GALGOTIAS
UNIVERSITY

(Established under Galgotias University Uttar Pradesh Act No. 14 of 2011)

SMART HEALTH PREDICTION SYSTEM
USING DATA MINING

A Report for the evaluation 3 of project 2

Submitted by:

TANSHI SHARMA (1713104105/17SCSE104107)

In partial fulfilment for the award of the degree

Of

Bachelors of computer applications

IN

Computer science and engineering

SCHOOL OF COMPUTING SCIENCE AND ENGINEERING

Under The Supervision Of

MR. S. SREEJI

Assistant Professor

APRIL/ MAY -2020

DECLARATION

I **TANSHI SHARMA** hereby declare that the Project work, which is presented in the report, entitled **"SMART HEALTH PREDICTION SYSTEM USING DATA MINING"** is duly prepared by me to be submitted to the department in partial fulfilment for the award of the degree of Bachelor of Computer Applications for the academic year 2019-2020.

TANSHI SHARMA

DEPARTMENT OF BACHELOR OF COMPUTER APPLICATION

GALGOTIAS UNIVERSITY, GREATER NOIDA, 201310

BONAFIED CERTIFICATE

Certified that this project report " **SMART HEALTH PREDICTION SYSTEM**

USING DATA MINING" is the Bonafide work of **" TANSHI SHARMA**

(1713104105) " who carried out the Project work under my supervision.

SIGNATURE   **SIGNATURE**   **SIGNATURE**

Mr. S. Sreeji   **Mr. U Samson**   **Mr. Priyanka Shukla**

Certify that the above mention Project Report has been duly carried out as per the norms of
the college and the statutes of the university.

Signature: Dr. Munish Sabharwal

ACKNOWLEDGEMENT

Every Project big or small is successful largely due to the effort of a number of wonderful people who have always given their valuable advice or lent a helping hand. I sincerely appreciate the inspiration, support and guidance of all those people who have been instrumental in making this project a success.

I wish to express sense of gratitude to my guide Mr. S. Sreeji, Mr. U .Samson (Project Coordinator), Ms. Priyanka Shukla (Panel In Charge) in Bachelor of Computer Application Department, Galgotias University, Greater Noida, to give me guidance at every moment during my entire report and giving valuable suggestions. They gives me unfailing inspirations and whole hearted co-opertions in carried out my Project work. Their continuous encouragement at each work through our grateful acknowledged.

I am also very grateful to my classmates, for their huge co-operation and valuable suggestion from time to time during my entire Project work. I also extend my gratitude to all the members of Department , without theirs support at various stages this report will not be materialized.

TANSHI SHARMA

TABLE OF CONTENTS

ABSTRACT:

In this paper, we introduce the techniques and applications of data mining in the clinical and informative components of Clinic Predictions. In the field of medical care, a large amount of data is becoming available due to the availability of computers. Such an enormous amount of data cannot be processed to make health predictions at the beginning of a phase and develop treatment schedules to find out who you are. Our aim is to explore data management strategies in the field of medical and healthcare to develop informed decisions. It also offers close-up of medical information processing techniques that can improve the various components of Clinical Prediction. It is the latest powerful technology that is very exciting for the computer world. It uses existing data in several databases to redesign it to new research and results. From large data sets, extracting new patterns and information related to these mining indicators is used by machine learning and data management. Mostly the task is to get the data in automatic or automated ways. The various parameters included in the data processing include integration, forecasting, path analysis and prediction analysis.

Index terms: Data mining, clinical forecasts, machine learning, integration, forecast analysis, forecasting

FIGURE INDEX

CHAPTER-1
INTRODUCTION

1. INTRODUCTION

Sometimes we need immediate medical attention, but for some reason they are not available. In our project we propose a user-friendly system to get guidance on health issues quickly through an online health care system. In recent years, with a focus on Bayesian statistics and previous distributions the problems are largely resolved. At present, Bayesian statistics are effective in economics, social sciences and a few different fields. In the field of medicine, international students have solved some medical issues that have difficulty solving classical mathematics with Bayesian classification. The Naive Bayes is among the most common separation methods introduced by Rev. Thomas Bayes. Without further details, classification rules are generated by samples trained by them.

The Smart Health Prediction System is a revolutionary medical information tool to help and expand

(1) The administration of health services

(2) Care clinics

(3) Medical analysis

(4) Training

It is the responsibility of the computing and communications technologies to enhance health science databases in assortment, storage, availability of power effectively at the right time and place a structured program is used particularly by all the people where data privacy and respect is very important. PC assisted data recovery can help improve quality in resolving options and away from personal cracks. Think of a doctor should check the five patient records; easily deliver. If the quantity of records grows with the time problem, it is limited that the accuracy with which the

expert produces the results will not be high because the fact that an expert has obtained five records should be checked.

1.2 MODULES:

Patient Module

Patient Login: - Patient login into the system using their ID and password.

Patient Registration: -If the patient is a new user they will enter their information and will receive an ID and password that can log in to the system.

Disease Prediction: - The patient will clarify the symptoms caused by his illness. The program will ask specific questions about the illness and then predict the disease according to the patient's symptoms and the program will also identify the disease-based doctors.

Search Doctor: - A patient can search for a doctor by specifying a name, address or type. Answer: -The developer will provide the answer and this will be reported to the manager.

Doctor Module

Doctor Login: - The doctor will log into the system using his / her User ID and password.

Patient Details: The doctor may view personal information.

Note: The doctor will receive a notification of how many people have access to the program and what all the diseases are predicted by the program. In order to use the system described above, the following diagrams help us to understand the proposed structure

1.3 SCRIPTURES

1) This causes the disease to be more effectively predicted.

2) In addition, the proposed system contains many suggestions such as doctor's details and instructions.

3) There is a specialist designated for each predicted disease. The details of each doctor and their location for each disease will be provided.

4) The cost of consultation with a physician can be avoided in the first phase of the prescribed medication is shown in detail.

1.4 DETAILS

The system is not fully automatic; requires doctors for a full diagnosis.

CHAPTER-2
RELATED WORK

2.1 RELATED WORK

Decision algorithms include CART (Separation and Destruction Tree), ID3 (Iterative Dichotomized 3) and C4.5 [8]. These algorithms are unique in selecting fragments, when it determines a location from a partition, and class allocation in a non-partitioned environment

Organizational Rule: The basic function of an organization's minerals is to find binary variability sets that interact regularly in transaction data [8], while the purpose of the feature selection problem is to find the groups that best match each other with a specific target variable. Corporate governance has several algorithms such as: Apriori, CDA, DDA, interest rate etc.

A smart shirt is selected for [10]. The shirt can measure the electrocardiogram (ECG) and acceleration signals of continuous monitoring and real-time monitoring of the patient. The shirt mainly consists of sensors and fabrics that allow you to get a physical signature. The measured physical signals are transmitted to the base station and the PC server via the IEEE 802.15.4 network. Critical devices consume less power and are small enough to fit a shirt. The noise reduction associated with the ECG signal has been proposed by the filtering method used in this work.

In the future home, some devices will include life-related information about the patient (eg, heartbeat, blood pressure), while other devices inside and home will provide information about the

patient's environment (e.g., humidity, temperature, carbon monoxide level). These physiological and environmental information will be integrated to assess the patient's health status and identify external factors that may influence that condition. In some cases, device groups will have sufficient interconnection awareness to work independently based on sensor information (e.g., a carbon monoxide detector can detect levels above the safety limit, establish window openings, sound the alarm, and activate vital signs of individual sensors in the home.

We can analyze this with a variety of smart health care forecasting systems

2.2EXISTING SYSTEM:

Everyone is a patient lately, and we all need serious treatment. We accept that holistic professionals are medical professionals and that there is considerable research behind every decision they make. That may not always be the case. They will not be able to concentrate on all the information they need in each situation, and may try not to get it right away. Even if they are able to access the multidisciplinary steps expected to look at the effects of treatment on every illness they experience, they may need time and skill to analyze that data and incorporate it into the patient's treatment profile. Either way, this kind of internal and external research and quantitative testing has outperformed the doctor's practice. They need an expert to interview them, be sensitive to what they say and give them guidance on how to show signs of progress and secure their progress over time. When in doubt, desiring an answer helps in the longing for praise. A reversal of the current framework would be that patients need to visit a specialist face-to-face and do not receive appropriate treatment, as experts are not prepared to foresee an accurate diagnosis. Human error

can be filed with the help of a PC that has helped senior leadership. It is not bad when there are many measures of information to be collected. Also, the quality and accuracy of decisions will be reduced when people are put under pressure and in a bad job. Identify a specialist who needs to look at five patient records; one will experience it with diligence. In any case, if the number of records increases with the time of incarceration, it is of the utmost importance that the accuracy the expert transmits his or her results will not be as high as that achieved while having only five records to be excluded.

2.3 Hardware:

1. Processor: Minimum 2.0GHz requires.

2. Ram: 2 GB.

3. Hard Disk: 100 GB.

4. Input device: Standard Keyboard and Mouse.

5. Output device: VGA and High Resolution Monitor.

2.4 Software:

1. Operating System: Windows 7.

2. Language: Java

3. Database: MySQL 5.0 & Above.

4. Tool: JDK 1.5 & Above, Eclipse IDE.

5. Server: Tomcat

2.6. Technologies Used

Eclipse IDE: From a user perspective, Eclipse IDE (Integrated Development Environment) will be used to design a Graphical User Interface (GUI). Java:

Java will be used to connect various components of user experience to the data system. Navicat MySQL: MYSQL is used as a database on a web server. In this program, the server used is the tomcat server. Diseases, Doctor and Patient database was created with the help of Navicat MySQL. It provides a visual and powerful interface for data management, development, and maintenance
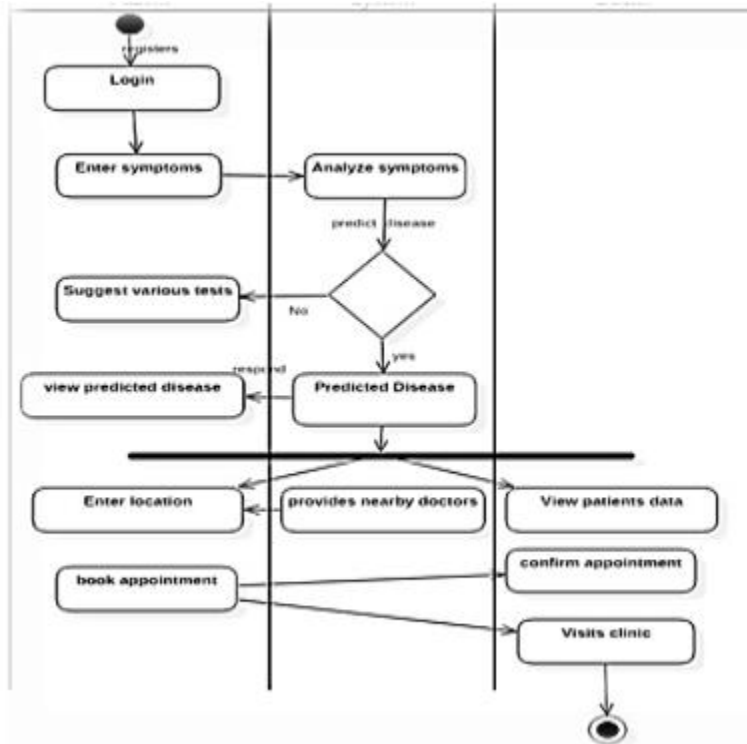
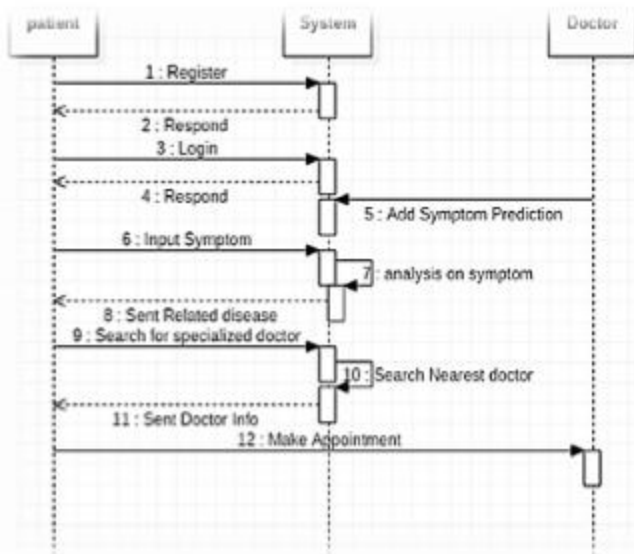# CHAPTER-3

PROPOSED MODEL
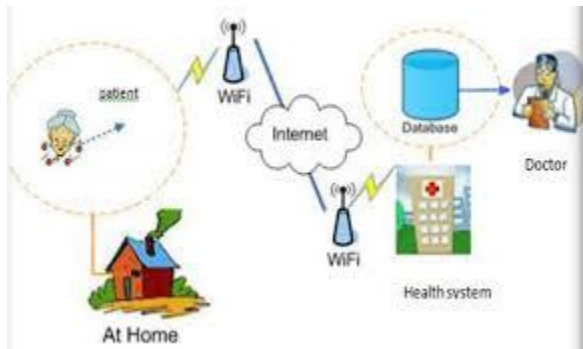
3.1 PROPOSED MODEL:

To break down the existing framework we have created a smart health forecasting system. We created a professional framework called Smart Health Prediction, which is used to improve professional work. The framework assesses the patient at a primary level and identifies potential diseases. It starts with obtaining specific information about the patient's manifestations, in which case the framework can distinguish the relevant illness, then proposing a specialist available to the patient in the closest possible location. On the off chance that the framework is not sure enough, it poses a few questions to patients, with the possibility that the framework is uncertain; at which time it will show a few tests in the patient. By looking at the overall information that is accessible, the outline will show the result. Here we use some clever mining techniques to detect potential abnormalities that may be related to the patient's appearance and based on the database of a few

patients' retraining records, the calculations (Naive Bayes) are linked to mapping out adverse effects on mysterious diseases. This framework improves professional practice and assists patients by providing valuable assistance where good planning can be done.

ACTIVITY DIAGRAM FOR HEALTH PREDICTION SYSTEM

SEQUENCE DIAGRAM FOR HEALTH PREDICTION SYSTEM

CHAPTER-4

ARCHITECTURE

4.1Data Mining Architecture:

Data mining used in the field of medicine may violate the hidden examples that exist in conventional medical data that may be left unfamiliar. The term Access to Papers, or KDD in short, refers to the broad process of accessing information learning, and proposes the use of a "standard" approach to certain data mining techniques.

4.2 Naive Bayes Algorithm:

The proposed framework uses the "Naive Bayes classifier" data mining strategy to develop an expectation framework. This framework comprises a very large number of data indicators and signals that are collectively collected from data on expert signal processing. The "Naive Bayes or Bayes" Rule is the reason for certain, AI and data mining techniques. The scale is used to make models with presidence capabilities. It is gained from "evidence" by finding connections between purpose (e.g., subordinate) and other objects.
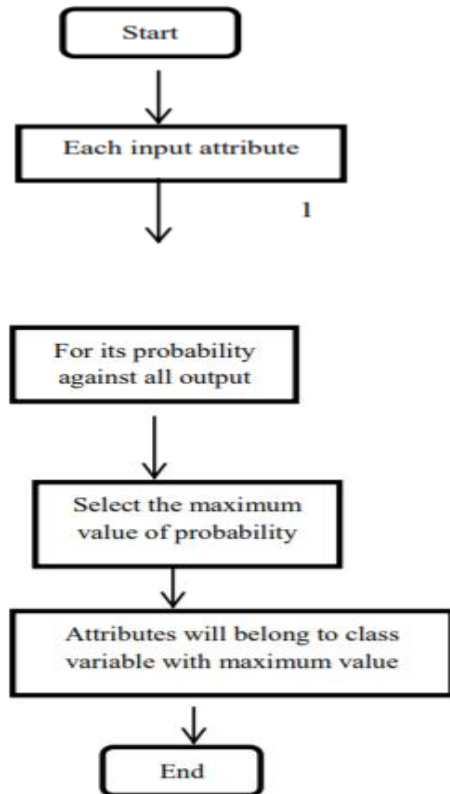
Naïve Bayes Algorithm: Following advances are actualized in Bayes calculation:-

Bayes' Theorem:

$$P(c \mid x) = P(x \mid c) P(c) / P(x)$$

Where, $P(c \mid x)$ = Posterior Probability

$P(c)$ = Prior probability

$P(x \mid c)$ = probability of predictor

$P(x)$ = Predictor's prior probability

Most cases taken in this program are adjusted by the data guides so that the number of classes that have knowledge of each condition is determined. The result was collected from the data and when diagnostic information was provided we found that the possible outcome classification results were given based on where we showed the patient to enter the class with the highest visibility. So is the command of the Naïve Bayes. By using this graphic design we can without thinking much whether the patient has been diagnosed with a particular disease. We are going to test this from a purely unbiased estimator to very large. We begin to look at all the possibilities that may be translated into the objective quality of a particular disease that contains all aspects of the disease. Subscribe the probability given that in all cases you choose p divided into two sides and one to Y and the second to N. Subsequently, in the case where the impact of P1 exposure is higher than that of P2, in which case the patient is not 'Ill

CHAPTER-5

LITERATURE SURVEY:

Tanshi Sharma introduced an update on the implementation of the Apriori Algorithm into a database that uses machine learning tool to position. Ruijuan Hu argues that conceptual information on things is usually a two-step process using the Apriori algorithms and the Rules of the Organization. This speaks to a new advanced method called the improved Apriori Algorithm for eliminating the Apriori algorithm. Gitanjali J, et.al is proposing a study of big data in various fields and finding useful information. These methods are helpful in diagnosing and providing the right solution in the same way. Krishnaiahet.al. it aims to calculate the various methods of data mining in decision-making applications and provides a detailed discussion of physicians. Data mining strategies can improve various angles of health prediction. Dan A. Simovici suggested that the rules of the association represent information in the data set as outputs and are directly related to the calculation of common item sets. Mohammed Abdul Khaleel argues that data mining as a concept that learns large amounts of data and produces patterns that can be transformed into useful information.

In this paper, we look at identifying an efficient algorithm for mining results. We can create various plans for the pharmaceutical industry to complete by using all of these analytics and data mining methods.

1. This tells you how the Naïve Byes algorithm is used to find common data objects and compare them with existing algorithms.

2. Data mining techniques are used to enter data sets with a wide range of optimizing health solutions.

3. Electronic health records and other medical history information that can prove miraculous if used for the right purpose.

4. Large amounts of complex data compiled by the health care industry include information on diseases, patients, diagnostic methods, electronic patient information.
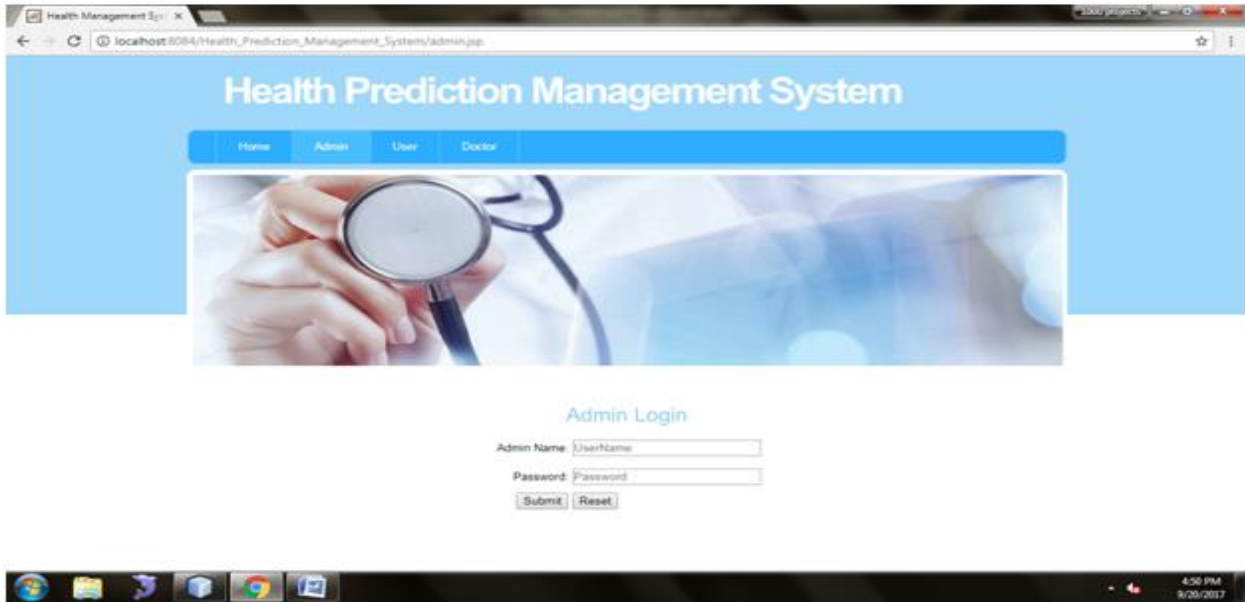
CHAPTER-6

OUTPUT:

USER REGISTRATION PAGE

ADMIN LOGIN PAGE

HOME PAGE

CHAPTER-7
FUTURE SCOPE

Hidden information will be extracted from the historical data in the proposed system, in preparation for the data details using the Apriori algorithm. Health prediction can only be made if the system responds that way. This information will be matched with questions and the final report will be produced using Association Rule Mining. Since this proposed method will work on real historical data, it will provide accurate and efficient results, which will help patients, get a diagnosis faster. More work can be done in the future using more data

Set up heart related diseases and use various data reduction techniques to improve the category. For better accuracy and prediction of heart disease the data to be used must be displayed qualitatively and with no interference from outside and missing values. This web app can also be upgraded to an Android app. This will be available to users by mobile and its use can be greatly expanded. Also, include like getting a doctor online in chat so that patients can talk directly to the doctors involved. Cancer diagnostic modules can be combined to find out how close a person is to cancer. This will make this web application predictable in a true sense.

CONCLUSION:

Data mining can help in the field of retrieval space. The protection, security and improper signing of the record are major automated problems that can be ignored and properly resolved. It presents a proposal for a crossover mining model for separating different diagnostic schedules in a clinical selection framework and also outlines the framework for the different devices used for investigations. Usually there is a situation where you need professional help immediately, but it is not available for some reason. Through our activities, we have organized another fantastic weather outline, online framework, and various patients from any area can see us. Our framework includes the basic elements, for example, silent logging, poses adverse effects on the system, and recommends medications, suggests a close professional. The app participates in multidisciplinary patient presentations, performs assessment of included side effects, and provides realistic patient expectations. Our framework enables clients to receive diagnostic tests that are offered while awaiting the diagnosis. Sometimes there is a situation where you need professional help immediately, but they are not available for some reason. Along these lines, it helps clients evaluate the side effects they give while waiting for the infection they are experiencing

REFERENCE:

1. Palaniappan S et al. Dept. for IT, heart disease prediction using data mining techniques, IEEE apr 4.

2. "Research on Dining Mining Algorithms in Disease Prediction", Kirubha V et al. JCTT, 2016

3. "Effectiveness of data mining techniques", Abdelkrim Haqiq et al. International Database Management Systems (IJDMS) International Journal, June 2016.

4. "A study on data mining techniques in the medical domain", Ridd, th e parvathi international journal of computer science and information technology, 2014.

5. Narendra To. J lambodar, distributed data classification algorithms with predictions for CKD, National journal of emerging management and technology research

6. "Machine learning algorithms are used to assess the risk of heart disease", Sitar-Taut, V.A., et al., Journal of Computer Applied Computer Science & Mathematics, 2009

7. Ghosh Joydepp, S Michael et al. Top 10 algorithms in data mining, 2007

8. G Attardi et al., Automatic section of a web page with link and content analysis, 1999.

9. "Data Mining and Data Warehousing", Ho, T. J. Prentice Hall, 2005.

10. "Developing Various Naive Naives a Smoothing Approach for Short Classization", Cor, Nadia et al. ACM, 2012