Multiple disease

Abstract: Using Machine learning, our project proposes disease prediction system. For small issues, the users need to go in person to the hospital for check-up that is longer intense. Also handling the telecom entails appointments is kind of agitated. Such a tangle may be solved by Disease prediction application by giving correct steerage relating to healthy living. Over the past decade, the utilization of the particular disease prediction tools alongside the regarding health has been magnified because of a range of diseases and fewer doctor-patient magnitude relation. Thus, during this system, we have a tendency to area unit concentrating on providing immediate and correct disease prediction to the users concerning the symptoms they enter alongside the severity of disease expected. Best appropriate rule and doctor consultation are given during this project. For prediction of diseases, totally different machine learning algorithms area unit wont to guarantee fast and correct predictions. In one channel, the symptoms entered are crosschecked with the information. Further, can be preserved within the information if the symptom is new that its primary work is and therefore the different channel will offer severity of disease expected. A web/android application is deployed for user for straightforward moveableness, configuring and having the ability to access remotely wherever doctors cannot reach simply. usually users don't seem to be privy to all the treatment relating to the actual disease, this project additionally appearance forward to providing medication and drug consultation of disease expected. Therefore, this arrangement helps in easier health management. Keywords: Machine Learning, KNN algorithm, SVM, Decision Tree Algorithm, Naïve Bayes Algorithm, Django, Python, etc.

I. INTRODUCTION The Earth is passing through a violet patch of technology, wherever there's increasing demand of intelligence and accuracy behind it. Today’s individuals area unit more likely addicted to web however they're not involved about their personal health. during this twenty first Century humans area unit encircled with technology as they're the constituent of our day to day life cycle. With this we tend to area unit invariably specializing in the health for ourselves and our earned valuables severally. individuals avoid to travel in hospital for tiny downside which can become a significant malady in future. Establishing question answer forums is turning into a straightforward thanks to answer those queries instead of browsing through the list of probably relevant document from the net. Our basic plan is to develop a system which can predict and provides the main points of the disease foreseen together with its severity that as symptoms area unit given as input by the user. The system can compare the symptoms with the datasets provided within the information. If the symptom matches the datasets then it ought to raise alternative relevant symptoms specifying the name of the symptom. If not, the symptom entered ought to be notified as wrong symptom. once this a prompt can come back up asking whether or not you would like to still save the symptom within the information. If you click on affirmative, it'll be saved within the information, if not it'll attend the recycle bin. the most feature are the machine learning, during which we'll be mistreatment algorithms like Naïve Thomas Bayes algorithmic rule, K-Nearest algorithmic rule, Decision Tree algorithmic rule, Random Forest algorithmic rule and Support Vector Machine, which can predict correct disease and additionally, can realize that which algorithmic rule provides a quicker and economical result by comparatively comparing.

IV. PROPOSED SYSTEM This paper proposing such a system which will flaunt a simple and elegant User Interface and also be time efficient. In order to make it less time consuming we are aiming at a more specific questionnaire which will be followed by the system. Our aim with this system is to be the connecting bridge between doctors and patients. The main feature will be the machine learning, in which we will be using algorithms such as Naïve Bayes Algorithm, K-Nearest Algorithm, Decision Tree Algorithm, Random Forest Algorithm and Support Vector Machine, which will help us in getting accurate predictions and Also, will find which algorithm gives a faster and efficient result by comparatively-comparing. Another feature that our system will comprise of is Doctor’s Consultation. After delivering the results, our system will also suggest the user to get a doctors consultation on this report

VII. METHOD DESCRIPTION In this paper, We are using total five algorithms for predicting the disease. Which are Naïve Bayes, SVM, KNN, Decision Tree and Random Forest Algorithm. NAÏVE BAYES. Naive mathematician classifiers are a set of classification algorithms supported Bayes’ Theorem. it's not one formula however a family of algorithms wherever all of them share a standard principle, i.e. each try of options being classified is freelance of every different. SVM. Support Vector Machine or SVM is one among the foremost standard supervised Learning algorithms, that is employed for Classification further as Regression issues. However, primarily, it's used for Classification issuesin Machine Learning. The goal of the SVM rule is to make the simplest line or call boundary that may segregate n-dimensional house into categories in order that we will simply place the new datum within the correct class within the future. This best call boundary is termed a hyperplane. SVM chooses the acute points/vectors that facilitate in making the hyperplane. These extreme cases square measure known as support vectors, and therefore rule is termed as Support Vector Machine. KNN. K-nearest neighbors (KNN) rule could be a variety of supervised metric capacity unit rule which may be used for each classification yet as regression prophetical issues. However, it's primarily used for classification prophetical issues in trade. the subsequent 2 properties would outline KNN well − • Lazy learning algorithm − KNN could be a lazy learning rule as a result of it doesn't have a specialised coaching section and uses all the information for coaching whereas classification. • Non-parametric learning algorithm − KNN is additionally a non-parametric learning rule as a result of it doesn’t assume something regarding the underlying knowledge. Decision Tree Algorithm. In medicative selection like arrangement, diagnosis there are varied circumstances wherever selection should be created with success and reliably. cheap easy basic leadership models with the probability of programmed learning ar the foremost fitting for playacting such undertakings. selection trees are a solid and successful basic leadership procedure that furnish high grouping accuracy with an easy portrayal of accumulated learning and that they are utilised as an area of varied zones of restorative basic leadership. Random Forest Algorithm. Random Forest could be a learning methodology that operates by constructing multiple call trees. the ultimate call is created supported the bulk of the trees and is chosen by the random forest. There square measure heaps of advantages to victimisation Random Forest algorithmic rule, however one amongst the most benefits is that it reduces the chance of overfitting and therefore the needed coaching time. in addition, it offers a high level of accuracy. Random Forest algorithmic rule runs with efficiency in massive information bases and produces extremely correct predictions by estimating missing data. VIII. CONCLUSION This paper offers analysis of multiple researches exhausted this field. Our planned System aims at bridging gap between Doctors and Patients which can facilitate each categories of users in achieving their goals. this technique provides support for multiple sickness prediction mistreatment completely different Machine Learning algorithms. this approach of the many systems focuses solely on automating this method that lacks in building the user’s trust within the system. By providing Doctor’s recommendation in our system, we have a tendency to guarantee user’s trust aspect by aspect making certain that the Doctor’s won't feel that their Business is obtaining affected because of this technique.

diabetes

Diseases are one of the most serious challenges in both developing and developed countries. According to the International Diabetes Federation, there are 285 million ill people worldwide. This total is expected to rise to 380 million within 20 years. Due to its importance, a design of classifier for the detection of symptoms of disease with optimal cost and better performance is the need of the age. The Pima Indian database at the UCI machine learning laboratory has become a standard for testing data mining algorithms to see their prediction accuracy in data classification. The proposed method uses Support Vector Machine (SVM), a machine learning method as the classifier for diagnosis of disease. The machine learning method focus on classifying disease from high dimensional medical data set. The experimental results obtained show that support vector machine can be successfully used for diagnosing disease.

Diabetes is one of the common and rapidly increasing diseases in the world. It is a major health problem in most of the countries. Diabetes is a condition in which your body is unable to produce the required amount of insulin needed to regulate the amount of sugar in the body. This leads to various diseases including heart disease, kidney disease, blindness, nerve damage and blood vessels damage. There are two general reasons for diabetes: The pancreas does not make enough insulin or the body does not produce enough insulin. Only 5- 10 % of people with this have this form of the disease (Type-1). Cells do not respond to the insulin that is produced (Type-2). Insulin is the principle hormone that regulates uptake of glucose from the blood into most cells (muscle and fat cells). If the amount of insulin available is insufficient, then glucose will not have its usual effect so that glucose will not be absorbed by the body cells that require it. Disease being one of the major contributors to the mortality rate. Detection and diagnosis of diabetes at an early stage is the need of the day. Disease diagnosis and interpretation of the data is an important classification problem. A classifier is required and to be designed that is cost efficient, convenient and accurate. Artificial intelligence and Soft Computing Techniques provide a great deal of human ideologies and are involved in human related fields of application. These systems find a place in the medical diagnosis. A medical diagnosis is a classification process. A physician has to analyze lot of factors before diagnosing the diabetes which makes physician‟s job difficult. In recent times, machine learning and data mining techniques have been considered to design automatic diagnosis system for diseases.

The Pima are a group of Native Americans living in Arizona. A genetic predisposition allowed this group to survive normally to a

diet poor of carbohydrates for years. In the recent years, because of a sudden shift from traditional agricultural crops to processed

foods, together with a decline in physical activity, made them develop the highest prevalence of disease and for this reason they

have been subject of many studies.

A. Dataset

The dataset includes data from 768 women with 8 characteristics, in particular:

1) Number of times pregnant (opction in this project)

2) Plasma glucose concentration a 2 hours in an oral glucose tolerance test

3) Blood pressure (mm Hg)

4) Triceps skin fold thickness (mm)

5) 2-Hour serum insulin (mu U/ml)

6) Body mass index (weight in kg/(height in m)^2)

7) Diabetes pedigree function

8) Age (years)

The Problem The type of dataset and problem is a classic supervised binary classification. Given a number of elements all with certain characteristics (features), we want to build a machine learning model to identify people affected by type 2 diabetes. To solve the problem we will have to analyse the data, do any required transformation and normalisation, apply a machine learning algorithm, train a model, check the performance of the trained model and iterate with other algorithms until we find the most performant for our type of dataset.

CONCLUSION The development of software or website includes so many people like user system developer, user of system and the management, it is important to identify the system requirements by properly collecting required data to interact with supplier and customer of the system. Proper design builds upon this foundation to give a blue print, which is actually implemented by the developers. On realizing the importance of systematic documentation all the processes are implemented using a software engineering approach. Working in a live environment enables one to appreciate the intricacies involved in the System Development Life Cycle (SDLC). We have gained a lot of practical knowledge from this project, which we think, shall make us stand in a good state in the future.

Parkinsons

ABSTRACT: The Parkinson’s disease is progressive neuro degenerative disorder that affects a lot only people significantly affecting their quality of life. It mostly affect the motor functions of human. The main motor symptoms are called "parkinsonism" or "parkinsonian syndrome". The symptoms of Parkinson’s disease will occur slowly, the symptoms include shaking, rigidity, slowness of movement and difficulty with walking, Thinking and behavior change, Depression and anxiety are also common. There is a model for detecting Parkinson’s using voice. The deflections in the voice will confirm the symptoms of Parkinson’s disease. This project showed 73.8% efficiency. In our model, a huge amount of data is collected from the normal person and also previously affected person by Parkinson’s disease. these data is trained using machine learning algorithms. From the whole data 60% is used for training and 40% is used for testing. The data of any person can be entered in db to check whether the person is affected by Parkinsons disease or not. There are 24 columns in the data set each column will indicate the symptom values of a patient except the status column. The status column has 0's and I's.those values will decide the person is effected with Parkinsons disease. I's indicate person is effected, 0's indicate normal conditions.

1. INTRODUCTION: Parkinson’s disease is•a disorder of the central nervous system affecting movement and inducing tremors and stiffness a neurodegenerative disorder affecting dopamine neurons in brain. Parkinson’s disease is difficult to diagnose. Common diagnostic criteria require the medication before. In this model, the huge data is collected from previously affected person and then by using machine learning algorithm will process the user input data with previous data to check he/she affected.

heart

Abstract. In the last few years, cardiovascular diseases have emerged as one of the most common causes of deaths worldwide. The lifestyle changes, eating habits, working cultures etc, has significantly contributed to this alarming issue across the globe including the developed, underdeveloped and developing nations. Early detection of the initial signs of cardiovascular diseases and the continuous medical supervision can help in reducing rising number of patients and eventually the mortality rate. However with limited medical facilities and specialist doctors, it is difficult to continuously monitor the patients and provide consultations. Technological interventions are required to facilitate the patient monitoring and treatment. The healthcare data generated through various medical procedures and continuous patient monitoring can be utilized to develop efficient prediction models for cardiovascular diseases. The early prognosis of cardiovascular illnesses can aid in making decisions on life-style changes in high hazard sufferers and in turn lessen the complications, which may be an outstanding milestone inside the field of medicine. This paper studies some of the most widely used machine learning algorithms for heart disease prediction by using the medical data and historical information. The various techniques are discussed and a comparative analysis of the same is presented. This report comparesfive common strategiesfor predicting the chanceof heart attack that havebeen published in the literature. KNN, Decision Tree, Gaussian Naive Bayes, Logistic Regression, and Random Forest are some of the approaches used. Further, the paper also highlights the advantages and disadvantages of using the various techniques for developing the prediction models. Keywords- Machine Learning, Heart disease prediction, Logistic regression, Decision Tree, Random Forest, Gaussian Naïve Bayes, KNN, Cross-Validation.