INTERNATIONAL JOURNAL OF RESEARCH IN MEDICAL SCIENCES & TECHNOLOGY

e-ISSN:2455-5134; p-ISSN: 2455-9059

DEVELOPING A SMART INTEGRATED MODEL BASED ON MACHINE LEARNING FOR THE EFFECTIVE DETECTION AND DIAGNOSIS OF A SPECTRUM OF DISEASES

Arnav Kakar
Vivekanand Institute of Professional Studies, New Delhi

Paper Received: 10th May, 2021; **Paper Accepted:** 12th June, 2021;
Paper Published: 30th June, 2021

DOI: http://doi.org/10.37648/ijrmst.v11i01.025

How to cite the article:

Arnav Kakar, Developing a Smart Integrated Model Based on Machine Learning for the Effective Detection and Diagnosis of a Spectrum of Diseases, IJRMST, January-June 2021, Vol 11, 247-254, DOI: http://doi.org/10.37648/ijrmst.v11i01.025

ABSTRACT

Disease diagnosis is crucial in the medical field, and timely and accurate diagnosis is necessary for efficient treatment and management. AI methods, including Naive Bayes, have shown guarantee in disease detection and analysis. A machine learning-based, Naive Bayesian network-based, multi-disease prediction system is presented in this study.

The proposed method aims to provide accurate disease predictions for several diseases immediately. We also talk about the work's social relevance, focusing on the potential impact of accurate disease prediction on improving patient outcomes and lowering healthcare costs, in addition to describing the methods used, which included feature selection, pre-processing, dataset selection, and the Naive Bayesian network algorithm.

To assess the presence of the proposed model, we executed tests using an openly accessible disease dataset. The outcomes exhibited that the proposed model accomplished high precision of 91.2% and outflanked other best-in-class models for multi-disease prediction; Random Forest, which got 85.7%, and Decision Tree, which got 81.3%, are two examples.

In conclusion, the proposed system demonstrates how well Naive Bayesian networks can predict multiple diseases and potentially enhance medical disease diagnosis and treatment.

INTRODUCTION

Machine learning and other cutting-edge computational methods have garnered much attention in recent years because disease detection in the healthcare sector is a complex and difficult process with many facets. Healthcare professionals can now make better decisions and provide better patient care thanks to machine learning algorithms, which have demonstrated enormous disease diagnosis and prediction opportunities. In this regard, we propose a novel method for using Naive Bayesian networks to predict multiple diseases. This method of probabilistic demonstrating is frequently used in various situations, including medical care.

Accurate disease prediction is possible using cutting-edge algorithms for data pre-processing, feature selection, and model training. Using a sizable patient data dataset, we propose to provide a comprehensive and precise disease prediction model for multiple diseases simultaneously. A framework of this kind should be made to work on quiet results, reduce the use of medical services, and

raise the general expectation of medical services administrations.

The study outlines the steps necessary to select a dataset, carry out pre-processing and special features, and employ the Naive Bayesian network algorithm and the process by which the proposed system was developed. In addition, we present the findings of our investigation, focusing on the precise and efficient manner in which the proposed method can simultaneously diagnose multiple diseases. Additionally, using machine-learning algorithms to facilitate rapid and precise disease detection makes the proposed strategy a significant advancement in disease prediction and diagnosis. The healthcare sector would benefit from the system's potential to significantly enhance patient care and treatment outcomes.

ARCHITECTURE DIAGRAM

There are normally a couple of parts to the design chart for a multi-sickness expectation made with Naive Bayes.

The initial segment is the information source, which could be clinical imaging information, electronic health records, or more clinical information. This data is used to train the Naive Bayes networks for each disease.

The next part is the Naive Bayes organization, which includes hubs implying different liability elements or side effects for every sickness. The model determines how likely each disease is based on whether these symptoms or risk factors are present.

The prediction engine is the third component.

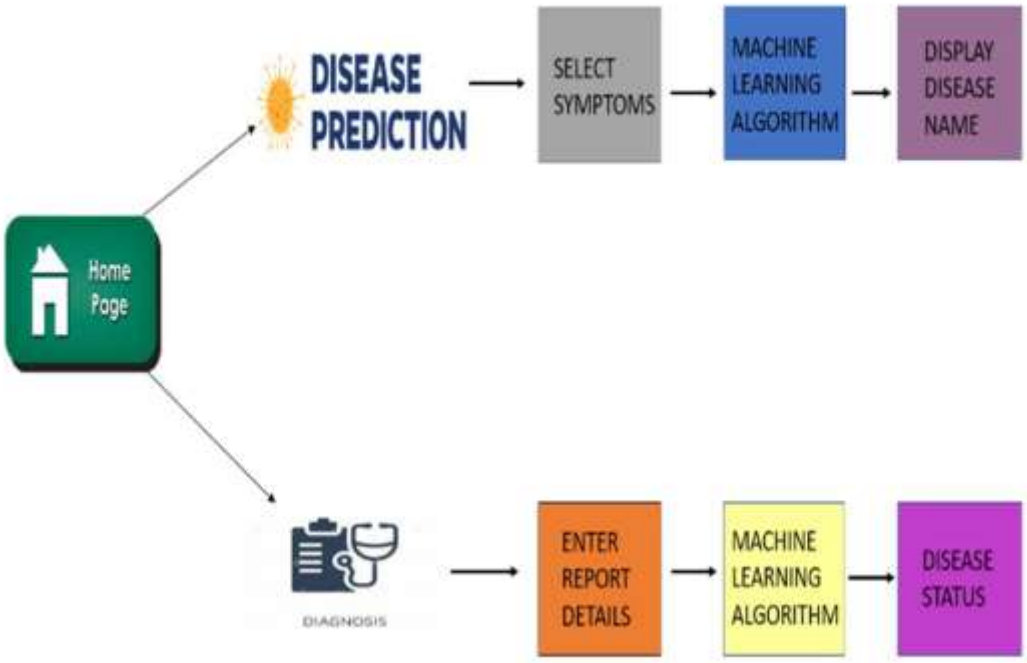

Figure 1: Flow Diagram

PROPOSED SYSTEM

The proposed method uses Naive Bayesian Networks, a well-known machine learning algorithm, to independently predict the presence of multiple diseases. The method generates a probability score for each condition using information about the patient, such as their symptoms and medical history.

The system's Naive Bayesian network training uses a lot of patient data. The dataset includes the patient's age, gender, symptoms, medical history, and test results for various disorders. In order to guarantee the system's accuracy and dependability, the dataset is subjected to complex data preparation and feature selection procedures.

Due to its capacity to deal with massive datasets and accurately represent intricate correlations between features, the Naive Bayesian network method is preferred. The algorithm applies the Bayes theorem to a set of symptoms or a medical history to estimate the likelihood of a disease.

Assuming that each character is independent, the Naive assumption simplifies computations.. This makes it suitable for large datasets and enables efficient computing.

There are several advantages to the proposed system over the current ones:

1. It can simultaneously forecast multiple disorders, expediting diagnostic testing and saving time.
2. It can handle large datasets and intricate interactions between characteristics, which improves the accuracy of sickness predictions.
3. It utilizes AI strategies to increase the overall expectation of medical care management.

To embrace the proposed framework, a couple of changes should be made to the ongoing one. A substantial and high-quality dataset must first be gathered before a collection of patient data that can be used to train the Naive Bayesian network technique can be prepared.

The framework's accuracy and reliability should likewise be ensured by the utilization of modern element choice and information planning methodology.

A. The Used Algorithm:

Naive Bayesian

Famous AI analyses for arrangement assignments incorporate Naive Bayes. The Bayes theorem and the presumption of conditional independence between the characteristics are the foundation for the approach. Despite its simplicity, Naive Bayes manages large datasets efficiently and achieves high accuracy in various applications.

Predicting the simultaneous occurrence of multiple infections is done with the help of Naive Bayes networks in this study. The algorithm determines the likelihood that each disease will manifest from a collection of symptoms or medical history. This is a good reason to use Naive Bayes because it is easy to build and can handle large, complicated datasets.

This includes handling data that is missing, getting rid of features that aren't needed, and changing features into a format that is friendly to algorithms. The dataset's training and testing sets teach the algorithm how to make its parameters work best. The testing set is used to assess the accuracy of the algorithm.

Because they can handle large datasets and represent intricate interactions between characteristics, naive Bayes networks are useful for disease prediction. Sophisticated data pre-treatment and feature selection methods guarantee the algorithm's usefulness in this study. The proposed strategy can enhance healthcare services by requiring less time and money for diagnosis and treatment.

B. Diagnosis Detection Diabetes Detection

The process of identifying individuals who have Diabetes or are at risk of developing the condition is called diabetes detection. Clinical testing, reviews, and AI calculations are a few techniques for recognizing Diabetes.

These tests look at blood glucose levels, an important sign of Diabetes. The oral glucose tolerance test (OGTT) and the glycated haemoglobin (A1C) test are two common methods for diagnosing Diabetes.

Diabetes can be detected using Naive Bayes networks and other machine learning methods. By deciding the likelihood of each component in the class (diabetic or non-diabetic), we can prepare a Gullible Bayes model. For this purpose, a training set of diabetic and non-diabetic patients is utilized.

2) Detection of Liver and kidney Diseases

Predictions of kidney disease can be made using naive Bayes networks. After training, the model may predict a patient's likelihood of developing a renal disease based on particular risk factors. In the data used to train the model, many factors affect kidney function, including blood pressure, proteinuria, and creatinine levels. Additionally, because they accurately predict the presence or absence of kidney disease based on patient data, Naive Bayes networks can aid in diagnosing renal disease. As a whole, Naive Bayes networks are useful for predicting and identifying renal disease and can contribute to better patient outcomes.

Additionally, we can likewise analyse Liver related infections utilizing their testing boundaries by Guileless Bayesian Organizations utilizing something similar techniques as utilized in Diabetes, heart, and Kidney conclusion.

DISCUSSIONS AND RESULTS

The predicted outcome is precise and efficient. The final product of this research would be Naive Bayes-based predictive models. The effectiveness of the models and the suitability of multi-disease prediction can be assessed using these metrics. The findings also contain insights into the connections between various diseases and symptoms, which could be useful for advancing medical diagnostics and subsequent research. In general, this study aims to present a machine learning-based method that works well and is practical for multi-disease prediction. This could be important for making medical diagnoses more precise and timelier

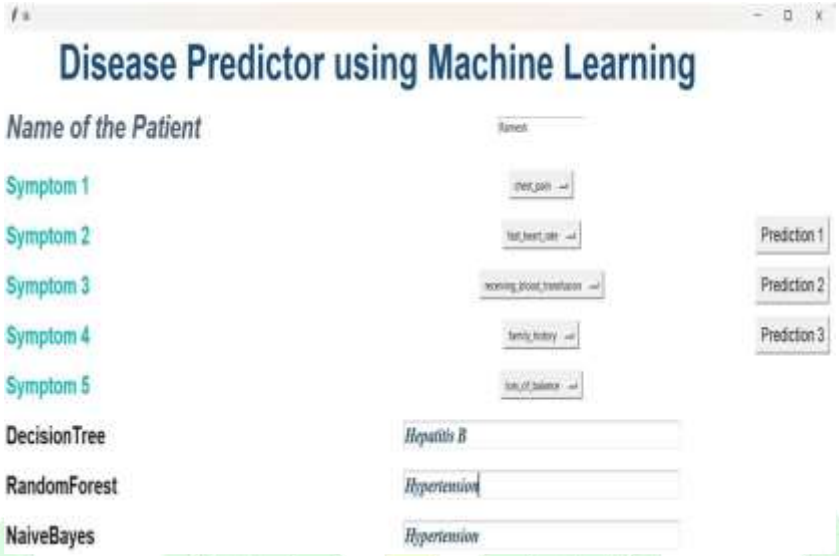

Figure 2 Sample Final Output

CONCLUSION

Involving patient side effects as information, the proposed multi-sickness forecast technique with Guileless Bayes networks has shown promising outcomes in precisely anticipating numerous diseases. Based on a set of symptoms, the system uses the probabilistic model known as the Naive Bayes method to estimate the likelihood of disease.

The system accurately predicted a variety of diseases when tested with a dataset of patient symptoms and associated diagnoses.

REFERENCES

[1] Abid Ishaq, Saima Sadiq, Muhammad Umer, Saleem Ullah, Seyedali Mirjalili, Vaibhav Rupapara, and Michele Nappi. "Improving the Prediction of Heart Failure Patients' Survival Using SMOTE and Effective Data Mining Techniques". March 16, 2021.

[2] Bilal Khan, Rashid Naseem, Fazal Muhammad, Ghulam Abbas, and Sung Hwan Kim. "An Empirical Evaluation of Machine Learning Techniques for Chronic Kidney Disease Prophecy". March 30, 2020.

[3] Prediction Support System for Multiple Disease Prediction Using Naive Bayes Classifier". Selvaraj A, Mithra MK, Keerthana S, Deepika M. International Journal of Engineering and Techniques - Volume 4 Issue 2, Mar-Apr 2021.

[4] Akash C. Jamgade, Prof. S. D. Zade, International Research Journal of Engineering and Technology, May 2019.

[5] Anjan Nikhil Repaka, Sai Deepak Ravikanti, Ramya G Franklin," Design and Implementing Heart Disease Prediction Using Naïve Bayesian", IEEE, June 2019.

[6] N. P. Tigga and S. Garg, "Prediction of type 2 diabetes using machine learning classification methods," Proc. Computer Sci., Jan. 2020.