

PAPER • OPEN ACCESS

An Expert Health Care System Providing Recommendations.

To cite this article: R V Argiddi and R M Yaldandi 2021 *J. Phys.: Conf. Ser.* **1854** 012021

View the [article online](#) for updates and enhancements.

You may also like

- [Preface](#)
- [Analysis of TCM Diagnosis and Treatment of Thyroid Diseases Based on Data Mining](#)
Kankan She and Caiyan Dai
- [Reply to 'Comments on Hereditary Effects of Radiation'](#)
K Sankaranarayanan and N E Gentner



ECS
The
Electrochemical
Society
Advancing solid state &
electrochemical science & technology

DISCOVER
how sustainability
intersects with
electrochemistry & solid
state science research

An Expert Health Care System Providing Recommendations.

R V Argiddi¹ and R M Yaldandi²

¹ Assistance professor of Walchand Institute of Technology, Solapur

² Student of Walchand Institute of Technology, Solapur

¹rvargiddi@gmail.com, ²yaldandi.rajani@gmail.com

Abstract. Today's health information management system collects a huge quantity of healthcare data and information, that hold complex information related to patients and their medical conditions. In data mining, there has been an increase in the use of techniques on medical to determine usage patterns that are used in the analysis. Nowadays human disease is the main reason for human death, one person may suffer from more than one disease. This work is to predict disease by adding symptoms related to patients' condition and behavior. In this developed system we enter symptoms, as per patients' condition, the system analyzes symptoms and predicts diseases using data mining algorithm. The number of symptoms can be added in the search engine as per symptoms disease can be changed. This proposed work is developed to analyze the symptoms given by the user or patient for checking health condition, knowing their health conditions it predicts the disease without the need of a doctor, this work is implemented to save the time of patients'. This proposed system will not only predict the diseases also recommend the appropriate doctors from the doctor list based on a particular disease. In the existing system heart, kidney disease is predicted, whether the patient's heart, kidney are healthy, or not using the K-Nearest Neighbors algorithm. This system is developed for predicting the disease based on the patient's symptoms by using a data mining algorithm.

Keywords. Data Mining, Disease, Prediction algorithm.

1. Introduction

In Today's lifestyle, many changes are happening in people, such as in diet, pollution, environment, and stress that humans being affected by diseases. Today's health care management system provides medical care for patients to protect them from diseases. This work can be used in the medical sector. Many people lose their lives due to diseases that are not detected within time. The proposed work aims to develop a system that maintains a huge amount of datasets, the system analyzes symptoms given by the patient and predicts diseases thus reducing patients' time and effort, this developed system helps patient for knowing their health conditions without the need of a doctor and system also recommends appropriate doctors based on the disease if the patient needs the next treatment.

1.1 Data mining

Data mining is the process of data that can be extracted where a large amount of data stored. Data mining also is known Knowledge Discovery database, which is an iterative process where new data can be merge and modify to get different and more proper results. Different types of data mining techniques are used in technology also in the medical sector.



1.2 Introduction to Diseases

Diseases are the medical conditions that are connected with specific symptoms and signs. A disease is a particular uncommon condition that affects the structure of part of the body. In humans, the disease is carried that causes pain, stress, social problems, or any internal problem to the person. Diseases can affect people physically and mentally, as living with a disease can change patient life. In today's world, many more diseases occur. Many more humans and other mammals are died because of diseases like cancer.

1.2.1 Types of the disease are

- Noncommunicable disease: This disease is spread slowly and long-lasting.
- Metabolic defects: A metabolic defects understanding the impact of non-communicable disease of the human body. It occurs when uncommon reactions in your body distract this process of the body organ.
- Environmental hazard: Metabolic abnormalities result from the effects of external environmental factors. This disease occurs in an area where seafood use.
- Communicable disease: Communicable diseases are transmitted from one body to another body.

2. Objective

The objective of this system is to analyze the symptoms given by the patients and predict the diseases. Based on the predicted disease proposed system recommends a doctor which suitable for a particular disease. In this work, the k-means clustering prediction algorithm is used to predict the disease.

3. Module Description

Diseases Shortlisted Using LAMSTAR Network.

Large Memory Storage and Retrieval (LAMSTAR) is a neural network methodology. A LAMSTAR network performs service for a dynamic neural network. LAMSTAR has been used and applied to many domains, including medical prediction. LAMSTAR network is a faster learning speed and a lower error rate. LAMSTAR network is a network system that analyzes input words for the searching and retrieval of relevant information. The system selects a module of a self-organizing map (SOM) that contains the same dimension of classification as a selected input and where neurons are interconnected horizontally and vertically by arrays of link weights.

In this Module, the correct disease is displayed by adding symptoms, and depending on the disease doctor shortlisted for taking the necessary tests. The final report is computing to obtain the correct symptoms.

The added symptoms were compared with the original symptoms entered in the symptoms dataset.

If the input is matched with any stored keyword, the system updates weights according to the following procedure:

- $W_{i,m}(t+1) = W_{i,m}(t) + \alpha_i (X_i(t) - W_{i,m}(t))$, for $m: \epsilon_{min} < \epsilon(\text{const})$.
- Where $W_{i,m}(t+1)$ = Modified weights in module I for neuron m;
- α_i = Learning coefficient for a module I;
- ϵ_{min} = Minimum error of all weights vectors W_i in a module.

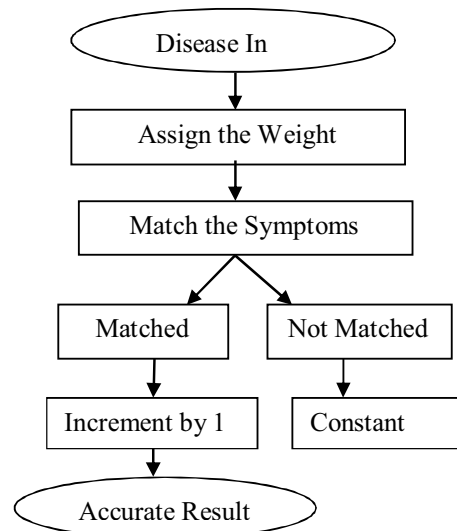


Figure 1. Disease Shortlisted

3.1 Algorithm

The K-Mean clustering prediction algorithm.

K-Means algorithm is a typical clustering algorithm in data mining and which is widely used for clustering the large set of data. In 1967, firstly proposed the K-Means algorithm, it was one of the most simple and non-supervised learning algorithms which were used to solve the problem. This algorithm is used to classify given data objects into K different clusters through the iterative, converging to a local minimum. So the results of created clusters are compact and independent. This algorithm includes two separate phases. In the first phase, selects K-centers randomly, where the value K is fixed in advance. The next phase or step is to take each data object to the nearest center. Euclidean distance is generally considered to determine the distance between each data object and the cluster centers when all data objects are included in some clusters. The first step is completed and the grouping was done. Recalculating the average of the previously formed clusters. This process is repeated continuously until the criterion function becomes minimum. Suppose that the target object is X, X_i indicates the average of cluster C_i , criterion function is defined as follows.

$$kE = \sum \sum |x - x_t| \quad 2 \quad i=1 \quad x \in c_i \quad (1)$$

The sum of the squared error of all objects in the database is E. The distance of the criterion function is Euclidean distance; Euclidean distance is used to determining the nearest distance between each data object and the center point of the cluster. The Euclidean distance d is between vector one is $X = (X_1, X_2, \dots, X_n)$ and vector two is $Y = (Y_1, Y_2, \dots, Y_n)$. The distance d (X_i, Y_i) can be obtained as follow:

$$nd(x_i, y_i) = \left[\sum (x_i - y_i)^2 \right]^{1/2} \quad i=1 \quad (2)$$

3.2 The process of K-Means clustering prediction algorithm as follows

Input: Number of clusters K and a database $D = \{d_1, d_2, \dots, d_n\}$ Containing data objects n.

Output: A set of K clusters.

Steps:

- 1) Randomly or regularly select K data objects from dataset D as an initial cluster center.
- 2) Repeat;
- 3) Calculate the distance between data object d_i ($1 \leq i \leq n$) and all K cluster center j ($1 \leq j \leq k$) and set data object d_i to the nearest cluster.
- 4) For each cluster j ($1 \leq j \leq k$), recalculate the cluster center.

5) Until no change in the center of clusters.

The K-means algorithm is used to partition the dataset into subgroups or clusters. The calculation of distance and center of the cluster is done whereas the loops are executed many times before merging the k-means algorithm. Where t is k-means iteration, the value of t changes at the starting cluster center. The data points are combined with new cluster center computational time complexity is $O(t \cdot n)$ is an iteration of k-means algorithm, n is a number of all data objects, k is a range of cluster, is $k \ll n$ and $t \ll n$.

4. Literature Review

1. In [1] Subasish Mohapatra, Prashanta Kumar Patra “Smart Health Care System using Data Mining” the system shows extracting hidden information or data for datasets. The smart health care management system is a system that allows a user to guide health issues through an online system. In this research, there is a requirement of well-organized methodologies for analyzing, predicting, and detecting diseases. To detecting and predicting diseases data mining applications are used for the management of smart healthcare. In this system heart, kidney diseases predicted, whether the patient's heart, kidney are healthy, or not using the K-Nearest Neighbors algorithm.
2. In [2] Nazim Razali1, Aida Mustapha1, Mohd Helmy Abd Wahab “A Data Mining Approach to Prediction of Liver Diseases”, The objective of this study is proposed a rule-based classification with the decision-making techniques to predict various types of heart disease. To get a better result different data mining algorithms are compared with previous liver disease prediction algorithm. All experiments have been implemented at the Azure Machine Learning tool.
3. In [3] Ria Arafiah1 and Fariani Hermin1 “Data mining for dengue hemorrhagic fever (DHF) prediction with naive Bayes method”, This expert system of dengue prevention by using ANFIS has predicted the weather and the number of sufferers. The number of data on DHF often cannot affect a person in making decisions. The use of data mining methods, able to build a database to support decision-makers' diagnoses. In this, the system predicts DHF with the method of Naive Bayes. The naïve Bayes method can be optimized by using Particle Swarm Optimization (PSO) to improve performance in this case.
4. In [4] Sarath Babu, Vivek EM, Famina KP “Heart Disease Diagnosis Using Data Mining Technique”, In this paper, the sequence of several attributes for effective heart disease prediction and its diagnosis is done using different algorithms in data mining. Decision Tree has good results after computing attributes, applying a genetic algorithm to decrease the existing size of data to get the optimal subset of attribute acceptable for heart disease.
5. In [5] Anne Dirkson, Suzan Verberne “Open Knowledge Discovery and Data Mining from Patient Forums”, In this paper, The researcher developed a methodology for mining the given experiences of the patient on health forums for novel clinical hypotheses. Here entities are extracted using a Conditional Random Fields (CRF).
6. In [6] Cincy Raju, Philip E “A Survey on Predicting Heart Disease using Data Mining Techniques”, Heart disease is a serious disease that will cause death. This disease attacks a person so instantly. For diagnosis, heart diseases such as algorithms are used like decision trees, neural networks, Bayesian classifiers, Support vector machines, Association Rule, K- nearest neighbor classification. Among these algorithms, the Support Vector Machine (SVM) technique gives the best result SVM is an efficient method for predicting heart disease. It also gives good accuracy.
7. In [7] Irum Matloob, Shoab Ahmed Khan “Sequence Mining, and Prediction-Based Healthcare Fraud Detection Methodology”, The fraud and abuse in healthcare systems are becoming a crucial problem nowadays. Healthcare needed to develop fraud detection models. In this proposed system, is validated by generating patient sequences from the last five years transactional data of a local hospital and identifies service procedures to patients for this proposed work using the Prefixspan algorithm and Compact prediction tree.
8. In [8] Mr.Sudhir M. Gorade, Prof. Ankit Deo “A Study of Some Data Mining Classification Techniques”, There are many classification techniques in data mining s like Decision Tree, KNearest Neighbor, Support Vector Machines, Naive Bayesian Classifiers, and Neural Networks, and each technique has its advantage and disadvantage. Decision tree classifiers, Bayesian classifiers,

classification by backpropagation, support vector machines, these techniques are eager learners they use training tuples to construct a model.

9. In [9] H K Shifali, Dr. B. Srinivasu “Mining of Medical Data to Identify Risk Factors of Heart Disease Using Frequent Itemset”, Data mining techniques are used in medicine for various purposes. In this proposed work, a method has been implemented to identify risk factors of heart disease through frequent Itemset. These frequent itemsets are based on the heart disease dataset which is collected from the UCI repository Cleveland Clinical Foundation.

5. Methodology

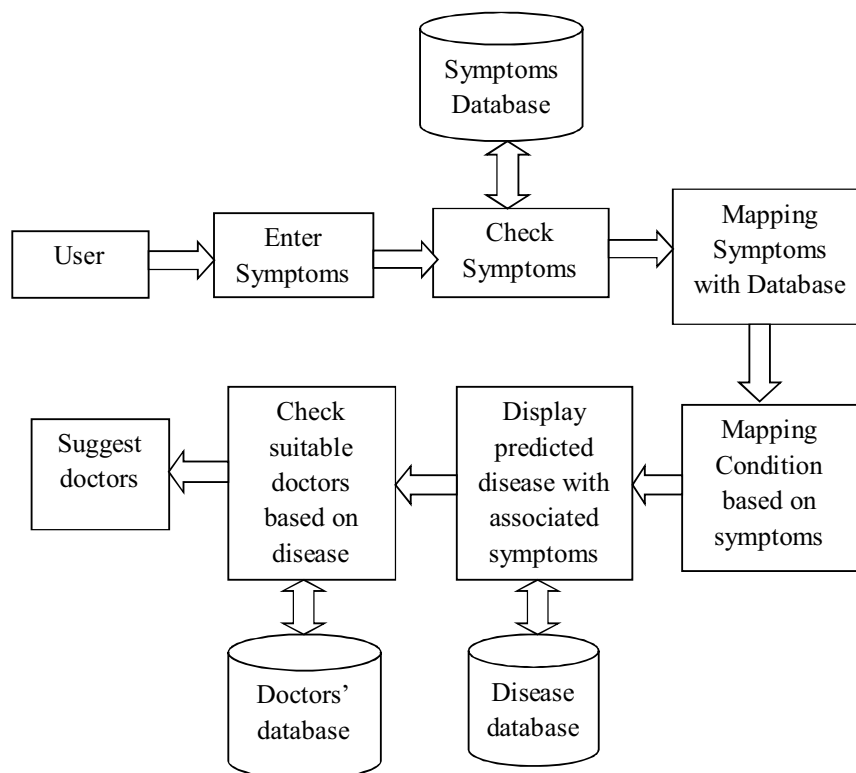
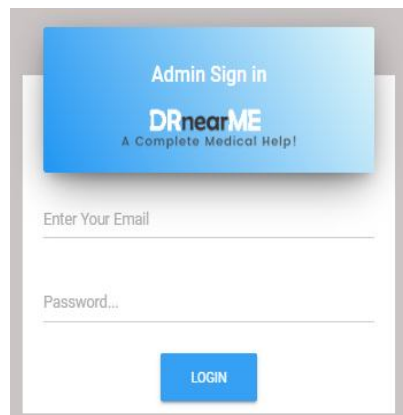


Figure 2. Working of Proposed system

In Figure2, in the healthcare system, a huge amount of data is stored some data is in the form of unstructured, noisy data. For analyzing disease the proposed system can be implemented. In this work user or patient enters symptoms as per the patient's condition. This entered symptoms compared with symptoms database based on this symptoms system mapping the current situation and checking the condition of the patient and it displays predicted disease with associated symptoms. In this proposed system we can enter several symptoms, as per the increasing number of symptoms system will show disease for that entered symptoms. After displaying the disease proposed system also check suitable doctor for a particular disease from the doctor database then the system will display the appropriate doctor for the next treatment.

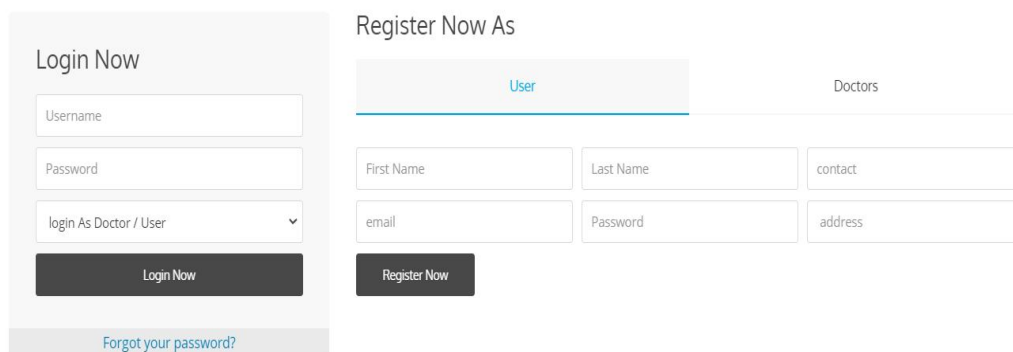
6. Results and Discussion



The image shows a web form titled "Admin Sign in" for "DRnearME - A Complete Medical Help!". It features two input fields: "Enter Your Email" and "Password...". Below these fields is a blue button labeled "LOGIN".

Figure 3. Admin Sign In

From Figure 1. Shows the Only Admin can fill the details of symptoms, diseases, and details of doctors depend on diseases. As per increasing, new diseases admin can fill symptoms, related diseases, and doctors' information depends on the disease.

The image displays two side-by-side forms. The left form is titled "Login Now" and includes fields for "Username", "Password", and a dropdown menu labeled "login As Doctor / User". It has a "Login Now" button and a link "Forgot your password?". The right form is titled "Register Now As" and has two tabs: "User" (selected) and "Doctors". The "User" tab has fields for "First Name", "Last Name", "email", and "Password". The "Doctors" tab has fields for "contact" and "address". Both tabs have a "Register Now" button.

Figure 4. Login Page

From Figure 2. Shows the user login page and Registration page for using the developed system. By using this developed system patients or users have to be registered on the registration page. Once a patient or user registered the page they will be using this developed system. The patient can be login to the page by using a username and password.

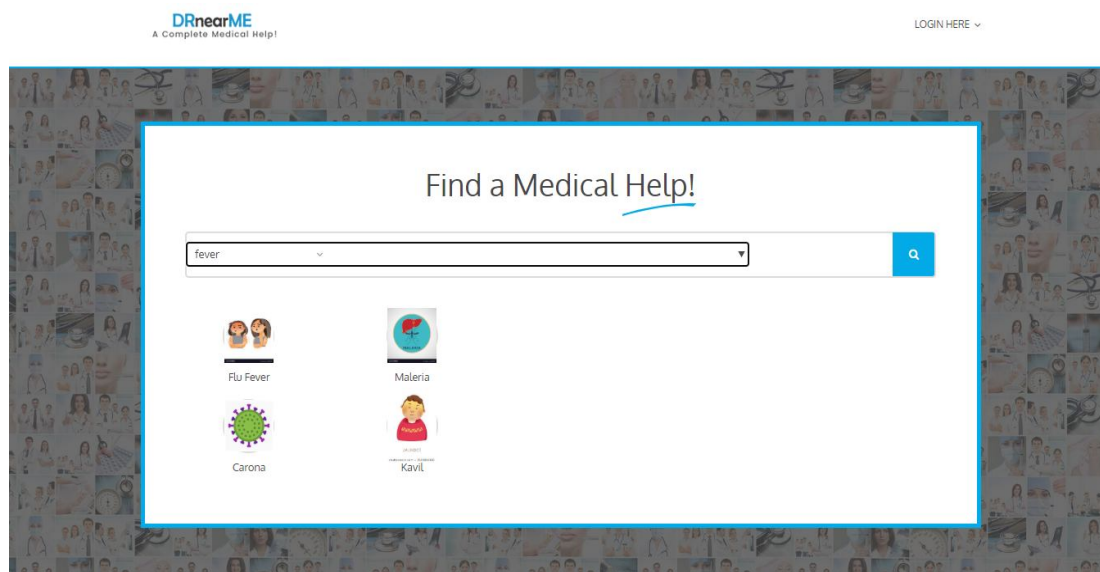


Figure 5. Add symptoms

From Figure 3. It shows the display for adding symptoms when the user adds symptom fever it displayed disease related to fever-like Flu fever, Malaria, corona, kavit. It will display disease by adding the next symptoms with a comma this disease can be changed as per additional symptoms and clicking on a particular disease it will show the doctor.

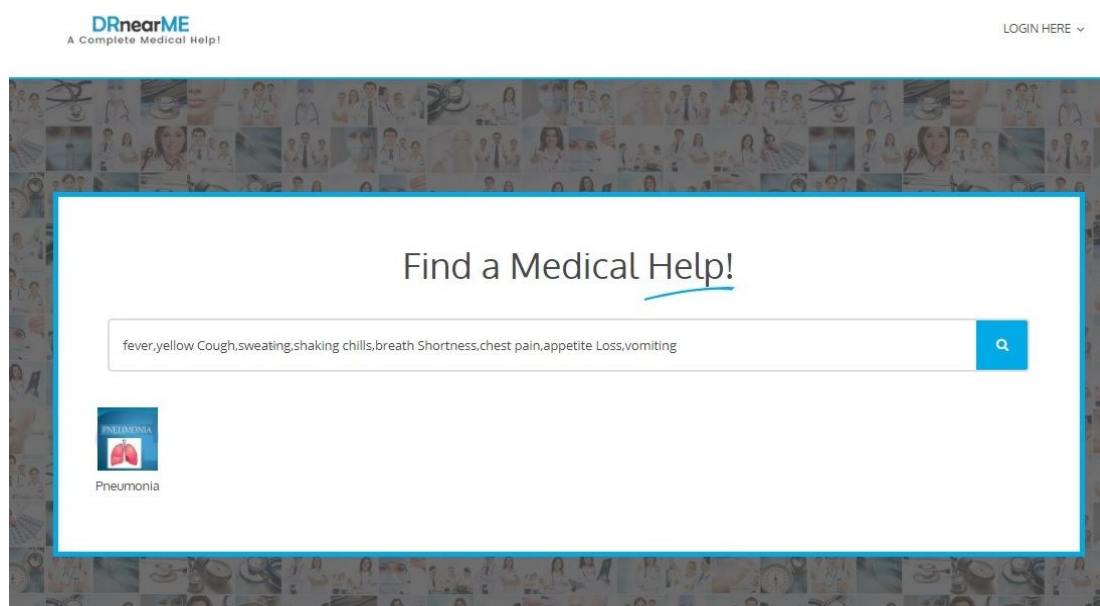


Figure 6. Symptoms of Pneumonia

From Figure 4. These symptoms are basic of pneumonia, by adding the symptoms of the patient system shows Pneumonia disease based on those symptoms. For knowing the particular doctor depend on the disease click on the disease icon which is displayed.

Specialist Doctors

	<p>Dr. Shailesh Doshi BHMS, PGNAHI - Pneumonia ☎ 9823230010 / ✉ shailesh@gmail.com 🏆 17 years of Experience Address: Tembhorni Road, Landmark: Opposite Bushan Lodge, Solapur</p>		<p>Dr. Dr. Sachin Doshi General physician doctors - Pneumonia ☎ 9867453432 / ✉ doshi.sachin@gmail.com 🏆 7 of Experience Address: Civil Chowk Solapur</p>	
				<p>Dr. Dr. Nitin Toshniwal Cardiologists - Pneumonia ☎ 9856070897 / ✉ toshniwal.nitin@gmail.com 🏆 9 of Experience Address: Balivas chowk solapur</p>

Figure 7. Doctors for Pneumonia

From Figure 5. After knowing the disease of the patient and clicking on the disease icon doctor's name is displayed which depends on the particular diseases. Doctors' names are displayed which is used for the next treatment for the needed patient.

7. Conclusion

The proposed work, the system analyzes symptoms given by the patient and predicts diseases thus reducing patients' time and effort, the system helps the patient for knowing their health conditions without the need of a doctor and the system also recommends appropriate doctors based on the disease if the patient needs the treatment. In this developed system disease can be predicted by using a clustering prediction algorithm. If patients were not able to go to the hospital for a checkup then they use this developed system for knowing disease.

8. References

- [1] Subasish Mohapatra, Prashanta Kumar Patra 2018 ICIT *Smart Health Care System using Data Mining* (Bhubaneswar, India)
- [2] Nazim Razali, Aida Mustapha 2019 IOP publication JICETS *A Data Mining Approach to Prediction of Liver Diseases* (Johar, Malaysia)
- [3] Ria Arafiah and Fariani Hermin 2018 IOP Publication *Data mining for dengue hemorrhagic fever (DHF) prediction with naive Bayes method* (Jl. Rawamangun Muka East Jakarta Indonesia.)
- [4] Dist. et al Sarath Babu, Vivek EM, Famina 2017 IEEE KP *Heart Disease Diagnosis Using Data Mining Technique* (Malappuram Kerala)
- [5] Anne Dirkson, Suzan Verberne 2018 IEEE *Open Knowledge Discovery and Data Mining from Patient Forums* (Leiden, the Netherlands)
- [6] Cincy Raju, Philip E 2018 IEEE *A Survey on Predicting Heart Disease using Data Mining Techniques* (Kerala, India)
- [7] Irum Matloob, Shoab Ahmed Khan 3 August 2020 IEEE *Sequence Mining and Prediction-Based Healthcare Fraud Detection Methodology* (Islamabad)
- [8] Mr. Sudhir M. Gorade, Prof. Ankit Deo April 2017 (IRJET) *A Study of Some Data Mining Classification Techniques* (India)
- [9] H K Shifali, Dr. B. Srinivasu 2017 IRJET *Mining of Medical Data to Identify Risk Factors of Heart Disease Using Frequent Itemset* (Telangana - Hyderabad, India.)
- [10] Sangeeta Oswal, Gokul Shah 2017 *A Study on Data Mining Techniques on Healthcare Issues and its uses and Application on Health Sector* (Mumbai, India)

- [11] et al Byju N B, Alexander G, Devanand P, Ravindra Kumar R 2009 IEEE *A Knowledge framework to Search Similar Disease Patterns using Data Mining* (Thiruvananthapuram-695033, India.)
- [12] Shusaku Tsumoto, Haruko Iwata 2012 IEEE *Exploratory Temporal Data Mining Process in Hospital Information Systems*

