

Prediction of Liver Disease using Machine Learning Algorithms



Samiksha H. Zaveri, Kamini Solanki

Abstract: *There are different machine learning techniques widely used in medical field to diagnosis and to predict the liver disease. To endorse the analysis of high and multi dimensional data in health care industry we have reviewed various research papers in which we have focused on various Data mining methods for making use of data in regard to this we come out with, assessment for chosen research papers. Hence, the Objective of this study is to improve the diagnosis and Prediction of the liver disease with the machine learning algorithms. In our paper we suggested that hybrid of Decision Tree and Navie Bayes can give better result with good accuracy.*

Keywords: *Decision Tree, Liver Disease, Machine Learning Algorithms, Navie Bayes.*

I. INTRODUCTION

In a Health care industry the data has been collected from various places. We required new tools and methods to help society for extracting helpful information from big amount of data. There are different data mining methods like classification, clustering, association, and pattern matching and data visualization. This paper surveys various data mining technique development through literature review.

II. DATA MINING

The process used to find different data patterns and knowledge among the vast data is known as Data Mining. In medical field it has been very useful and helpful different data mining algorithms. There are different algorithms like Association, Classification, and Clustering. To find out a model that can allow distinct data classes is Classification algorithm. The classification techniques are Decision Tree, Support Vector Machine, K-NN, Neural Network, Navie Bayes [1]. If Liver is not working properly the whole body might affected and day by day problems may arise more and more. [4]. Jaundice consist of yellow discoloration of the eyes and skin owing more bilirubin in the mature blood cell that contains hemoglobin to carry oxygen to the bodily tissues; a biconcave disc that has no nucleus.

In the human body RBC (Red Blood Cell) carry oxygen to all the other parts and bilirubin is created. Once RBC is dying after 120 days new RBC has been created. When the old RBC breaks down, the bilirubin gets created and moves towards From the liver, it is eliminated from body into the digestive fluid duct, and keep aside in the gallbladder. Then, the bilirubin is released in small quantities as digestive fluid, into the small intestine. When the liver is not metabolizing bilirubin it is supposed to jaundice occurs. In india Hepatitis and other type of jaundice are more commonly seen in male and female as well[26].

III. RELATED WORK

We utilized two classification algorithms in the proposed methodology to predict a disease affected to liver by the examining the rendering accuracy of Navie Bayes and Decision Tree classification algorithms. Test data set used to test the model and Training data set used to train the data using K-fold cross-validation for data partitioning.

A. Decision Tree

It Contains collection of nodes that form a rooted tree, that has no incoming edges. By using Decision Tree we can make certain decision and make analysis of all non terminal nodes through which we can perform testing or specifying the branch and the follow the set of rules [2].

B. Navie Bayes

This Classifier is properly useful where little quantity of training data set to calculate approximate parameters are essential. It can solve many complex real world problems with very highly scalable model building. Zero conditional probability problems can be solved with the help of this classifier because it can consider the entire attribute in each class [2].

IV. LITERATURE REVIEW

Nazmun Nahar and Ferdous Ara [1] they look into expected time predication of disease affected to liver by using unusual decision tree techniques like Random tree, LMT, Random Forest, J48, Hoeffding Tree and Decision Stump. The dataset is used for this research consist of parameters like age, gender, total bilirubin, albumin, direct bilirubin, proteins and globulin ratio of liver disease. Their study says that Decision Stump gives the highest accuracy among other methods.

Revised Manuscript Received on July 30, 2020.

* Correspondence Author

Ms. Samiksha H. Zaveri*, MCA, Parul University, Vadodara, India.
E-mail: patel.samiksha@gmail.com

Dr. Kamini Solanki, Ph. D., MCA, Parul University, Vadodara, India.
E-mail: kamini.solanki@paruluniversity.ac.in

© The Authors. Published by Blue Eyes Intelligence Engineering and Sciences Publication (BEIESP). This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>)



A Saranya, G.Seenuvasan [2] studied many kinds of liver diseases that are carried out their methodology, the advantages and drawbacks of the studied articles and also intended to thoroughly investigate the symptoms, cause and medical reports of liver disorders. It also analyze the data of liver diseases for manipulating the liver disorder using the classification technique.

V.V. Ramalingam, A. Pandian, R. Ragavendran[3] approached various liver disease related datasets to predict result using Data mining methodologies. The calculation are perform using Machine learning are helpful to measure certain king of examination of patients reports.

M. Banu Priya P. Laura Juliet, P.R. Tamilselvi[4], implemented a PSO feature model , and the accuracy is calculated using root mean Square value, root mean error value. Using PSO features selection method J48 gives better results with 95.04% accuracy.

Joel Jacob, Joseph Chakkalakal Mathew, Johns Mathew, Elizabeth Issac[5] they have used different classification methods to identify the liver patients from individuals on the basis their results to serve in the healthcare industry. They design graphical model using python and utilized by doctors to diagnose the patient.

M. Kiran Kumar, M.Sreedevi, Y.C.A. Padmanabha Reddy[6] compared different supervised algorithms in healthcare industry and some algorithms utilized for liver disorder.

Dr. S. Vijayarani, Mr. S. Dhayaand[7] used Navie Bayes and Support Vector Machine (SVM) to predict liver disease and make comparison. On the basis of accuracy and execution time SVM gives good result.

S Pushpalatha, Dr. Jagdesh Pandya[8]used various data mining techniques to diagnosis hepatitis disease and gives the different result which were implemented.

D. Sindhuja, R. Jemina Priyadarsini[9] analyzing liver disease disorders using classification technique and made survey on it.

Shambel Kefelegn, Pooja Kamat[10]they used confusion matrices for measurement of accuracy for prediction and analyzing liver diseases disorder among various classification techniques.

Meherwar Fatima , Maruf Pasha[11] they come out with which machine learning algorithms and tools are performed.

Kemal Akyol, Yasemin Gultepe[16] found dataset was balanced by using sampling method to get more accuracy. Attribute selection is done using the Stability Selection method .They show that the combination of Stability Selection and Random Forest methods improve the performance.

Dewan Md.Farid,Nouria Harbi, and Mohammad Zahidur Rahman[17] used navie bayes classifier and decision tree to perform balance detections and to reduce missing values. They gets better result with their new algorithm with existing and prove new algorithm gets better output.

Bendi Venkata, Ramana, Prof. M.Surendra Prasad Babu, Prof. N.B. Venkateswarlu[18] checked accuracy, Precision, sensitivity and specificity on liver patient using different classifications algorithms.

Fadl Mutaheer Ba-Alwi, Houzifa M. Hintaya [19] compared Hepatitis prognostic data using various machine learning

algorithm and among them Navie bayes produce good accuracy and consume less time to build a model.

V. RESULT AND DISCUSSION

Through the literature review it is found that the accuracy of machine learning depends on the size of training and testing data set. Data mining is the promising area through which we can predict the liver disease using Hybrid model of Decision Tree and Navie Bayes. The common characteristics of healthcare datasets are highly imbalanced and missing values. The sample size of data is mostly small in size. It is found that a proper data mining method resolve all the issues is not yet developed.

VI. CONCLUSION

To Support medical services and decision in healthcare industry machine learning algorithms are used for detection and diagnosis of liver disease using different supervised learning algorithms. In this paper we had provided the summary of machine learning techniques used for medical data set. Among supervised machine learning algorithms hybrid approach of Decision Tree and navie Bayes can give better result. S.Dhamodharan[12] he focused on Navie Bayesian classifier because prediction is more effective and easily possible with it. Dinu A.J., Ganesan R, Joseph and Balaji V.[13] shows significant impacts in the detection and diagnosis in healthcare industry using machine learning.

Sumedh Sontakke, Jay Lohokare, Reshul Dani[14] they identified the patient details with the help of different methods and make improvise diagnostic of liver disease.

Ashwani Kumar, Neelam Sahu[15] achieved 79.22% accuracy using Random Forest classification technique.

REFERENCES

1. Nazmun Nahar and Ferdous Ara, Liver Disease Prediction Using Different Decision Tree Techniques, International Journal of Data Mining & Knowledge Management Process Vol.8 No.2(March 2018).
2. A Saranya, G.Seenuvasan,A Comparative Study Of Diagnosing Liver Disorder Disease Using Classification Algorithm, International Journal Of Computer Science and Mobile Computing, Vol. 6 Issue 8mpage no 49-54(August 2017).
3. V.Ramalingam, A. Pandian, R. Ragavendran, Machine Learning techniques on Liver Disease- A Survey, International Journal Of Engineering & Technology, Vol. 7,page no. 485-495(2018).
4. M. Banu Priya P. Laura Juliet, P.R. Tamilselvi,Performance Anaysis Of Liver Disease Prediction using Machine learning Algorithms, International Reserach Journal Of Engineering and Technoogy, Vol. 05 Issue:01(Jan. 2018).
5. Joel Jacob,Joseph Chakkalakal Mathew,Johns Mathew, Elizabeth Issac,Diagnosis Of Liver Disease Using Machine Learning Techniques, International Research Journal Of Engineering and Technoogy, Vol.05 Issue:04)Apr.-2018).
6. M. Kiran Kumar, M. Sreedevi, Y. C. A. Padmanabha Reddy, Survey on Machine Learning Algorithms for Liver Disease Diagnosis and Predication, International Journal Of Engineering and Technoogy, Vol. 7 page No.99-102(2018).
7. Dr. S. Vijayarani, Mr. S. Dhayanand, Liver Disease Predication Using SVM and Naïve Bayes Algorithms, International Journal Of Science, Engineering and Technology Research, Vol.4 Issue 4,(April 2015).
8. S Pushpalatha Dr. Jagdesh Pandya, Comparison Of Data Models For Hepatitis Diagnosis- Data Mining Techique, , International Journal Of Advanced Research In Computer Science And Software Engineering, Vol.4, Issue 1(Jan. 2014).



9. D. Sindhuja, R. Jemina Priyadarsini, In Survey On Classification Techniques In Data Mining For Analyzing Liver Disease Disorder, International Journal Of Computer Science And Mobile Computing, Vol. 5, Issue 5, Page No 483-488(May 2016).
10. Shambel Kefelegn, Pooja Kamat, Predication And Analysis Of Liver Disorder Diseases by Using Data Mining Technique : Survey, International Journal Pure And Applied mathematics, Vol. 9, Page No 765-770(2018).
11. Meherwar Fatima , Maruf Pasha, Survey Of Machine Learning Algorithms For Disease Diagnostic, Journal Of Intelligent Learning Systems And Applications, Vol.9, Page No 1-16(2017).
12. S.Dhamodharan, Liver Disease Predication Using bayesian Classification, An International Journal Of Computing Technology(May 2014).
13. Dinu A.J., Ganesan R, Felix Joseph and Balaji V., A Study On Deep machine Learning Algorithms For Diagnosis Of Diseases, International Journal Of Applied Engineering Research ,ISSN 0973-4562, Vol 12, Page N 6338-6346(2017).
14. Sumedh Sontakke, Jay Lohokare, Reshul Dani, Diagnosis Of Liver Disease using Machine Learning , International Conference On Emerging Trends & Innovation In ICT, Feb 3-5(2017).
15. Ashwani Kumar, Neelam Sahu, Categorization Of Liver Disease Using Classification Techniques, International Journal For Research In Applied Science & Engineering Technology, Vol. 5, Issue 5(May 2017).
16. Kemal Akyol, Yasemin Gultepe, A Study On Liver Disease Diagnosis based On Assessing the Importance Of Attributes, IJ. Intelligent systems And Applications, Vol. 11, Page No 1-9(2017).
17. Dewan Md. Farid, Nouria Harbi, and Mohammad Zahidur Rahman , Combining naïve bayes And Decision tree For Adaptive Intrusion Detection, International Journal Of Network Security & Its Application, Vol 2 (April 2010).
18. Bendi Venkata Ramana, Prof. M.Surendra Prasad Babu, Prof. N. B. Venkateswarlu, A Critical Study of Selected Classification Algorithms For Liver Disease Diagnosis, International Journal Of Database Management Systems, Vol 3, (May 2011).
19. Fadl Mutaheer Ba-Alwi, Houzifa M. Hintaya ,Comparative Study For Analysis the Prognostic in Hepatitis Data: Data Mining Approach, International Journal Of Scientific & Engineering Research , Vol 4, Issue 8 ISSN 2229-5518(2013).
20. <https://www.practo.com/health-wiki/jaundice-symptoms-complications-and-treatment/79/article> (visited on 15/03/2019).
21. R. Agrawal and G. Psaila, "Active data mining," Current, pp. 3-8(1995).
22. S. H. Liao, P. H. Chu, and P. Y. Hsiao, "Data mining techniques and applications - A decade review from 2000 to 2011," Expert Syst. Appl., vol. 39, no. 12, pp. 11303-11311(2012).

AUTHORS PROFILE



Ms. Samiksha H. Zaveri M.Phil-MCA. In Computer Science and Application. She had published 9 Research Papers in National & International Journals and 2 Paper Presented in conferences. She has published 4 books.



Dr. Kamini Solanki Ph.D. in Computer Science. She has guided 8 Ph.D students from Parul University .She had published 14 Research Papers in National & International Journals and 6 Paper Presented in conferences. She got "Academic Performance Award 2018" based on ARC (Academic Review Committee) in 2017-18 from Parul University, Vadodara. She is also been awarded as Young scientist Award 2018" from IJRULA foundation by IDAMAS LEARNING CENTER (Malaysia) She also got award of "Academic Performance Award 2018" based on ARC (Academic Review Committee) in 2019 from Parul University - Teacher Innovation Award by ZIIEL. One Research Project going on SSIP under her guidance SOIC 2020.