ISSN: 2278-4632
Vol-12 Issue-01 No.01: 2022

Liver disease prediction using Machine Learning

N.Anil Kumar[1], G.Madhumitha[2] , K.Chandu Priya[3] , A.S M Kamal [4] , MD.Ayman[5]

T.H M Pradeep[6]

[1]Ass.Professor, [2,3,4,5,6]Students

Dept of Computer Science And Engineering

Sri Vasavi Institute Of Engineering And Technology, Pedana, A.P, India.

ABSTRACT:

Recently, Information Systems and strategic tools are being incorporated as additional means to aid the process of diagnosis of diseases in medical research. The liver, an essential organ is crucial in enzyme activation, bile production, metabolism of fats and storage of vitamins, glycogen and minerals. Liver diseases are difficult to diagnose and hence are often neglected due to the lack of proper symptoms at the initial stages. One of the most common symptoms of most liver diseases is hyperbilirubinemia which is hard to distinguish in early determination. In any case, this isn't quite certain and the perception of enzyme level is required to distinguish and affirm the nearness of liver illness.. Various machine learning techniques have been used in the prediction of liver diseases. In this research, we propose the usage of , Random Forest Algorithm Machine techniques in the prediction of liver disease by Binary Classification of the dataset into two given categories of patient experiencing liver sickness or not. The dataset contains information about patient attributes such as Total Bilirubin, Alanine Aminotransferase, Direct Bilirubin, Aspartate Aminotransferase, Age, Gender, Albumin, Total Proteins, Alkaline Phosphatase, Albumin and Globulin Ratio and the Result. The prediction from the above mentioned algorithms are compared on the parameters of Accuracy and various error calculations to determine the best suited algorithm.

INTRODUCTION:

The liver is a large, meaty organ that sits on the right side of the belly. Weighing about 3 pounds, the liver is reddish brown in color and feels rubbery to the feel. The liver has two large sections, called the right and the left lobes. The gallbladder sits below the liver, along with parts of the pancreas and intestines. The liver and these organs behavior together to digest, absorb, and process food. The liver's main job is to strain the blood coming from the digestive tract, before passing it to the rest of the body. The liver also detoxifies chemicals and metabolizes drugs. As it does so, the liver hides bile that ends up back in the intestines. The

DOI : 10.36893.JK.2022.V12I01N01.927-933

ISSN: 2278-4632
Vol-12 Issue-01 No.01: 2022

liver also makes proteins important for blood clotting and other functions [1]. Liver disease is any trouble of liver function that causes sickness. The liver is responsible for many dangerous functions within the body and should it become diseased or damaged, the loss of those functions can cause significant injury to the body. Liver disease is also referred to as hepatic disease. Liver disease is a large term that covers all the potential problems that cause the liver to fail to perform its designated functions. Usually, more than 75% or three quarters of liver tissue needs to be affected before a decrease in function occurs .

LITERATURE SURVEY:

1.Nazmun Nahar , Ferdous Ara, "Liver disease prediction by using different Decision tree techniques", International Journal of Data Mining & Knowledge Management Process (IJDKP) Vol.8, No.2, March 2018

2.Joel Jacob, Joseph Chakkalakal Mathew, Johns Mathew, Elizabeth Issac, "Diagnosis of Liver Disease Using Machine Learning Techniques", International Research Journal of Engineering and Technology (2018)

[3] Sumedh Sontakke, Jay Lohokare, Reshul Dani, “Diagnosis of Liver Diseases using Machine Learning”, International Conference on Emerging Trends & Innovation in ICT (ICEI) , 2017.

for the predictions and this has indicated atomic science approach to enhance the predictive capacity. The outcomes obtained from the investigations in

[4] Shambel Kefelegn, Pooja Kamat, "Prediction and Analysis of Liver Disorder Diseases by using Data Mining Technique: Survey", International Journal of Pure and Applied Mathematics, 2018

namely Precision, Recall, Mean Absolute Error, Runtime and Accuracy are used and the results show that the Decision Stump has the highest accuracy. [2] has proposed the use of classification algorithms namely Logistic Regression, SVM, ANN and KNN with back propagation which consists of 10 input neurons layers and the study shows that ANN is comparatively efficient. In an experiment, genetic microarrays in addition to the neural systems has been proposed by [3] for the predictions and this has indicated atomic science approach to enhance the predictive capacity. The outcomes obtained from the investigations in [4] have demonstrated that Random Forest calculation isn't appropriate because of the issue of overfitting and recommendations towards the usage of oversampling techniques have been made to address this issue..

DOI : 10.36893.JK.2022.V12I01N01.927-933

ISSN: 2278-4632
Vol-12 Issue-01 No.01: 2022

Proposed System:

Machine learning is understandably one of the most extensively utilized paradigms of big data management where a significantly high set of distinct raw data can be collated effectively to make appropriate inferences and eventually to come up with a usual collection of contextually useful collection of integrative information. With the onset of the exponential technological explosion in the field of medicine, there is a felt need to handle a colossal set of data, thereby managing and utilizing the same to make effective and informative inferences for the doctors and patients. The system being proposed here uses concept of machine learning, and the models are first trained, then tested. Finally the most accurate model will predict the final result.

At first, the system asks you to enter your details including age, gender, total Bilirubin, direct Bilirubin, total proteins, Alkaline Phosphatases, Alamine Aminotransferase, Aspartate Aminotransferase Albumin, AlbumiGlobulin ratio.Values of last eight parameters mentioned here, can be known by blood test report of the user. After taking these inputs from the user, the system compares the data input with the training dataset of most accurate model and then predicts the result accordingly as risk or no risk.

Advantages of the Proposed System

- The performance classification of liver-based diseases is further improved
- Time complexity and accuracy can be measured by various machine learning models, so that we can measures different parameters, owing to the needs of the user
- Different machine learning having high accuracy of the result
- Risk factors can be predicted early by machine learning models.

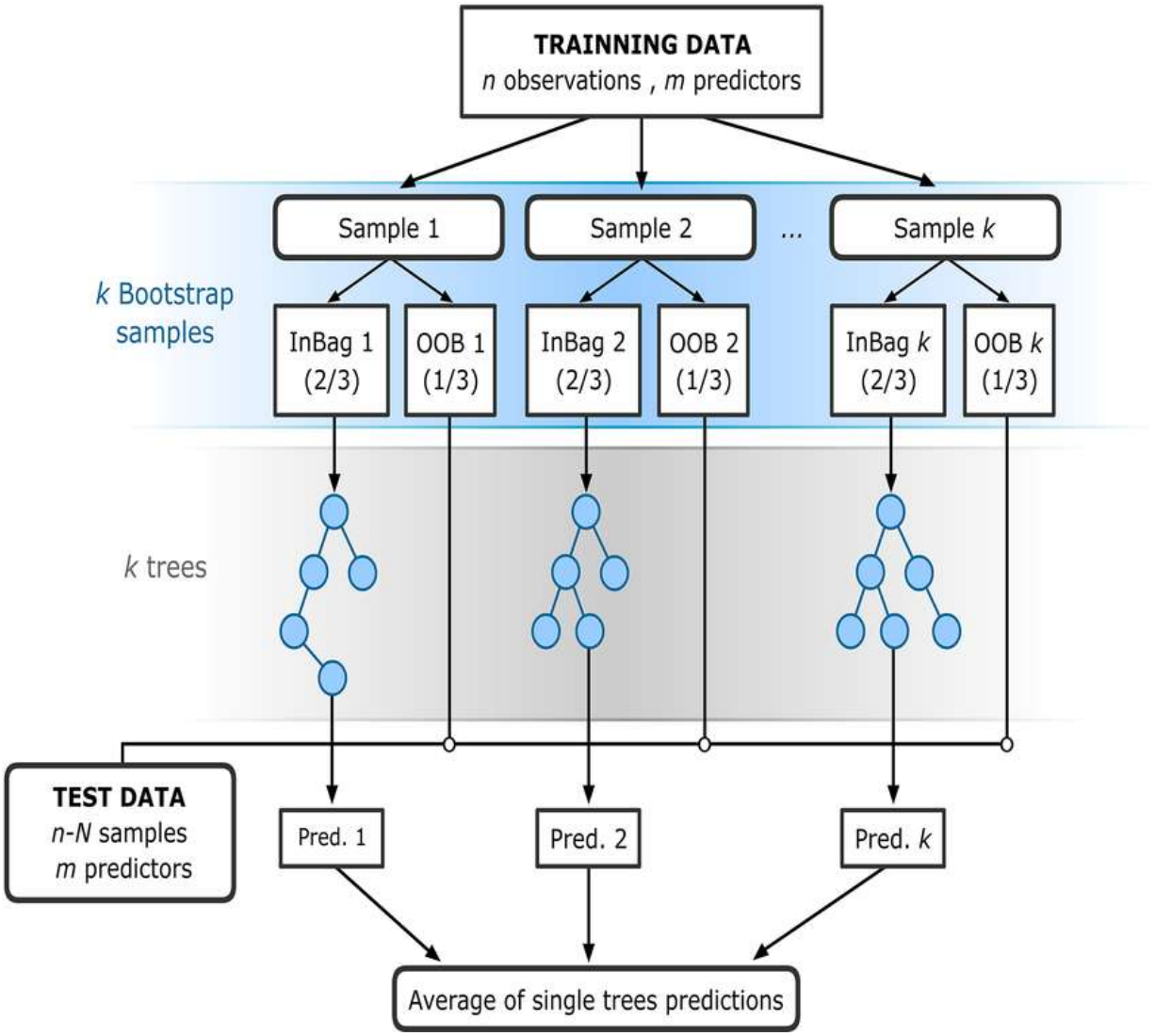
MODULES

- Data gathering
- Pre processing
- Processing
- Interpretation

DOI : 10.36893.JK.2022.V12I01N01.927-933

ISSN: 2278-4632
Vol-12 Issue-01 No.01: 2022

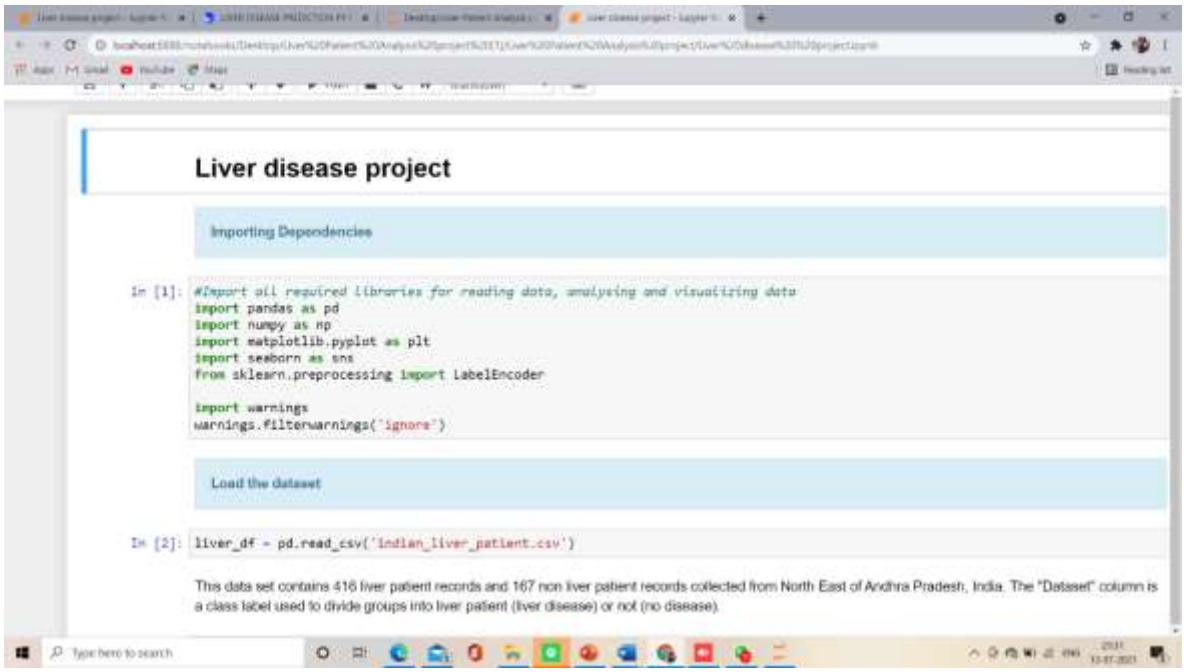
Architecture(methodology):

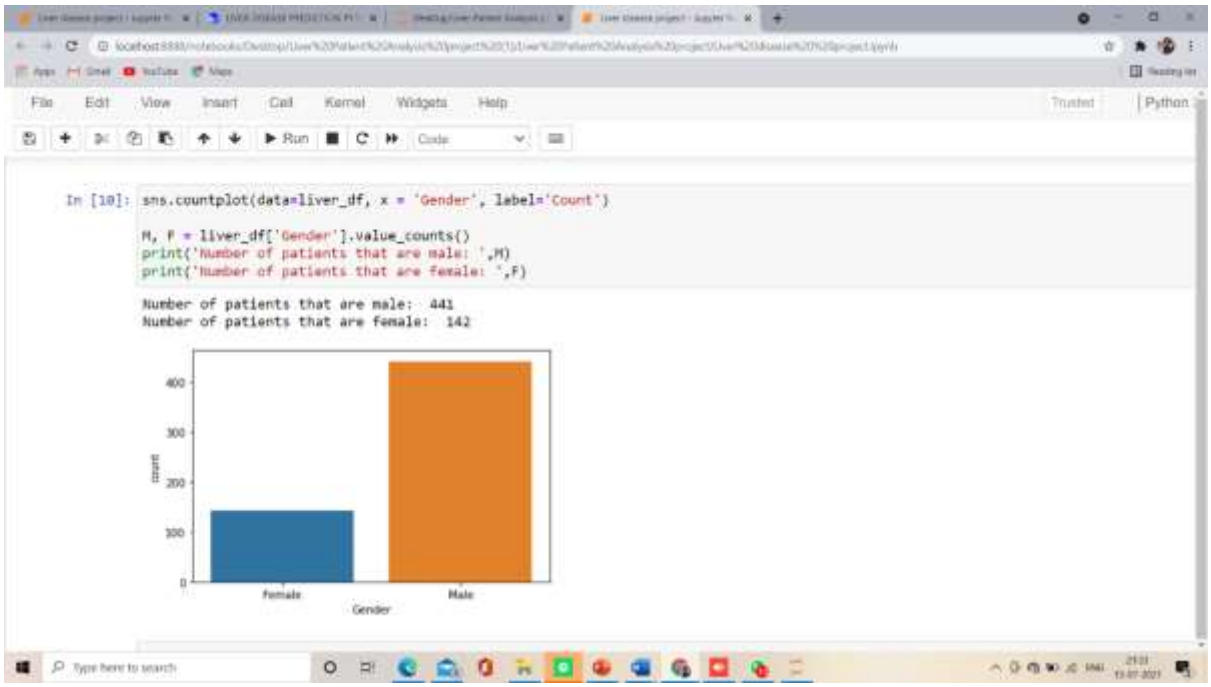


DOI : 10.36893.JK.2022.V12I01N01.927-933

ISSN: 2278-4632
Vol-12 Issue-01 No.01: 2022

Dataset:





DOI : 10.36893.JK.2022.V12I01N01.927-933

ISSN: 2278-4632
Vol-12 Issue-01 No.01: 2022

CONCLUSION:

Diseases related to liver and heart are becoming more and more common with time. With continuous technological advancements, these are only going to increase in the future. Although people are becoming more conscious of health nowadays and are joining yoga classes, dance classes; still the sedentary lifestyle and luxuries that are continuously being introduced and enhanced; the problem is going to last long.

In this project, we have proposed methods for diagnosing liver disease in patients using machine learning technique. The system was implemented using Random Forest, Logistic Regression and Naive Bayes and their performance was evaluated. Based on the performance we conclude that the most suitable algorithm for the liver disease prediction is to be the "Random Forest". Through this we acquire high accuracy and less time complexity and predict the probability of occuring liver disease with the given information.

REFERENCES:

-P. Mell and T. Grance, "The NIST definition of cloud computing," Nat. Inst. Stand. Technol., vol. 53, no. 6, pp. 50–50, 2009.

- H. T. Dinh, C. Lee, D. Niyato, and P. Wang, "A survey of mobile cloud computing: Architecture, applications, and approaches," Wireless Commun. Mobile Comput., vol. 13, no. 18, pp. 1587–1611, 2013.

-H. Li, W. Sun, F. Li, and B. Wang, "Secure and privacy-preserving data storage service in public cloud," J. Comput. Res. Develop., vol. 51, no. 7, pp. 1397–1409, 2014.

-Y. Li,T.Wang,G.Wang,J.Liang,andH.Chen,"Efficientdatacollection in sensor-cloud system with multiple mobile sinks," in Proc. Adv. Serv. Comput., 10th Asia-Pac. Serv. Comput. Conf., 2016, pp. 130–143.

-L. Xiao, Q. Li, and J. Liu, "Survey on secure cloud storage," J. Data Acquis. Process., vol. 31, no. 3, pp. 464–472, 2016.

DOI : 10.36893.JK.2022.V12I01N01.927-933

ISSN: 2278-4632
Vol-12 Issue-01 No.01: 2022

-R. J. McEliece and D. V. Sarwate, “On sharing secrets and reed-solomon codes,” Commun. ACM, vol. 24, no. 9, pp. 583–584, 1981.

] J. S. Plank, “T1: Erasure codes for storage applications,” in Proc. 4th USENIX Conf. File Storage Technol., 2005, pp. 1–74.

R.Kulkarni,A.Forster,andG.Venayagamoorthy,“Computationalintelligenceinwirelesssensornetworks:Asurvey,” IEEE Commun. Surv. Tuts., vol. 13, no. 1, pp. 68–96, First Quarter 2011.

-Z. Xia, X. Wang, L. Zhang, Z. Qin, X. Sun, and K. Ren, “A privacypreserving and copy-deterrence content-based image retrieval scheme in cloud computing,”

IEEE Trans. Inf. Forensics Security, vol. 11, no. 11, pp. 2594–2608, Nov. 2016.

-J. Shen, D. Liu, J. Shen, Q. Liu, and X. Sun, “A secure cloud-assisted urban data sharing framework for ubiquitous-cities,” Pervasive Mobile Comput., vol. 41, pp. 219–230, 2017.

-Z.Fu,F.Huang,K.Ren,J.Weng,andC.Wang,“Privacy - preserving smart semantic search based on conceptual graphs over encrypted outsourced data,” IEEE Trans. Inf. Forensics Security, vol. 12, no. 8, pp. 1874–1884, Aug. 2017.

DOI : 10.36893.JK.2022.V12I01N01.927-933