ISSN 2395-1621

MEDICAL PRESCRIPTIVE ANALYSIS USING MACHINE LEARNING

[#1]Rohan Bhagat, [#2]Sushrut Deore, [#3]Sumit Deshmukh, [#4]Shreyas Dhorje,
[#5]Prof. Mily Lal

[1]rohankbhagat.rb.5@gmail.com
[2]sushrutdevre@gmail.com
[3]sumitdeshmukh9922@gmail.com
[4]djshreyas98@gmail.com
[5]milylike@gmail.com

[#12345]Department of Computer Engineering,
Dr. D.Y. Patil Institute of Engineering, Management and Research, Akurdi
Savitribai Phule Pune University, Pune, India

ABSTRACT

This is project for medical prescriptive analysis by using machine leaning technology. The basic idea is that user can pass on the health details and on processing health details, Application can predict the disease. This application also helps user to find out the Medical centre for treatment of disease. It also helps in monitoring of health of user as for many people it is not possible to carry out routine follow ups. This application suggest the list of medical practitioners which are best in market related to user's disease. In proposed system, it provides machine learning algorithms for effective prediction of various disease occurrences. This also carries all medical history of patient, so it is easy for doctor and also for user to maintain reports.

ARTICLE INFO

Article History

Received: 26th February 2020

Received in revised form :

28th February 2020

Accepted: 2th March 2020

Published online :

2th March 2020

Keywords: Medical Prescription Analysis, Classification, Machine Learning

I. INTRODUCTION

Health care system built using machine learning is developed with huge number of patient records. With enormous growth over the machine learning in recent years made it possible to manage huge collection of record sets, it is easier for organizations to get a grip on the large volumes of data being generated each day, but at the same time can also create problems related to security, data access, monitoring, high availability and business continuity. So with these huge record sets a machine can be built which can learn those data, cluster them, classify them and extract them whenever in need. So by using various machine learning algorithms we can train a System with those data. So it does something more than a normal database system which can only be referred rather than making a study over those data. Preventable medical errors (PMEs), which means errors which can be avoided by proper monitoring of the disease. A recent survey has said that around 250,000 errors are done in major countries around the world. The important fact about this survey is that these are preventable. So effective measures must be taken to avoid these errors. So a proper system which can maintain the records of patients, so that it can be an effective measure to avoid errors and to concern it in any case whether in the case of disease prediction or in drug prediction. But how a system can be trusted entirely for the prediction of the human diseases and prescribing solution.

So the entire activities of the system must be under the surveillance of an expert Doctor who have a detailed knowledge in the field of medicine.

II. MOTIVATION

Machine learning technologies and algorithms along with their applications have attracted significant attention over the past few years. An increasing number of enterprises invention machine learning analytics and try to exploit their potential in order to obtain useful insights about their performance and gain a competitive advantage.

III. LITERATURE SURVEY

[1] Meta learning algorithms for diabetes disease diagnosis has been discussed by Sen and Dash. The employed data set is Pima Indians diabetes that is received from UCI Machine Learning laboratory. WEKA is used for analysis. CART, Adaboost, Logiboost and grading learning algorithms are used to predict that patient has diabetes or not. Experimental results are compared on the behalf of correct or incorrect classification. CART offers 78.646% accuracy. The Adaboost obtains 77.864% exactness. Logiboost offers the correctness of 77.479%. Grading has correct classification rate of 66.406%.

CART offers highest accuracy of 78.646% and misclassification Rate of 21.354%, which is smaller as compared to other techniques.

[2] An experimental work to predict diabetes disease is done by the Kumari and Chitra
Machine learning technique that is used by the scientist in this experiment is SVM. RBF kernel is used in SVM for the purpose of classification. Pima Indian diabetes data set is provided by machine learning laboratory at University of California, Irvine. MATLAB 2010a are used to conduct experiment. SVM offers 78% accuracy.

[3] Ephzibah has constructed a model for diabetes diagnosis. Proposed model joins the GA and fuzzy logic. It is used for the selection of best subset of features and also for the enhancement of classification accuracy. For experiment, dataset is picked up from UCI Machine learning laboratory that has 8 attributes and 769 cases. MATLAB is used for implementation. By using genetic algorithm only three best features/attributes are selected. These three attributes are used by fuzzy logic classifier and provide 87%
Accuracy. Around 50% cost is less than the original cost.

IV. PROPOSED SYSTEM

This is project for medical prescriptive analysis by using machine leaning technology. The basic idea is that user can pass on the health details and on processing health details, application can predict the disease. This application also helps user to find out the Medical centre for treatment of disease. This application suggests the list of medical practitioners which are best in market related to user's disease. In proposed system, it provides machine learning algorithms for effective prediction of various disease occurrences. This also carries all medical history of patient, so it is easy for doctor and also for user to maintain reports. User classes and characteristics are:

[1] User/Patient: Use the application for getting the disease predicted and use the list of suggested medical practitioner by system.

[2] Doctor: It has information regarding medical practitioner and their clinics. It can also fix the patient appointments.

[3] Database: It stores all information regarding patient's data, doctor's data, reports provided by user as well as generated by system.

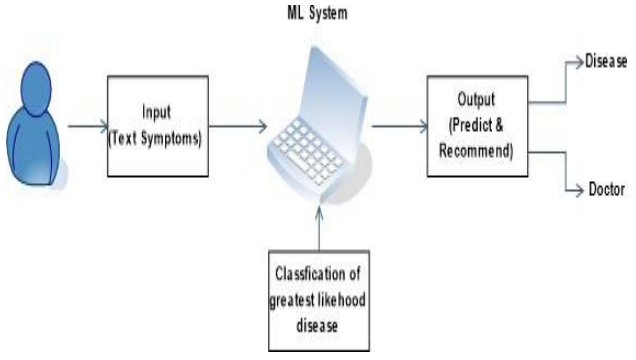

Fig. Architecture Diagram

V. METHODOLOGIES

[1] Machine Learning: Machine learning is an application of artificial intelligence (AI) that provides systems the ability to automatically learn and improve from experience without being explicitly programmed. Machine learning focuses on the development of computer programs that can access data and use it learn for themselves. The process of learning begins with observations or data, such as examples, direct experience, or instruction, in order to look for patterns in data and make better decisions in the future based on the examples that we provide. The primary aim is to allow the computers learn automatically without human intervention or assistance and adjust actions accordingly.

Machine learning algorithms are often categorized as supervised or unsupervised.

a) Supervised machine learning algorithms can apply what has been learned in the past to new data using labelled examples to predict future events. Starting from the analysis of a known training dataset, the learning algorithm produces an inferred function to make predictions about the output values. The system is able to provide targets for any new input after sufficient training. The learning algorithm can also compare its output with the correct, intended output and find errors in order to modify the model accordingly.

b) In contrast, unsupervised machine learning algorithms are used when the information used to train is neither classified nor labelled. Unsupervised learning studies how systems can infer a function to describe a hidden structure from unlabelled data. The system doesn't figure out the right output, but it explores the data and can draw inferences from datasets to describe hidden structures from unlabelled data.

c) Reinforcement machine learning algorithms is a learning method that interacts with its environment by producing actions and discovers errors or rewards. Trial and error search and delayed reward are the most relevant characteristics of reinforcement learning. This method allows machines and software agents to automatically determine the ideal behaviour within a specific context in order to maximize its performance. Simple reward feedback is required for the agent to learn which action is best; this is known as the reinforcement signal.

[2] Classification: In machine learning and statistics, classification is a supervised learning approach in which the computer program learns from the data input given to it and then uses this learning to classify new observation. This data set may simply be bi-class (like identifying whether the person is male or female or that the mail is spam or non-spam) or it may be multi-class too. Some examples of classification problems are: speech recognition, handwriting recognition, bio metric identification, document classification etc. Here we have the types of classification algorithms in Machine Learning: Linear Classifiers: Logistic Regression, Naive Bayes Classifier, Nearest Neighbour, Support Vector Machines, Decision Trees, Boosted Trees, Random Forest, Neural Networks.

[3] Naïve Bayes: It is a classification technique based on Bayes' Theorem with an assumption of independence among predictors. In simple terms, a Naive Bayes classifier assumes that the presence of a particular feature in a class is unrelated to the presence of any other feature. For example, a fruit may be considered to be an apple if it is red, round, and about 3 inches in diameter. Even if these features depend on each other or upon the existence of the other features, all of these properties independently contribute to the probability that this fruit is an apple and that is why it is known as 'Naive'. Naive Bayes model is easy to build and particularly useful for very large data sets. Along with simplicity, Naive Bayes is known to outperform even highly sophisticated classification methods. Bayes theorem provides a way of calculating posterior probability P (c|x) from P(c), P(x) and P (x|c). Look at the equation below:

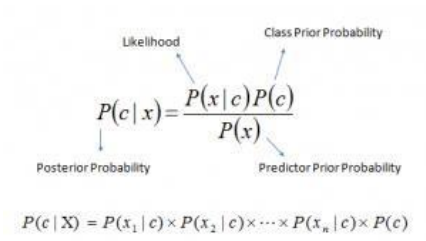

Fig. Naïve Bayes Formula

Above,

- P(c|x) is the posterior probability of class (c, target) given predictor (x, attributes).
- P(c) is the prior probability of class.
- P(x|c) is the likelihood which is the probability of predictor given class.
- P(x) is the prior probability of predictor.

[4] Support Vector Machine: "Support Vector Machine" (SVM) is a supervised machine learning algorithm which can be used for both classification and regression challenges. However, it is mostly used in classification problems. In this algorithm, we plot each data item as a point in n-dimensional space (where n is number of features you have) with the value of each feature being the value of a particular coordinate. Then, we perform classification by finding the hyper-plane that differentiate the two classes very well (look at the below snapshot).

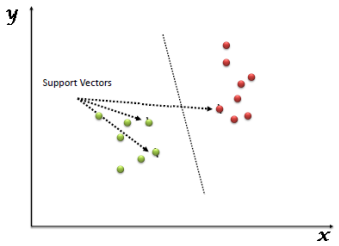

Fig. Support Vector Machine

Support Vectors are simply the co-ordinates of individual observation. Support Vector Machine is a frontier which best segregates the two classes (hyper-plane/ line).

VI. ACKNOWLEDGEMENT

I wish to express my profound thanks to all who helped us directly or indirectly in making this paper. Finally I wish to thank to all our friends and well-wishers who supported us in completing this paper successfully. I am especially grateful to our guide ***Prof. Mily Lal*** for her time to time, very much needed, and valuable guidance. Without the full support and cheerful encouragement of my guide, the paper would not have been completed on time.

VII. CONCLUSION

In this project we are using machine learning technology to predict the disease to user and also suggesting medical centre for treatment. This application enables the user to undergo frequent health checkups and focuses on user's health.

REFERENCE

[1] Centers for Disease Control and Prevention, "National Ambulatory Medical Care Survey: 2012 State and National Summary Tables," pp. 1–40, 2015.

[2] S. H. Kaplan, S. Greenfield, and J. E. Ware, "Kaplan, Assessing_the_Effects_of_Physician_Patient, Medical Care, 1989.pdf," Med. Care, vol. 27, no. 3 Suppl, pp. S110-27, 1989.

[3] M. Stewart, J. Brown, A. Donner, I. McWhinney, and J. Oates, "The Impact of Patient-Centered Care on Outcomes: The Journal of Family Practice," J. Fam. Pract., vol. 49, no. 9, pp. 796–804, 2000.

[4] D. Angus, B. Watson, A. Smith, C. Gallois, and J. Wiles, "Visualising conversation structure across time: Insights into effective doctor-patient consultations," PLoS One, vol. 7, no. 6, 2012.

[5] B. C. Wallace, M. B. Laws, K. Small, I. B. Wilson, and T. A. Trikalinos, "Automatically annotating topics in transcripts of patient-provider interactions via machine learning.," Med. Decis. Making, vol. 34, no. 4, pp. 503–12, 2014.

[6] T. Althoff, K. Clark, and J. Leskovec, "Identifying differences in physician communication styles with a log-linear transition component model," Proc. Natl. Conf. Artif. Intell., vol. 2, no. January, pp. 1314–1320, 2014.

[7] S. C. Ratzan, "Health literacy: communication for the public good.," Health Promot. Int., vol. 16, no. 2, pp. 207–214, 2001.

[8] C. Zinn, "Doctors told to use positive language in managing pain," BMJ, vol. 326, p. 301, 2003.

[9] C. J. Hutto and E. Gilbert, "Vader: A parsimonious rule-based model for sentiment analysis of social media text," Eighth Int. AAAI Conf. Weblogs …, pp. 216–225, 2014

[10] Hiransha Ma, Gopalakrishnan E.Ab , Vijay Krishna Menonab, Soman K.P- "NSE Stock Market Prediction Using Deep-Learning Models".