LITERATURE SURVEY

Visualizing and Predicting Heart Diseases with an Interactive Dashboard

Abstract

In most developed countries, heart disease is the leading cause of death among young adults and the elderly. Even after spending a significant amount of money on treatment, the patient may be unable to be saved owing to their medical condition. Doctors and hospitals are on the verge of developing new strategies and methods for treating patients with appropriate procedures. With the advent of new medical equipment, diagnosis has become increasingly data-driven. The patient's previous readings are documented, and appropriate diagnoses are given to maximize the likelihood of survival. The heart disease datasets contain well-defined heart disease-based readings.

Introduction

The leading cause of death in the US is heart disease, and early detection of the illness's symptoms is essential for effective treatment and prevention. In this study, we assess visualization methods for coronary artery disease detection. Endothelial shear stress is a crucial physical parameter of medical interest (ESS). Sites of low ESS have been linked to The development of lesions and quick disease progression in the coronary arteries. Effective ESS data visualizations are essential for a cardiologist to do a swift and complete non-invasive assessment. We provide a hemodynamics task taxonomy based on preliminary user research with subject-matter specialists. Using the findings of this investigation, we created an interactive visualization.

Literature Surveys

In the 2018 paper, "Predicting the Risk of Heart Failure With EHR Sequential Data Modeling," Bo Jin, Chao Che, and colleagues suggested a neural network-based approach. [1]This essay utilized the real-world datasets derived from electronic health record (EHR) data connected to congestive heart failure to complete the

experiment and foretell cardiac disease in advance. We tend to utilize word vectors and one-hot encryption when modeling the cardiac failure events predicted by the diagnosis using a long memory's fundamental tenets against its network theory. Analyzing the outcomes, we frequently find the significance of honoring clinical procedures' logical sequence documents. In the literature, methods for identifying cardiac disease include waveform analysis, time-frequency analysis, Neuro-Fuzzy RBF ANN, and Total Least Square-based Prony modeling algorithms. However, Marshall et al investigation's found (Marshall et al 1991), This method's classification accuracy (up to 79%) was poor, and choosing the best model required considering a range of modifications.was still adequate. They also showed how well neural networks function at detecting heart attacks (acute myocardial infarction) by contrasting the multilayer perceptron with the Boltzmann perceptron classifier, two neural network classifiers. Most of these methods have to do with diagnosis, not the comprehension of fundamental concepts.By putting up a computerized diagnostic model to support standardization and wide use of traditional Chinese medicine (TCM) diagnosis, Huiyan Wang has sought out a niche in this sector (Huiyan Wang 2008). The system works as follows. First, a database of instances incorporating mutual information theory is used to learn the Bayesian network topology and choose the symptoms. The target variable's Markov blanket is chosen as the symptom set in the structure. Second, naïve Bayesian classifiers are used to map the link between symptom sets and diagnostic outcomes.

In this paper," Heart Attack Prediction and Visualization of Contributing Factors Using Machine Learning" Megha Banerjee, Reetodeep Hazra, Suvranil Saha, Megha Bhushan, Subhankar Bhattacharjee Heart disease, alternatively known as cardiovascular disease, encases various conditions that impact the heart and is the primary basis of death worldwide over the span of the past few decades. [2] It associates many risk factors in heart disease and a need for time to get accurate, reliable, and sensible approaches to make an early diagnosis to achieve prompt management of the disease. Data analysis and machine learning are the most commonly used techniques for processing enormous data in the healthcare domain. Researchers apply several data mining and machine learning techniques to analyze huge complex medical data, helping healthcare professionals to predict heart disease. This research paper presents various attributes like age, gender, chest pain,

cholesterol, etc which are used to predict heart attack, and the model is trained using 4 machine learning algorithms namely- Logistic Regression, Gaussian Naïve Bayes, Decision tree, and Random Forest algorithm. It uses the existing dataset from the UCI Heart Disease Data set of heart disease patients. The dataset comprises 303 instances and 76 attributes. This research paper aims to envision the probability of developing heart attacks in patients. The results portray that the highest accuracy score is achieved with Logistic Regression.

This research paper"A novel approach for heart disease prediction using strength scores with significant predictors" talks about CVDs are disorders of the heart and blood vessels and include coronary heart disease, cerebrovascular disease, and other conditions. [3] Heart attacks and strokes are the main causes of mortality in cardiovascular disease in which the rate nears one out of three. With the high rate of mortality, diagnosis and prevention measures need to be performed effectively and efficiently. Many data mining techniques have been used to help address these issues. Most of the past Heart Attack Prediction and Visualization of Contributing Factors Using Machine Learning research looked into identifying features that contribute to better heart prediction accuracy. However, very little research looked into the relationships that exist between these features. The association between each feature that contributes to heart disease prediction can be obtained using the Associative Rule Mining (ARM) technique. The ARM technique is popular in transactional and relational datasets. The hidden knowledge in large datasets such as business transactions developed the interest of many business owners to understand the patterns that can help them to improve their business decisions (Agarwal and Mithal). For instance, discovering the frequently bought items by customers in market basket analysis. This analysis looks at the various items found in customers' shopping carts and identifies the associations between them. A good example would be if customers were looking to purchase milk, they were likely to purchase bread on the same trip to the supermarket. This approach is also widely used in the healthcare industry specifically in privacy preservation of healthcare data, predicting cancer-associated protein interactions, predicting obstructive sleep apnea, and predicting co-diseases in Thyroid patients.

In this paper “Heart Disease Risk Prediction Using Machine Learning Classifiers with Attribute Evaluators”- Karna Vishnu Vardhana Reddy, Irraivan Elamvazuthi, Azrina Abd Aziz, Sivajothi Paramasivam, Hui Na Chua, and S. Pranavanand. The annual death toll from cardiovascular diseases (CVDs) is roughly 20.5 million. Early detection can assist people in altering their habits and, if necessary, ensuring that they receive the right medical care. Ten machine learning (ML) classifiers from various categories, such as Bayes functions, were used in this study. Using the complete set of data, lazy, meta, rules, and trees were trained for effective heart disease risk prediction. Cleveland Heart dataset attributes and the best attribute sets were derived from three attribute evaluators[4]. An evaluation of the algorithms' performance was made using a 10-fold cross-validation test option. The hyperparameter number of nearest neighbors, specifically, was tuned last. Instance-based (IBC) classifier's 'k'. SMO, which stands for sequential minimum optimization, produced an accuracy of 86.468% and 85.148%, respectively, when all the attributes are used the greatest accuracy score utilizing the chi-squared attribute evaluator's best attribute set, or set of attributes. However, the meta Using both classifier bagging and logistic regression (LR), the ROC area with the greatest value of 0.91 was obtained. The complete and ideal attribute sets that the ReliefF attribute evaluator produced. In general, the SMO Compared to other methods, the classifier provided the most accurate prediction. IBC attained a by setting the hyperparameter 'k' to 9, and using the chi-squared attribute, accuracy is increased by 8.25% set.

Our Inference:

All living species have a vital organ called the heart, which is responsible for pumping blood via the circulatory system's blood capillaries into the other organs. The brain will suffer, the heart will cease beating entirely, and death will result from poor blood circulation throughout the body. The heart's ability to beat properly is essential to life. Heart and cardiovascular system disorders are referred to as "heart disease." some factors are:

- Family background
- Age
- Smoking
- poor diet

- high blood pressure
- high blood cholesterol
- obesity
- physical inactivity
- high blood pressure

Diagnoses are frequently made using the patient's most recent test results and the doctor's experience. As a result, making a diagnosis is a difficult undertaking that calls for great expertise and experience. IBM Cognos employs Data Visualization and machine learning algorithms on patient health data to forecast the likelihood of heart disease. Machine Learning Approach for multi-classification.

Public health and community organizations tackling CVD mortality can benefit from the Local Trends in Heart Disease and Stroke Mortality Dashboard in several ways. Users can identify counties with high or rising mortality using county-level visualizations, and they can then develop policies and programs for CVD prevention and treatment that are specific to the requirements of those regions. As a result of the spatiotemporal aspect of the dashboard, regions that could benefit from initiative living species have a vital organ called the heart, which is responsible for pumping blood via the circulatory system's blood capillaries into the other organs. The brain will suffer, the heart will cease beating entirely, and death will result from poor blood circulation throughout the body. The heart's ability to beat properly is essential to life. Heart and cardiovascular system disorders are referred to as "heart disease." some factors are:

- Obtain information about patients with known heart conditions.

- This dataset includes details regarding heart conditions, such as a person's blood sugar, cholesterol, and other medical data.
- Using the data, develop a multi-classification neural network model.
- Utilize that model to forecast the chance of heart disease in additional people for whom we have access to their medical history or other relevant data.

Diagnoses are frequently made using the patient's most recent test results and the doctor's experience. As a result, making a diagnosis is a difficult undertaking that calls for great expertise and experience.

PREDICTION OF THE DISEASE:

Let's determine whether they have heart disease or not. Here, target = 1 suggests that the person has heart disease, whereas target = 0 suggests that the person is not afflicted. One of the most prevalent diseases today is heart disease (HD), which has several contributing variables, including high blood pressure, diabetes, fluctuating cholesterol levels, tiredness, and many others. Since a long time ago, it has been sought after to identify some of the early symptoms of HD. To this end, numerous data analytics technologies have been deployed. Numerous, trustworthy tests can be carried out on prospective patients to help them take extra steps to lessen the effects of having such an illness. Finding ways to anticipate HD in its early phases, like the approaches suggested in this research, can be essential for saving lives. Numerous Machine Learning (ML) algorithms, including Naive Bayes, Stochastic Gradient Descents (SGD), Support Vector Machine (SVM), K-Nearest Neighbor (K-NN), Adaboost, JRip, Decision tree J48, and others, were used to classify and predict HD datasets, and many encouraging results were published in the literature.

REFERENCES:

[1]"Predicting the Risk of Heart Failure With EHR Sequential Data Modeling," Bo Jin, Chao Che.

[2]"Heart Attack Prediction and Visualization of Contributing Factors Using Machine Learning" by Megha Banerjee, Reetodeep Hazra, Suvranil Saha, Megha Bhushan, Subhankar Bhattacharjee.

[3]"A novel approach for heart disease prediction using strength scores with significant Predictors" by Armin Yazdani, Kasturi Dewi Varathan, Yin Kia Chiam, Asad Waqar Malik, and Wan Azman Wan Ahmad.

[4]"Heart Disease Risk Prediction Using Machine Learning Classifiers with Attribute Evaluators" by Karna Vishnu Vardhana Reddy, Irraivan Elamvazuthi, Azrina Abd Aziz, Sivajothi Paramasivam, Hui Na Chua, and S. Pranavanand