IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

PREDICTION OF LIVER CHRONIC DISEASE FOR HEALTH CARE SERVICES

[1]Dr.J.Sridhar, [2]K.Rajitha, [3]K.Prathyusha, [4]P.Punith Varma, [5]V.Revanth

[1]Associate Professor, [2345]UG Student, Department of Computer Science and Engineering, Siddharth Institute Of Engineering & Technology, Puttur , Andhra Pradesh, India.

ABSTRACT:

Due to the quick rise in liver illness caused by excessive alcohol use, contaminated gas inhalation, drug use, tainted food, and pickled food packaging, a doctor can make an automatic forecast with the aid of a medical expert system. Early liver disease prediction is now attainable because to the consistent advancements in machine learning technology, allowing for simple early identification of the fatal condition. This will make healthcare more beneficial, and a medical expert system can be employed in a remote location. The liver is vital to life and promotes the body's ability to rid itself of poisons. Early detection of the condition is therefore crucial for recovery. For the diagnosis of liver illness, several machine learning methods, such as Random Forest, KNN, Logistic Regression, Decision Tree, etc., are utilised. These algorithms have varying degrees of accuracy, precision, and sensitivity.

Keywords: *Liver Diagnostic, Machine Learning, Supervised Learning.*

INTRODUCTION

According to the World Health Organization's most recent survey report, which was published in 2017, death from liver disease accounts for 2.95% of all deaths, placing India at 63rd place globally. The liver is our body's largest internal organ. The liver is reddish-brown in hue and weighs about 3 pounds. Under the liver is where the gallbladder is situated. The liver's primary function is to filter out dangerous and toxic compounds from the blood before it is sent to various bodily regions. One of the deadliest and most dangerous diseases in the world is liver disease. The causes of liver disease are detailed in the list below. Hepatitis infection, liver cirrhosis, fatty liver, excessive alcohol use, toxic substances, and genetic disorders. If liver is 100% fail there is not option to recover but only one treatment that is liver transplantation. Early diagnosis of liver illness is important for effective treatment and a quick recovery. Even when liver tissue has only somewhat been damaged, liver disease is exceedingly difficult to diagnose in its early stages, making it challenging for many medical expert systems. This results in treatment and drug failure. Early diagnosis is essential to providing the patient with the right care and saving their life. Different symptoms of chronic liver disease include digestive issues such as stomach pain, dry mouth, constipation, and internal bleeding, dermatological problems such as yellow skin colour, spider veins, and redness on the feet, and abnormalities of the brain and nervous system such as memory loss, numbness, and fainting. So, some preventative measures for liver disease include seeing the doctor often, getting immunised, avoiding soda and alcohol, exercising frequently, and maintaining a healthy weight. According to the current medical expert system, diagnosing liver illness has been beneficial to society. In addition, using the expert system makes it simple to detect and anticipate the condition. Several forms of machine learning algorithms have been developed as a result of continued advancements in artificial intelligence, which will assist to increase the precision and quality of liver disease diagnosis and prediction.

LITERATURE SURVEY

The use of information systems and tactical tools in the medical fields is continually expanding. Automated medical models are useful tools for doctors to use when making decisions about patient care since they enable quick and accurate diagnosis or even prognosis. Several statistical mining and machine learning tools can be utilised to make use of the knowledge or even at the early phases of knowledge collection. One of the worries is, for example, determining if the patient with the Hepatitis C virus also has liver fibrosis or not. If the outcome of the forecast is accurate, What stage of fibrosis is it? A completely integrated system is required to easily get this knowledge without expensive diagnostic regular laboratory procedures. To predict the level of liver fibrosis in individuals, we therefore utilised a machine learning approach model based on decision tree classifier in this work. The accuracy of the decision tree classifier, according to the results, is 93.7%, which is greater than the range reported in recent studies conducted under comparable circumstances.

The technique of removing valuable information from sizable databases is known as data mining. With the use of machine learning tools, data mining plays a critical role in the medical industry in predicting and diagnosing disease at an early stage. The liver, the largest internal organ in the human body, plays a crucial role in a significant role in the human body and performing various essential tasks. Jaundice, a propensity to bruise easily, ascites, impaired cognitive function, and general health decline are all possible indications of liver disease. The person who consumes more alcohol quickly develops liver damage. Acute liver failure, hepatitis, liver cancer, and cirrhosis are a few examples of liver diseases. In India, liver disease issue afflicted a large number of men.

In order to bypass the limitations of biopsy, non-invasive machine learning techniques have lately been applied as an alternate staging method for chronic liver illnesses. By combining serum biomarkers and clinical data to create classification models, this study seeks to examine several machine learning approaches in the prediction of advanced fibrosis. Methods: Based on the METAVIR score, a prospective cohort of 39,567 individuals with chronic hepatitis C was split into two groups: one group was classified as having mild to moderate fibrosis (F0-F2), while the other group was classified as having advanced fibrosis (F3-F4). Advanced fibrosis risk prediction models using decision trees, genetic algorithms, particle swarm optimization, and multi-linear regression were created. Analysis of the receiver operating characteristic curve was done to gauge how well the suggested models worked. Results: It was discovered that advanced fibrosis was statistically correlated with age, platelet count, AST, and albumin. The advanced fibrosis in HCC patients may be predicted by the machine learning algorithms under research with an accuracy of between 66.3 and 84.4 percent and an AUROC of 0.73 to 0.76.

RECENT WORKS:

Around 2.4% of deaths in India occur as a result of liver disease each year. Due to its modest symptoms, liver disease can be challenging to diagnose in its early stages. Frequently, the symptoms show up only when it's too late. This study examines two techniques for identification—patient parameters and genome expression—in an effort to enhance the diagnosis of liver disorders. The study also outlines drawbacks and examines the computational techniques that can be used to the aforementioned methodology. It suggests ways to increase these algorithms' effectiveness. This demonstrates how machine learning is crucial in making an early diagnosis of the liver.

Due to increased alcohol use, drug use, tainted food, and pickling of foods, liver disease sufferers have been rising quickly. Hence, a doctor will receive automatic prediction assistance from the medical expert system. Early liver disease prediction is now attainable because to the consistent advancements in machine learning technology, allowing for simple early identification of the fatal condition. This will make healthcare more beneficial, and a medical expert system can be employed in a remote location. The liver is vital to life and promotes the body's ability to rid itself of poisons. Early detection of the condition is therefore crucial for recovery. Various machine learning techniques, including supervised, unsupervised, and semi-supervised reinforcement learning, can be used to diagnose liver disease. These techniques include SVM, KNN, K-Mean clustering, neural networks, decision trees, and others. The goal of this study is to provide an overview and comparative analysis of all machine learning approaches currently being utilised in the medical field for the diagnosis and prediction of liver disease. The analysis is based on accuracy, sensitivity, precision, and specificity. In this study, methodologies for liver diagnosis using supervised and unsupervised machine learning algorithms are discussed.

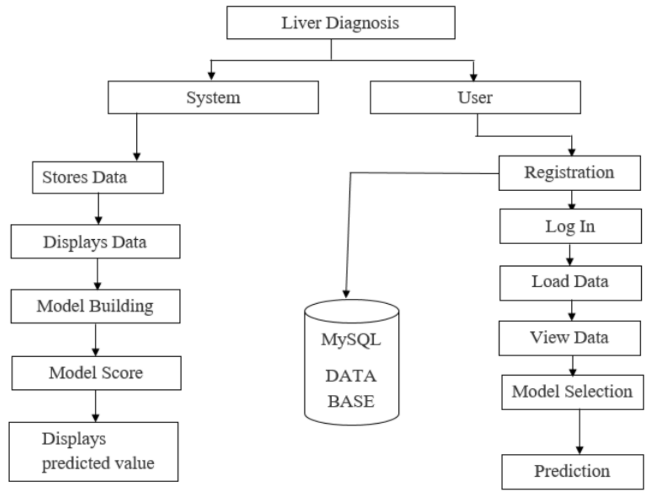
SUGGESTIVE METHODOLOGY

Machine learning is becoming more and more prevalent, with classical and machine learning approaches used in computer science. The related efforts of liver diagnosis are discussed in this section, along with how machine learning techniques outperform conventional ones. The project's current approach uses the KNN algorithm to diagnose the liver. However, it is not very reliable, and the outcome is inaccurate.

We employ machine learning methods in the proposed system to draw conclusions from the intricate patterns in the data. This method's straightforward architecture makes computations with it cheap.

CONTRIBUTIONS

- Very high precision.
- Saving time.
- Low computational cost.
- Minimal complexity.

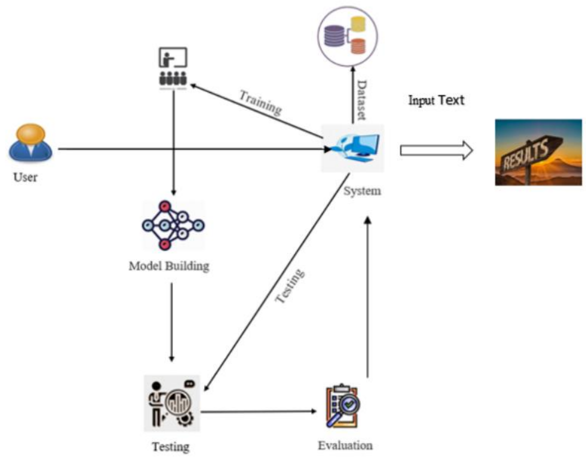


METHODOLOGY AND ALGORITHMS:

K Nearest Neighbors: Based on the supervised learning technique, K-Nearest Neighbor is one of the simplest machine learning algorithms. The K-NN algorithm makes the assumption that the new case and the existing cases are comparable, and it places the new instance in the category that is most like the existing categories. A new data point is classified using the K-NN algorithm based on similarity after all the existing data has been stored. This means that utilising the K-NN method, fresh data can be quickly and accurately sorted into a suitable category. Although the K-NN approach is most frequently employed for classification problems, it can also be utilised for regression. Since K-NN is a non-parametric technique, it makes no assumptions about the underlying data.

It is also called a lazy learner algorithm because it does not learn from the training set immediately instead it stores the dataset and at the time of classification, it performs an action on the dataset. KNN algorithm at the training phase just stores the dataset and when it gets new data, then it classifies that data into a category that is much similar to the new data. Example: Suppose, we have an image of a creature that looks similar to cat and dog, but we want to know either it is a cat or dog. So for this identification, we can use the KNN algorithm, as it works on a similarity measure. Our KNN model will look for similarities between the new data set's features and those in the photos of cats and dogs, and based on those similarities, it will classify the new data set as either cat- or dog-related. Why is a K-NN algorithm necessary? If there are two categories, Category A and Category B, and we have a new data point, x1, which category does this data point belong in? We require a K-NN algorithm to address this kind of issue. K-NN makes it simple to determine the category or class of a given dataset. Take a look at the diagram below: Deep Learning domain How does K-NN function? The following algorithm can be used to describe how the K-NN works: Step 1: Decide on the neighbours' K-numbers. Compute the Euclidean distance between K neighbours in step two. Step 3: Based on the determined Euclidean distance, select the K closest neighbours. Step 4: Count the number of data points in each category among these k neighbours. Step 5: Allocate the fresh data points to the category where the neighbour count is highest. Step six: Our model is complete. Let's say we need to classify a new data point in order to use it. Think on the photo below: Now, we'll decide on the number of neighbours; we'll go with k=5. The Euclidean distance between the data points will then be determined. The distance between two points is known as the Euclidean distance. which our geometry class has already covered. It can be determined as follows: By calculating the Euclidean distance, we were able to determine who our closest neighbours were: three in group A and two in category B. Think on the photo below:

Architecture Diagram



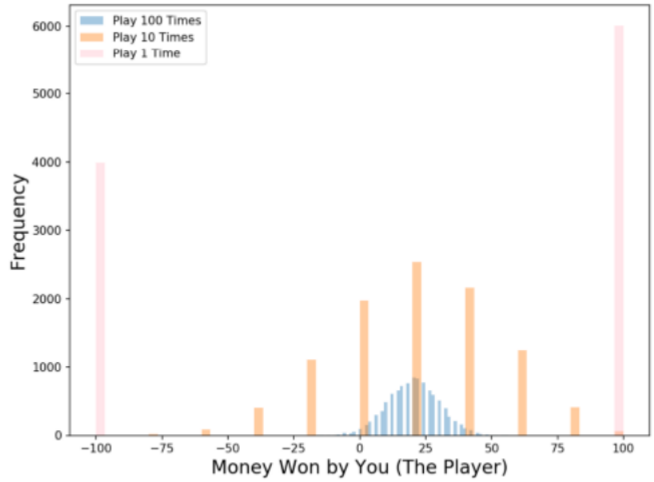
Algorithm 1 Causal KNN Algorithm

procedure CAUSALKNN(k, y, w, x)
  Input:
    $k \rightarrow$ number of nearest neighbours
    $y \rightarrow$ vector with outcome values
    $w \rightarrow$ vector with treatment status
    $x \rightarrow$ matrix with covariates

  Begin
  1. select nearest neighbours for each observation i based on the covariates x
  2. separate k treated $(w_i = 1)$ and k untreated $(w_i = 0)$ nearest neighbours
  for each observation i **do**
    calculate mean of $y_k(w = 1)$ and $y_k(w = 0)$
    $uplift_i$ = mean of $y_k(w = 1)$ - mean of $y_k(w = 0)$
  End

RESULTS

The comprehensive findings from the two approaches that were compared are presented in this section. We discuss the outcomes of this first, and then the results of this second. The described dataset is used for both the training and testing. We have created an application that may take the patient's existing characteristics and suggest a specific fibrosis stage in order to assess the effectiveness of the suggested diagnosis formula. The main GUI form, which is depicted in Figure, allows doctors to enter details about each patient. Additionally, the inference procedure can be transparently linked to the patient's medical file to continuously monitor the patient's health.



CONCLUSION

In this, we've put up a foundation for creating a FIS that uses this method. In order to choose the most pertinent features, we made use of information gained. There are many different input features. As a result, we developed the original system using the subtractive clustering method. The Liver Fibrosis Prediction in Hepatitis Patients training procedure has used the hybrid optimization method. We trained a number of models with varying numbers of epochs and varied configurations of SC radii and squash factors in order to determine which architecture was the most appropriate. The same dataset was used to train all models and test them. When choosing the best architecture for this project, overfitting was taken into account. Using RMSE, We have chosen the most reliable and understandable model. This produced results with a high testing accuracy of 93.3%. Comparing our chosen methods to the medical approach and the other two FIS techniques, WM and FDT, revealed that they performed better by 0.366, 0.171, and 0.066, respectively. We will expand on this work in the future to address additional medical issues. We'll also make an effort to include different kinds of knowledge, like clinical practise recommendations, into the decision-making process.

REFERENCES

[1] Jain, A.K., Ross, A., Prabhakar, S.: 'An introduction to biometric recognition', IEEE Trans. Circuits Syst. Video Technol., 2004, 14, pp. 4–20

[2] Jain, A.K., Flynn, P., Ross, A.A.: 'Handbook of biometrics' (Springer, 2008)

[3] Jain, A.K., Nandakumar, K., Nagar, A.: 'Biometric template security', EURASIP J. Adv. Signal Process., 2008, 2008, pp. 1–17

[4] Cavoukian, A., Stoianov, A.: 'Biometric encryption in encyclopaedia of biometrics' (Springer Verlag, 2009)

[5] Rathgeb, C., Uhl, A.: 'A survey on biometric cryptosystems and cancelable biometrics', EURASIP J. Inf. Secur., 2011, 2011, (3), doi: 10.1186/1687-417X-2011-3

[6] Uludag, U., Pankanti, S., Prabhakar, S., et al.: 'Biometric cryptosystems: issues and challenges', Proc. IEEE, 2004, 92, (6), pp. 948–960

[7] Ratha, N., Connell, J., Bolle, R.: 'Enhancing security and privacy in biometrics-based authentication systems', IBM Syst. J., 2001, 40, (3), pp. 614–634

[8] Juels, A., Rivest, R.L.: 'Honeywords: Making password-cracking detectable'. Proc. ACM SIGSAC Conf. on Computer and Communications Security, 2013, pp. 145–160