**Abstract:**

The healthcare domain is one of the prominent research fields in the current scenario with the rapid improvement of technology and data. It is difficult to handle the huge amount of data of the patients. It is easier to handle this data through Big Data Analytics. There are a lot of procedures for the treatment of multiple diseases across the world. Machine Learning is an emerging approach that helps in prediction, heart of a disease. This project depicts the heart prediction of disease based on symptoms using machine learning. Machine Learning algorithms such as Naive Bayes, Decision Tree and Random Forest are employed on the provided dataset and predict the disease. Its implementation is done through the python programming language. The research demonstrates the best algorithm based on their accuracy. The accuracy of an algorithm is determinedly the performance on the given dataset.

**Introduction:**

At present, when one suffers from particular disease, then the person has to visit to doctor which is time consuming and costly too. Also if the user is out of reach of doctor and hospitals it may be difficult for the user as the disease can not be identified. So, if the above process can be completed using a automated program which can save time as wells money, it could be easier to the patient which can make the process easier. There are other Heart related Disease Prediction System using data mining techniques that analyzes the risk level of the patient. Heart Disease Predictor is a web based application that predicts the heart disease of the user with respect to the symptoms given by the user. heart Disease Prediction system has data sets collected from different health related sites. With the help of heart Disease Predictor the user will be able to know the probability of the disease with the given symptoms. As the use of internet is growing every day, people are always curious to know different new things. People always try to refer to the internet if any problem arises. People have access to internet than hospitals and doctors. People do not have immediate option when they suffer with particular disease. So, this system can be helpful to the people as they have access to internet 24 hours.

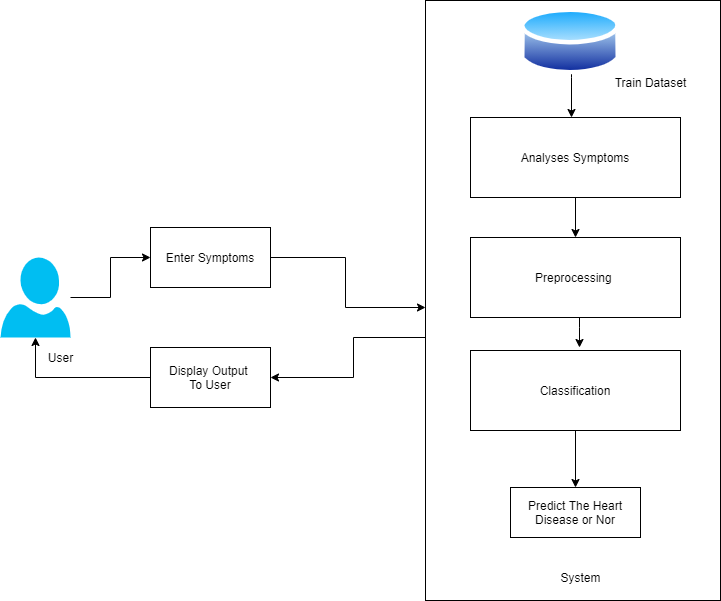
**Problem Statement:**

There are many tools related to disease prediction. But particularly heart related diseases have been analyzed and risk level is generated. But generally there are no such tools that are used for prediction of heart diseases. So Dis-ease Predictor helps for the prediction of the heart diseases.

**Motivation:**

The current systems working on heart disease prediction works on a small dataset. The aim of our system is to work on a larger dataset to increase the efficiency of the overall system. Our system is easy to heart predict disease give output quickly the heart Disease prediction functioning depends on Natural language processing that helps users to submit their problem about the health.

**System Architecture:**

****

**Literature Survey**

**1.Paper Name:** Application of Machine Learning in Disease Prediction

**Authors:** Pahulpreet Singh Kohli,Shriya Arora

**Abstract:** The application of machine learning in the field of medical diagnosis is increasing gradually. This can be contributed primarily to the improvement in the classification and recognition systems used in disease diagnosis which is able to provide data that aids medical experts in early detection of fatal diseases and therefore, increase the survival rate of patients significantly. In this paper, we apply different classification algorithms, each with its own advantage on three separate databases of disease (Heart, Breast cancer, Diabetes)available in UCI repository for disease prediction. The feature selection for each dataset was accomplished by backward modeling using the p-value test. The results of the study strengthen the idea of the application of machine learning in early detection of diseases.

**2.Paper Name:**Design And Implementing Heart Disease Prediction Using Naives Bayesian

**Authors:**Anjan Nikhil Repaka,Sai Deepak Ravikanti.

**Abstract:** Data mining, a great developing technique that revolves around exploring and digging out significant information from massive collection of data which can be further beneficial in examining and drawing out patterns for making business related decisions. Talking about the Medical domain, implementation of data mining in this field can yield in discovering and withdrawing valuable patterns and information which can prove beneficial in performing clinical diagnosis. The research focuses on heart disease diagnosis by considering previous data and information. To achieve this SHDP (Smart Heart Disease Prediction) is built via Navies Bayesian in order to predict risk factors concerning heart disease. The speedy advancement of technology has led to remarkable rise in mobile health technology that being one of the web application. The required data is assembled in a standardized form. For predicting the chances of heart disease in a patient, the following attributes are being fetched from the medical profiles, these include: age, BP, cholesterol, sex, blood sugar etc... The collected attributes acts as input for the Navies Bayesian classification for predicting heart disease. The dataset utilized is split into two sections, 80\% dataset is utilized for training and rest20\% is utilized for testing. The proposed approach includes following stages: dataset collection, user registration and login(Application based), classification via Navies Bayesian, prediction and secure data transfer by employing AES (Advanced Encryption Standard). Thereafter result is produced. There search elaborates and presents multiple knowledge abstraction techniques by making use of data mining methods which are adopted for heart disease prediction. The output reveals that the established diagnostic system effectively assists in predicting risk factors concerning heart diseases.

**3.Paper Name:** Efficient Heart Disease Prediction System using Decision Tree

**Authors:**Purushottam,Prof. (Dr.) Kanak Saxena,Richa Sharma

**Abstract:** Cardiovascular disease (CVD) is a big reason of morbidity and mortality in the current living style. Identification of Cardiovascular disease is an important but complex task that needs to be performed very minutely, efficiently and the correct automation would be very desirable. Every human being can not be equally skillful and so as doctors. All doctors cannot be equally skilled in every subspecialty and at many places we don’t have skilled and specialist doctors available easily. An automated system in medical diagnosis would enhance medical care and it can also reduce costs. In this study, we have designed a system that can efficiently discover the rules to predict the risk level of patients based on the given parameter about their health. The rules can be prioritized based on the user’s requirement. The performance of the system is evaluated in terms of classification accuracy and the results shows that the system has great potential in predicting the heart disease risk level more accurately.

**4.Paper Name:** Diabetes Disease Prediction Using Data Mining

**Authors:**Deeraj Shetty,Kishor Rit,Sohail Shaikh

**Abstract:** Data mining is a subfield in the subject of software engineering. It is the methodical procedure of finding example sin huge data sets including techniques at the crossing point of manufactured intelligence, machine learning, insights, and database systems. The goal of the data mining methodology is to think data from a data set and change it into a reasonable structure for further use. Our examination concentrates on this part of Medical conclusion learning design through the gathered data of diabetes and to create smart therapeutic choice emotionally supportive network to help the physicians. The primary target of this examination is to assemble Intelligent Diabetes Disease Prediction System that gives analysis of diabetes malady utilizing diabetes patient's database. In this system, we propose the use of algorithms like Bayesian and KNN (K-Nearest Neighbor) to apply on diabetes patient’s database and analyze them by taking various attributes of diabetes for prediction of diabetes disease.

**5.Paper Name:** Neurodegenerative disease prediction based on gait analysis signals acquired with force-sensitive

**Authors:**ger Selzler 1, James R. Green 1, Rafik Goubran

**Abstract:** Neurodegenerative diseases such as Parkinson’s Disease (PD), Huntington Disease (HD), and Amyotrophic Lateral Sclerosis (ALS) affect the lives of thousands of people around the world. One of the consequences of such diseases occurs in the motor neurons of the patients, resulting in problem sin movement, causing a change in gait pattern. Force sensitive resistors can be used to measure the force/pressure between the shoe and the patient’s foot, providing information about the gait dynamics when the patient walks. This project uses signals from the Gait Dynamics in Neuron-Degenerative Disease database to extract features for classification of neurodegenerative diseases(NDD). Manually labeled features from the database are used for comparison with previous studies. Time series signals is also used, where algorithms for signal reliability, feature extraction and feature selection are implemented, allowing real-time signal processing and classification. Multiple feature sets are used for classification with algorithms such as K-nearest neighbor, Support Vector Machines, and Decision Trees, and the performance of these algorithms are then reported. This study presents a real-time system with accuracy exceeding 82\% for the aforementioned diseases. Finally, a discussion about possible improvements for future studies are presented.\

**System Requirements:**

**Hardware Requirements:**

* System : Intel I5 Processor.
* Hard Disk : 500 GB.
* Monitor : 15 VGA Colour.
* Mouse : Logitech.
* Ram : 8 GB

**Software Requirements:**

* Operating system : Windows 10.
* Coding Language : Python
* IDE : Spyder
* Database : Xampp

**Conclusion:**

This project aims to predict the heart disease on the basis of the symptoms. The project is designed in such a way that the system takes symptoms from the user as input and produces output i.e. predict heart disease. Average prediction accuracy probability of 55% is obtained. heart Disease Predictor was successfully implemented using grails framework.

**References**

1. D. Tian, J. Zhou, Y. Wang, Y. Lu, H. Xia, and Z. Yi, “A dynamic andself-adaptive network selection method for multimode communicationsin heterogeneous vehicular telematics,” IEEE Transactions on IntelligentTransportation Systems, vol. 16, no. 6, pp. 3033–3049, 2015.

2. M. Chen, Y. Ma, Y. Li, D. Wu, Y. Zhang, C. Youn, “Wearable 2.0: EnableHuman-Cloud Integration in Next Generation Healthcare System,” IEEECommunications, Vol. 55, No. 1, pp. 54–61, Jan. 2017.

3. M. Chen, Y. Ma, J. Song, C. Lai, B. Hu, ”Smart Clothing: ConnectingHuman with Clouds and Big Data for Sustainable Health Monitoring,”ACM/Springer Mobile Networks and Applications, Vol. 21, No. 5, pp.825C845, 2016

4. M. Chen, P. Zhou, G. Fortino, “Emotion Communication System,” IEEEAccess, DOI: 10.1109/ACCESS.2016.2641480, 2016.

5. J. Wang, M. Qiu, and B. Guo, “Enabling real-time information service ontelehealth system over cloud-based big data platform,” Journal of SystemsArchitecture, vol. 72, pp. 69–79, 2017.

6. K. Lin, J. Luo, L. Hu, M. S. Hossain, and A. Ghoneim, “Localizationbased on social big data analysis in the vehicular networks,” IEEETransactions on Industrial Informatics, 2016.

7. K. Lin, M. Chen, J. Deng, M. M. Hassan, and G. Fortino, “Enhancedfingerprinting and trajectory prediction for iot localization in smartbuildings,” IEEE Transactions on Automation Science and Engineering,vol. 13, no. 3, pp. 1294–1307, 2016.

8. B. Qian, X. Wang, N. Cao, H. Li, and Y.-G. Jiang, “A relative similaritybased method for interactive patient risk prediction,” Data Mining andKnowledge Discovery, vol. 29, no. 4, pp. 1070–1093, 2015.

9. S. Bandyopadhyay, J. Wolfson, D. M. Vock, G. Vazquez-Benitez,G. Adomavicius, M. Elidrisi, P. E. Johnson, and P. J. O’Connor, “Datamining for censored time-to-event data: a bayesian network model forpredicting cardiovascular risk from electronic health record data,” DataMining and Knowledge Discovery, vol. 29, no. 4, pp. 1033–1069, 2015.