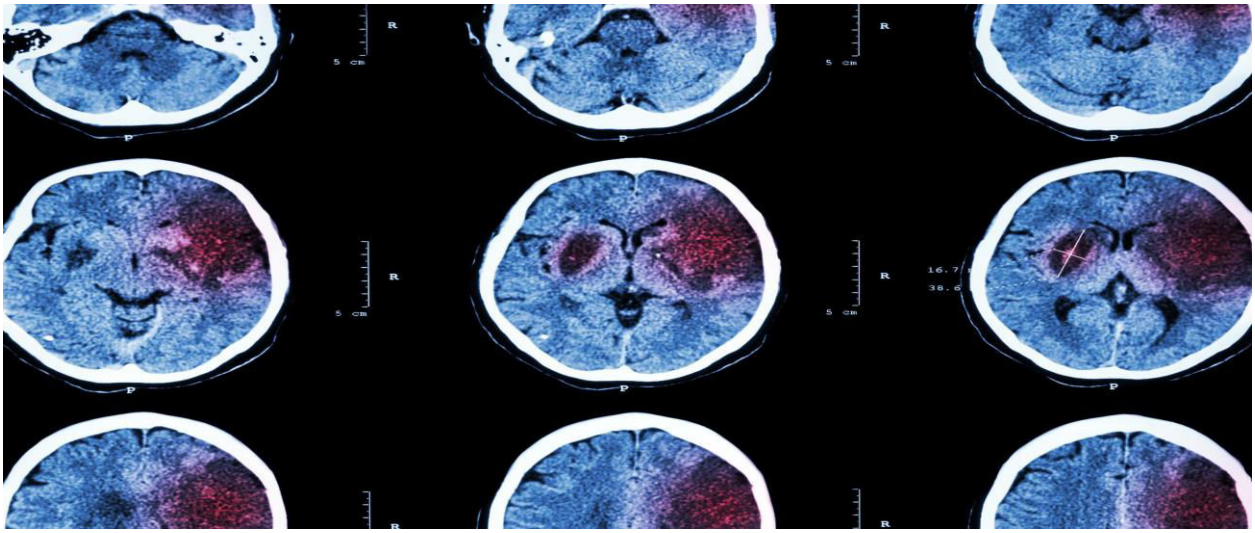

Fig 1: Machine learning model accurately predicts stroke

This paper kicks off with an in-depth exploratory data analysis (EDA), setting the stage for understanding the dataset's structure and characteristics, which is vital for subsequent analytical phases. After meticulous data cleaning, the study examines a rich array of health indicators ranging from demographics like age and sex to specific medical measures such as chest pain type, blood pressure, cholesterol, glucose levels, and ECG results. The study further involves blood pressure, cholesterol, and glucose levels, recognized as conventional markers of heart health, and uses electrocardiographic outcomes and maximum heart rate during stress tests as indicators of potential cardiac issues. The research employs `dplyr` and `ggplot2` for data manipulation and visualization, presuming access to an existing dataframe or the necessity to import data for analysis.

II. RELATED WORK

The word stroke, similar to the term heart attack, has a forbidding connotation in our minds. It delivers up the image of a life threatening event, with dismal consequences that include disability, paralysis or even death. Magnify this with the national statistics on stroke and the concerns can only intensify – given that stroke is one of the leading causes of death and disability in India. And that the incidence (new cases per year) and prevalence (total number of cases in the community) of stroke in India actually exceed those in developed, high-income countries.

Before we despair, there is good news. Medical technology and science has advanced to the point where stroke is treatable if diagnosed early – the emphasis being on early. New age Artificial intelligence (AI) algorithms can help doctors diagnose stroke within minutes as the patient is wheeled out of the CT room. To understand how AI can help, one must know what a stroke is. Stroke's end result is damage to brain tissue. One form of stroke is the "ischemic" stroke in which there is lack of oxygen supply to brain cells due to a blocked artery. The other is "hemorrhagic" stroke, where there is bleeding within the brain.

It is also important to realize that the early diagnosis of stroke is key to its effective treatment and to prevention of its progression. The Golden Hour concept in stroke care refers to the fact that if treatment is instituted promptly, the chances of survival and avoiding long term brain damage is significantly decreased. According to a study from the American Heart Association, every minute in which a stroke is untreated, the average patient loses 1.9 million neurons. Therefore, as the saying among Stroke Care teams goes, Time is Brain, and Time saved is Brain saved.

Once the patient is brought into the hospital the key diagnostic test is a CT scan of the brain. The CT scan performs a very critical role of showing the Doctors whether the stroke is hemorrhagic or ischemic. The reason this is so important is that the treatment is diametrically opposite for the two conditions. In ischemic stroke, treatment consists of administering blood thinning medications (such as TPA or Tissue Plasminogen Activator) to dissolve the blood clot that has formed. On the other hand, in hemorrhagic stroke such blood thinners would

actually worsen the bleeding and should not be used. Advanced imaging can also help identify the blocked vessel (CT Angiography), and estimate the amount of brain tissue at risk of damage (Perfusion CT).

The challenge with this advanced imaging is that while extremely accurate, it produces a very large amount of data that needs to be analyzed in a very short period of time by a radiologist. And the greater challenge is that radiologists today are in short supply worldwide which means that the likelihood of having a radiologist on-site at a hospital at the same time when a patient with an acute stroke is wheeled in is not always guaranteed.

Enter the technology solutions of Teleradiology and Artificial Intelligence. Teleradiology brings the scan to the radiologist, who may be anywhere, to ensure that it is immediately interpreted. It is possible to have a stroke CT scan performed in Boston reported within 15 minutes from across the globe. From our reading center in Bangalore, we have reported thousands of such stroke CT scans over the past two decades from emergency rooms across the United States. Such a network across India could similarly ensure immediate analysis by an expert of a stroke CT scan performed at any center anywhere in India.

III. METHODS

The improved model was trained on data from patients from age 35 to 70 years old. When developing medical prediction models for stroke, heart disease, and mortality from circulatory disease, researchers look to measure the architecture of the arteries and veins in the retina, the area of the vessels and the degree to which they were curved. After the initial scan, they researchers used the AI models to analyze the retinal pictures of of over 9,500 patients who participated in the research. The participants' assigned risk was compared to established cardiovascular risk prediction models.

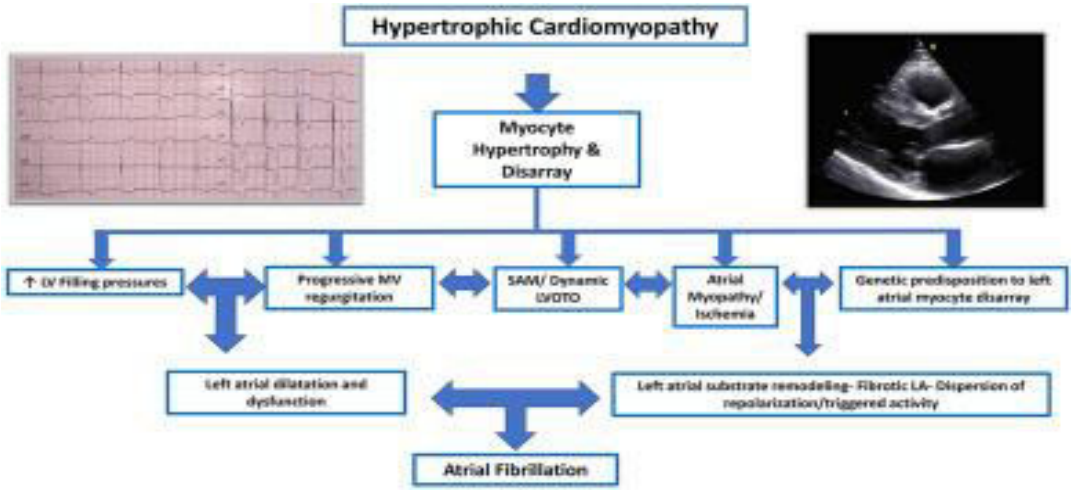

Fig 2: Atrial fibrillation in hypertrophic

The overall health of each participant was monitored for a period of five to seven years, on average. Once the trial was over researchers discovered that the breadth, curviness, and width variation of veins and arteries in the retina were strong predictors of mortality from cardiovascular illness. This was the case, particularly in men, and variations in artery area and breadth and vein curviness and width contributed to risk prediction in females. The artificial intelligence model also used the individuals' medical history data, including smoking history, medications used to manage high blood pressure, and prior comorbidities.

Researchers have developed an new AI powered eye exam that can determine the likelihood of that person developing heart disease in under a minute. The discovery paves the way for a very accurate test, that does not require any intrusive procedures, and does not even require a visit to the doctors office. By scanning a person's eyes with an artificial intelligence tool from an App on a cell phone doctors can will be able to precisely determine a person's risk of developing heart disease in the future. After some of the most extensive research ever conducted, the innovative new test may make it possible for ophthalmologists and other general physicians to conduct cardiovascular screening using a simple camera. This non-invasive method would eliminate the requirement for blood tests, stress tests or blood pressure exams.

IV. RESULT ANALYSIS

The prevalence of heart failure with preserved ejection fraction (HFpEF) is increasing with the aging population.1,2 However, developing tools to diagnose left ventricular diastolic dysfunction (LVDD) as the underlying mechanism for HFpEF in the older adult population is difficult due to age-related physiological changes in cardiac function.

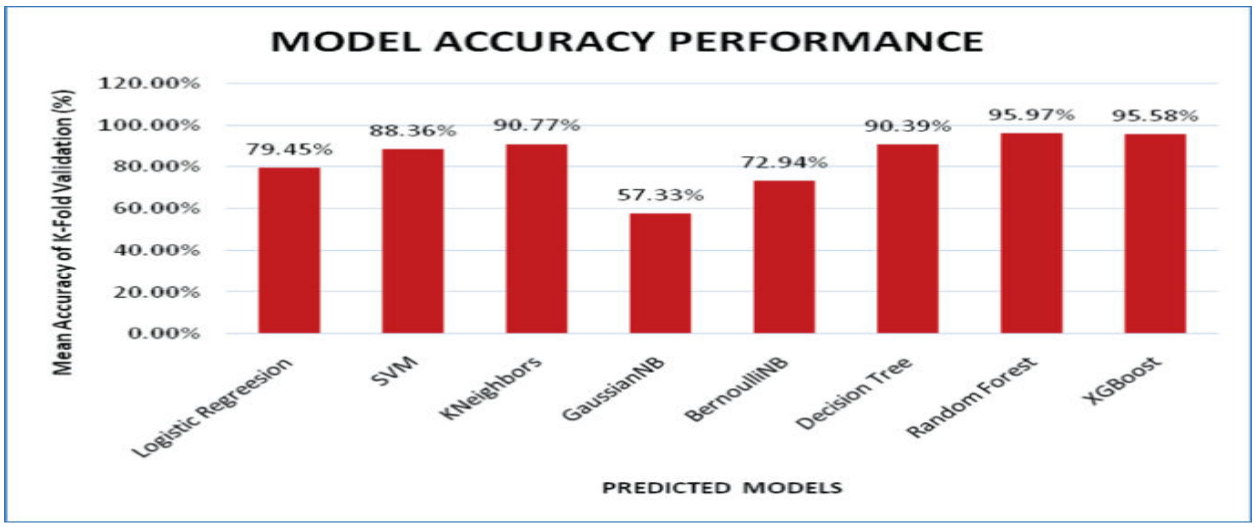

Fig 3: Result analysis of An Efficient Detection of Brain Stroke

Deep Learning Patient-Similarity Framework (DeepNN) is a data-driven tool that can explore the multivariate echocardiographic relationships of diastolic dysfunction to identify patients at risk of future HF events.6,7 We previously described how a DeepNN classifier can be used to characterize the severity of LVDD and identify a specific subgroup of patients with HFpEF who have elevated LV filling pressures, biomarkers of myocardial injury and stress, and adverse events and those who are more likely to respond to therapeutic interventions.7 In this study, we sought to assess whether the DeepNN that was developed in younger cohorts could also be applied to the older population to identify subsets of patients with a heightened risk of developing HF since the DeepNN may be capable of automatically adjusting for age-related changes in LVDD.

V. CONCLUSION

The exploration and analysis of the heart disease dataset underscore the critical role of data-driven insights in advancing medical research. Through rigorous EDA, correlation analysis, cluster analysis, and the development of predictive models, this study offers valuable insights into the factors contributing to heart disease. These findings have the potential to inform clinical practices, guide further research, and facilitate the development of effective prevention and intervention strategies, contributing to the fight against heart disease. This study has demonstrated the power of correlation analysis and machine learning in understanding and predicting heart disease risk. The findings emphasize the multifactorial nature of heart disease and the importance of a holistic approach to its prediction and management. By continuing to leverage data analytics and machine learning, healthcare professionals can develop more effective early detection and intervention strategies, ultimately leading to better prevention and treatment of heart disease. This research not only contributes to the academic discourse on heart disease prediction but also highlights the critical role of data analytics in advancing healthcare outcomes.

REFERENCES

1. Lloyd-Jones DM, Larson MG, Leip EP, et al. Lifetime risk for developing congestive heart failure: the Framingham Heart Study. *Circulation* 2002;106:3068-72.
2. Forman DE, Fleg JL, Wenger NK. Cardiovascular disease in the elderly. In Zipes DP, Bonow RO, Mann DL, Tomaselli GF, Braunwald E, eds. Braunwald's Heart Disease: A Textbook of Cardiovascular Medicine. Philadelphia, PA: Elsevier; 2019.

3. Kitzman DW, Scholz DG, Hagen PT, Ilstrup DM, Edwards WD. Age-related changes in normal human hearts during the first 10 decades of life. Part II (Maturity): a quantitative anatomic study of 765 specimens from subjects 20 to 99 years old. *Mayo Clin Proc* 1988;63:137-46.
4. Cheng S, Xanthakis V, Sullivan LM, et al. Correlates of echocardiographic indices of cardiac remodeling over the adult life course: longitudinal observations from the Framingham Heart Study. *Circulation* 2010;122:570-78.
5. Nagueh SF, Smiseth OA, Appleton CP, et al. Recommendations for the evaluation of left ventricular diastolic function by echocardiography: an update from the American Society of Echocardiography and the European Association of Cardiovascular Imaging. *Eur Heart J Cardiovasc Imaging* 2016;17:1321-60.
6. Tokodi M, Shrestha S, Bianco C, et al. Interpatient similarities in cardiac function: a platform for personalized cardiovascular medicine. *JACC Cardiovasc Imaging* 2020;13:1119-32.
7. Pandey A, Kagiyama N, Yanamala N, et al. Deep-learning models for the echocardiographic assessment of diastolic dysfunction. *JACC Cardiovasc Imaging* 2021;14:1887-1900.
8. Shah AM, Cheng S, Skali H, et al. Rationale and design of a multicenter echocardiographic study to assess the relationship between cardiac structure and function and heart failure risk in a biracial cohort of community-dwelling elderly persons: the Atherosclerosis Risk in Communities study. *Circ Cardiovasc Imaging* 2014;7:173-81.

International Journal of Advanced Research in Arts, Science, Engineering & Management (IJARASEM)

| Mobile No: +91-9940572462 | Whatsapp: +91-9940572462 | ijarasem@gmail.com |

www.ijarasem.com