



ISSN: 2454-9940



**INTERNATIONAL JOURNAL OF APPLIED
SCIENCE ENGINEERING AND MANAGEMENT**

E-Mail :
editor.ijasem@gmail.com
editor@ijasem.org

www.ijasem.org

COMPARATIVE ANALYSIS OF LIVER DISEASES BY USING MACHINE LEARNING TECHNIQUES

¹ K VENKATESH, ²M.ASHA JYOTHI

¹(Assistant Professor), MCA, DANTULURI NARAYANA RAJU COLLEGE(A) PG COURSES,
BHIMAVARAM ANDHRA PRADESH

²MCA, scholar, DANTULURI NARAYANA RAJU COLLEGE(A) PG COURSES, BHIMAVARAM
ANDHRA PRADESH

ABSTRACT

In a human body function of the liver is important. Many persons are suffering from liver disease, but they don't know it. The identification of liver diseases in the early stage helps a patient get better treatment. If it is not diagnosed earlier, it may lead to various health issues. To solve these issues, physicians need to examine whether the patient has been affected by liver disease or not, based on the multiple parameters. In this paper, we classify the patients who have liver disease or not by using different machine learning algorithms by comparing the performance factors and predicting the better result. The liver dataset is retrieved from the Kaggle dataset.

Patients with liver problems that are difficult to detect in the early stage will help to continue their function normally even if they are partially damaged. There are chances for a patient surviving a liver disease to be better if they are diagnosed early.[1] The liver is an important organ that performs many functions energy storage, linked to metabolism, and waste cleansing. It also aids in the digestion of food, the change of food into energy, and the storing of energy until needed. It also helps in the removal of potentially dangerous compounds from our bloodstream. The disease may be a general term that refers to any condition affecting the liver

A. Functions of Liver:

These are some of the functions of the Liver:

1.INTRODUCTION

- It produces a component in the immune system that can combat illness.
- Producing the proteins that aid in blood coagulation.
- Red blood cells that are old or damaged are broken down.
- Excess blood sugar is stored as glycogen

The liver and its activities can be harmed by a variety of disorders. Some people respond well to treatment, while others do not. Fig. 1 shows the condition of Normal and affected liver with diseases. Some of the common conditions that affect the liver are discussed below:

1) Autoimmune hepatitis

The immune system of the body attacks itself and destroys healthy liver tissue in this disease. Cirrhosis and other liver damage can result from autoimmune hepatitis.

2) Cirrhosis

In this healthy liver, tissues are affected and changed as scar tissue due to chronic hepatitis, Long-term excessive alcohol consumption, and rare hereditary

disorders such as Wilson's disease are all examples that might cause this problem.

3) Hemochromatosis

An overabundance of iron builds up in the body as a result of this disorder. The liver might be harmed by too much iron.

4) Hepatitis A

viral infection which causes swelling in the liver is known as Viral hepatitis. There are some types of hepatitis, like A, B, C, D, and E. Each has its own set of causes and consequences. Hepatitis A is most common in underdeveloped nations with poor sanitation and access to clean drinking water. Hepatitis A is usually treatable without causing liver damage or long-term consequences

5) Hepatitis B

It can be an infection either short-term or long-term. It can also be contracted by sharing the needles with others or inadvertently injecting oneself with a contaminated needle. These serious complications, which include a cause of liver failure and cancer, can occur as a result of illness. There is a vaccine available to prevent the sickness.

6) Hepatitis C

Hepatitis C is a viral infection that can be either acute or persistent. It's disseminated most usually through coming into touch with hepatitis C virus-infected blood, such as by using dirty needles to inject drugs or apply tattoos. Due to this liver failure, and liver cancer are all possible.

7) Non-alcoholic fatty liver disease and NASH

The excess fat that builds up in the liver will damage the liver, which causes swelling. Fatty liver disease may cause scarring or fibrosis due to non-alcoholic steatohepatitis. Type 2 diabetes-related diseases may cause due to this problem.

B. Symptoms of liver conditions

There are so many types of liver disorders, which show symptoms like flu- and cause more serious damage in the liver which includes jaundice and dark-colored urine.

The following are some of the signs and symptoms of liver disease:

- fatigue

- a decrease in appetite
- vomiting
- Pain in joint
- stomach ache or discomfort
- Bleeds in the nose
- aberrant blood vessels on the surface of the skin (spider angiomas)

Symptoms that are more severe include:

- The skin and eyes turn a yellowish color (jaundice)
- bloating in the stomach (ascites)
- Leg swollenness (edema)
- Gynecomastia is a term used to describe a condition in which a man develops (when males start to develop breast tissue)
- Enlarged liver (hepatomegaly)

2.LITERATURE SURVEY

Nazmun Nahar, et al. implemented by using various decision trees techniques like LMT, J48, Hoeffding Tree, Decision Stump, and Random tree. for calculating expected time predication of disease affected to liver

finally, the Decision Stump gives the highest accuracy results among other techniques.

A Saranya, et al. explained the applications in data mining techniques and also used Medical Data Mining (MDM) to diagnose liver diseases. This technique includes prediction in the early stage, the existence and also complexity of the disease which helps partial assistance to the physicians.

S. Dhamodharan considers three major liver diseases like cirrhosis, hepatitis, and liver cancer. The fundamental purpose of this forecast is to find the type of disease by using classifications techniques such as cirrhosis, hepatitis, liver cancer, and "no disorders." Then compare the accuracy of the FT and Naive Bayes tree algorithms and shows that the Naive Bayes algorithm accuracy is significantly higher than that of the other methods

Kemal Akyol, Yasemin Gultepe by using the dataset which has shown a balanced result by using sampling technique for getting accuracy. the Stability Selection technique is used for selection based on attributes. For improving the performance, a blend of Stability Selection and Random Forest methods is used.

Shambel Kefelegn, Pooja Kamat for getting better results different data mining classification techniques are compared with the earlier liver prediction methods. The accuracy is measured with the help of confusion matrices for getting the better performance of the accuracy

Fadl Mutaher Ba-Alwi, et al.using various machine learning algorithms compared the Hepatitis prognostic data among them. In that Naive Bayes, technique gave good accuracy and also takes less time to build a model.

K. Thirunavukkarasu, et al. Used different classification techniques for predicting liver diseases. They compare the results of accuracy score and confusion matrix with Logistic Regression, SVM, and K-Nearest Neighbour. Bendi Venkata, et al. used different classifications algorithms, they checked the accuracy, precision, sensitivity, and specificity on liver datasets.

Tapas Ranjan Baitharua et al. has proposed an Intelligent medical decision support system to help physicians diagnose liver disorders through a learning pattern technique. In this, several classification techniques are used to compare the effectiveness, correction rate, and also

accuracy for the data is analysed with different scenarios. A diagnostic support system was developed with the support of a number of models with help of neural, which is helped to the physicians for diagnosis on the liver in the medical field.

M. Banu Priya P, et al. Using a root mean error value, root mean square value, the accuracy is calculated, and better accuracy is produced with the support of the PSO features selection technique. Dietterich, Thomas G states that the ensemble learning technique produces a better performance than the other single classifier techniques with the Bayesian averaging, error-correcting output coding, boosting, and bagging. In this paper, the author analyses existing ensemble approaches with some novel experiments to determine why Adaboost does not overfit quickly.

3. EXISTING SYSTEM

Prior to the implementation of the Liver Disease Classification System, healthcare professionals relied primarily on traditional diagnostic methods for identifying liver diseases. These methods often included a combination of medical history review, physical examination, liver function tests, imaging studies (such as ultrasound, CT

scan, or MRI), and liver biopsy in some cases. However, these methods may have limitations in terms of accuracy, early detection, and efficiency.

Disadvantages

- An existing methodology doesn't implement an effective Machine Learning Classifiers.
- The system not implemented Decision Trees which leads complexity in testing and training datasets.

3.1 PROPOSED SYSTEM

To analysis about liver disease, the data is retrieved from the Indian Liver Patient from the Kaggle. Here the patient has been characterized based on diseases as either 1 or 2. The values used in the dataset are given in Table I. The gender attribute is transformed to an integer value during the data pre-processing stage (0 and 1). The overall procedure of the proposed system is depicted in this system. The workflow of the proposed system is as follows a collection of data sets, Handling Categorical values, Splitting the data for Training, and Testing.

Perform feature selection and apply the machine learning techniques and compare the predicted result and find better accuracy.

Advantages

- It is a supervised Machine learning technique applied for both classification and regression kinds of problems, but it is used for classification types of problems. This model is applied to predict the categorical dependent variable with support of independent variables the output should be 0 or 1.
- This Classifier technique is effective when only a small amount of training data is required to derive approximation parameters. With highly scalable model creation, it can tackle a wide range of challenging real-world problems.
- This is a pattern recognition system that involves the training datasets for finding the k closest relatives in new conditions. When using k-NN for classification, we must calculate the location data within the nearest neighbor's category.

4. OUTPUT SCREENS

User Login



User registration



Prediction of liver disease



User profile



View trained and tested accuracy in bar charts



All remote users



View trained and tested accuracy results



Liver disease prediction pie char



Liver disease found ratio details



Liver disease prediction line chart



Liver disease prediction type ratio results



View Liver disease prediction



5. CONCLUSION

In the above system, we have studied some classification algorithms like SVM, Random Forest, KNN, LR, Naive Bayes, Kernel SVM, Decision tree to predict the patient has a Live disease or not. The early prediction gives the physicians to take the necessary steps to save the life of the patient. In this, the Kernel SVM approach produces better accuracy results than the other techniques. But we need to predict more accurate results. To get better results we need to use some advanced techniques for predicting the diseases

6. REFERENCES

- [1] K. Sumeet, J.J. Larson, B. Yawn, T.M. Therneau, W.R. Kim, Underestimation of liver-related mortality in the United States. *Gastroenterology*; (2013) 145:375–382, e371–372.
- [2] A.A. Mokdad, A.D. Lopez, S. Shahraz, R. Lozano, A.H.

Mokdad, J. Stanaway, et al, Liver cirrhosis mortality in 187

countries between 1980 and 2010: a systematic analysis.

BMC Med 2014; 12:145.

[3] Byass, Peter, The global burden of liver disease: a challenge for methods and for public health. *BMC medicine* 12.1 (2014); 159.

[4] L. A. Auxilia, Accuracy Prediction Using Machine Learning

Techniques for Indian Patient Liver Disease. 2018 2nd

International Conference on Trends in Electronics and Informatics (ICOEI). IEEE (2018).

[5] Hashem, M. Esraa, S. Mai, A study of support vector machine algorithm for liver disease diagnosis. *American Journal of Intelligent Systems* 4.1 (2014); 9-14.

[6] P. Sajda, "Machine learning for detection and diagnosis of disease." *Annu. Rev. Biomed. Eng.* 8 (2006); 537-565.

[7] UCI Machine Learning Repository. ILPD (Indian Liver Patient Dataset) Data Set.

[https://archive.ics.uci.edu/ml/datasets/ILPD+\(Indian+Liver+Patient+Dataset\)](https://archive.ics.uci.edu/ml/datasets/ILPD+(Indian+Liver+Patient+Dataset))

[8] Logistic Regression, Retrieve from: [HTTPS://WWW.SAEDSAYAD.COM/LOGISTIC_REGRESSION.HTM](https://www.saedsayad.com/LOGISTIC_REGRESSION.HTM), LAST Accessed: 5 October, 2019

[9] L. Breiman, Random Forests. *Machine Learning*, 45(1),

- (2001); 5–32.
<https://doi.org/10.1023/A:1010933404324>
- [10] Decision Trees, Retrieve from:
<https://dataaspirant.com/2017/01/30/how-decision-tree-algorithm-works/>, Last Accessed: 5 October,2019
- [11] Support vector machine, Retrieve from:
<http://www.statsoft.com/textbook/support-vector-machines>, Last Accessed: 5 October,2019
- [12] V. Vapnik, I. Guyon, T. H.-M, Learn, and undefined 1995.
Support vector machines. statweb.stanford.edu (1995).
- [13] Zhang M, Zhou Z, "ML-KNN: A lazy learning approach to multi-label learning." *Pattern recognition*40.7:(2007); 2038-2048.
- [14] G. Guo, H. Wang, D. Bell, Y. Bi, K. Greer, KNN Model-Based Approach in Classification (pp. 986–996). Springer, Berlin, Heidelberg (2003).
https://doi.org/10.1007/978-3-540-39964-3_62
- [15] Naive Bayes, , Retrieve from:
<https://www.geeksforgeeks.org/naive-bayes-classifiers/>,
Last Accessed: 5 October,2019
- [16] S. M. Mahmud, et al. "Machine Learning Based Unified Framework for Diabetes Prediction." *Proceedings of the 2018 International Conference on Big Data Engineering and Technology*. ACM (2018).
- [17] S. Safavian, D. Landgrebe, A survey of decision tree classifier methodology. *IEEE transactions on systems, man, and cybernetics*, 21(3), (1991); 660-674.
- [18] A. Chervonenkis, Early history of support vector machines. In *Empirical Inference* (pp. 13-20). Springer, Berlin, Heidelberg (2013).
- [19] K.M. Leung, Naive bayesian classifier. Polytechnic University Department of Computer Science/Finance and Risk Engineering (2007).