Web based Approach for Diagnosing Liver Diseaseusing Deep Learning

Vallabhaneni Sri Harsha Sai¹
SCOPE
Vellore Institute of Technology Vellore, 632014, India

Talasila Bhanuteja²
SCOPE
Vellore Institute of Technology Vellore, 632014, India

Prof. Sureshkumar N³
SCOPE
Vellore Institute of TechnologyVellore, 632014, India

---***---

Abstract—Liver Diseases are becoming very common these days and can be dangerous. Liver diseases are increasing all over the world due to different factors such as excess alcohol consumption, drinking contaminated water, eating contaminated food, and exposure to polluted air. Liver disease is regarded as the disturbance in the proper functioning of the tasks related to liver .Liver is involved in many functions related to human body and if not functioned properly can effect the other parts too. Diagnosis of the disease at an earlier stage can help reduce the risk of severity. This disease can be treated with proper nutrition and health care. Machine Learning and AI have made a major impact on many domains across the world and the usage of it is rapidly increasing all over. In this paper, we analyze and predict whether the person is having liver disease or not using Machine Learning and Artificial Intelligence, there are various factors which influence the probability of having the disease .Eachand every factor is taken into consideration. We have integrated the model into a website which would help the ease of detection.

Index Terms—**Liver disease, machine learning, artificial intelli- gence**

I. INTRODUCTION

Liver diseases in India cause a huge 22.2 deaths/100,000 population attributed to cirrhosis by the Global Health Obser- vatory data from the World Health Organization. According to recent studies around 27.7 percent people diagnosed with cirrhosis die within 1 Year of getting diagnosed. As this can become such a deadly disease, diagnosing it becomes that much more important. We show in this paper how to detect Liver Disease using modern ML algortithms.

Liver is one of the vital organs of our body. It is responsible for many critical functions, without which a body cannot func-tion. Around 75 percent of the liver tissue needs to be affected before a decrease in its functioning. The critical functions of liver include production of bile, storing excess glucosein our body, producing blood clotting elements, processing iron, converting waste products of body metabolism into urea and many more. Cirrhosis is a term that describes permanent scarring of the liver. In cirrhosis, the normal liver cells are replaced by scar tissue that cannot perform any liver function. Acute liver failure may or may not be reversible, meaning that on occasion, there is a treatable cause and the liver may be able to recover and resume its normal functions.

For the training purposes of the ML algorithm to be usedwe are using the dataset provided by the UCI, ILPD (Indian Liver Patient Dataset). This dataset has 10 variables that are age, gender, total Bilirubin, direct Bilirubin, total proteins, albumin, A/G ratio, SGPT, SGOT and Alkphos. We have used multiple algorithms and found that the best and most accurate results are obtained using algorithm. There are approx2 million deaths worldwide per year, caused by liver diseases. This number causes grave concern and thus the need for its diagnosis increases further. In this study we explain the methodologies used, and the process of web app creation.//

1) Age: It is considered one of the major factors relating to the level of risk concerned to a person.As the age increases the level of the disease will increase. The gender with respect to age is also a important factor.
2) Bilirubin: Bilirubin is a yellow substance which is formed when a red blood cell breaks,this is a useless substance for body so it needs to removed ,if functioningof the liver is abnormal,the bilirubin might not move away from the body and this one of the major factor influencing liver disease.
3) Alkaline Phosphotase: The level of alkaline phosphotase plays an major role in identifying the functioning of liver ,if the level of alkaline phosphotase is more then we caninfer that there is some problem with the liver.
4) Alamine Aminotransferase,Aspartate Aminotransferase: The level of these two are directly propotional to the disfucntioning of the liver.
5) Protein: Protein is a body builder helps in reconstruction of damaged part,they help in stopping building up of fatty acids and damage to liver ,so this is also an important factor.
6) Albumin: It is the plasma protein produced by liver to fight different diseases and is very much essential.

Currently there are many medical oriented methods to find whether a person is having liver disease or not.Methods include magnetic resonance imaging,which is shortly called MRI,ultra sound and ct scan.Further people also get their liver tissue sample checked to get to know about any damage thisis called biopsy.Now a days people have become busy in their technical life and have been neglecting their health ,the above

factors in the dataset can be found by dataset and thus the liver disease can be diagnosed by simple method of machine learning techniques, at an earlier stage the disease can bedetected.

II. LITERATURE SURVEY

Applying Machine Learning in Liver Disease and Transplantation: Ashley Spann Angeline Yasodhara Justin Kang Kymberly Watt Bo Wang Anna Goldenberg Mamatha Bhat proposed ML creates prediction models using reasoning, it creates a better model by gathering a lot of information. Because of this it can be applied in hepatology on both clinical and atomic information. ML has already been widely applied in many fields in hepatology research. This paper gave a chance to study many such ML techniques that are accessible and useful in hepatology. k. Thirunavukkarasu, A.

S. Singh, M. Irfan and A. Chowdhury,in their paper gave an emphasis on new technologies stating, recent past, we have gathered a lot of medical data and it can now be used for machine learning calculations. Also liver problems have been on quite a rise and need faster treatment and detection. They have used logistic regression, KNN and SVMs to anticipate liver diseases. P.Rajeswari,G.Sophia Reena in their paper gave results of data classification obtained by using Naive Bayes, FT tree and KStar algorithms. In the whole testing, FT Tree algorithm on liver disease datasets, took less time to run the data. The result is fast and accurate when compared to other algorithms with accuracy of 97.10%. The experimental results the classification accuracy is found to be better using FT Tree algorithm.P.Thangarajul, R.Mehala have used Particle Swarm Optimization (PSO) with K star Classification to enhance the performance of the accuracy of the previous classifications. They further stated that the PSO-Kstar algorithm is really suit-able for liver disease classification. This algorithm is said to have good understandability and transformability with respect to data mining.Tapas Ranjan Baitharua, Subhendu Kumar Panib proposed focus on the feature of Medical diagnosis by learning by using the collected data of Liver disorder to create smart medical decision support systems to help the physicians. They have used multiple classification algorithms and com- pared their effectiveness in classification. After analysing theyconcluded that multilayer perceptrons had the best accuracy of71.59%. Dr.S.Vijayarani, Mr.S.Dhayanand have used multiple classification algorithms on liver diseases. The algorithms are SVM and Naive Bayes. They compared performance, accuracy and execution time. SVM had the best accuracy, whereas Naive Bayes had the better execution time.S.Dhamodharan has proposed using clinical consideration of the doctors to help anticipating the sickness. The objective was to classifythe disease into the

class type, i.e. malignancy, cirrhosis, hepatitis and no disease. Naive Bayes and FT trees are used and compared in this paper and Naive Bayes was concludedto be better. Onwodi Gregory paper focused on building arrangement models to anticipate liver determination. Many mining calculations were used and the classifiers were tested against each other based on accuracy, exactness and review.

FT tree algorithm showed 78% accuracy, 77.5% precision, 86.4% sensitivity and 38.2% specificity. Anju Gulia, Dr. Rajan Vohra, Praveen Rani have developed a hybrid model for improved accuracy. It has three core phases, first is applying classification algorithms on the dataset, second is feature selection to select core, significant attributes and then applying classification algorithms, and third phase has the comparison of the first and second phase. The results of this analysis showed that random forest is the best to use with feature selection. Ebenezer Obaloluwa Olaniyi, Khashman Aadnan have proposed the use of neural networks, specifically, back propagation and radial basis function. These networks were compared to SVM, Naive Bayes and CART on an accuracy and efficiency basis. The radial basis neural network was found to be the most optimal for diagnosis of diseases.

III. METHOD AND METHODOLODY

A. Dataset

he data-set consists of hundreds of rows with data related to the factors influencing liver disease. The dataset has been taken from the UCI machine learning repository. There are 11 columns in the dataset, the first ten being the factors that the disease depends on and the last one which determines whether there is disease or not. The dataset consists of 10 parameters. The below is the nippet of info regarding the database taken.

```
df.info()

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 583 entries, 0 to 582
Data columns (total 11 columns):
Age                         583 non-null int64
Gender                      583 non-null object
Total_Bilirubin             583 non-null float64
Direct_Bilirubin            583 non-null float64
Alkaline_Phosphotase        583 non-null int64
Alamine_Aminotransferase    583 non-null int64
Aspartate_Aminotransferase  583 non-null int64
Total_Protiens              583 non-null float64
Albumin                     583 non-null float64
Albumin_and_Globulin_Ratio  579 non-null float64
Dataset                     583 non-null int64
dtypes: float64(5), int64(5), object(1)
memory usage: 50.2+ KB
```

Fig. 1. Snippet of test dataset

B. Flask

The web based architecture which we proposed is flask
.Flask is a python web framework mostly used for deploying Machine learning models.Models not involving very large data use flask for deploying.It does not have any pre defined third party libraries and distributed database layer. In our proposed method we have used flask for deploying the deep learning model.It is used to input data from the user and this data is fed to the model then the output received is displayed to the user.

C. Workflow

There are different factors that influence the liver disease,considering many factors in the given dataset.First we take the values of each parameter required for the diagnosis,then this data is send to the model for calculating whether the person having the disease or not.After evaluation of the model the output is sent back on the screen using the Web frame work flask.

D. Neural Network

A neural network is a progression of calculations that attempts to perceive underlying relationships in a bunch of information through an interaction that imitates the manner in which the human cerebrum works. The neural network used in the proposed model is Convolutional Neural network.Cnn is widely used for deep learning problems.The data is trained is trained on many parameters so that the accuracy of the model is very high.A sample snippet of how the neural network looks like is given below.We have used the LeNet architecture for the training of the model.Usually for a CNN architecture there will be an input ,few convolutional layers ,and few fully connected layer and output,but coming to LeNet architecture it is modified on the basic template of CNN.LeNet architecture has input ,four convolutional layers ,two fully connected layer and then an output layer.
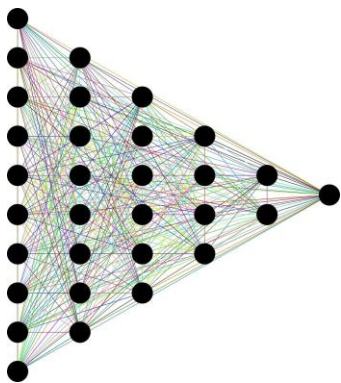
E. Workflow

The below is the workflow chart fig(3).The model is first made and then moved into a pickle file .We first pre-process the data and clean it for the further process.After pre process- ing the data is is split into train and test data and use the Cnn and advanced works to train,test and validate the data is done.The the model is pushed to the pickle file.The pickle file is used in the flask framework to evaluate the problem and estimate the results.The neural networks are used ,the architecture used is the Lenet architecture.

IV. Implementation

V. Results

The below are the images of the way the web ui looks and how the liver disease prediction works.The accuracy of
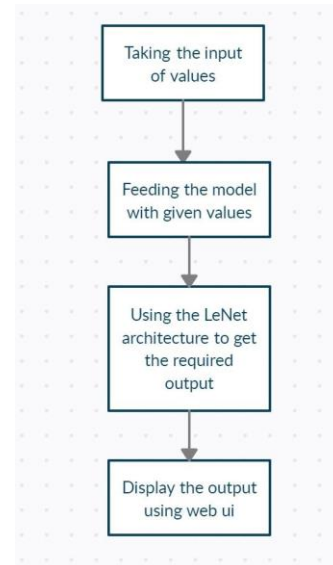

Fig. 3. Workflow

Fig. 2. How a neural network looks like

Fig. 4. Data Pre processing

Fig. 5. Data Splting

```
[13]:   history=model.fit(X_train, y_train, batch_size = 20, epochs = 50,validation_data = (X_test,y_test), verbose = 1)
```

Fig. 6. Epoches

the model trained is around 67.6.The graphs below show the accuracy and loss function.When we give the values of the features to the model it finally the result is given whether the person is effected with the liver disease or not.

The below images are the screenshots of the web hosting, it gives a clear example of how the model works

VI. CONCLUSION

In this paper we have codified an effective process for diagnosing the liver disease using deep learning giving it an web based approach. The model attained an accuracy of 67.6 percentage and this model predicts that whether the user is
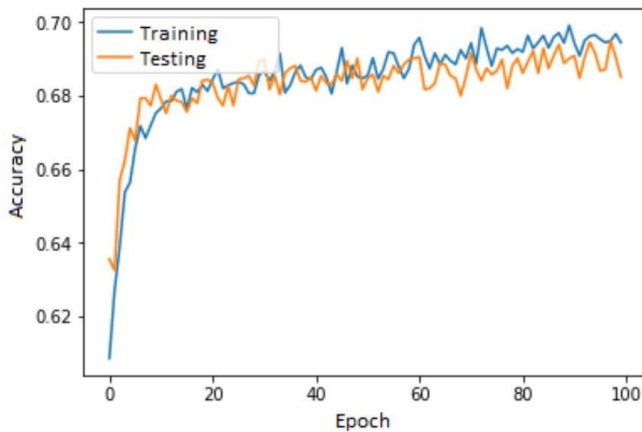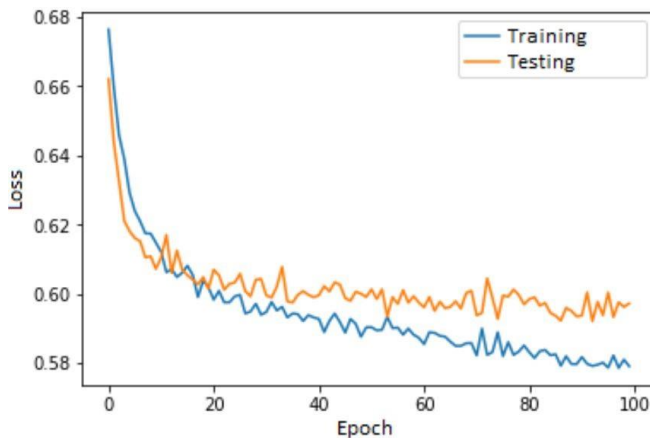

Fig. 7. Accuracy Graphs

Fig. 8. Accuracy Graphs

Fig. 9. with Values

Fig. 10. Result

having liver disease or not.Coming to the merits ,this way can be a very good method to find the liver dis functionality and help the people in understanding their health.It is far more economical than the original tests.The web configuration eases up the job to calculate things.The use of neural network makes the model very strong .Coming to the demerits, the accuracy is a bit lower than the medical accuracy and is needed to be improved we are working on using ensemble techniques to improve the model working.More of feature extraction to be done ,so the accuracy of the model gets better.The advancement of technology in the field of medical terms helps many in easing up things and help diagnosis of diseases easier and economical.

REFERENCES

[1] Spann, Ashley Yasodhara, Angeline Kang, Justin Watt, Kymberly Wang, Bo Goldenberg, Anna Bhat, Mamatha. (2020). Applying Machine Learning in Liver Disease Transplantation: A Comprehensive Review. Hepatology. 71. 10.1002/hep.31103.

[2] Kannapiran, Thirunavukkarasu Singh, Ajay Irfan, Md Chowdhury, Abhishek. (2018). Prediction of Liver Disease using Classification Algorithms. 1-3. 10.1109/CCAA.2018.8777655.

[3] Rajeswari, P Reena, G. (2010). Analysis of Liver Disorder Using Data mining Algorithm. Global Journal of Computer Science and Technology.10.

[4] Sharma, Tanu Kumawat, Gaurav Chakrabarti, Prasun Poddar, Sandeep Chakrabarti, Tulika Kamali, Ali Kateb, Babak Nami, Mohammad. (2021). Using Artificial Neural Network and Machine Learning Algo- rithms to Scrutinize Liver Diseases. 10.21203/rs.3.rs-324049/v1.

[5] Mohan, Vijayarani. (2015). KIDNEY DISEASE PREDICTION USINGSVM AND ANN ALGORITHMS.

[6] Dhamodharan. S, Liver Disease Prediction Using Bayesian Classifica- tion, Special Issue, 4th National Conference on Advanced Computing, Applications Technologies, May 2014, page no 1-3.

[7] Karthik. S, Priyadarishini. A, Anuradha. J and Tripathi. B. K, Classi- fication and Rule Extraction using Rough Set for Diagnosis of Liver Disease and its Types, Advances

in Applied Science Research, 2011, 2 (3): page no 334-345

[8] Omar S.Soliman, Eman Abo Elhamd, Classification of Hepatitis C Virus using Modified Particle Swarm Optimization and Least Squares Support Vector Machine, International Journal of Scientific Engineering Research, Volume 5, Issue 3, March-2014 122

[9] Tiruveedhula, Sajana Burramukku, Chandana Aishwaryagade, rama narasingarao, Manda. (2018). A COMPARISON OF UNSUPERVISED LEARNING TECHNIQUES IN JAUNDICE DIAGNOSIS. Indian Jour-

nal of Pure and Applied Mathematics. 116.

[10] Onwodi Gregory, "Prediction Of Liver Disease (Biliary Cirrhosis) Using Data MiningTechniques", IJETR, Vol.4, pp.12-20, 2015.

[11] Anju Gulia, Dr. Rajan Vohra, Praveen Rani, "Liver Patient Classification Using Intelligent Techniques", (IJCSIT) International Journal of Com- puter Science and Information Technologies, Vol. 5 (4), 2014, 5110- 5115,

[12] Olaniyi, Ebenezer Khashman, Adnan. (2015). Liver Disease Diagnosis Based on Neural Networks.