JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

AI-Based Multi Disease Detection Using Random Forest Approach

Ms. Pooja Vishwakarma, Mr. Rajneesh Pachouri, Mr. Anurag Jain
Research Scholar, Assistant Professor, Assistant Professor
Department Of Computer Science & Engineering , Adina Institute Of Science And Technology, Sagar, India

Abstract : One of the most significant issues facing society today is healthcare for humans. In order to provide patients with the necessary care, it looks for the most accurate, reliable, and efficient disease diagnosis as soon as feasible. Since this detection is frequently a challenging endeavour, the medical industry needs assistance from other fields, such computer science and statistics. These fields are challenged to investigate novel approaches that go beyond the conventional ones. An extensive overview that steers clear of very specific features is required due to the multitude of evolving techniques. In order to do this, we suggest conducting a thorough analysis of machine learning as it relates to medical diagnostics. This review focuses on contemporary methods for developing machine learning that are used in the medical industry to diagnose human disorders in order to find intriguing patterns, generate meaningful predictions, and aid in decision-making. This work can thus assist researchers in learning about and, if needed, evaluating the application of machine learning techniques in their respective fields of expertise. One of the main causes of death in the world is heart disease, however mortality rates can be lowered with early detection. Prominent research has demonstrated that the most recent artificial intelligence (AI) can be utilised to assess a person's risk for heart disease. However, in order to obtain the most performance out of these AI models in the event that the number of users increases, earlier research did not take dynamic scalability into consideration. Our solution to this issue is Health FaaS, an AI-powered smart healthcare framework that uses a serverless computing environment and the Internet of Things (IoT) to reduce heart disease-related mortality and prevent financial losses by minimising misdiagnoses.

***IndexTerms*:** human disease; machine learning, artificial intelligence; big data, smart healthcare

I. INTRODUCTION

The widespread use of artificial intelligence (AI) in a variety of industries, including robotics and medicine, has become more common with the introduction of breakthroughs in machine learning (ML)-based approaches. The availability of large-scale clinical and multi-omics datasets has contributed significantly to the recent boom in the application of AI to multimedicine research. These applications are offering a fresh viewpoint for a deeper comprehension of multiple diseases, which may be applied to the creation of innovative treatment and diagnostic plans. Research has demonstrated, for instance, that machine learning (ML) holds great promise for the early detection of many diseases, the prediction of unfavourable illness outcomes like heart failure, and the creation of more individualised and innovative treatments.

Precision medicine in actual practice depends on early diagnosis and precise multi-disease classification. We sought to investigate the application of various machine learning approaches in the field of multi health through the projects conducted as part of this thesis research, but from a fresh angle of employing gut microbiome data as a new feature to categorise people with and without multi disease. This decision is justified by the fact that, in addition to heredity and environmental variables, in recent times, The gut microbiome has become a novel component that influences

The heart, a muscular organ that pumps blood into the body, is the primary part of the body's circulatory system, which also includes the lungs. Capillaries, arteries, and veins make up the cardiovascular system's blood vessel network. These blood arteries carry blood throughout the body. Heart problems, commonly known as cardiovascular diseases (CVD), are brought on by deviations from the heart's regular blood flow. Globally, heart problems are the leading cause of death. The World Health Organization (WHO) assessment indicates that heart attacks and strokes account for 17.5 million deaths worldwide. The majority of cardiovascular disease deaths—more than 75% of them—occur in middle-class and lower-class nations.

Moreover, 80% of deaths from CVD are caused by heart attacks and strokes. Early identification of cardiac abnormalities and the advancement of technology to predict heart disorders can thereby prevent many deaths and help medical practitioners design effective treatment regimens that reduce the mortality rate from cardiovascular diseases.

As a result of the advancement of sophisticated healthcare systems, a vast amount of patient data—known as big data in electronic health record systems—is currently accessible and can be utilised to create predictive models for cardiovascular illnesses. A finding technique called data mining, also known as machine learning, is used to analyse large amounts of data from multiple angles and condense it into meaningful knowledge. "Extraction of implicit, unknown, and potentially useful information about data is a non-trivial task known as data mining." Today's healthcare sectors produce enormous amounts of data about patients, diseases, and other related topics. Numerous strategies are available in data mining to uncover hidden patterns or similarities in data.

Thus, in order to create a heart disease prediction system, a machine learning technique is suggested in this research and verified on two public heart disease prediction datasets. The computer-based method of obtaining valuable information from massive datasets is known as data mining. Because exploratory analysis requires nontrivial knowledge from vast volumes of evidence, data mining is particularly useful in this type of investigation. Investigating the mysterious patterns in the clinical domain's data sets can be greatly aided by medical data mining.

These patterns can be used to diagnose medical conditions. Nonetheless, the raw medical data that are currently accessible are dispersed broadly, abundant, and diverse.It is necessary to gather this data in an organised manner. After gathered, this data can be combined to create a health information network. A user-oriented method for finding new and hidden patterns in data is provided by data mining. In the world of healthcare, data mining techniques and technologies are helpful in forecasting different diseases and providing answers to business queries. In data mining, disease prediction is important. This study uses a classification algorithm to analyse the forecasts of cardiac diseaseIn healthcare data, these imperceptible patterns can be used to diagnose medical conditions.

An effective method for addressing the most recent and enduring trends in the data is provided by data mining technologies. Healthcare administrators can use the identified information to obtain better services. The most common cause of death for victims in nations likes the US and India was heart disease. With the help of categorization algorithms, we are forecasting cardiac illness in this research. Several heart-related issues are investigated using machine learning approaches such classification algorithms as Yolo Classifications and logistic regression.

1.2 Motivation

Ensuring the provision of quality services at reasonable rates is a significant problem for healthcare organisations, including hospitals and medical centres. Accurate patient diagnosis and successful treatment delivery are prerequisites for providing quality service. Inadequate clinical judgement can have terrible results, which is unacceptable. Hospitals also need to keep clinical test costs as low as possible. By utilising the proper computer-based information and/or decision support technologies, they can accomplish these goals.

Hospital information systems are used by the majority of hospitals nowadays to handle patient data and healthcare. These systems often produce enormous volumes of data in the form of text, numbers, photos, and charts. Sadly, clinical decision making rarely uses these evidence to support it. These data provide a plethora of undiscovered hidden information. This brings up a crucial query: "How can we transform data into actionable knowledge that empowers medical professionals to make wise clinical decisions?" This is the primary driving force of the study.

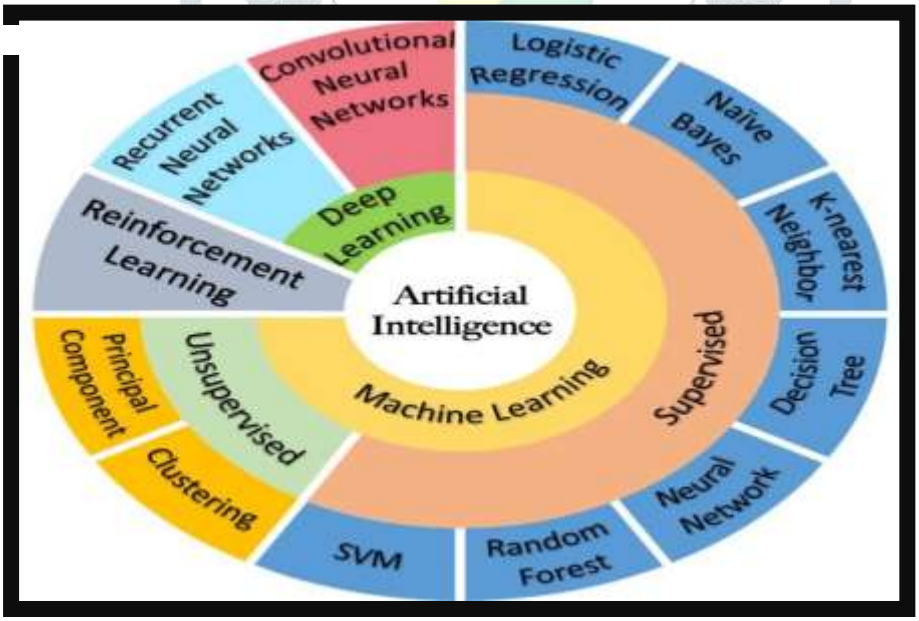

Figure 1 various algorithms of artificial intelligence applied in Multi medicine.

II. LITERATURE SURVEY

Muhammed Golec , Sukhpal Singh Gill , Ajith Kumar Parlikad , and Steve Uhlig "HealthFaaS: AI-Based Smart Healthcare System for Heart Patients Using Serverless Computing", IEEE Internet Of Things Journal, Vol. 10, No. 21, 1 November 2023 [1]

Heart disease is the most significant ailment that has increased as a result of the modern lifestyle's altered dietary and physical activity habits [1]. Heart failure, cardiac defects, and heart attacks are only a few of the numerous disorders with circulatory roots that fall under the umbrella term "heart disease" [2]. The leading cause of death in the UK and around the world is heart disease [3]. Although rates of heart disease have decreased recently due to efforts to decrease obesity and raise public knowledge of the causes of heart disease, the condition is still quite common.

"The top 10 causes of death." Dec. 2020. [Online]. Available: http www.who.int/news-room/fact-sheets/detail/the-top-10-causes-of-death.[4]

The World Health Organisation (WHO) has issued a list of the top four causes of mortality globally [4], with ischemic heart disease coming in first place with 8.9 million fatalities. Chest pain is the most common symptom among individuals with heart

disease. A expert diagnoses cardiac illness in a patient based on results from Thallium testing, electrocardiography (ECG), and hematocrit [2]. Consequently, it is challenging to identify heart illness using conventional techniques (physical examination, etc.).

S. S. Gill et al., "AI for next generation computing: Emerging trends and future directions," Internet Things, vol. 19, Aug. 2022, Art. no. 100514. [5]

In fact, it functions as a silent killer, and most patients are not even aware that they are impacted by it. Because of this, it's critical to identify cardiac conditions at an early stage and to begin therapy as soon as feasible. Research that provides early diagnosis is predicted to lower the frequency of fatal instances, given current advancements in artificial intelligence (AI) model approaches and the Internet of Things (IoT) [5]. As a result, new systems utilising IoT and sensors are required to detect heart problems.

S. Price. "Avoidable deficiencies in heart failure cost NHS £21m." Mar. 2020. [Online]. Available: https://www.healtheuropa.eu/avoidabledeficiencies- in-heart-failure-cost-nhs-21m/98696/. [6]

A. **Motivation and Our Contributions:** Smart healthcare systems have seen some promising breakthroughs as a result of AI advancements. Millions of pounds are lost financially in the health sector as a result of misdiagnosing heart disease and the high expense of testing for it [6]. IoT devices can be used to gather patient data, and the most recent machine learning (ML) methods can be applied to successfully assess the risk of heart disease at a level that is acceptable [7]. As a result, early and correct diagnosis can prevent needless medical service costs and heart disease-related deaths.

R. Singh and S. S. Gill, "Edge AI: A survey," Internet Things Cyber Phys. Syst., vol. 3, pp. 71–92, Feb. 2023 [8] Based on patients' medical conditions, the most effective machine learning model must be identified in order to forecast cardiac disease early. In order to do this, it is required to compare the performance of several recent machine learning models that have been effectively proven in [8] in order to determine which model has the highest accuracy. In order to do this, we have chosen models such as support vector machine (SVM), artificial neural network (ANN), gradient boosting machine (GBM), extreme gradient boosting (XGBoost), and light gradient boosting machine (LightGBM) because previous research [2] has demonstrated that these models are best suited for examining the emergence of environmental-related cardiovascular disease and the demand for healthcare.

H. Raj, M. Kumar, P. Kumar, A. Singh, and O. P. Verma[9] Furthermore, these models' performances were evaluated in terms of accuracy, precision, recall, F-Score, and AUC. It was shown that LightGBM performed better with an accuracy of 91.8%. Research that uses machine learning (ML) models to identify diseases using IoT, fog, and cloud computing is highly interested due to the recent advancements in AI [9],

S. Iftikhar et al [10]. In this section, we review the research that use machine learning models from the UCI Heart Disease data set [11] to estimate the risk of heart disease, as employed in our work.

K. Polat and S. Güne¸s Data sets are converted from feature space to kernel space using linear and radial functions in the first study outlined by Polat and Gunes utilising machine learning [12]. The ML models are used to generate the F-score. Furthermore, in order to avoid adversely affecting the accuracy rates of machine learning models, they employed a technique known as Kernel F-score feature selection (KFFS). As a result, they have an 83.70% accuracy rate in identifying the risk of heart disease.

R. Spencer, F. Thabtah, N. Abdelhamid, and M. Thompson In order to find specific features in their work and utilise them to predict heart disease, Spencer et al. recommended utilising the chi-squared feature evaluator. The BayesNet algorithm and chi-squared feature evaluator worked together to diagnose heart disease with 85% accuracy in their study [13].

Y. Khourdifi and M. Bahaj [14], "Heart disease prediction and classification using machine learning algorithms Optimized by particle swarm optimization and ant colony optimization," Int. J. Intell. Eng. Syst., vol. 12, no. 1, pp. 242–252, 2019.Nonetheless, disparate feature groups might also produce contradictory outcomes. employed fast feature selection in a different study to eliminate unnecessary variables from the data set prior to using ML models. They obtained the best performance rates with the k-nearest neighbours (KNN) algorithm by optimising their models using the fast correlation-based feature selection (FCBF), particle swarm optimisation (PSO), and ant colony optimisation (ACO) techniques.

N. Satyanandam and C. Satyanarayana Optimal multi-nom logistic regression (OMLR) was employed in a different study [15] to assess the severity of the cardiac condition. This has produced a 92% accuracy rate. Several research have attempted to improve the prediction accuracy for heart disease risk by integrating multiple models.

One of these studies from the literature makes use of many data mining techniques.

Tarawneh and Embarak proposed [16]. An accuracy rate of 89.2% has been attained in this work. One of the most recent studies used WEKA and open-source knowledge extraction based on evolutionary learning (KEEL) techniques to diagnose cardiac disease. In a different study [17], principle component analysis (PCA) and fuzzy logic are combined, and by reducing the size of the features in the data set, the accuracy rate was raised. The accuracy rate achieved with the hybrid models employed in the study was as high as 94%.

P. H. Vilela, J. J. Rodrigues, P. Solic, K. Saleem, and V. Furtado A fog computing-based study for real-time and latency-sensitive healthcare applications was proposed by Vilela et al. [18]. The authors examined energy consumption and delay indicators independently in cloud and fog environments using IoT and a few medical sensors. In comparison to the cloud, fog computing offers benefits in terms of data security, latency, and network consumption because patient health data is analysed locally. The proposed work (HealthFaaS) and previous works are compared in Table I. To the best of our knowledge, no research utilising IoT, serverless computing, and artificial intelligence within a single framework has been done to detect cardiac disease.

Users can rapidly submit data to this study using wearables and the Internet of Things. Furthermore, prior studies have solely examined the efficacy of machine learning in identifying heart disease. We implemented ML models on Serverless Computing, broadening the scope of our work. The ML model on the server receives real-time data from users as a result. In the event that an illness is detected using HealthFaaS, it is suggested that the treatment process begin with the transmission of information to health centres. In this sense, it is anticipated that heart disease-related fatalities will decline and cost savings up to millions of pounds will occur.

Data Source

Clinical databases have collected a significant amount of information about patients and their medical conditions. Records set with medical attributes were obtained from the Cleveland Heart Disease database. With the help of the dataset, the patterns significant to the heart attack diagnosis are extracted. The records were split equally into two datasets: training dataset and testing

dataset. A total of 303 records with 76 medical attributes were obtained. All the attributes are numeric-valued. We are working on a reduced set of attributes, i.e. only 14 attributes.

All these restrictions were announced to shrink the digit of designs, these are as follows:

1) The features should seem on a single side of the rule.
2) The rule should distinct various features into the different groups.
3) The count of features available from the rule is organized by medical history of people having heart disease only.

The following table shows the list of attributes on which we are working.

Table 1 List of attributes

S no	Attribute Name	Description
1	Age	age in years
2	Sex	(1 = male; 0 = female)
3	Cp	Chest Pain
4	Trestbps	resting blood pressure (in mm Hg on admission to the hospital)
5	Chol	serum cholesterol in mg/dl
6	Fbs	(fasting blood sugar > 120 mg/dl) (1 = true; 0 = false)
7	Restecg	resting electrocardiographic results
8	Thalach	maximum heart rate achieved
9	Exang	exercise induced angina (1 = yes; 0 = no)
10	Oldpeak	ST depression induced by exercise relative to rest
11	Slope	the slope of the peak exercise ST segment
12	Ca	number of major vessels (0-3) colored by flourosopy
13	Thal	3 = normal; 6 = fixed defect; 7 = reversible defect
14	Target	1 or 0

Arti Gupta, Maneesh Shreevastava IJETAE, 2011. **Medical Diagnosis using Back Propagation Algorithm.[19]** In this paper, feed forward Back Propagation algorithm is described which is used as a classifier to distinguish between infected and non-infected person in medical diagnosis. The back propagation algorithm presented in this paper used for training depends on a multilayer neural network with a very small learning rate, especially when using a large training set size**.**

Shraddha Subhash Shirsath, Prof. Shubhangi Patil IJIRSET, June 2018. Disease Prediction using Machine Learning over Big Data [20]. This paper discusses about machine learning algorithm which is used for the accurate disease prediction. Here to achieve the incomplete data latent factor model is used. Yolo algorithm is used for clarification of large volume of data from hospital and then Convolutional Neural Network Based Multimodal Disease Prediction (CNN-MDRP) algorithm helps to provide result of a disease prediction.

Nikita Kamble, Manjiri Harmalkar, Manali Bhoir, Supriya Chaudhary, IJSRCSEIT, 2017. Smart Health Prediction System Using Machine Learning[21]. The paper presents an overview of the Machine Learning techniques with its applications, medical and educational aspects of Clinical Predictions. In medical and health care areas, due to regulations and due to the availability of computers, a large amount of data is becoming available. Such a large amount of data cannot be processed by humans in a short time to make diagnosis, and treatment schedules. A major objective is to evaluate Machine Learning techniques in clinical and health care applications to develop accurate decisions. It also gives a detailed discussion of medical Machine Learning techniques which can improve various aspects of Clinical Predictions. It is a new powerful technology which is of high interest in computer world. It is a sub field of computer science that uses already existing data in different databases to transform it into new researches and results. It makes use of machine learning and database management to extract new patterns from large data sets and the knowledge associated with these patterns. The actual task is to extract data by automatic or semi-automatic means. The different parameters included in Machine Learning include clustering, forecasting, path analysis and predictive analysis.

Nilesh Borisagar, Dipa Barad, Priyanka Raval, Conference paper (PICCN), April 2017. Chronic Kidney Disease Prediction using Back Propagation Neural Network Algorithm [22]. In this paper, various training algorithms like Levenberg, Bayesian regularization, Scaled Conjugate and Resilient back propagation algorithm are discussed. After neural network is trained using back propagation algorithms, this trained neural network system is used for detection of kidney disease in the human body. The back propagation algorithms presented here have capacity for distinguishing amongst infected patients or non-infected person.

Sellappan Palaniappan, Rafiah Awang IEEE, 2008**.** Intelligent Heart Disease Prediction System Using Machine Learning Technique [23]. This paper discusses about the development of prototype using Machine Learning techniques, namely, Yolo

Classification, Yolo and Neural Network. It can answer complex "what if "queries which traditional decision support system cannot.it is web-based, user-friendly, scalable, reliable and expandable.

III. SYSTEM ANAYLSIS

a) EXISTING SYSTEM

Rather of relying on the wealth of knowledge concealed in the database, physicians frequently depend their clinical decisions on their experience and intuition. Unwanted prejudices, mistakes, and exorbitant medical expenses are caused by this practice, which lowers the standard of care given to patients. A medical misdiagnosis can occur in a variety of ways. A misdiagnosis of a major illness can have extremely detrimental implications, regardless of whether the medical staff or a doctor is at blame. According to data from the National Patient Safety Foundation, 42% of medical patients believe they have had a medical error or missed diagnosis. Sometimes, other considerations, like the price of medical tests, medications, and operations, negligently take precedence over patient safety. Medical misdiagnosis poses a severe threat to the healthcare industry. If they persist, people will come to fear receiving medical care at a hospital. By educating the public and bringing legal action against the negligent medical professionals, we can eradicate medical misdiagnosis.

Disadvantages:

- It is impossible to predict in the early stages.
- It takes time to use acquired data practically under the current system.
- Any mistakes made by the physician or hospital employees in foreseeing might result in catastrophic events.
- Before beginning treatment, a very costly and time-consuming procedure must be carried out to determine the patient's future risk of developing heart disease.

b) PROPOSED SYSTEM

This section presents an overview of the suggested system and demonstrates all of the parts, methods, and resources that were employed in the system's development. An effective software tool is required to train large datasets and compare several machine learning algorithms in order to create a heart disease prediction system that is both intelligent and easy to use. The robust algorithm with the best accuracy and performance metrics will be selected, and it will then be used to construct a smartphone application that will identify and forecast the risk level of heart disease. To construct the continuous patient monitoring system, hardware elements such as an Arduino/Raspberry Pi, several biological sensors, a display monitor, a buzzer, etc. are required.

c) ALGORITHMS

Logistic Regression

Logistic regression is a widely used statistical method for predicting binomial outcomes (y = 0 or 1). Categorical outcomes (binomial or multinomial values of y) are predicted by logistic regression. The probability of an event occurring are the predictions made by Logistic Regression (referred to as LogR in this article). i.e. the probability of y=1, given certain values of input variables x. Thus, the results of LogR range between 0-1.

The conventional logistic function, often known as the sigmoid curve and represented by the equation: S, is used by LogR to model the data points.

$$\frac{1}{1+e^{-x}} = \frac{e^{x}}{1+e^{x}}$$

Logistic Regression Assumptions:

- Logistic regression requires the dependent variable to be binary.
- For a binary regression, the factor level 1 of the dependent variable should represent the desired outcome.
- Only the meaningful variables should be included.
- The independent variables should be independent of each other·
- Logistic regression requires quite large sample sizes.
- Even though, logistic (**logit**) regression is frequently used for binary variables (2 classes), it can be used for categorical dependent variables with more than 2 classes.
- In this case it's called Multinomial Logistic Regression.

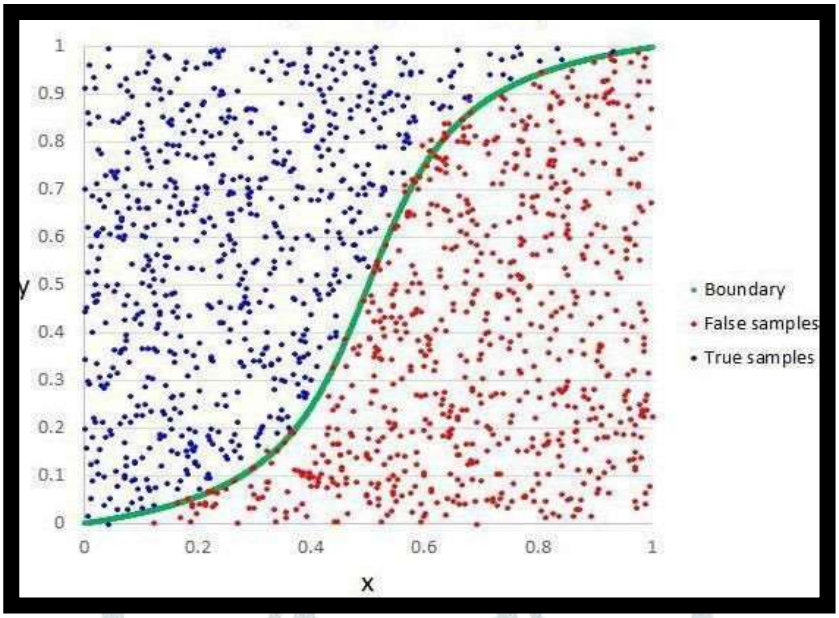

Figure 2 logistic regression

d) Yolo Classifications

Yolo Classifications is an approach for supervised learning that may be applied to regression and classification. However, categorization difficulties are the primary application for it. Since trees are what make up a forest, as we all know, a forest with more trees will be more robust.

In a similar manner, Yolo Classifications builds decision trees using data samples, gathers predictions from each one, and then uses voting to determine which solution is the best. Because the ensemble method averages the results, it reduces over-fitting, making it superior to a single decision tree.

Working of Yolo Classifications with the help of following steps:

- This algorithm will create a decision tree for each sample after first electing random samples from a given dataset. After that, it will receive each decision tree's prediction outcome.
- Voting will be done for each anticipated outcome in this step.
- Finally, decide which forecast outcome received the most votes to be the winning one.

The following diagram will illustrates its working-

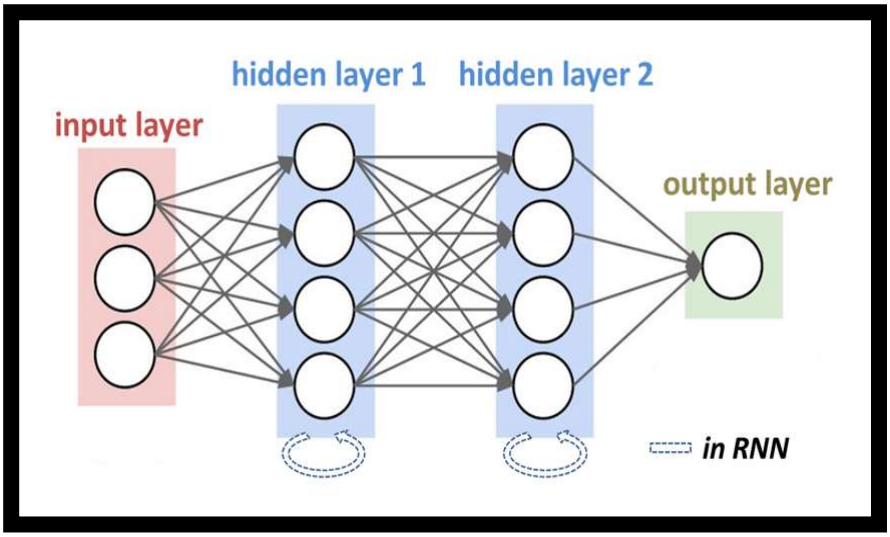

Figure 3 Yolo Classifications

IV. RESULT AND IMPLEMENTATION

Purpose

Automation of the diagnosis process by using machine learning modelling tools and techniques, enable accurate diagnosis of the disease and hence hold the promise of delivering reliable health-care to resource-scarce areas. The purpose of the project is to provide exploratory analysis of how different features affect the chances of being infected by life ceasing diseases such as breast

cancer, PIMA diabetes, liver disease and heart diseases. Furthermore, the project furnishes sophisticated machine learning algorithms to enable prediction of whether the person is infected or not with the above-mentioned diseases.

Aims

Aim 1: PIMA Diabetes Prediction

Diabetes is a chronic disease or group of metabolic diseases where a person suffers from an extended level of blood glucose in the body, which is either the insulin production is inadequate, or because the body's cells do not respond properly to insulin. The objective is to make use of significant features, design a prediction algorithm using Machine learning and find the optimal classifier to give the closest result comparing to clinical outcomes.

Aim 2: Heart Disease Prediction

A study shows that from 1990 to 2016 the death rate due to heart diseases have increased around 34 per cent from 155.7 to 209.1 deaths per one lakh population in India. Thus, preventing Heart diseases has become more than necessary. Good data-driven systems for predicting heart diseases can improve the entire research and prevention process, making sure that more people can live healthy lives.

Aim 3: Liver Disease Prediction

Patients with liver disease have been continuously increasing because of excessive consumption of alcohol, inhale of harmful gases, intake of contaminated food, pickles and drugs. Thus, correct and early diagnosis is necessary to prevent the damage and to reduce the burden on doctors. The objective is to evaluate prediction algorithms in an effort to determine which patients have liver disease and which ones do not.

Aim 4: Breast Cancer Prediction

Breast cancer (BC) is one of the most common cancers among women worldwide, representing the majority of new cancer cases and cancer-related deaths according to global statistics, making it a significant public health problem in today's society. Accurate classification of benign tumours can prevent patients undergoing unnecessary treatments. Thus, the correct diagnosis of BC and classification of patients into malignant or benign groups is the subject of much research.

FEATURE IMPORTANCE

Feature importance rates how important each feature is for the decision a tree makes. It is a number between 0 and 1 for each feature, where 0 means "not used at all" and 1 means "perfectly predicts the target." The feature importances always sum to 1.

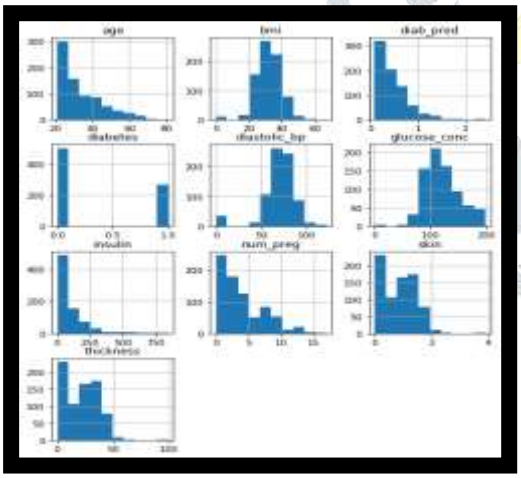

Figure 4 Diabetes Prediction Analytics

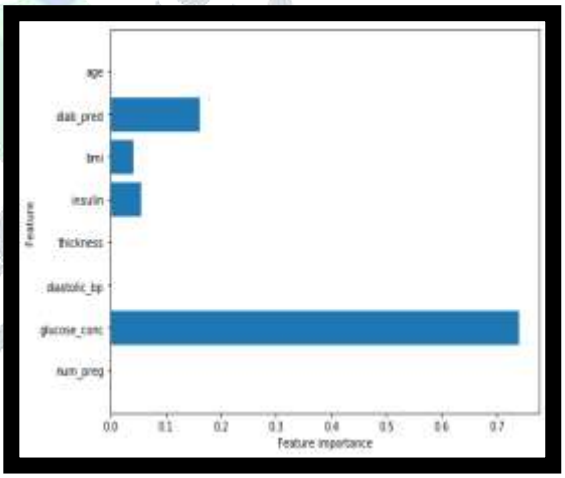

Figure 5 Decision Tree

Feature "Glucose" is by far the most important feature.

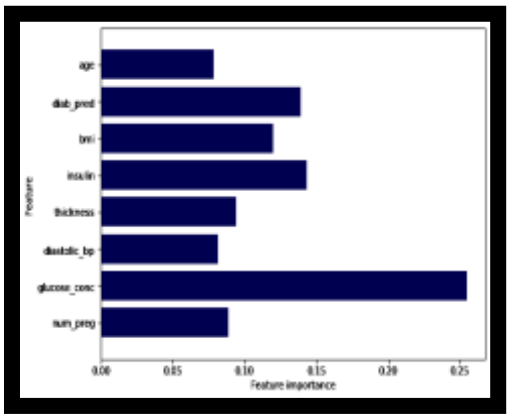

Figure 4 Random Forest

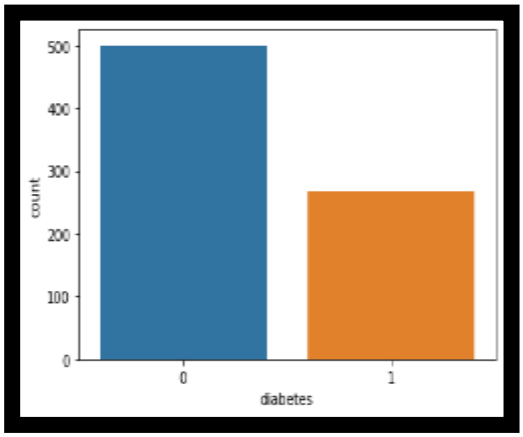

Figure 5 Dataset Outcome Analysis

Similarly to the single decision tree, the random forest also gives a lot of importance to the "Glucose" feature, but it also chooses "BMI" to be the 2nd most informative feature overall.

DATASET OUTCOME ANALYSIS

Outcome 0 means No diabetes, outcome 1 means diabetes. Out of these 768 data points, 500 are labeled as 0 and 268 as 1

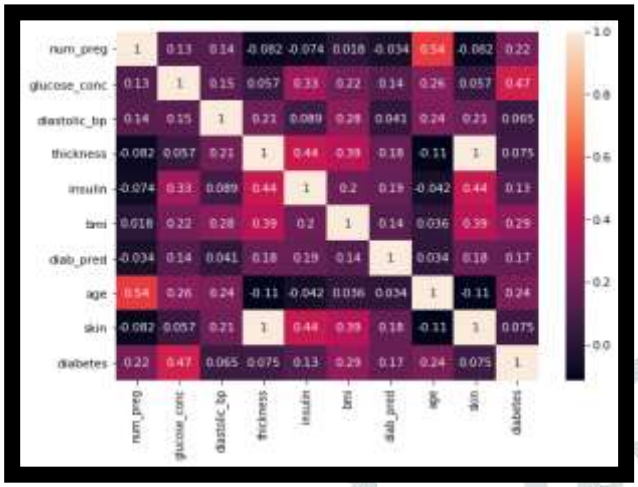

Figure 8 Heat Map

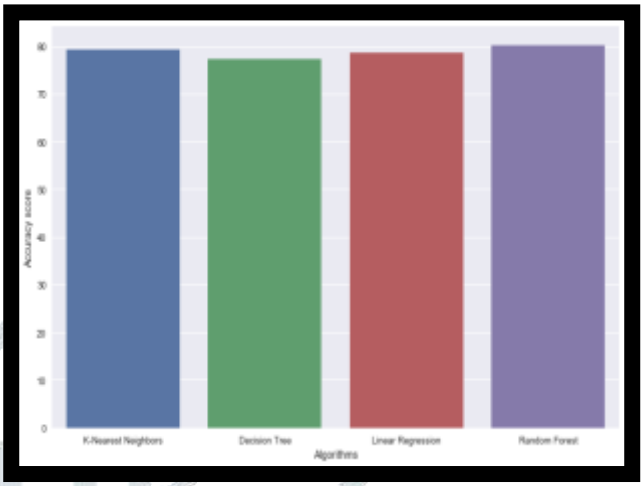

Figure 9 Model Accuracy

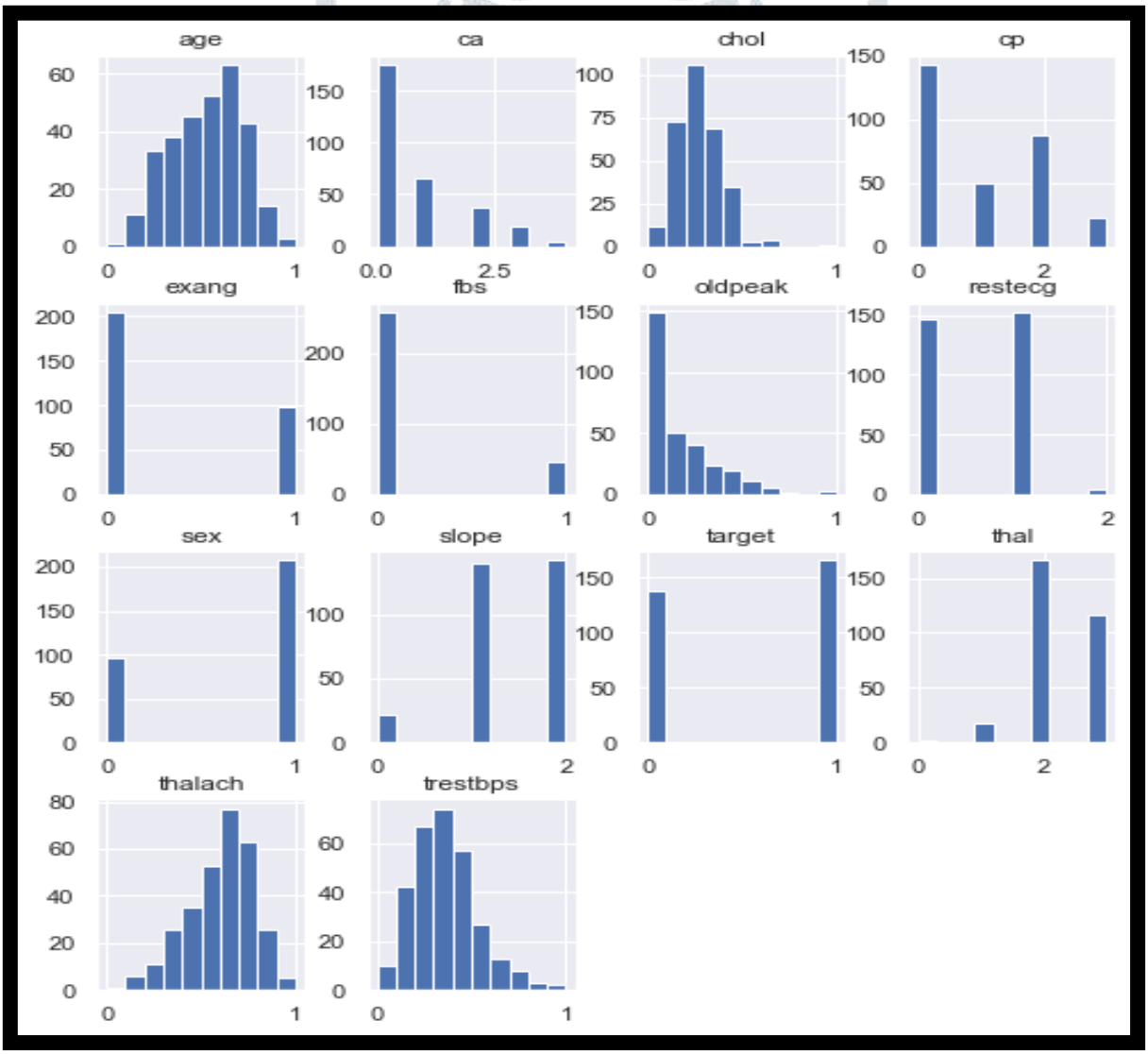

Figure 10 Heart Disease Prediction Analytics

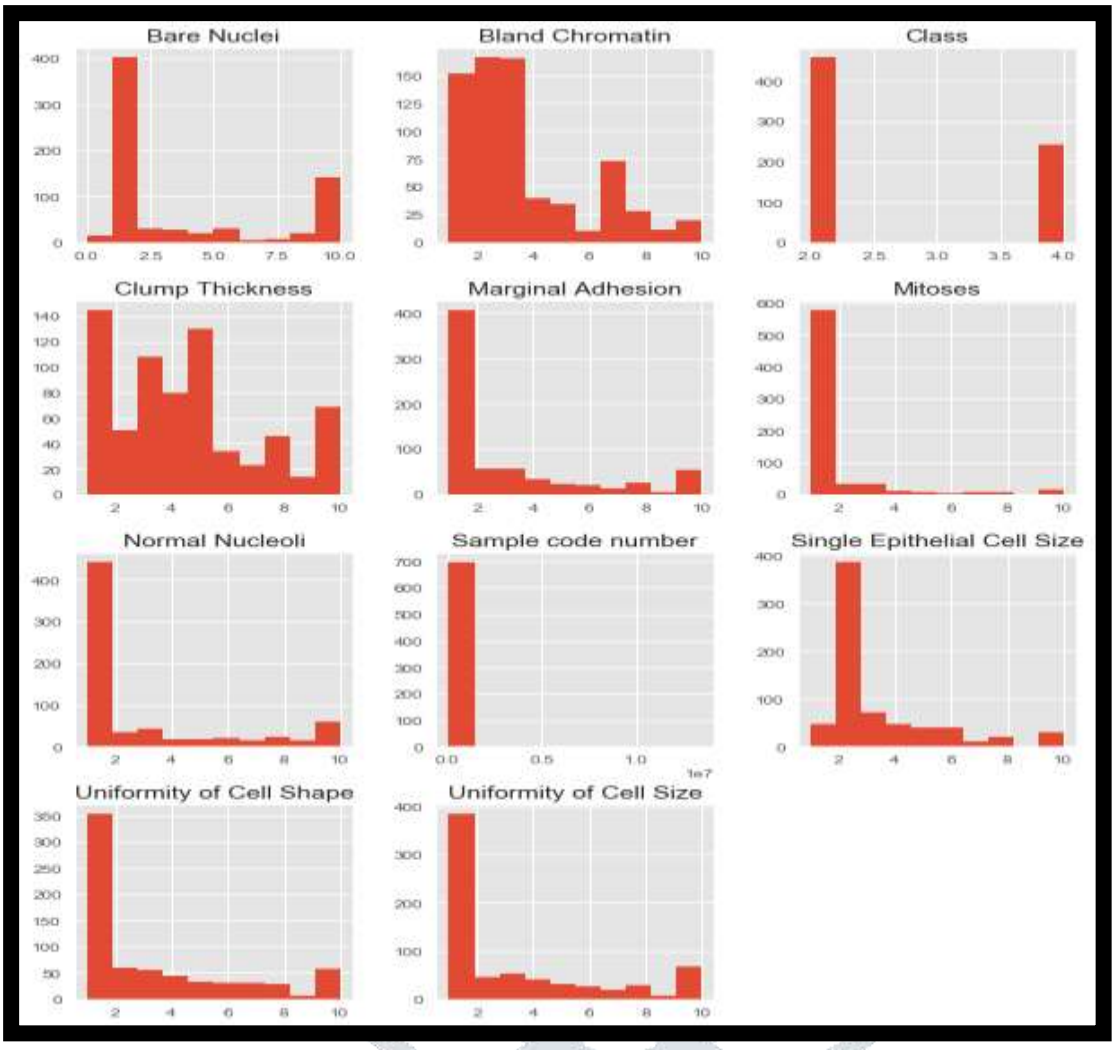

Figure 11 Breast Cancer Classification Analytics Histogram

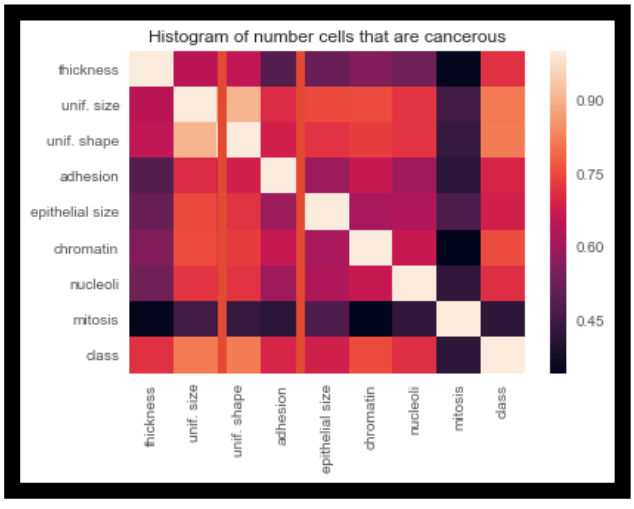

Figure 16 Heat Map

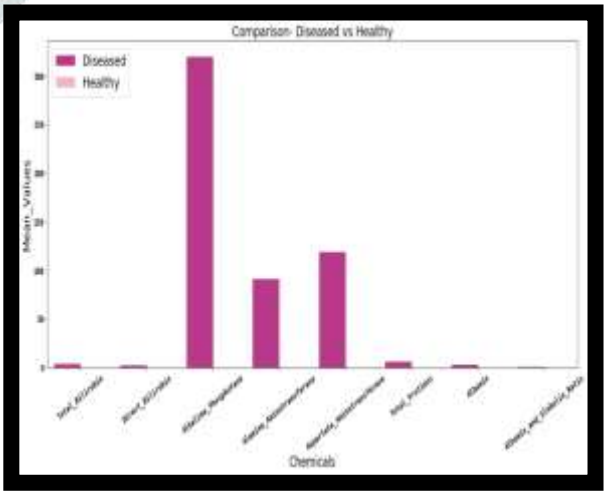

Figure 7 Percentage Of Chemicals In Unhealthy People

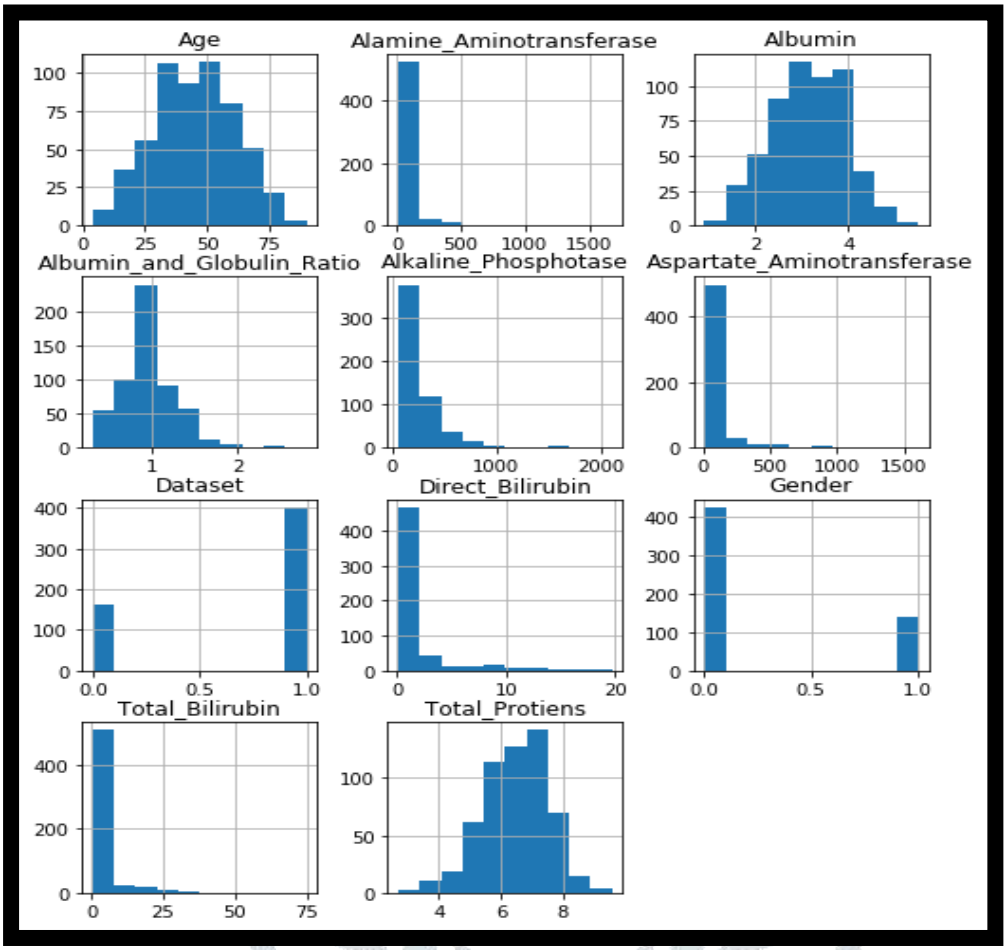

Figure 14 Liver Disease Prediction Analytics

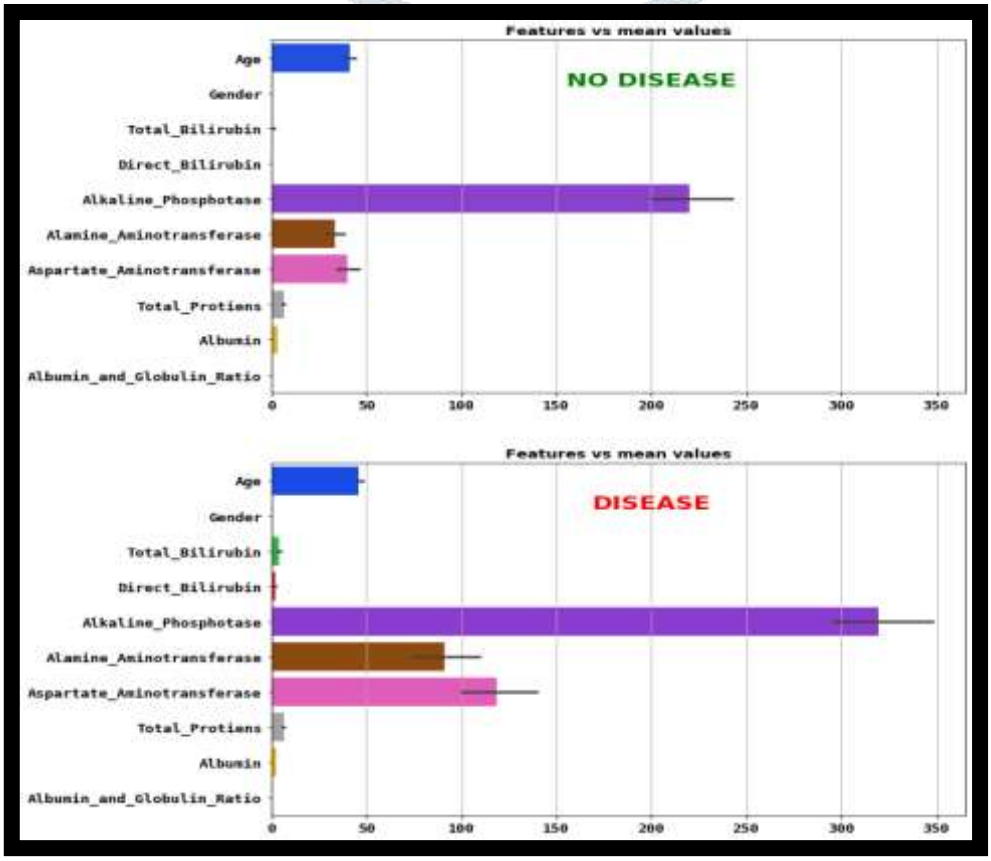

Figure 15 Exploratory Data Analysis

VISUALISING THE DIFFERENCES IN CHEMICAL IN (healthy/unhealthy) PEOPLE

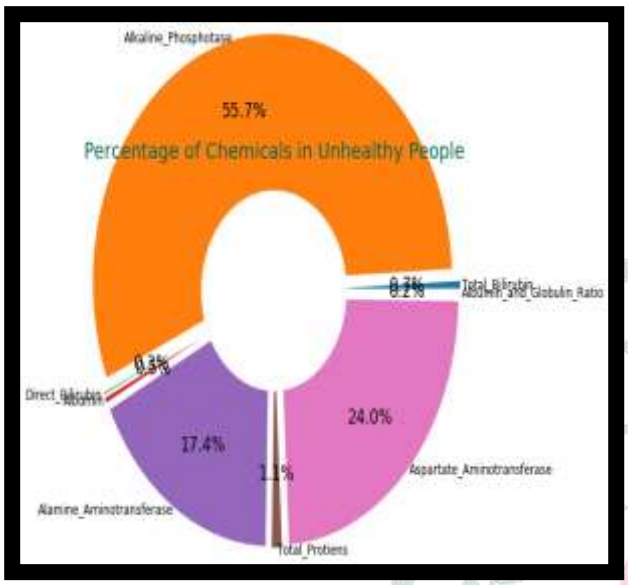

Figure 16 Dataset Outcome Analysis

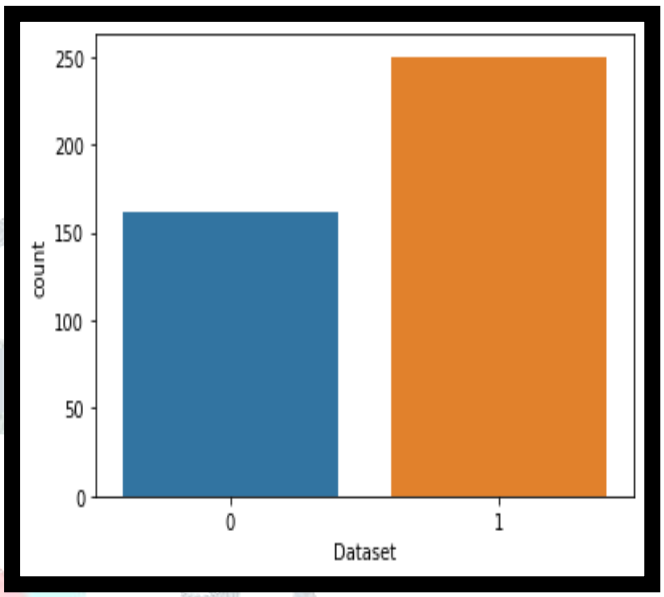

Figure 17 Data Set

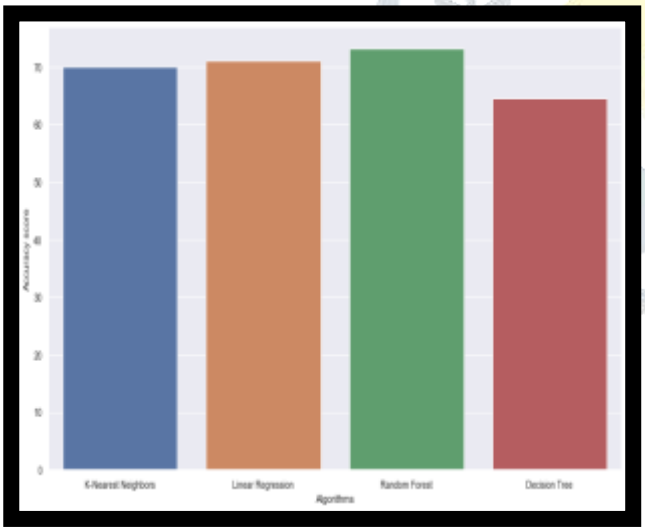

Figure 18 Model Accuracy

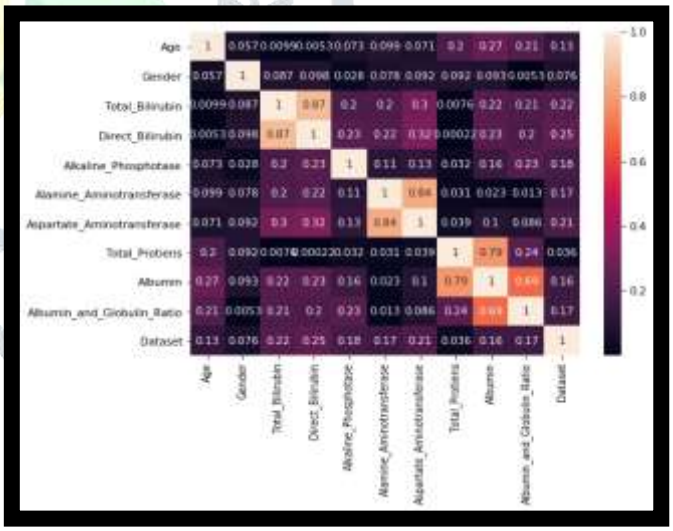

Figure 8 Heat Map

IV. CONCLUSION & FUTURE SCOPE

We present the heart disease prediction system in this project, which uses various classifier algorithms to forecast the onset of heart disease. Yolo Classifications and Logistic Regression are the methodologies used. Our analysis shows that Yolo Classifications are more accurate than Logistic Regression. By eliminating extraneous and irrelevant features from the dataset and selecting only the most informative ones for the classification task, we hope to enhance the performance of the Yolo Classifications.

As previously demonstrated, any practitioner can use the system as a clinical assistant. Any internet user can access the system using a web browser and comprehend the risk of heart disease because the disease prediction based on risk variables can be hosted online. You can use the suggested model for any real-time application. Other forms of cardiac illness can also be identified with the help of the suggested model. It is possible to distinguish between many heart conditions, including inflammatory heart disease, hypertensive heart disease, ischemic heart disease, and cardiovascular disease.

This suggested concept can be used to create other health care systems that will detect ailments early on. To execute the suggested model in real time, a processor that is both efficient and has a well-designed memory arrangement is needed. The suggested paradigm can be used in many other contexts, such as robotic modeling, cloud computing, grid computing, etc. In the future, we will focus on combining two methods, Yolo Classifications and Adaboost, to improve the performance of our classifier. The combination of these two methods will yield excellent performance.

REFERENCES

[1]. Muhammed Golec , Sukhpal Singh Gill , Ajith Kumar Parlikad , and Steve Uhlig "HealthFaaS: AI-Based Smart Healthcare System for Heart Patients Using Serverless Computing", IEEE Internet Of Things Journal, Vol. 10, No. 21, 1 November 2023

[2]. H. Qiu, L. Luo, Z. Su, L. Zhou, L. Wang, and Y. Chen, "Machine learning approaches to predict peak demand days of cardiovascular admissions considering environmental exposure," BMC Med. Inform. Decis. Making, vol. 20, no. 1, p. 83, 2020.

[3]. M. Roberts. "Unhealthy Britain: Nation's five big killers." Mar. 2013. [Online]. Available: https://www.bbc.com/news/health-21667065.

[4]. "The top 10 causes of death." Dec. 2020. [Online]. Available: https:// www.who.int/news-room/fact-sheets/detail/the-top-10-causes-of-death.

[5]. S. S. Gill et al., "AI for next generation computing: Emerging trends and future directions," Internet Things, vol. 19, Aug. 2022, Art. no. 100514.

[6]. S. Price. "Avoidable deficiencies in heart failure cost NHS £21m." Mar. 2020. [Online]. Available: https://www.healtheuropa.eu/avoidabledeficiencies- in-heart-failure-cost-nhs-21m/98696/.

[7]. A. Sitar-T˘aut, D. Zdrenghea, D. Pop, and D. Sitar-T˘aut, "Using machine learning algorithms in cardiovascular disease risk evaluation," Age, vol. 1, no. 4, p. 4, 2009.

[8]. R. Singh and S. S. Gill, "Edge AI: A survey," Internet Things Cyber Phys. Syst., vol. 3, pp. 71–92, Feb. 2023.

[9]. H. Raj, M. Kumar, P. Kumar, A. Singh, and O. P. Verma, "Issues and challenges related to privacy and security in healthcare using IoT, fog, and cloud computing," in Advanced Healthcare Systems: Empowering Physicians With IoT-Enabled Technologies. Hoboken, NJ, USA: Wiley, 2022, pp. 21–32.

[10]. S. Iftikhar et al., "AI-based fog and edge computing: A systematic review, taxonomy and future directions," Internet Things, vol. 21, Apr. 2023, Art. no. 100674.

[11]. "Heart disease data set." 1988. [Online]. Available: https://archive.ics. uci.edu/ml/datasets/Heart+Disease

[12]. K. Polat and S. Güne¸s, "A new feature selection method on classification of medical datasets: Kernel F-score feature selection," Expert Syst. Appl., vol. 36, no. 7, pp. 10367–10373, 2009.

[13]. R. Spencer, F. Thabtah, N. Abdelhamid, and M. Thompson, "Exploring feature selection and classification methods for predicting heart disease," Digit. Health, vol. 6, Mar. 2020, Art. no. 205520762091477.

[14]. Y. Khourdifi and M. Bahaj, "Heart disease prediction and classification using machine learning algorithms Optimized by particle swarm optimization and ant colony optimization," Int. J. Intell. Eng. Syst., vol. 12, no. 1, pp. 242–252, 2019.

[15]. N. Satyanandam and C. Satyanarayana, "Heart disease detection using predictive optimization techniques," Int. J. Image, Graph. Signal Process., vol. 11, no. 9, pp. 18–24, 2019.

[16]. M. Tarawneh and O. Embarak, "Hybrid approach for heart disease prediction using data mining techniques," in Advances in Internet, Data and Web Technologies. Cham, Switzerland: Springer, 2019, pp. 447–454.

[17]. F. Z. Abdeldjouad, M. Brahami, and N. Matta, "A hybrid approach for heart disease diagnosis and prediction using machine learning techniques," in The Impact of Digital Technologies on Public Health in Developed and Developing Countries (Lecture Notes in Computer Science 12157). Cham, Switzerland: Springer, 2020, pp. 299–306.

[18]. P. H. Vilela, J. J. Rodrigues, P. Solic, K. Saleem, and V. Furtado, "Performance evaluation of a fog-assisted IoT solution for e-health applications," *Future Gener. Comput. Syst.*, vol. 97, pp. 379–386, Aug. 2019.

[19]. Arti Gupta, Maneesh Shreevastava, "Medical Diagnosis using Back propagation Algorithm" International Journal of Emerging Technology and Advanced Engineering Website: www.ijetae.com (ISSN 2250-2459, Volume 1, Issue 1, November 2011)

[20]. Shraddha Subhash Shirsath, Prof. Shubhangi Patil, "Disease Prediction Using Machine Learning Over Big Data" International Journal of Innovative Research in Science, Engineering and Technology, Vol. 7, Issue 6, June 2018

[21]. Nikita Kamble, Manjiri Harmalkar, Manali Bhoir, Supriya Chaudhary,. "Smart Health Prediction System Using Machine Learning." IJSRCSEIT, 2017, Volume 2, Issue 2, March-April-2017.

[22]. Nilesh Borisagar, Dipa Barad, Priyanka Raval,. "Chronic Kidney Disease Prediction using Back Propagation Neural Network Algorithm". Conference paper (PICCN), April 2017,

[23]. Sellappan Palaniappan, Rafiah Awang IEEE, 2008. Intelligent Heart Disease Prediction System Using Machine Learning Technique. SP-108, EP-115 SN- 978-1-4244-1967-8.