International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
www.ijarst.in
ISSN: 2457-0362

EVALUATION BASED APPROACHES FOR LIVER DISEASE PREDICTION USING MACHINELEARNING ALGORITHMS

T.ANILBABU[1], Dr.V.Bhaskara Murthy[2]
[1]MCA Student, B V Raju College, Kovvada, Andhra Pradesh, India.
[2]HOD & Professor, B V Raju College, Kovvada, Andhra Pradesh, India.

ABSTRACT:
Diagnosis of liver disease at a preliminary stage is important for better treatment. It is a very challenging task for medical researchers to predict the disease in the early stages owing to subtle symptoms. Often the symptoms become apparent when it is too late. To overcome this issue, this project aims to improve liver disease diagnosis using machine learning approaches. The main objective of this research is to use classification algorithms to identify the liver patients from healthy individuals. This project also aims to compare the classification algorithms based on their performance factors. To serve the medicinal community for the diagnosis of liver disease among patients, a graphical user interface will be developed using python. The GUI can be readily utilized by doctors and medical practitioners as a screening tool for the liver disease.

Keywords: GUI, CDC, SVM, liver disease.

1. INTRODUCTION:

Classification techniques are much popular in medical diagnosis and predicting diseases. Michael J Sorich reported that SVM classifier produces best predictive performance for the chemical datasets. Lung-Cheng Huang reported that Naïve Bayesian classifier produces high performance than SVM and C 4.5 for the CDC Chronic fatigue syndrome dataset. Paul R Harperreported that there is not necessary a single best classification tool but instead the best performing algorithm will depend on the features of the dataset to be analyzed. To identify the liver patients from healthy individuals. In this study, FOUR classification algorithms Logistic Regression, Support Vector Machines (SVM), K Nearest Neighbor (KNN) and artificial neural networks (ANN) have been considered for comparing their performance based on the liver patient data. Further, the model with the highest accuracy is implemented as a user-friendly Graphical User Interface (GUI) using Tkinter package in python.

PROBLEM STATEMENT

There are about 100 types of liver infections. Therefore, developing a machine that will enhance in the diagnosis of the disease will be of a great advantage in the medical field. Given a dataset containing biological and diagnostic data of Indian patients, this project aims to identify a suitable machine learning algorithm which is capable of identifying whether a person has liver disease or not. This is a binary classification problem to be solved using supervised learning. We have ten features for each data point and a label which identifies whether the patient is suffering from liver disease or not. In order to arrive at the solution, our aim should be to train various supervised learning models on this dataset so that we have a well performing model which is able to classify any new data point as a positive or negative with a reasonable degree of accuracy and perform better than the benchmarks.

PROJECT OBJECTIVE

Problems with liver patients are not easily discovered in an early stage as it will be functioning normally even when it is partially damaged. An early diagnosis of liver problems will increase patient's survival rate. Liver failures are at high rate of risk among Indians. It is expected that by 2025 India may become the World Capital for Liver Diseases. The widespread occurrence of liver infection in India

is contributed due to deskbound lifestyle, increased alcohol consumption and smoking. There are about 100 types of liver infections. Therefore, developing a machine that will enhance in the diagnosis of the disease will be of a great advantage in the medical field. These systems will help the physicians in making accurate decisions on patients and also with the help of Automatic classification tools for liver diseases (probably mobile enabled or web enabled), one can reduce the patient queue at the liver experts such as endocrinologists.

2. LITERATURE SURVEY:

Title: An intelligent model for liver disease diagnosis. Author Name: Michael J Sorich. Description: Liver disease, the most common disease in Taiwan, is not easily discovered in its initial stage; early diagnosis of this leading cause of mortality is therefore highly important.

The design of an effective diagnosis model is therefore an important issue in liver disease treatment. This study accordingly employs classification and regression tree (CART) and case-based reasoning (CBR) techniques to structure an intelligent diagnosis model aiming to provide a comprehensive analytic framework to raise the accuracy of liver disease diagnosis.

Title: A review and comparison of classification algorithms for medical decision making Author Name: Paul R. Harper Description: Within a health care setting, it is often desirable from both clinical and operational perspective to capture the uncertainty and variability amongst a patient population, for example to predict individual patient outcomes, risks or resource needs. Homogeneity brings the benefits of increased certainty in individual patient needs and resource utilization, thus providing an opportunity for both improved clinical diagnosis and more efficient planning and management of health care resources. A number of classification algorithms are considered and evaluated for their relative performances and practical usefulness on different types of health care datasets. The algorithms are evaluated using four criteria: accuracy, computational time, comprehensibility of the results and ease of use of the algorithm to relatively statistically naive medical users. The research has shown that there is not necessarily a single best classification tool, but instead the best performing algorithm will depend on the features of the dataset to be analyzed, with particular emphasis on health care data, which are discussed in the paper.

EXISTING SYSTEM:

Problems with liver patients are not easily discovered in an early stage as it will be functioning normally even when it is partially damaged. An early diagnosis of liver problems will increase patient's survival rate. Liver failures are at high rate of risk among Indians. It is expected that by 2025 India may become the World Capital for Liver Diseases. The widespread occurrence of liver infection in India is contributed due to deskbound lifestyle, increased alcohol consumption and smoking.

There are about 100 types of liver infections.

The poor performance in the training and testing of the liver disorder dataset as resulted from an insufficient data in the dataset.

PROPOSED SYSTEM:

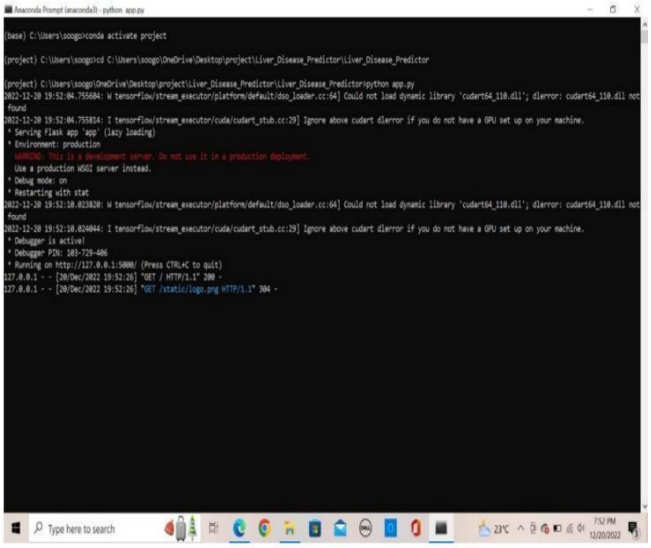
The main objective of this research is to use classification algorithms to identify the liver patients from healthy individuals. In this study, FOUR classification algorithms Logistic Regression, Support Vector Machines (SVM), K Nearest Neighbor (KNN) and artificial neural networks (ANN) have been considered for comparing their performance based on the liver patient data. The dataset used is The Indian Liver Patient Dataset (ILPD) which was selected from UCI Machine learning repository for this study. It is a sample of the entire Indian population collected from Andhra Pradesh region and comprises of 585 patient data. The diagnoses of the liver diseases to be more effective and efficient by preventing

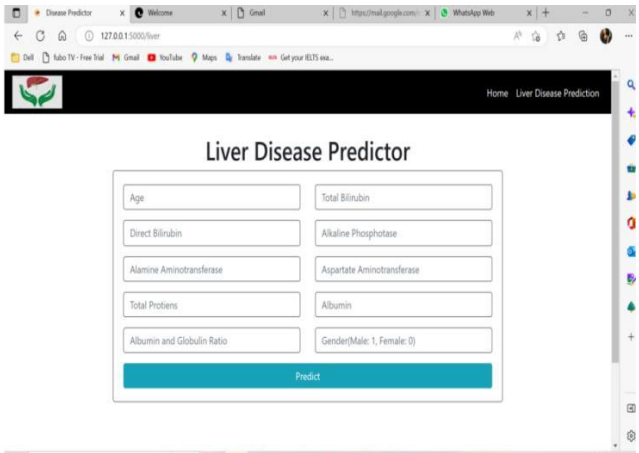
International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
www.ijarst.in
ISSN: 2457-0362

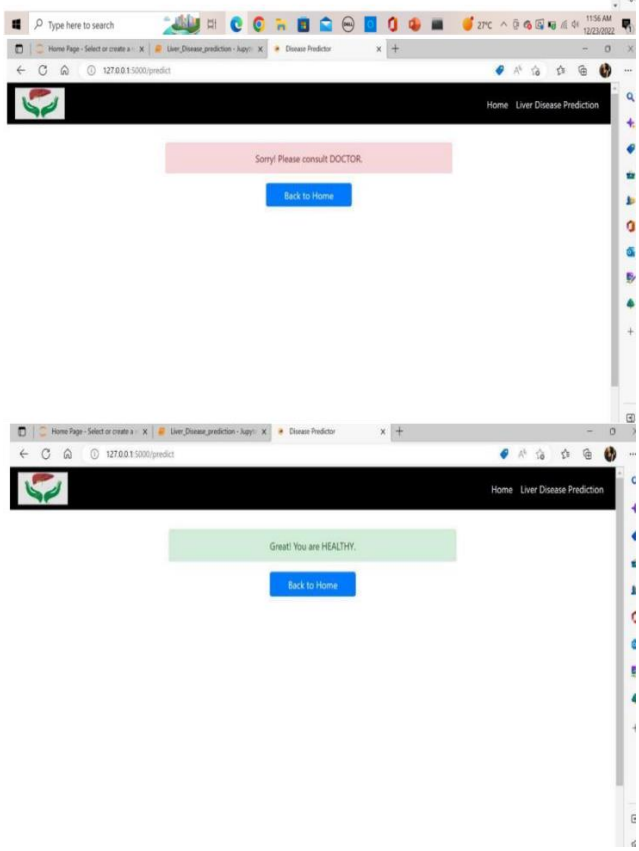
misdiagnosis of the liver disorder. Developing a system with better performance than the previous works will help in preventing misdiagnosis of the disease and help in providing the best and required medication for the patient.

3. METHODOLOGY

The model that gave the maximum accuracy for the test data was the artificial neural network. So, Artificial neural network is used for creating the GUI. The GUI is created using Tkinter package in python. Two GUIs are created, one for predicting and the other for training new data. The GUI contains input fields for all attributes in the dataset. The system will predict whether the patient has liver disease or not based on the trained model. The GUI will be a useful tool for medical staff in the early diagnosis of liver disease in patients. A picture of the developed GUI is shown below.







CONCLUSION

In this project, we have proposed methods for diagnosing liver disease in patients using machine learning techniques. The five machine learning techniques that were used include Random Forest, SVM, Logistic Regression, KNN and Artificial Neural Network. The system was implemented using all the models and their performance was evaluated. Performance evaluation was based on certain performance

metrics. Random Forest was the model that resulted in the highest accuracy with an accuracy of 84.75%. Comparing this work with the previous research works, it was discovered that Random forest proved highly efficient. A GUI, which can be used as a medical tool by hospitals and medical staff was implemented using Random Forest.

REFERANCES

[1] Alves G, Forsaa EB, Pedersen KF, Dreetz Gjerstad M, Larsen JP (2008) Epidemiology of Neurological disorder. J Neurol 255 Suppl 5: 1832.

[2] Vu TC, Nutt JG, Holford NH (2012) Progression of Motor and NonMotor Features of Neurological disorder and Their Response to Treatment. Br J Clin Pharmacol.

[3] High Accuracy Detection of Early Neurological disorder through Multimodal Features and Machine Learning. R Prashant, Sumantra Dutta Roy, Pravat K. Mandal, Shantanu Ghosh.

[4] Rustempasic, Indira, & Can, M. (2013). Diagnosis of Neurological disorder using Fuzzy C-Means Clustering and Pattern Recognition. SouthEast Europe Journal of Soft Computing, 2(1).

[5] Amit S., Ashutosh M., A. Bhattacharya, F. Revilla, (2014, March). Understanding Postural Response of Neurological disorder Subjects Using Nonlinear Dynamics and Support Vector Machines. Austin J, Biomed

[6] L. Silveira-Moriyama, J. Carvalho Mde, R. Katzenschlager, A. Petrie, R. Ranvaud, E.R. Barbosa, A.J. Lees, The use of smell identification tests in the diagnosis of Neurological disorder in Brazil, MovDisord, 23 (2008)2328-2334.

[7] Neurological disorder detection using olfactory loss and REM sleep disorder features. R. Prashanth-

[8] EMBS Member, S. Dutta Roy-IEEE Member, and P. K. Mandal, S. Ghosh Olfactory dysfunction in Parkinson disease Richard L. Doty

[9] M.M. Ponsen, D. Stoffers, J. Booij, B.L. van Eck-Smit, E.C. Wolters, H.W. Berendse, Idiopathic hyposmia as a preclinical sign of Neurological disorder, Ann Neurol, 56 (2004) 173-181.

[10] The REM Sleep Behaviour Disorder Screening Questionnaire A new Diagonostic Instrument, Karin Stiasny- Kolster, Geert Mayer, Hephata Klinik Schwalmstadt-Treysa, Germany,Sylvia Schfer,Wolfgang H Oertel.