ISSN: 2366-1313

LIVER DISEASE DETECTION USING MACHINE LEARNING AND DEEP LEARNING

[1]Mrs.D.Deepthi Sri,[2]Bairapaka Nithisha,[3]Rayabarapu Akhil Sai,[4]Tanda Rajesh Kumar Goud

[1]Assistant Professor, Dept. of CSE-Cyber Security, Geethanjali College of Engineering and Technology Cheeryal (V), Keesara (M), Medchal(D), Hyderabad, Telangana 501301,

ddeepthi.cse@gcet.edu.in

[2, 3, 4, BTech] Student, Dept. of CSE-Cyber Security, Geethanjali College of Engineering and Technology Cheeryal (V), Keesara (M), Medchal(D), Hyderabad, Telangana 501301,

nithisha.bairapaka@gmail.com,akhilsai0207@gmail.com ,tandarajeshgoud@gmail.com

ABSTRACT:

Liver is a vital organ in the body and works to filter blood from the digestive tract before passing it on to the rest of the body. Liver diseases are varied and may be assessed by liver function tests including ALT. The main objectives of this study were to use neural network analysis to predict liver disease, and to identify the relative contribution of liver disease predictors. A dataset of Indian liver patients posted on Kaggle was used to be analyzed for liver disease prediction. The dataset included 583 subjects among whom 71.4% had liver disease. Study predictors included age, gender, ALT, AST, bilirubin, albumin, total protein, albumin/globulin ratio, and alkaline phosphatase. The prediction model was effective in 79.6% predicting the liver disease. The most important predictor was ALT, and the least important predictor was alkaline phosphatase. Taken together, using neural network analysis is effective in predicting liver disease from one side and from another side, it can be improved to give more accurate results.

Keywords: Deep Learning,Machine Learning

I INTRODUCTION

The liver is a big, meaty organ found on the right side of the stomach. The liver is reddish brown in color and rubbery to the touch, weighing around 3 pounds. The liver

is divided into two main portions, known as the right and left lobes. The gallbladder, as well as sections of the pancreas and intestines, are located beneath the liver. To digest, absorb, and process food, the liver and these organs work together. The liver's primary function is to filter blood from the digestive tract before passing it on to the rest of the body. Chemicals are detoxified and medicines are metabolized by the liver. The liver does this by storing bile that eventually returns to the intestines. Proteins required for blood coagulation and other activities are also produced by the liver (medicine net, 2021). Any problem with the liver's function that produces illness is referred to as liver disease. The liver is in charge of several dangerous tasks in the body, and if it becomes diseased or injured, the loss of such functions can result in serious injury to the body. Hepatic disease is a term used to describe liver disease. Liver disease is a broad phrase that encompasses all potential issues that cause the liver to fail to perform its intended activities. Before a decline in function occurs, more than 75 percent of the liver tissue, or three quarters of the liver, must be compromised (medicine net, 2021). Reena et al (2010) suggested a data classification system based on liver disease. The training dataset is made up of 345 instances with seven different attributes that were gathered from the UCI repository. The findings of Nave Bayes algorithms in the realm of data classification are discussed in this work. When FT Tree algorithms and KStar algorithms were evaluated on liver disease datasets, the time taken to run the data for results was rapid when compared to other algorithms, with an accuracy of 97.1%. According to the findings of the experiments, the classification accuracy of the FT Tree algorithm is superior to that of other algorithms. Data mining is an important aspect of automated disease diagnosis and prediction. It includes algorithms and techniques to analyze medical data.

Liver disorders have increased excessively in the past decade, and in several countries, liver disease has become one of the most fatal diseases (Priya et al., 2018). Jeyalakshmi and Rangaraj (2021) conducted a study to offer a technique that uses deep learning to assure accurate and reliable liver disease prediction. The Modified Convolutional Neural Network based Liver Disease Prediction System.

II. LITERATURE SURVEY

1. Liver Disease Detection Using Machine Learning Techniques Publication: September 2022 The Computing and

Information Technology Research and Education New Zealand (CITRENZ) 2022 Authors: DeepikaBhupathi, Christine Nya-Ling Tan, SreenivasSremathTirumula, and Sayan Kumar Ray

Around a million deaths occur due to liver diseases globally. There are several traditional methods to diagnose liver diseases, but they are expensive. Early prediction of liver disease would benefit all individuals prone to liver diseases by providing early treatment. As technology is growing in health care, machine learning significantly affects health care for predicting conditions at early stages. This study finds how accurate machine learning is in predicting liver disease. This present study introduces the liver disease prediction (LDP) method in predicting liver disease that can be utilized by health professionals, stakeholders, students and researchers. Five algorithms, namely Support Vector Machine (SVM), Naïve Bayes, K-Nearest Neighbors (K-NN), Linear Discriminant Analysis (LDA), and Classification and Regression Trees (CART), are selected. The accuracy is compared to uncover the best classification method for predicting liver disease using R and Python. From the results, K-NN obtains the best accuracy with 91.7%, and the auto encoder network achieved 92.1% accuracy, which is above the acceptable level of accuracy and can be considered for liver disease prediction.

2.Liver Disease Prediction and Classification using Machine Learning Techniques Publication: (IJACSA) International Journal of Advanced Computer Science and Applications, Vol. 14, No. 2, 2023 Authors: SrilathaTokala, KoduruHajarathaiah, Sai Ram PraneethGunda, PathipatiNagamanohar, SatishAnamalamudi, Murali Krishna Enduri. Recently liver diseases are becoming most lethal disorder in a number of countries. The count of patients with liver disorder has been going up because of alcohol intake, breathing of harmful gases, and consumption of food which is spoiled and drugs. Liver patient data sets are being studied for the purpose of developing classification models to predict liver disorder. This data set was used to implement prediction and classification algorithms which in turn reduces the workload on doctors. In this work, we proposed apply machine learning algorithms to check the entire patient's liver disorder. Chronic liver disorder is defined as a liver disorder that lasts for at least six months. As a result, we will use the percentage of

patients who contract the disease as both positive and negative information we are processing Liver disease percentages with classifiers, and the results are displayed as a confusion matrix. We proposed several classification schemes that can effectively improve classification performance when a training data set is available. Then, using a machine learning classifier, good and bad values are classified. Thus, the outputs of the proposed classification model show accuracy in predicting the result.

3. Software-based Prediction of Liver Disease with Feature Selection and Classification Techniques. Publication: International Conference on Computtaional Intelligence and Data Science(ICCIDS 2019) Authors: JagdeepSingha, SachinBaggab, RanjodhKaurc.

Today's health care is very important aspect for every human, so there is a need to provide medical services that are easily available to everyone. In this paper, the main focus is to predict the liver disease based on a software engineering approach using classification and feature selection technique. The implementation of proposed work is done on Indian Liver Patient Dataset (ILPD) from the University of California, Irvine database. The different attributes like age, direct bilirubin, gender, total bilirubin, Alkphos, sgpt, albumin, globulin ratio and sgotetc, of the liver patient dataset, are used to predict the liver diseases risk level. The various classification algorithms such as Logistic Regression, SMO, Random Forest algorithm, Naive Bayes, J48 and k-nearest neighbor (IBk) are implemented on the Liver Patient dataset to find the accuracy. The comparison different classifier results are done of feature selection and without using feature selection technique. The development of intelligent liver disease prediction software (ILDPS) is done by using feature selection and classification prediction techniques based on software engineering mode is done by using feature selection and classification prediction techniques based on software engineering model.

III SYSTEM ANALYSIS

EXISTING SYSTEM

• In Existing system they used Decision tree, Random Forest and Naïve Bayes algorithms to predict a disease on the basis of systems and to enable synchronized and well-versed medical systems ensuring maximum patient satisfaction

ISSN: 2366-1313

• Liver Disease Prediction with Machine Learning Approaches made use of Logistic Regression and Random Forest algorithms for prediction of heart disease with proper data processing and implementation of Machine Learning algorithm with different parameters and among all Machine Learning algorithms, the highest accuracy is achieved by Logistic Regression with 87%.

• "Liver Disease Prediction By Using Machine Learning Techniques" has compared various Machine Learning models with the help of performance metrics and to detect heart related problems with highest accuracy of 89.34% by Logistic Regression.

• "Disease Prediction Using Machine Learning over Big Data" has proposed a CNN algorithm which combines structured and unstructured data and proved that CNN is more accurate than previous prediction algorithm

Disadvantages

• Existing system works for only single disease and for each type of disease use should visit multiple websites to get predict disease

• Existing methods use same algorithms for all disease predictions.

PROPOSED SYSTEM

• We are proposing such a system that will flaunt a simple, cost effective, elegant User Interface and also be time efficient.

• Our proposed system bridges the gap between doctors and patients which will help both classes of users to achieve their goal. This system is used to predict liver disease.

• In this proposed system we are going to take down liver disease dataset from the Kaggle website and evaluate them by applying algorithms such as Logistic Regression, Random Forest, Voting Classifier and CNN which will help in getting accurate prediction.

• Our system will explore and merge more datasets which includes large diversity of population to get more effective results and thus our system will improve and enhances the accuracy of the results. Along with the increased accuracy rate, we will proliferate the reliability of our system for this job and can gain the trust of patient in this system. Hence this system will contribute in easier health management with better satisfaction to the users

Advantages

ISSN: 2366-1313

• Single website is used to predict live type of diseases.

• Each disease is trained with different algorithms

• Time taken for training and prediction is less with high actuary.

IV IMPLEMENTATION

Architecture:

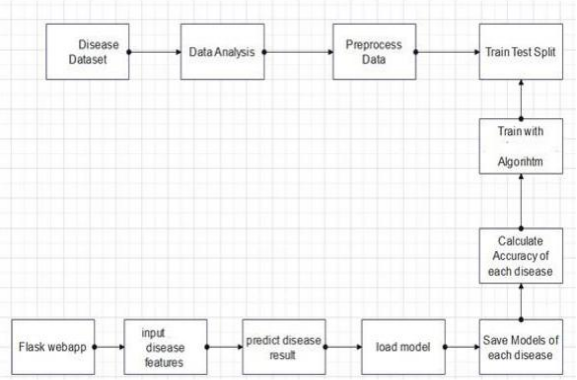


Fig-1. Architectures of the system model

A system architecture or systems architecture is the conceptual model that defines the structure, behavior, and more views of a system. An architecture description is a formal description and representation of a system. Organized in a way that supports reasoning about the structures and behaviors of the system.

MODULES

1. Data Collection

Considered liver datasets are collected form Kaggle website which are in the form of csv and image format. These datasets have features and labels based on type of disease dataset we are using features and labels are changed.

2. Understanding features of dataset.

Age of the patient, Gender of the patient, tot bilirubin, Total Bilirubin direct bilirubin, Direct Bilirubin, Alkaline Phosphotase, Alamine Aminotransferase, Aspartate Aminotransferase, Total Proteins, Albumin and Globulin Ratio ,Selector field used to split the data into two sets.

3. Pre-Processing

In this stage data analysis of each dataset is performed to check relation between features and labels with graphical representation. Null values are removed from the dataset and balanced dataset is prepared for all diseases datasets.

4. Split Data Data

Set is split in to two parts using test train split function (80 and 20) as test and train datasets. Train features are called as train x

ISSN: 2366-1313

and labels as train y. These values are used to train algorithm and test data is used to check accuracy of each disease dataset.

5. Apply ML algorithms

In this stage pre -processed dataset is taken as input of liver disease dataset and trained features and labels are given as input to fit function to train model and model is saved in to system in the form of pkl file. The model is used in web application for prediction results based on user given input.

6. Accuracy results

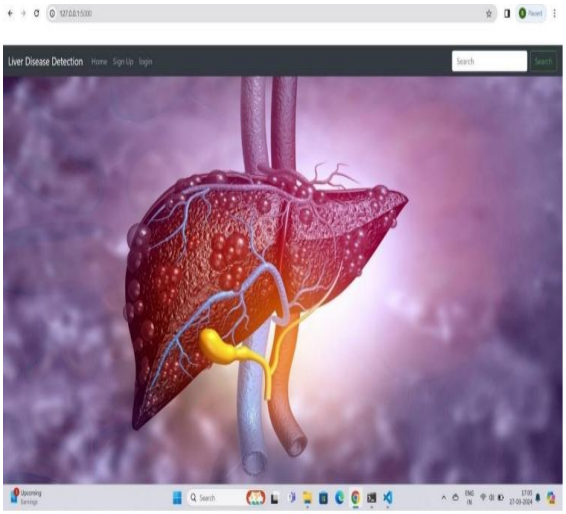
After training is done test set is given and input to algorithm to test accuracy of each dataset.

7. Flask Web framework

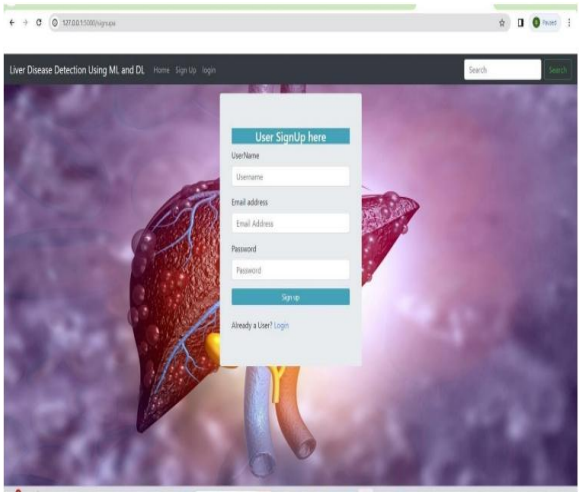
For this project web application is developed using flask framework which takes trained model as input and HTML, CSS for web page design. Using this application own input is given to webpage and disease is predicted.

V RESULT AND DISCUSSION

Home Page

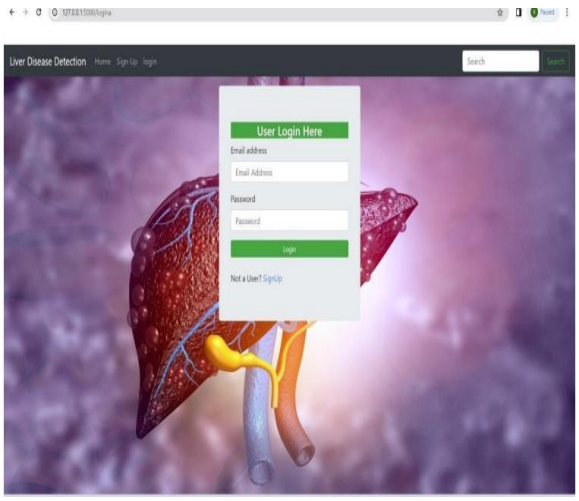


SignUp for registration

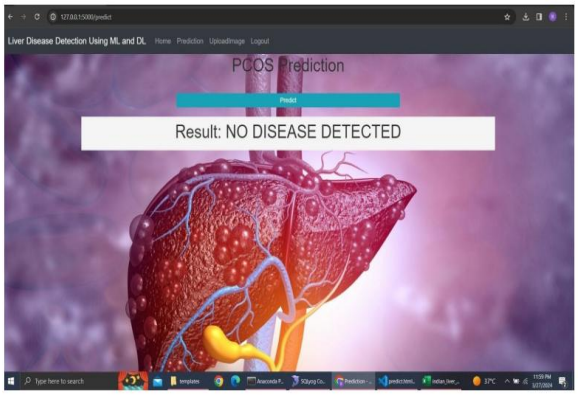


Login for prediction

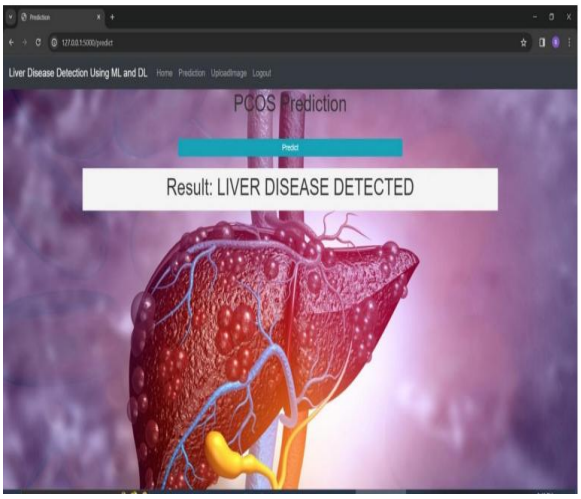
ISSN: 2366-1313



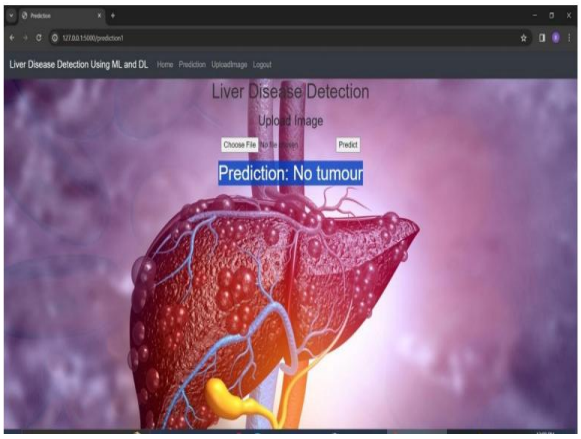
Prediction No Disease Detected



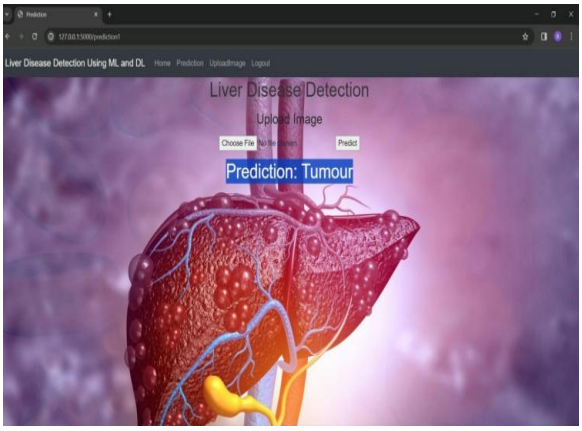
Liver Disease Detected



Detection for No Tumour



Detection for Tumour



VI CONCLUSION

There are several problems that restrict the use of deep learning methods to research the liver. First, for most liver research, there is a lack of standard datasets, and in many studies, researchers use non-public data, which makes it very difficult to evaluate deep learning methods. Second, among all the topics, only liver tumors have publicly available datasets, and these datasets have a small amount of data and have not been

ISSN: 2366-1313

updated for a long time. Third, for the same topic, there are different data modalities for research; however, for the majority of deep learning in liver research, researchers only focus on one of the data modalities, which reduces the practicability of the method. In the Conclusion, Liver disease prediction model is used to predict multiple diseases at a time. Here based on the user input disease will be predicted. The choice will be given to user. If the user want to predict particular disease or if the user don't enter any disease type then based on user entered inputs corresponding disease model will be invoked and predicted. The advantage of multi disease prediction model in advance can predict the probability of occurrence of liver disease and also can reduce mortality ratio.

FUTURE ENHANCEMENT

Attention Mechanisms:

Integrating attention mechanisms within deep learning architectures to focus on relevant regions or features within medical images or clinical records. Attention mechanisms can help in capturing important diagnostic cues while disregarding irrelevant information, potentially leading to more accurate predictions.

Real-time Monitoring and Feedback: Developing systems for real-time monitoring of patients' health parameters and providing timely feedback based on predictive models. Such systems can assist healthcare providers in early detection of liver diseases, personalized treatment planning, and monitoring disease progression.

Ethical and Regulatory Considerations:

As AI technologies are integrated into healthcare systems, ethical considerations such as patient privacy, data security, bias mitigation, and regulatory compliance (e.g., GDPR, HIPAA) need to be prioritized. Responsible AI frameworks and guidelines should be followed to ensure safe and ethical deployment of machine learning solutions in liver disease diagnosis and treatment.

Collaborative Research and Data Sharing: Collaboration among researchers, healthcare institutions, and industry stakeholders is crucial for advancing liver disease detection using machine learning. Initiatives for data sharing, benchmark datasets, and standardized evaluation metrics can facilitate reproducibility and comparison of different algorithms and approaches.

VII REFERENCES

[1] Khurana, Sarthak . , Jain, Atishay ., Kataria ,Shikhar. ,Bhasin ,Kunal . , Arora ,Sunny . ,& Gupta , Dr.Akhilesh . Das. (2019). Disease Prediction System.International Research Journal Of Engineering and Technology , 6(5) , 5178-5184.

[2] Kamboj ,Mgha. (2020).Heart Disease Prediction with Machine Learning Approaches.International Journal Of Science and Research , 9(7) , 1454-1458.

[3] Ware,Miss.Sangya . , Rakesh,Mrs.Shanu. K.,&Choudhary,Mr.Bharat . (2020). Heart Attack Prediction By Using Machine Learning Techniques. International Journal Of Recent Technology and Engineering , 8(5), 1577- 1580. Turkish Journal of Computer and Mathematics Education 4023 Research Article Vol.12 No.6 (2021), 4013-4023

[4] Shirsath ,Shraddha.Subhash .,&Patil , Prof. Shubhangi . (2018).Disease Prediction Using Machine Learning over Big Data .International Journal Of Innovative Research in Science and Technology, 7(6), 6752-6757

[5] Marimuthu , M. , Abinaya, M. ,Hariesh,K.S., Madhan,K.,&Pavithra, Kumar. V.(2018). A Review of Heart Disease Prediction Using Machine Learning and Data Analytics Approach .International Journal of Computer Application , 181(18), 20-25.

[6] Battineni ,Gopi. , Sagaro,Getu.Gamo. ,Chinatalapudi, Nalini. ,&Amenta,Francesco. (2020). Application Of Machine Learning Predictive Models in the Chronic Disease .International of PersonalisedMedicine , 10(21), 1- 11

[7] Prasadu Peddi (2015) "A review of the academic achievement of students utilising large-scale data analysis", ISSN: 2057-5688, Vol 7, Issue 1, pp: 28-35.

[8] Prasadu Peddi (2015) "A machine learning method intended to predict a student's academic achievement", ISSN: 2366-1313, Vol 1, issue 2, pp:23-37.

[7]. Parsa, A.B.; Chauhan, R.S.; Taghipour, H.; Derrible, S.; Mohammadian, A. Applying Deep Learning to Detect Traffic Accidents in Real Time Using Spatiotemporal Sequential Data. arXiv 2019, arXiv:1912.06991.

AUTHORS

Mrs. D.Deepthi Sri,Assistant Professor Dept. of CSE-Cyber Security,Geethanjali College of Engineering and Technology Cheeryal (V), Keesara (M), Medchal(D), Hyderabad, Telangana 501301.

Email: ddeepthi.cse@gcet.edu.in

Miss. Bairapaka Nithisha, Dept. of CSE-Cyber Security, Geethanjali College of Engineering and Technology Cheeryal (V), Keesara (M), Medchal(D), Hyderabad, Telangana 501301.

Email: nithisha.bairapaka@gmail.com

Mr.Rayabarapu Akhil Sai, Dept. of CSE-Cyber Security, Geethanjali College of Engineering and Technology Cheeryal (V), Keesara (M), Medchal(D), Hyderabad, Telangana 501301.

Email: akhilsai0207@gmail.com

Mr. Tanda Rajesh Kumar Goud Dept. of CSE-Cyber Security,Geethanjali College of Engineering and Technology Cheeryal (V), Keesara (M), Medchal(D), Hyderabad, Telangana 501301.

Email: tandarajeshgoud@gmail.com