INTERNATIONAL JOURNAL OF PROGRESSIVE RESEARCH IN ENGINEERING MANAGEMENT AND SCIENCE (IJPREMS)

e-ISSN : 2583-1062

Impact Factor : 5.725

www.ijprems.com
editor@ijprems.com

Vol. 04, Issue 02, February 2024, pp : 505-508

DISEASE PREDICTION USING NAÏVE BAYES - MACHINE LEARNING ALGORITHM

Divya shree T[1]

[1]M.Sc., Department of Computer Science, Fatima College, Madurai, India.

ABSTRACT

The web application that will be developed as part of the smart healthcare system will be able to anticipate diseases based on a user's symptoms and also provideas an online health advisor for a range of illnesses. Our Smart Health Care System is an expert system designed to facilitate physicians' work. A machine performs a cursory examination of a subject and suggests potential illnessesFirst, it asks about the patient's symptoms; if the device can identify the pertinent ailment, it then suggests a physician who is close to the patient. The result will be displayed by the system based on the collected and available data. Here, we'll make advantage of some ingenious data mining methods. We employ an algorithm (Naive Bayes) to map the symptoms with prospective diseases based on a database of several patients' medical records, and we apply a number of clever data mining approaches to determine the most accurate illness that may be linked to a patient's symptoms. This system assists patients by ensuring they receive the care they require as soon as possible, in addition to making doctors' duties easier.

Keywords: Disease Prediction, Naïve Bayes, Machine Learning Algorithm, Smart Healthcare System

1. INTRODUCTION

Anyone who has contracted a disease already needs to see a doctor, which is costly and time-consuming. The user may also experience difficulties if the sickness cannot be identified and they are unable to get in touch with a physician or a hospital. Therefore, the patient's experience may be enhanced if the aforementioned treatment can be completed more quickly and affordably by using automated software. Some Heart Disease Prediction Systems analyze the patient's risk level using data mining techniques.

A web-based tool that forecasts a user's disease based on their symptoms is known as a smart healthcare system. For the smart healthcare system, data sets from numerous websites pertaining to health have been assembled. Using the symptoms this technique provides, the customer could assess the possibility of a sickness.

2. BRIEF OVERVIEW OF SMART HEALTH CARE

2.1 Objectives

The key goals of a smart healthcare system are twofold.

a.) Create a Nave Bayes Classifier that classifies the disease based on the user's feedback.

b.) To build a web interface framework for disease prediction.

2.2 Purpose and Scope

Purpose

There are numerous approaches to illness prediction. Nonetheless, research has been done on heart-related conditions, and a risk assessment has been determined. These techniques aren't typically applied to illness prediction, though. Thus, the intelligent healthcare system facilitates the prognosis of common illnesses. In some situations, you or a family member may require immediate medical attention, but doctors are not available because of unanticipated events, or we may not be able to find the right physician for the treatment. In order to solve this issue, we will make an effort to include a smart healthcare system that is online in this project. Through the use of this web-based tool, people can receive prompt counsel regarding their health issues.

Scope

The aim of this project is to develop a web platform that can predict disease occurrences based on various symptoms. The consumer will search for diseases based on their probabilistic figures by selecting different symptoms.

Limitations

The Smart Healthcare System has the following limitations:

a.) All information on this project is intended for personal educational and medical guidance only; patients should contact their physicians for accurate diagnosis and treatment.

b.) The patient's medical history has not been taken into account. c. Doesn't allow you to create data mining dimensions.

INTERNATIONAL JOURNAL OF PROGRESSIVE RESEARCH IN ENGINEERING MANAGEMENT AND SCIENCE (IJPREMS)

e-ISSN : 2583-1062

Impact Factor : 5.725

www.ijprems.com
editor@ijprems.com

Vol. 04, Issue 02, February 2024, pp : 505-508

Algorithm Used: Naive Bayes Algorithm

The Naive bayes algorithm is a classification algorithm that uses Bayesian techniques and is based on the Bayes theorem in predictive modelling. Some algorithms are more computationally intensive than this one. As a result, it can be used to quickly generate mining models to find relationships between input columns and predictable columns. The Naive Bayes algorithm is primarily used in the creation of classifiers. The categorical class labels are predicted using certain classifiers. Classifiers, on the other hand, are used to determine which class the given inputs belong to.

For the advancement of the expectation framework, the proposed smart healthcare system framework employs a data mining technique known as "Naive Bayes classifier.

This system contains a greater number of data indexes and characteristics that are genuinely collected from expert data in order to determine the exact symptom expectation. Some AI and data mining techniques are based on the "Naive Bayes or Bayes" Rule. The norm is used to create models of precognitive abilities. It benefits from the "proof" by determining the relationship between the goal (i.e., subordinate) and other variables. The Bayes theorem and Bayes approximation are used in the Naive Bayes algorithm. Bayes theorem P(c|x) = P(x|c) P(c)/P(x), where, P(c|x) is the posterior probability of class (target) given predictor. P(c) is the prior probability of class. P(x|c) is the probability of predictor given class likelihood. P(x) is the prior probability of predictor.Naive Bayes algorithm is a simple technique which is used for developing the models that are used to assigns class labels to problem instances. The class labels are drawn from finite set. Naïve Byes algorithm is not a single algorithm, but it is a family of algorithm based on a common principle.

This principle states that the value of each feature is independent of values of other features of all Naive Bayes classifiers. There are many probability models, out of which the naïve byes algorithm is efficiently trainee The Naive Bayes is a straightforward method for creating templates for assigning class labels to optimization problems. The class labels are chosen from a finite set of options.

The Nave Bayes method is a family of algorithms based on a general concept, rather than a single algorithm. The value of each feature of all naive bayes classifiers is proportional to the value of the other features, according to this theory. There are a variety of probability models, but the nave byes algorithm is one of the most effective in supervised training

3. PROPOSED SYSTEM

For a number of reasons, you can find yourself in a scenario where you require emergency medical attention but the doctor is not accessible. We developed a framework for smart health prediction to overcome the limitations of existing technologies. This technique aids users in having the ailment they are experiencing diagnosed based on an analysis of their symptoms.

Our Smart Healthcare System is an expert system designed to facilitate doctors' work. A machine performs a cursory examination of a subject and suggests possible ailments. It starts by asking the patient about their symptoms; if the device can identify the underlying ailment, it will then suggest a nearby doctor. In cases when the system lacks certainty, it will display the outcome predicated on the gathered and accessible data.

We employ a number of analytical data mining methods to determine the most precise sickness that might only be connected to the patient's condition, and a Naive Bayes algorithm is used to map the symptoms with probable illnesses based on a database of records for various disease symptoms.

Justification of Platform

a) Html is used to create effective web pages, CSS is used to create effective presentations, and java script is a clientside language used for dynamic actions.

b) PHP (hypertext pre-processor) is an open source and freely accessible server-side scripting language that is primarily used to create static and interactive websites and web applications.

c) MySQL is a free open-source relational database management system that allows us to store large amounts of data in a safe manner while still allowing us to retrieve data quickly.

d) XAMPP is a portable and simple solution that can run on a variety of platforms (windows, Linux and Mac). Xampp stands for cross platform, Apache, MySQL, PHP, and Perl, and it enables you to build web applications on your computer's local web server.

INTERNATIONAL JOURNAL OF PROGRESSIVE RESEARCH IN ENGINEERING MANAGEMENT AND SCIENCE (IJPREMS)

e-ISSN : 2583-1062

Impact Factor : 5.725

www.ijprems.com
editor@ijprems.com

Vol. 04, Issue 02, February 2024, pp : 505-508

ER Diagram

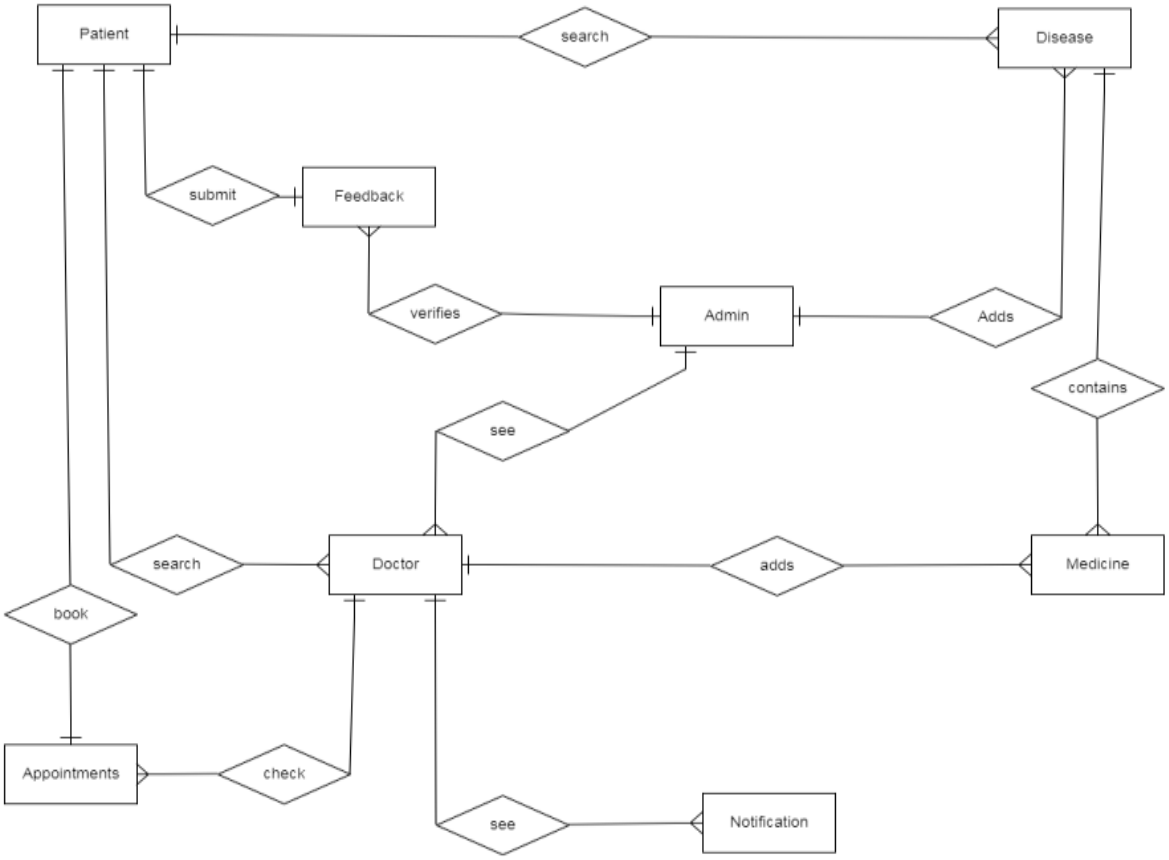

Fig. 1: Entity Relationship diagram for smart healthcare system

4. DATA FLOW DIAGRAMS

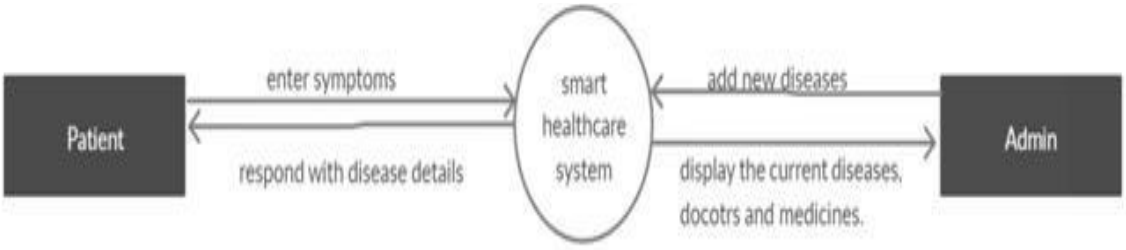

Fig. 2. Level 0 data flow for smart healthcare system.

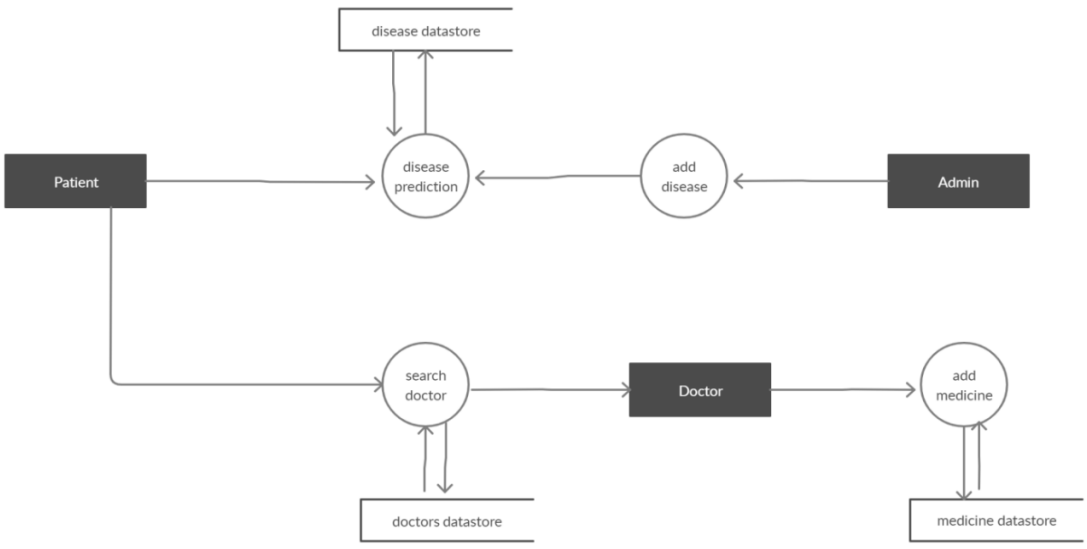

Fig. 3. Level 1 data flow for smart healthcare system.

5. CONCLUSION

The project's main goal is to develop a system that will enable users to receive prompt advice on health-related matters. The patient enters a variety of symptoms into the application, which analyzes the data and provides an accurate disease prediction.

With our technology, users can receive an analysis of the symptoms they report in order to determine which condition they are likely to have. We have developed a new online health prediction system that is accessible to different patients from anywhere.

Data mining has applications in the medical area. However, if they are not adequately addressed and rectified, privacy, security, and the Inability to log into the account are the major issues that need to be addressed.

Acknowledgement: None

Conflict of Interest: None

Source of Funding: None

6. REFERENCES

[1] Irshad Ullah, “Data Mining Algorithms And Medical Sciences”,IJCSIT, Vol 2, No 6, December 2010

[2] Amit Kumar Das, Aman Kedia,“Data mining techniques in Indian Healthcare: A Short Review”, 2015 International Conference on Man and Machine Interfacing (MAMI), 978-1-5090-0225-2/15

[3] T. Devasia, T. P. Vinushree and V. Hegde, "Prediction of students performance using Chandrasekhar Rao Jetti et.al. Disease prediction using Naïve Bayes - Machine learning algorithm International Journal of Science and Healthcare Research (www.ijshr.com) 22 Vol.6; Issue: 4; October-December 2021 Educational Data Mining", Proceedings of IEEE International Conference on Data Mining and Advanced Computing (SAPIENCE), pp. 50-100, 2016.

[4] W. Song, C. H. Li and S. C. Park, "Expert Systems with Applications Genetic algorithm for text clustering using ontology and evaluating the validity of various semantic similarity measures", Expert Syst. Appl, vol. 36, no. 5, pp. 9090-9110, 2009.

[5] Seung-Jin Lim, Yiu-Kai Ng and Xiaochun Yang, "Integrating HTML tables using semantic hierarchies and meta-data sets", International Database Engineering and Application Symposium (IDEAS), 2002.

[6] Zurkiewicz, M Milosz, “Selecting a PHP Framework for a web application projectthe method and case study”, 2013.

[7] R. Lawrence, "Integration and Virtualization of Relational SQL and NoSQL Systems Including MySQL and MongoDB", Proceedings of International Conference on Computational Science and Computational Intelligence (CSCI), pp. 250- 300, Mar. 2014.

[8] Sylviane Levi, "Quality Requirements for Multimedia Interactive", Journal of Software Engineering and Applications Sci Res, 2013.