

Disease Prediction System Using Machine Learning

Dhvale Nikita¹, Goverdhane Damini², Handge Shweta³, Chinchkhede Vijay⁴

^{1,2,3,4} Dept of Computer Engineering

^{1,2,3,4} Sandip Institute of Technology and Research Centre, Nashik

Abstract- Intelligent health care system is successful internet application of data mining conception like e-commerce application and a few different applications for industries. The medical field is info made however information extraction is weak. The undefeated application of information mining in extremely visible within the fields of e-business, selling and retail and has semiconductor diode to its applications in different industries and sectors. Among these sectors the recently rising and necessary sector is health care. There are wealth of knowledge out there with the health care systems. However, there's an absence of effective analysis to get hidden relationships and trends in data. This project intends to supply a survey of current techniques of information discovery in databases exploitation data processing techniques that are in use in today's medical analysis significantly in Disease Prediction. Variety of experiments is conducted to check the performance of predictive data processing technique on identical data set and also the experiments renders to following conclusions:

1. Decision Tree outperforms and sometime Bayesian classification has similar accuracy as of decision tree however different predictive ways like K-Means, and Classification supported cluster aren't performing well.
2. The accuracy of the decision Tree and Bayesian Classification more improves when applying genetic algorithmic rule to scale back the particular data size to urge the optimum set of attribute sufficient.

Keywords- Machine Learning, Cancer, Diabetes, Prediction

I. INTRODUCTION

Machine learning is not new concept for diseases prediction and detection. K-nearest neighbouring (KNN) and K-means have been used in Diseases prediction and detection for nearly 20 years. Today machine learning methods are being used in a wide range of applications for detecting and classifying diseases. However, these papers are concerned with using machine learning methods to identify, classify, detect, or distinguish diseases. In other words machine learning has been used primarily as an aid to diseases diagnosis and detection . Scientists applied different methods, such as screening in early stage, in order to find types of diseases before they cause symptoms. They have developed

new strategies for the early prediction of diseases. The accurate prediction of a disease is one of the most interesting and challenging tasks for the physicians. As a result, Machine learning methods have become a popular tool for medical researchers. These techniques can identify patterns and relationships between them, from complex datasets, while they are able to effectively predict future outcomes of diseases. The basic goals of diseases prediction are different from the goals of diseases detection and diagnosis. In diseases prediction one is concerned with three things 1) the prediction of diseases susceptibility (i.e. risk assessment); 2) the prediction of diseases recurrence and 3) the prediction of Diseases survivability. In the first case, one is trying to predict the type of disease . In the second case one is trying to predict the redeveloping disease . In the last case one is trying to predict an outcome. The prediction is totally dependent on the success or quality of the diagnosis. However a disease prediction can only come after a medical diagnosis. Disease prediction typically involves multiple physicians from different specialties using different subsets of biomarkers and multiple clinical factors, including the age and general health of the patient, the location and type of diseases Even for the most skilled clinician, this is not easy to do. Similar challenges also exist for both physicians and patients alike when It comes to the issues of disease prevention and disease prediction. Family history, age, diet, weight (obesity), high-risk habits (smoking, heavy drinking) all play a role in predicting an individual's disease. Unfortunately environmental and behavioural parameters generally do not provide enough information to make robust predictions very specific molecular details about the patient's own genetic make-up is needed. As the number of parameters we measure grows, so too does the challenge of trying to make sense of all this information. However, with today's high-throughput diagnostic and imaging technologies we now find ourselves overwhelmed with dozens or even hundreds of molecular, cellular and clinical parameters. In these situations, human intuition and statistics don't generally work. Instead of that we must increasingly rely on non-traditional, intensively computational approaches such as machine learning. The use of computers and machine learning in disease prediction is part of a growing trend towards personalized, predictive medicine .This movement towards predictive medicine is important, not only for patients but also for physicians as well as health economists and policy planners (in implementing large scale cancer prevention or

cancer treatment policies).The growing importance of predictive medicine and the growing reliance on machine learning to make predictions, we believed it would be of interest to conduct a detailed review of published studies employing machine learning methods in diseases prediction. The intent is to identify key trends with respect to the types of machine learning methods being used, the types of training data being integrated, the kinds of endpoint predictions being made, the types of diseases being studied and the overall performance of these methods in predicting patient outcomes .In comparing and evaluating the existing studies a number of general trends were noted and the number of common problems were detected. Some of the more obvious trends include a rapidly growing use of machine learning methods in diseases prediction Nevertheless, among the better designed and better validated studies it was clear that machine learning methods, relative to simple statistical methods, could substantially (15-25%) improve the accuracy of diseases susceptibility and diseases prediction. In other words, machine learning has an important role to play in diseases prediction and prognosis.

II. BACKGROUND & LITERATURE REVIEW

A. "Disease Predicting System Using Data Mining Techniques."

Description: The application of data mining in surprisingly visible fields like e-enterprise, commerce and trade has caused its application in other industries. The medical surroundings continues to be in formation rich however information weak. There may be a wealth of in-formation feasible with in the clinical structures. But, there is a loss of effective evaluation great perceive hidden relationships and tendencies in data. Disease prediction is a service that provides a solution to a large range of health care situations associated with health. Medical data mining strategies like association rule mining, category, clustering is implemented to analyse the unique types of heart based troubles. Data type is a vital hassle in data mining. Given a database incorporate collection of data, every with an unmarried class label, a classifier performs a short and clear definition for every class that maybe used to classify successive facts. Some of famous classifiers construct selection trees to generate magnificence models. The statistics type is based on MAFIA algorithms which bring about accuracy, the records is anticipated using entropy based totally cross validations and partition techniques and the effects are compared. C4.5 algorithm is used because the education set of rules to reveal rank of heart attack with the choice tree. The coronary heart disease database is clustered the usage of the k-manner clustering set of rules, if you want to put off the

information applicable to coronary heart attack from the database.

B. "Artificial Intelligence and Dietician."

Description: In this paper proposed system is a responsive website which contains the knowledge and data regarding the fitness of a person .Author sales or referred data required to develop the web-site, from gym exercise book which makes the website a unique one. This website consist the user interface which will be publicly displayed on the website i.e. the basic information regarding the fitness such as how to maintain good health by doing some workouts and by eating some food products which includes calories, proteins and carbohydrates.

C. "Predictive Data Mining for Medical Diagnosis."

In this paper authors have performed numerous experiment to evaluate the performance of predictive data mining method on the data set and the final results reveals that decision Tree outperforms and a while Bayesian classification is having comparable accuracy as of selection tree but different predictive strategies like KNN, Neural Networks, type primarily based on clustering are not performing properly. The second end is that the accuracy of the decision Tree and Bayesian category similarly improves after applying genetic algorithm to reduce the real data length to get the gold standard subset of attribute sufficient for disease prediction.

III. DESIGN ISSUES

Mathematical Model:

System Description:

Input: I = I1, I2, I3, I4, I5, I6

Where,

I1: User Information.
I2: Doctor Information.
I3: Admin information.
I4: Symptoms
I5: User Area
I6: Test Details.
{Process (P):

Where,

P = P1, P2
P1: Registration and Sign in

P2: Disease Prediction

{Rules (R):

Where,

R = R1, R2

R1: Web server should be always available.

R2: Browser should support the features of web app.

Output:

O = O1, O2, O3, O4

Where,

O1: Disease Predicted

O2: Doctors Suggested

O3: Appointment Confirmation

O4: Report

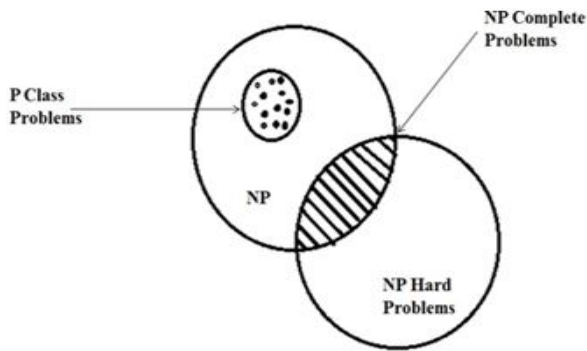


Figure 1.Venn Diagram

IV. SYSTEM ARCHITECTURE

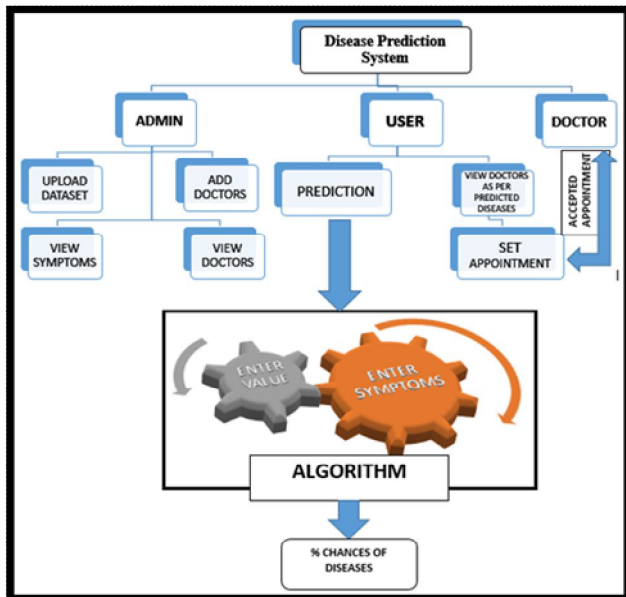
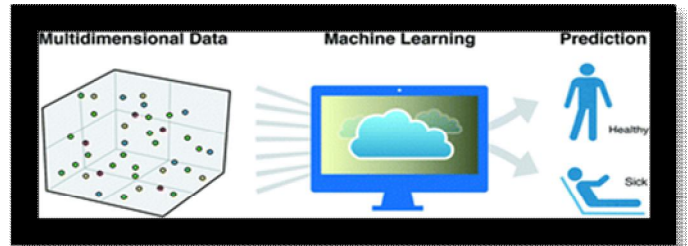


Fig. 1. System Architecture

V. RESULT & ANALYSIS

VI. IMPLEMENTATION

It is system which retrieves the data from stored data set and compares the user values with datasets. It helps in predicting diseases.



VII. CONCLUSION

Health care is moving closer to the patient, and a lots of technologies allow patients to play a greater role in their own care. In order to maximize the health benefit to the patient, these home-based systems must be smarter than former generations of health care systems because of the limited knowledge base of the typical patient. Conversely, the implementation of intelligent health care systems on line facilitate care that have not been fully explored Using this system we can predict the diseases with more accuracy. So, that people can take precautions.

REFERENCES

- [1] A. NisharaBanu, B.Gomathy , "Disease Predicting System Using Data Mining Techniques.",International Journal Of Technical Research and Applications.e-ISSN: 2320-8163, www.ijtra.com Volume 1, Issue 5 (Nov-Dec 2018), PP.41-45.
- [2] Mr.Shibakali Gupta, ArindamSarkar, IndraniPramanik, Banani Mukherjee, "Implementation Scheme for Online Medical Diag- nosis System Using Multi Agent System With JADE" ,Interna- tional Journal of Scientific and Research Publications,Volume 2, Issue 6, June 2017 1 ISSN2250-3153.
- [3] Industrial Science, "Artificial Intelligence and Dietician." ,Do- mal PratibhaArvind, GajulHiralalMurlidhar, DomalVishal Arvind, and InamdarFirdos Hamid ISSN 2347-5420 Volume-1 Issue-12, Aug -2015.
- [4] JyotiSoni,UjmaAnsari,DipeshSharma,SunitaSoni, "Predictive Data Mining for Medical Diagnosis: An Overview of Heart Disease Prediction",International Journal of Computer Applica- tions(0975 8887) Volume 17 No.8, March2017.
- [5] Hlaudi Daniel Masethe, Mosima Anna Masethe,

”Prediction of Heart Disease using Classification Algorithms”Proceedings of the World Congress on Engineering and Computer Science puter Science, 2014 VolII WCECS 2014, 22-24 October,2018.

- [6] Silvia Miksch, Kenneth Cheng, Barbara Hayes-Roth, ”An Intel- ligent Assistant for Patient HealthCare”.