Early Detection of Hepatic Encephalopathy in Liver Disease Through Machine Learning

Ms. G. Parimala
Assistant professor
Vignan's Foundation for Science Technology and Research

Dr.S.Siva Nageswara Rao
Professor
Narasaraopeta Engineering College

Mr R Kalyan Chakravarthy
TGT Mathematics
Dr B R Ambedkar Gurukulam

ABSTRACT: Food is the basic need for Human living. The quality of the food is falling day by day for the past few years. Consuming such a qualityless food which will cause health issues that might affect the functionality of the human organs. In the human body LIVER is one of the most important organs which plays a vital role in the human digestive system. By consuming such qualityless food humans would suffer from diseases like Hepatitis C, Cirrhosis, liver failure, Hepatic Encephalopathy, etc., good research was done on Hepatitis C, Cirrhosis. Extending the research on the other diseases like Hepatic Encephalopathy, Liver Enlargement, etc., research was conducted based on the Indian Liver Patient Dataset (ILPD) to predict liver disorders in the earlier stage based on the parameters like Bilirubin, Alamine Aminotransferase (ALT), Aspartate Aminotransferase (AST), etc., through the machine learning technique which is Support Vector Machine (SVM). Through this most accuracy can be obtained and could be able to prevent liver disorders in the earlier stage before getting the complicated situations to the human body.

KEYWORDS: Hepatic Encephalopathy, Cirrhosis, Bilirubin, Alamine Aminotransferase, Aspartate Aminotransferase, Support Vector Machine.

1. INTRODUCTION: In this project, we are predicting the Liver diseases like hepatitis C, Cirrhosis, and Hepatic Encephalopathy through the Machine learning Algorithms like SVM, Random Forest, Decision tree, Gaussian Naïve Bayes, and Logistic Regression. The objective of this paper is to tell about the disease occurring to a person through the parameters which are considered for the analysis of the disease. As we are dealing with multiple algorithms, we will get different outputs and results based on the input we give to different algorithms. In this, we are comparing the results of different datasets with which we will be getting different results based on the number of samples taken for the different datasets. The model we are working on is the SVM using kernel functions.

2. LITERATURE SURVEY: The diseases like Hepatitis C and Cirrhosis the disease can be identified with the help of tests done on the liver like the Liver Functionality Test (LFT), ALT test, AST test etc., [1] to define the exact properties of the disease we take the parameters like ALT, AST, etc., the dataset is taken for the experiment was named after the Discovery Cohort [1]. For more info on the patient's condition Blood samples were also taken to know more about the dissolved substances in the blood throughout the body.

Hepatitis C is an infectious sickness that influences more than 70 million individuals around the world, even killing 400 thousand of them yearly. To all the more likely to grasp this infection and its Prognosis, clinical specialists can exploit the Electronic Health Records (EHRs) of patients, which contains taken information that computer-based approaches based on insights and computational knowledge can process to uncover new disclosures, what's more, drifts in any case unnoticeable by doctors. In this study, we break down EHRs of 540 [2] healthy controls furthermore, 75 patients were determined to have hepatitis C and used AI classifiers to anticipate their analysis.

We utilize the top classifier (Random Forests) to distinguish the most symptomatic factors for hepatitis C the outcome being aspartate aminotransferase (AST) and alanine aminotransferase (ALT). These two enzyme levels are likewise utilized by doctors in the AST/ALT proportion, a traditional measure generally utilized in gastroenterology and hepatology [2]. We apply a similar way to deal with a validation dataset of 123 patients with hepatitis C and cirrhosis, and similar two factors emerged as generally important.

Acute Hepatic Encephalopathy (AHE) is a typical type of delirium, a state of confusion, disabled consideration, also, diminished excitement because of intense liver failure. Nonetheless, the neurophysiological instruments hidden in AHE are ineffectively perceived. To foster theories for systems of AHE, our work expands on a current brain mean-field model for comparative EEG designs in cerebral anoxia, the bursting Liley model. The model recommends that summed up intermittent releases, like the triphasic waves (TPWs) seen in severe AHE, emerge through three sorts of cycles a) expanded neuronal edginess; [3] b) damaged Brain energy metabolism prompting weakened synaptic transmission; [4] c) and improved postsynaptic restraint intervened by expanded GABA-ergic and glycinergic transmission [3]. We relate the model boundaries to human EEG information utilizing a Particle-filter based enhancement strategy that matches the TPW between occasion span circulation of the model with that seen in patients' EEGs [7]. In this manner, our model relates minuscule systems to EEG designs. Our model addresses a beginning stage for investigating the basic components of cerebrum elements in ridiculousness.

3. EXISTING SYSTEM: Liver disease is Hepatic Encephalopathy in the past the disease was described or analyzed based on the Image data through which the end cause of the disease could be found through brain MRI images [3]. Where the root cause of the disease cannot be found. So, from that, we have taken the problem statement as the prediction of the root cause of the disease. From the brain MR images we, get the data at the final stage which will have the perseverance of living for the human. But it is also essential data for instant diagnosis and medication. We should be looking for the initial identification of the disease before the poisonous chemical enters the blood and flows to the brain and damages or infects the brain, keeping a human at-risk situation. This situation gave hope to take the chance of discarding the risk of brain disease in the initial stage.

4. PROBLEM STATEMENT: Predicting the root cause of the disease with the help of the parameters which may alter the health condition of the liver which leads to the Liver disease Hepatic Encephalopathy which affects the brain functionality leads to Brain cancer, brain failure, etc. by predicting the disease in the earlier stage which will save a lot of lives.

5. PROPOSED SYSTEM: As we know, the imbalance in the body will cause diseases in different body parts, which implies the imbalances in the liver will cause liver failure and has a chance of leading to the cause of brain damage disease Hepatic Encephalopathy. The parameters like Total bilirubin, Direct Bilirubin, Alamine Aminotransferase, Aspartate Aminotransferase, Alkaline Phosphatase, Total Protein, Albumin, Albumin and Globulin Ratio etc., parameters will give information about the patient's health condition of the liver which plays a major role in different functionalities, like Blood purification, Digestion process, etc. if any one of the processes gets affected that might lead to bigger problems one of them is the Hepatic Encephalopathy a brain disease which causes by the failure of the Liver Functionality. From figure 1, the Data collected by the system through the Webpage will be sent to the Machine Learning system which processes the inputs given to the system and analysis the input based on the past training given to the machine learning system in which different algorithms were implemented and getting different accuracy results. Proposed model work flow is shown in figure 1.

Working Process:

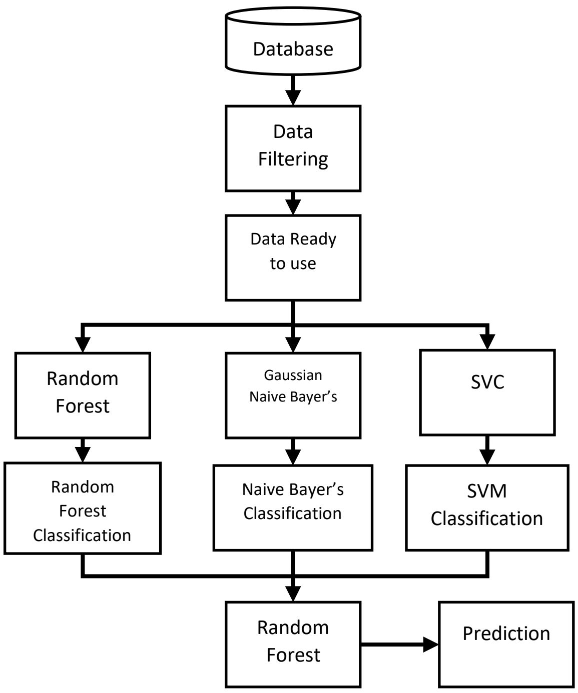

Fig 1: Flow chart of the Working Algorithm

The data which is stored in the CSV file will be taken to the Machine Learning Algorithm through the panda's library and then read the data in the CSV file and then the data taken from the CSV should be cleaned and filtered for the model execution. Then the data we have given will be sent into the 'fit' function where the data will be split for training and testing than the model will be trained and will be ready for the testing procedure then the model is trained and tested. Then the algorithm is ready for the Prediction operation. We must give a proper input to the algorithm with a correct number of inputs to the function.

6. MODEL ARCHITECTURE:

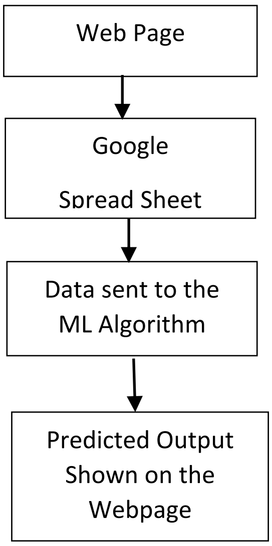

Fig 2: Architectural Flow chart of the System

The Data will be given to the system through the Webpage the data will be recorded in the spreadsheet which is attached to the webpage the data will be sent to the Machine Learning process in which the dynamical data is analyzed by the machine learning algorithm where the algorithm is trained with the numerous samples through which it predicts the output of the given input to the system. Architecture of the proposed method is shown in figure 2.

Kernel Functions/equations:

Standard Kernel Equation is shown in equation 1.

$$\boldsymbol{k(\overline{x}) = 1, if \|x\| \leq 1} \quad \text{------} \quad (1)$$

$$\boldsymbol{k(\overline{x}) = 0, Otherwise}$$

Radial Kernel Function is shown in equation 2.

$$\boldsymbol{k(x, y) = e^{-(\gamma\|x-y\|^2)}} \quad \text{------} \quad (2)$$

Sigmoid Kernel Function is shown in equation 3.

$$\boldsymbol{k(x, y) = tanh(\gamma \cdot x^T y + r)} \quad \text{-------} \quad (3)$$

Gaussian Kernel Function is shown in equation 4.

$$\boldsymbol{k(x, y) = e^{\left(-\frac{\|x-y\|^2}{2\sigma^2}\right)}} \quad \text{-------} \quad (4)$$

Polynomial Kernel Equation is shown in equation 5.

$$\boldsymbol{k(x, y) = tanh(\gamma \cdot x^T y + r)^d, (\gamma > 0)} - (5)$$

Confusion matrix

$$\boldsymbol{CM} = \begin{bmatrix} \boldsymbol{TP} & \boldsymbol{FP} \\ \boldsymbol{FN} & \boldsymbol{TN} \end{bmatrix}$$

TP = True Positive, FP = False Positive,

FN = False Negative, TN = True Negative

Accuracy: Equation 6 is a base metric used for model evaluation is often Accuracy, describing the number of correct predictions and overall predictions

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \qquad - (6)$$

Precision: Equation 7 is used to measure how many of the correct predictions made correct.

$$\text{Precision} = \frac{TP}{TP+FP} \qquad - (7)$$

Recall: Equation 8 is used to measures how many positive cases the classifier correctly predicted, overall, the positive cases in the data.

$$\text{Recall} = \frac{TP}{TP+FN} \qquad - (8)$$

F1-score: Equation 9 is used to measure F1-Score by combining both precision and recall. It is generally described as the harmonic mean of two.

$$\text{F1-score} = 2* \frac{Precision*Recall}{Precision+Recall} \qquad (9)$$

`

7. RESULTS:

Comparative Results:

	Accuracy	**F1-score**	**recall**	**Precision**
Logistic Regression	70	59	95	73
Random Forest	68.97	48	100	99
Gaussian Naïve Bayes	53.1	41	39	95
Support Vector Machine	73.1	46	100	71
Decision Tree	88.1	91	94	90

Table 1: Five Algorithms result scores

Table 1 explains the comparison study of the different machine learning algorithms with proposed method in terms of accuracy, F1-Score, Recall and Precision. According to table 1, our proposed method gives more accuracy compared to various existing methods.

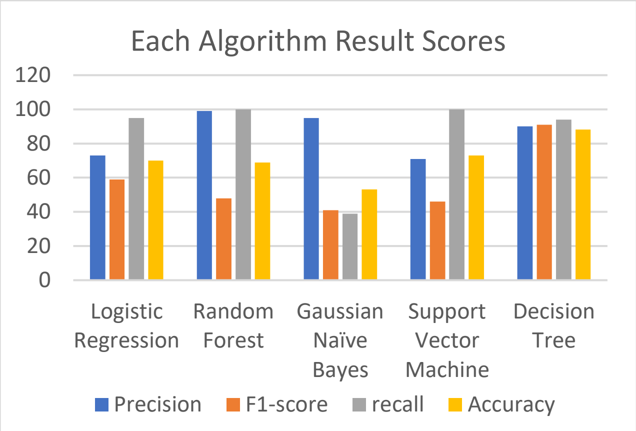

Fig 3: Each Algorithm Results graph

Figure 3 provides a visual comparison of the results obtained from various algorithms tested on different datasets. The precision and accuracy of these models largely depend on the quality and quantity of training data. A larger dataset enhances the model's ability to learn and accurately interpret input patterns, leading to more reliable analysis and predictions.

Final Results:

Figure 4, shows front end view of the proposed model to collect information from users. This web page collects basic information from patients about their deices and stores in database in excel form.

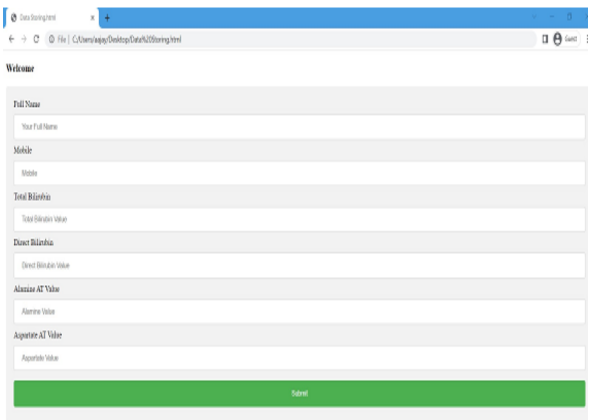

Fig 4: Web page to enter the Data.

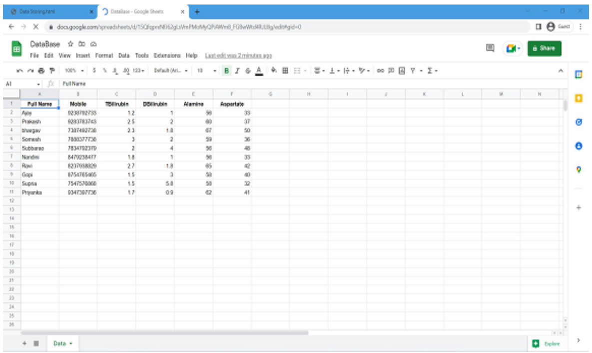

Fig 5: Online Excel Sheet to store the data.

The data will be given on the webpage that will send the data to the spreadsheet(figure 5) which will send the data to the Algorithms that will analyze the data which is given by the patient dynamically to the system. So, the patient will get the result instantly which is good to know the status of the health. That will help a human to take care of his/her health or even go the hospital to take a proper treatment for his/her health disorder.

8. FUTURE SCOPE: The data received from the patient can be stored and monitored by the hospital management for the future diagnosis and treatment of the patient. By implementing a Mobile app and then sending the data to the mobile from the sensor to the mobile for the further processing of the data. The data can be taken from the patient by attaching a smart device with a sensor which dynamically takes the data from the patient and sent to the concerned person mobile who takes care of the patient which helps the patient in the future diagnosis in the hospital.

9. CONCLUSION: The prediction of the Disease in the early stage will be more helpful for society in saving lives. By using a Machine Learning Algorithm, we can make a predicted future which will be more helpful for future generations and changes the lifestyle of Humans. When a long life is given to a human by using the latest technology it helps the human race to evolve into a new leap of leading life by increasing the chances of evolving into the new technologies.

REFERENCES

[1] Davide Chicco and Giuseppe Jurman, " An Ensemble Learning Approach for Enhanced Classification of Patients With Hepatitis", Volume 9, 2021 p.no 24485 to 24498, 2019 IEEE.

[2] Jiangling Song, Haoqi Sun, Jin Jing, Luis Carlos, Lingya Chao, Sydney S Cash, Rui Zhang and M.Brandon Westover, "A Mean Field Model of Acute Hepatic Encephalopathy", p.no 2366-2369,978-1-5386-3646-6/18 ©2018 IEEE.

[3] H. W. Syh, W. K. Chu and J. R. McConnell, "A Microcomputer Based System for MR Imaging Analysis of Brain for Hepatic Encephalopathy", p.no 130-134, CH2951-2/ 91/0000 /0130/0 1991 IEEE.

[4] Yehonatan Sela, Moti Freiman, Elia Dery, Yifat Edrei, Rifaat Safadi, Orit Pappo, Leo Joskowicz and Rinat Abramovitch, "fMRI-Based Hierarchical SVM Model for the Classification and Grading of Liver Fibrosis", Vol 58, no. 9, September 2011.

[5] Yun Jiao, Gao-Jun Teng and Xunheng Wang, "Predictive Model for Minimal Hepatic Encephalopathy Based on Cerebral Functional Connectivity", pno.541-4=545 978-1-4799-2761-6 /13/©2013 IEEE

[6] P. Hall and J. Cash, "What is the real function of the liver 'function' tests?" Ulster Med. J., vol. 81, no. 1, pp. 30-36, 2012.

[7] S. C. Nandipati, C. X. Ying, and K. K. Wah, "Hepatitis C virus (HCV) prediction by machine learning techniques," Appl. Model. Simul., vol. 4, pp. 89-100, Mar. 2020.

[8] Z. Wu, O. Matsui, A. Kitao, K. Kozaka, W. Koda, S. Kobayashi, Y. Ryu, T. Minami, J. Sanada, and T. Gabata, "Hepatitis c related chronic liver cirrhosis: Feasibility of texture analysis of MR images for classification of fibrosis stage and neuroinflammatory activity grade," PLoS ONE, vol. 10, no. 3, Mar. 2015, Art. no. e0118297.

[9] O. Akkaya, "Clinical significance of activity of ALT enzyme in patients with hepatitis C virus," World J. Gastroenterol., vol. 13, no. 41, p. 5481, 2007.

[10] P. Y. Kwo, S. M. Cohen, and J. K. Lim, "ACG clinical guideline: Evaluation of abnormal liver chemistries," Amer. J. Gastroenterol., vol. 112, no. 1, pp. 18-35, Jan. 2017.

[11] D. Manning and N. Afdhal, "Diagnosis and quantitation of fibrosis," Gastroenterology, vol. 134, no. 6, pp. 1670–1681, 2008.

[12] S. Bonekamp, I. Kamel, S. Solga, and J. Clark, "Can imaging modalities diagnose and stage hepatic fibrosis and cirrhosis accurately?" J. Hepatol., vol. 50, no. 1, pp. 17–35, 2009.

[14] M. Freiman, Y. Sela, Y. Edrei, O. Pappo, L. Joskowicz, and R. Abramovitch, "Multi-class SVM model for fMRI-based classification and grading of liver fibrosis," Proc. SPIE, vol. 7624, pp. 76240S-1–8, 2010.

[15] R. Bataller and D. Brenner, "Liver fibrosis," J. Clin. Invest., vol. 115, no. 2, pp. 209–218, 2005.

[16] https://towardsdatascience.com/a-look-at-precision-recall-and-f1-score-36b5fd0dd3ec