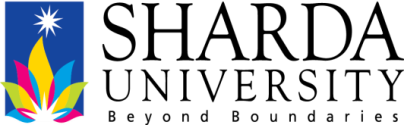
****

**PROJECT BASED LEARNING (PBL-3) LAB (CSP351)**

**HEALTH MONITORING SYSTEM USING MACHINE LEARNING  
B.TECH 3rd YEAR**

**SEMESTER: 6th**

**SESSION: 2021-2022**

**Submitted By:**

**Name:**

**Sayon Chakraborty**

**Shubham**

**Shreya Singh**

**Roll No / System ID:**

**2019606193**

**2019645826**

**2019004953**

**Section: C**

**Submitted To: Dr. Mandeep Kaur**

**DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING**

**SHARDA UNIVERSITY, GREATER NOIDA**

Table of Contents Page No

[Project Title 3](#_Toc78798320)

[Team / Group Formation: 3](#_Toc78798321)

[Technologies to be used 3](#_Toc78798322)

[Problem Statement 4](#_Toc78798324)

[Literature Survey 4](#_Toc78798325)

[Project Description 4](#_Toc78798326)

[Project Modules: Design/Algorithm 5](#_Toc78798327)

[Implementation Methodology 6](#_Toc78798328)

[Result & Conclusion 7](#_Toc78798329)

[Future Scope and further enhancement of the Project 8](#_Toc78798330)

[Advantages of this Project 8](#_Toc78798331)

[Outcome 9](#_Toc78798332)

[References 9](#_Toc78798333)

# Project Title

Our novel approach to create a HEALTH MONITORING SYSTEM USING MACHINE LEARNING that provide user analysis of their current state of health and what should do they do for being healthy by using machine learning techniques.

# Team / Group Formation:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **S. No** | **Student Name** | **Roll Number** | **System ID** | **Role** |
| **1** | **Sayon** | **190101274** | **2019606193** | **Implementation** |
| **2** | **Shubham** | **190101288** | **2019004953** | **Research and analysis** |
| **3** | **Shreya** |  | **2019645826** | **Searching and testing training of the dataset** |

# 

# Technologies to be used

Ssoftware Platform

1. VS CODE as a python based IDE
2. Python modules
3. Test and train a set of datsets

Hardware Platform

1. OS:- Windows or Mac
2. RAM:- 4GB or more

# Problem Statement

In the sphere of portable and smart IoT devices, the healthcare sector has seen significant advancement. The analysis of huge patient data sets by clustering and classification of the data is one of the ways that has been employed to offer this E-health monitoring system a new direction. Clustering is the process of creating a comprehensive dataset of patients with a specific ailment and then analysing their various data points.E-Health monitoring system, but none of them has shown to be effective in solving the problem. Many academics are analysing a large amount of data on patients in order to remedy the problem with the health-monitoring technologies using different machine learning approaches, adequate accuracy may be reached. Machine learning techniques are significantly more efficient for patient health monitoring and provide effective performance of “feature selection, classification, and clustering of data” by sifting through a massive database of patient information.

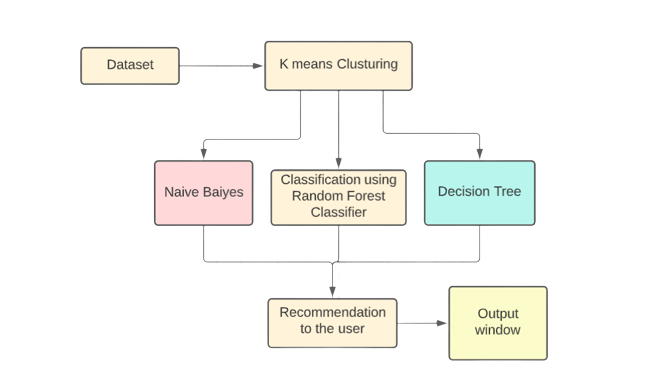
# Literature Survey

Health monitoring system is one of the most important I set in the future for the identification and prevention of illnesses like diabetes in nations like India, which has the highest number of diabetic patients in the world, utilising big data and machine learning methods. One of the most important innovations in the healthcare business is the usage of IoT Devices, which allows patients' vitals to be measured in a cost-effective manner. There are a variety of health monitoring systems available on the market. We're going through things like "activity tracking," "SpO2 level," "heartbeat monitoring," and so on. There is a large data set in the suggested system via which it can anticipate.

# Project Description

The proposed system has a large data set through which it can predict the disease criteria on which the patient will fall, as well as recommend a hilly diet and rate of exercise in the patient's daily life. The goal is to make this system more and more accurate by analysing a large patient data set. The first approach is to go through with the diabetes prediction and other diseases prediction that the patient will fall into. The system will attempt to forecast the outcome with increasing accuracy in order to provide a better consumer interactive experience. Including this it will assists the patient in understanding what foods are required and what steps and diet should be followed to be healthy. The system is analyzing the data that is available and through which it can predicts the disease and what are the steps and diet should one follow to get rid of it or be that in control. The available devices that are present in the market are basically based on IOT but very of them are precise and accurate which will cost the customer’s pocket. Beside this the devices need a lot of technical and not easy to use, The Nobel approach is to make a better user-friendly E-health Monitoring system that is precise and cost effective. [1] In today's Internet world, people are frequently confused about diseases, and a large amount of data on the internet is frequently unable to effectively answer the question of a patient's symptoms due to a lack of required dataset. In this proposed system, the symptoms are entered by the patient themselves, and the system predicts which disease the patient may be suffering from based on that information. The suggested system issue statement is based on the notion that individuals nowadays are highly interested in learning what type of sickness they have, even if their symptoms are moderate. We used to look for solutions all over the internet, but we never received the ideal answer; instead, we got a lot of unwelcome results, which made us feel even more anxious and afraid

# Algorithms



*Figure 1. Workflow of the system*

The dataset for the proposed system is collected as two csv files that contain informative information about various types of diseases and their severe symptoms, and by analysing the different symptoms and using classifier techniques, the prediction of the particular disease can be done. There are several machine learning prediction and classification techniques used to perform the prediction of the disease, which increases the system's accuracy. The following are some of the classification approaches that were utilised to create the prediction system:

1.1 **Content Filtering Method:-**In order to produce predictions, the content-based technique is a domain-dependent algorithm that emphasizesanalyzing the attributes of objects. When Pages, articles, and news are presented as files, and Content-based filtering is without a doubt the most profitable filtering method. In the event of an emergency, the recommendation is based on a content-based filtering method. Characteristics obtained from data in the objects are used to create human profiles. A person who has previously investigated. [2]

1.2 **Collaborative based Filtering Method** - Collaborative filtering is a domain-agnostic prediction method for media such as movies and music that isn't well represented by metadata. The collaborative filtering method saves user preferences for entries in a database (user-item matrix). In a newer, narrower definition, collaborative filtering is a method of collecting preferences or taste information from a large number of users in order to develop automatic predictions (filtering) about a user's interests (collaborating). The collaborative filtering technique's core premise is that if two persons A and B have the same viewpoint on one issue, A is more likely than a randomly selected individual to share B's viewpoint on a different topic.  **Memory-based Filtering Method-** The user's prior ratings are important in his hunt for a neighbour who shares his interests. When a person's neighbour is recognised, a variety of algorithms that take into account the preferences of friends may be utilised to provide suggestions. These tactics have produced considerable success in real-world circumstances due to their applicability.

1. **Model-based Filtering Method -** In this technique, models are constructed using various data mining and machine learning methods to predict user ratings of unrated items. There are a number of model-based CF approaches to choose from. Bayesian networks, clustering models, singular value decomposition, probabilistic latent semantic analysis, multiple multiplicative factors, latent Dirichlet allocation, and Markov decision process based models are all examples of latent semantic models.

**1.3 Hybrid Filtering Method -** Memory-based and model-based CF algorithms are used in a number of applications. Traditional CF techniques have limitations, and these methods increase prediction performance. [3]

**1.4 Naïve Baiyes -** The strategy of creating classifiers in models to assign through a representative vector value using a limited number of class levels is known as naive bias. There is no other approach in which a large number of algorithms are based on the same premise as the night bias classifier, which is considered to be independent of the features assigned to its class variable. [4]

**1.5 Decision Tree -** The decision tree is one of the most powerful and extensively used tools for classification and prediction of data in machine learning. It appears like a tree-like structure and creates a flow chart with the internal node and attribute asset, with each branch representing a different type of data. [5]

**Implementation**

In the proposed system, we use the Python IDE to import all of the training and testing datasets, and we use various types of classifiers in the needed labels to make illness predictions using various types of machine learning approaches. The following is a step-by-step guide to putting the system in place

**Step 1. Header Files Used**

* Pandas use for reading the dataset
* Numpy used to convert features into numpy and then perform different further operations accordingly.
* Tkinter is used for creation of the interface of the proposed system.
* K-Means is used for clustering.
* Train\_test\_split was used to divide the dataset into train and test portions to train and test the model.
* Random Forest classifier is used to predict the the diseases based on the symptoms.

**Step 2. Training and Testing the Data**

The data in the csv files that are to be used for training and testing of various diseases and their symptoms are being implemented on various methods of classifiers, such as the random forest, the naive bayes, and decision tree making. The data is being used to analyse the symptoms of various diseases and predictions are being made on the basis of that. The set of csv files training and testing helps to give outa a set of symptoms and a set of diseases that are present in the current situation by analyzing and predicting measures we can determine that the confusion of the people about any disease can be particularly resolved.

**Step 3. Analysing the data using Classifiers and calculating accuracy**

The data of the files will be tested by different classifiers techniques so that the prediction can be done on that basis. Not only one classifier technique will be used if one fails, the other two will respond accordingly. The use of the prediction system is necessary because people have been confused between different diseases and their symptoms for a long time. This system not only resolves that problem that most people have been having.

**Step 4. Adding of GUI interface**

The system must have a graphical user interface (GUI) through which the patient can enter his or her set of symptoms. The patient simply needs to select from a list of symptoms that are already listed in the symptoms set. Based on the symptoms, the prediction technique works, and with the help of various classifier techniques, it can predict which disease is causing the patient's symptoms.

**Result & Conclusion**

After all of the above steps have been completed successfully, the result should be in the form of a GUI interface created with Tkinter (a Python library for creating graphical user interfaces). The prediction techniques used in this system are quite accurate in themselves and help to predict the proper disease based on their training and testing dataset. The system can cluster the data and then train and test the data using various techniques to produce the best prediction. There is a conflict between the data in the training and testing phases because the system is still in the testing phase and the data set is much larger, so there is a chance that there might be clashes or incorrect predictions, but there is also a thing that the patient must keep in mind that the symptoms that are entered must be in appropriate order otherwise the prediction may come out incorrect in conflict situation the patient can go with the tested two output of the classifier techniques the accuracy is pretty good and there is a lot of scope of improvement can be done in this field. The larger the data we have and better features of algorithms if used this system must give out a plenty set of prediction that can be useful in the field of the health for the greater cause.

# Future Scope and further enhancement of the Project

The suggested method uses a vast data collection to anticipate which illness criteria the patient would meet, as well as suggest a hilly diet and activity pace in the patient's everyday life. The objective is to improve the accuracy of this system by analysing a huge set of patient data. The first method is to proceed with the diabetes and other ailments that the patient is likely to develop. In order to give a better customer interactive experience, the system will attempt to anticipate the outcome with increasing accuracy. The future scope of the proposed system is quite intriguing; if the proposed system can be connected to a set or a cloud of datasets with a large amount of training and testing data, then the outcome must be much more precise; additionally, the algorithms that are being used must be more specific and predictive, or to put it another way, more accurate; therefore, the outcome must be quite interesting. Becoming anxious as a result of online searches for their symptoms and receiving strange replies if we talk about the system's GUI interface, the system uses a python-based GUI called Tkinter, which is pretty old fashioned. The system can be connected to the frontend either through API to communicate with users in a better way.

# 

# Advantages of this Project

* Easy to use help patient to predict the disease one would have
* Accuracy based on the choose of symptoms by the user.
* GUI interface for better User Experience
* Potable and easily can be used in a wider platform with a better storage and prediction techniques.
* Consist of huge data set and a set of three prediction algorithms so the outcome of the prediction is mostly said to be accurate.

# Outcome

# The noble approach of this system is to create the system based interface that helps the patient the disease they might be suffering from or any other health issues. The outcome of the project is publishing the research paper from a respected journal.

# References

**[1]** Yong-Yuan Deng, Chin-Ling Chen, Woei-Jiunn Tsaur, Yung-Wen Tang, and Jung-Hsuan Chen, “Internet of things (IoT) based design of a secure and lightweight body areanetwork (BAN) healthcare system”, Sensors, vol. 17, no. 12, pp. 2919, 2017.

**[2]** Mr. Ravish Gupta, Dept. of Electronics and Communication Engineering Smart Health Monitoring and Management Using Internet of things Artificial Intelligence Cloud based processing 2017.

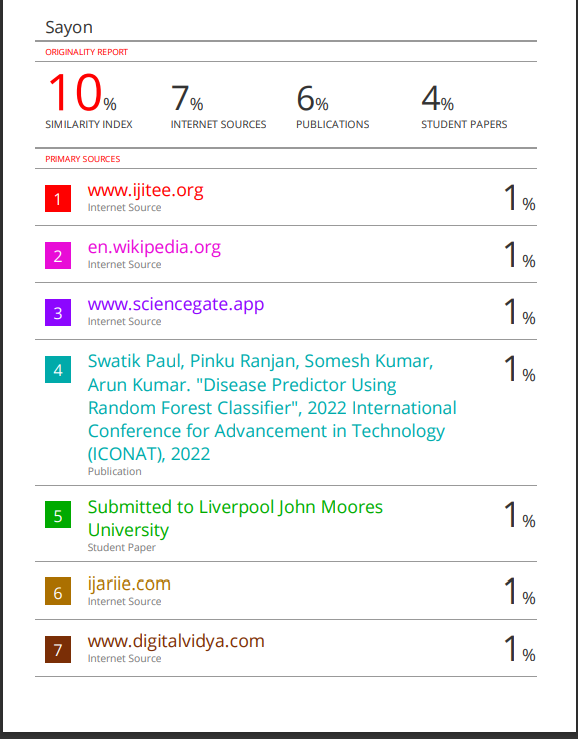
**[3]** Mr. Adars , Vinayaka Mission’s Kadakkal, Kollam, Kerala study on an iot based securehealthcare monitoring system using cryptography 2017.

**[4]** G Sahithi , P Vinayasree , Pallati Narsimhulu real time health monitoring using iot with integration of machine learning approach 2019.

**[5]** Md. Milon Islam , Ashikur Rahaman ,Md. Rashedul Islam Development of Smart Healthcare Monitoring System in IoT

Environment 2018.

**PLAGRASIM REPORT**

****