International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
ISSN:2147-6799 www.ijisae.org Original Research Paper

Prediction and Classification of Fatty Liver Disease Thesis- A Survey

Appanaboyina Sindhuja [*1], Dr. Seetharam Khetavath[2]

Submitted: 25/11/2023 **Revised**: 27/12/2023 **Accepted**: 06/01/2024

Abstract Liver illnesses have recently become the disorder with the highest mortality rate in a lot of countries. Consuming alcohol, breathing in toxic gases, eating food that has gone bad, and taking medications has all contributed to an increase in the number of persons diagnosed with liver disease. Research on patient data sets pertaining to the liver is being carried out in attempt to construct classification models that can anticipate liver disorder. This data set was utilised in the implementation of prediction and classification algorithms, which in turn reduced the amount of work that needed to be done by clinicians. In this research, we explore various machine learning techniques that can be used to evaluate a patient's liver condition comprehensively. The term "chronic liver disorder" refers to any condition that affects the liver and lasts for at least six months. As a consequence of this, we shall make use of the percentage of patients who become infected with the disease as both a positive and a negative piece of information. The percentages of liver disease are being discussed in this work with classifiers, and the findings are being presented in the form of a confusion matrix. We presented numerous classification strategies that, when used in conjunction with a data set for training, have the potential to significantly increase classification performance. Then, using a classifier that was learned by machine learning, the values are separated into good and poor categories. This article presents and discusses the various methods that can be used for the forecasting and categorization of fatty liver disease.

Keywords: *Fatty liver, Machine Learning, Neural Networks, Segmentation, Classification, Feature Extraction*

1. Introduction

The liver is the most important organ in the human body's overall structure. The liver is responsible for the breakdown of insulin. Glucuronidation in the liver is responsible for the breakdown of bilirubin, which in turn facilitates the passage of bile [1]. In addition to this, it is responsible for the breakdown and elimination of a great deal of waste products. It demonstrates a significant function in the transformation of hazardous substances. It demonstrates a significant function in the breakdown of certain pharmaceutical drugs. The term for this process is drug metabolism. The weight would be approximately 1.3 kilogrammes. The liver is divided into two enormous regions, which are known as the privileged portion and the left estimate, respectively. Below the liver and next to the pancreas is where you'll find the gallbladder. Along with the other organs mentioned, the liver contributes to both the consumption and the provision of nutrients. It is responsible for facilitating the movement of any foreign substances that may be present in the blood coming from the stomach, before delivering this blood to the remainder of the body. Diseases of the liver can manifest themselves when the liver's normal function is disrupted or when the organ sustains any kind of injury [2].

11Research Scholar, Department of ECE, Chaitanya Deemed to be University, Hanamkonda , Warangal ,Telangana ,India.
2Professor, Department of ECE, Chaitanya Deemed to be University, Hanamkonda , Warangal ,Telangana ,India.
seetharamkhetavath@gmail.com
** Corresponding Author Email: sindhu4b3@gmail.com*

The progression of liver problems [3] is a complex and changeable process that is controlled by a variety of different factors that together establish a person's susceptibility to disease. Some of these factors include gender, race, ethnicity, genetics, environmental factors (such as viruses, alcohol, nutrition, and toxins), body mass index (BMI), and the presence of other disorders such as diabetes. Liver illnesses, which can be fatal, are associated with a high mortality rate because they are potentially fatal conditions. The first step in determining the prognosis of liver problems is to do the standard tests on the urine and blood. In light of the symptoms that were seen, a liver function test (also known as an LFT) is advised for the patient [4].

Diseases of the liver are a major public health concern that impact millions of individuals all over the world. Better patient outcomes and less strain on the healthcare system can be achieved by earlier diagnosis of liver illnesses and more precise classification of the conditions that affect the liver. A major health problem is known as non-alcoholic fatty liver disease (NAFLD), and it affects one-third of adults and an increasing proportion of children and adolescents in wealthy countries [5]. The first symptom of the disorder is an abnormal accumulation of triglycerides in the liver. This abnormal buildup of triglycerides creates an inflammatory reaction in certain people, which in turn can progress to cirrhosis and liver cancer. Although there is a strong association between obesity, insulin resistance, and non-alcoholic fatty liver disease (NAFLD), the pathophysiology of NAFLD remains poorly understood, and treatment options are limited. While there is a correlation between obesity, insulin resistance, and

NAFLD, there is also a significant correlation between insulin resistance and NAFLD. However, techniques based on machine learning have shown positive results in predicting and classifying liver disorders based on patient data. These methods are able to recognise patterns and forecast outcomes because they make use of complex algorithms to do analysis on vast datasets and learn from their experiences. The application of techniques from machine learning in the prediction and categorization of liver disease is a vibrant area of research that is seeing continuous breakthroughs to improve accuracy and reduce the expenses of medical care.

II. Literature Review

A. **Overview of Liver Disease**

An disturbance in the function of the liver can lead to illness, which is what is meant by the term "liver disease" [4]. Because the liver is responsible for a large number of essential tasks within the body, any damage or infection that causes it to stop performing these duties can have a major negative effect on an individual's overall health. There is also a condition known as hepatic disorder that can be used to denote liver illness [6]. This all-encompassing word refers to a variety of potential conditions in which the liver is unable to carry out the functions for which it is responsible. Even if only one-fourth of the liver is still operating while the rest of the organ is damaged, the effectiveness of this organ will still be significantly diminished. The liver is the largest solid structure in the human body. It is classified as a gland due to the fact that, among its numerous functions, it is responsible for the production and secretion of bile. The rib cage provides protection for the liver, which is located in the more vertical portion of the abdomen. It possesses two main lobes that are interspersed with lobules of varying sizes. The cells that make up the liver have two different blood bases at their disposal. The hepatic artery is responsible for transporting oxygen-rich blood propelled by the heart, whereas the portal vein is responsible for bringing nutrients from the intestines. The vein's primary function is to transport blood from all of the body's other organs to the heart. However, the portal vein also carries nutrients from the digestive tract to the liver, where they are processed and purified before being returned to the bloodstream. This is an exception to the vein's primary function. The portal vein is responsible for the efficient delivery of the chemicals that liver cells need in order to produce the proteins, cholesterol, and glycogen that are necessary for the body to carry out its regular functions.

B. Common Risk Factors for Liver Disease: There are many different kinds of behaviours that might cause liver diseases [7]. The following are the categories:

1) Infection: The liver is susceptible to infection from parasites and viruses, which can result in inflammation or edoema and diminish the liver's ability to perform its function. The virus that causes damage to the liver is often transmitted through blood or sperm and is most commonly brought on by consuming tainted food, drinking contaminated water, or coming into contact with a person who is afflicted. People can get infections of the liver known as hepatitis A, hepatitis C, and hepatitis B.

2) Abnormalities of the immune system Certain diseases might cause the body's immune system to turn on other sections of the body and attack them. The liver is also impacted by this condition. There is a possibility that these disorders are autoimmune hepatitis. Additionally, primary biliary cholangitis and primary sclerosing cholangitis are two possible diagnoses for this condition.

3) Inheritance: A rare gene, which can be genetically inherited from either of your parents, can produce an accumulation of numerous substances in the liver, which can cause harm to the liver. There are a number of inherited conditions that can affect the liver, including Wilson's disease, hemochromatosis, and alpha-1 antitrypsin deficiency.

4) Cancer and other types of progressions: Liver adenoma, bile duct cancer, and liver cancer are all types of cancers that have the potential to cause liver disorders.

5) Other causes: The most common causes are chronic alcoholism, a buildup of fat in the liver (also known as NAFLD), the use of particular medications or over-the-counter remedies, and the consumption of specific herbal combinations.

6) Risk factors: Factors that might raise the risk of liver diseases include excessive consumption of alcohol, being overweight, having type 2 diabetes, having tattoos or body piercings, injecting drugs with used needles, receiving a blood transfusion, being exposed to foreign blood, having unprotected sexual encounters, being exposed to chemicals, and having a family history of liver disease.

The chemical industry Compounds in the Liver Chemicals such as bilirubin, albumin, alkaline phosphatase, aspartate aminotransferase, and globulin may be found in the liver, and each of them plays an important part in the healthy functioning of the organ on a day-to-day basis.

1) Bilirubin: Bilirubin is a yellowish complex that is produced in the typical catabolic pathway that breaks down heme in vertebrates. This pathway is known as the heme catabolic pathway. Bile and urine emit it. Diseases are caused when there is an excessive amount of bilirubin in the body. The pigment bilirubin is responsible for the yellow colour that develops on cuts and for the yellow staining that occurs in jaundice. The brown colour of faeces can be traced

back to the subsequent breakdown products of this compound, such as stercobilin. Urobilin, an additional byproduct of the breakdown process, is the primary component responsible for the straw-yellow colour of urine.

2) Alkaline phosphatase: In humans, alkaline phosphatase is present in all tissues throughout the body, although it is mostly concentrated in the liver, intestinal mucosa, bile duct, bone, kidney, and placenta. It is also present in small amounts in the spleen and testicles. The predominant types of alkaline phosphatase isozymes in the serum come from the skeletal and hepatic systems. During childhood, the majority of the alkaline phosphatase that is produced comes from the skeleton. The majority of animals, including humans, possess one or more of the following types of alkaline phosphatases:

• ALPI: It has a molecular mass of 150 kDa and is found in the intestinal tract.

• ALPL is a tissue-nonspecific protein that is predominantly found in the bone, liver, and kidneys.

• ALPP: This isozyme originates in the placenta and is also known by the name Regan.

• GCAP: This particular cell is a germ cell.

3) Aspartate aminotransferase: an enzyme of the aminotransferase family.

The levels of AST are highest in the liver and the heart. AST is present, albeit in lower concentrations, in both the kidneys and the muscles. It is present in very small amounts in human blood. The AST is released into the bloodstream if there is damage to the cells of the muscle or liver. As a result, the AST test will be helpful for monitoring or diagnosing liver damage or functioning.

4) Albumins are examples of proteins that are globular in structure. Albumins found in serum are extremely common and are the most important protein found in blood. It is able to bind thyroxine (T4), water, cations such as Ca2+ and Na+, hormones, bilirubin, fatty acids, and medicines. The primary function of this organ is to manage and maintain the blood's proper oncotic pressure. It is able to bind a number of fatty acids, as well as cations and bilirubin.

5) Globulin: These are globules of protein in the body. At the level of the molecules, they have a greater mass than albumin. It is not going to dissolve in water, but it will solvate in salt solutions that are very diluted. There are some globulins produced by the liver. The human blood of healthy individuals absorbs approximately 2.6-3.5 g/dL of globulin. There are many subtypes of globulins, the most common of which are beta, alpha 1, alpha 2, and gamma globulins. An imbalance in the body's chemistry, brought on by an excessive amount of one or more of these compounds, can be brought on by kidney illness. These are regarded as characteristics.

There is an infinite number of distinct types of liver diseases, each of which may be traced back to a certain ratio of these chemicals stored. Non-alcoholic fatty liver disease, often known as NAFLD, is the most common form of liver disease in many regions of the world [8]. NAFLD refers to the buildup of fat in the liver parenchyma that is not caused by the ingestion of alcohol. NAFLD comprises a range of disorders, ranging from simple steatosis to inflammatory steatohepatitis (NASH), with increasing levels of fibrosis, and finally cirrhosis [9]. Simple steatosis can be treated with diet and lifestyle changes, while NASH requires medical treatment. The number of people who have NAFLD is rapidly growing, which has made it a problem for public health all around the world [10]. There are estimates that the condition affects anywhere from 14% to 30% of the general population [11,12]. In addition, there is a rising prevalence across the world of both obesity and type II diabetes [13]. There is an estimate that the prevalence of NAFLD in Western and Asian countries is between 20% and 30%, and that prevalence is roughly 15% among the adult population [11,14,15]. According to the findings from the Dallas Heart Study, NAFLD affects thirty percent of individuals in the United States [16]. The reported prevalence of this disease in Iran ranges widely from 2.8% to 24.0% depending on the age group [17,18] and geographical region of the study. These percentages reach a high of 21.5 and 15.3 percent in southern Iran [19,20].

The accurate and fast diagnosis of NAFLD cases would produce considerable benefits, which is important due to the negative outcomes that would result from a delayed diagnosis of NAFLD. Accordingly, the development of a trustworthy clinical decision-making guide for its diagnosis is one of the primary focuses of a large number of clinical research [21]. Not only does the early identification of persons at high risk avoid damage to hepatocytes, but it also reduces the side consequences of NAFLD, such as heart failure, which is one of the leading causes of mortality in patients with fatty liver disease [22-25]. It is necessary to identify important risk factors and the linkages between them (21, 26), since this will allow for the most effective treatment to be selected for each individual patient. To this point, there have been no efforts made to realise these objectives through the utilisation of a classification tree (CT) strategy. A CT is a method of statistical learning that does not rely on parameters and can be applied to the creation of predictive models (21). This technique examines a large number of variables and organises them into categories based on the links between them and the significance of those relationships. A series of decision criteria can be used to recursively partition the data into subsets using predictors as the organising principle. CTs are frequently utilised as a solution to a variety of practical decision-making issues [13, 14] due to a number of essential

characteristics, including their flexibility and their hierarchical structure.

III. Materials

The current investigation was a cross-sectional study that was carried out in the city of Kavar, which is located in the province of Fars in southern Iran, from January to August of 2013. The city has a population of 75,000 people and a temperature similar to that of the Mediterranean; it is situated 35 kilometres from the city of Shiraz. A combined total of 1,600 people were chosen to participate in the study through the use of a stratified approach as well as a multiple-stage cluster random selection method. This means, first and foremost, that urban areas and rural communities were believed to be stratified. The town was subdivided into sections (called clusters) depending on the location of the hospitals. Some of these centres were chosen at random, and the participants were selected at random using the family registration data that is accessible at these centres. The samples were chosen at random from within the specified communities using the information gathered from the health houses. Researchers looked at thirty different factors, including demographic and clinical variables, in order to determine whether or not NAFLD was present in a person (with or without NAFLD). Participants in the study who had a history of liver cirrhosis, underlying liver disease, or hepatobilliary malignancies, as well as those who consumed more than 20 grammes of alcohol per day or who were using anti-thyroid medication, were not allowed to participate. For the purpose of this investigation, a questionnaire was developed to collect demographic information (such as age, gender, marital status, level of education, etc.), as well as medical history and health-related behaviours, such as smoking habits. The waist-hip ratio (WHR) of the participants was determined by taking their waist circumference and dividing that number by their hip circumference. The body mass index (BMI) was determined by taking their weight in kilogrammes and dividing that number by their height in square metres.

According to the body mass index (BMI) criteria established by the World Health Organisation (WHO), the subjects were divided into the following categories: 18.5 (underweight), 18.5 - 24.9 (normal), 25 - 29.9 (overweight), and over 30 (obesity) for both the male and female participants [27]. After an overnight fast, each of the volunteers was required to report to the clinic the next day. A group of nurses and doctors conducted interviews, filled out questionnaires, and drew intravenous blood samples from each participant in order to determine the levels of triglycerides (TG), cholesterol (CHO), high-density lipoprotein (HDL), alanine aminotransferase (ALT), and aspartate aminotransferase (AST), glucose (GLU), HBSAG, and HBSAB, as well as albumin (AL). In addition, the systolic blood pressure (SBP) and diastolic blood pressure (DBP) of each participant were taken three times, and the mean of these readings was recorded as a reference. Because liver enzyme levels are sensitive indicators for the diagnosis of NAFLD [27], we decided to test them as one of our clinical factors.

We discovered that the level of GGT in the serum was usually raised in people who had NAFLD. This was associated with advanced fibrosis and increased mortality in people who had NAFLD [28]. GGT was able to accurately predict advanced fibrosis with a sensitivity of 83% and a specificity of 69% when using a cutoff serum GGT value of 96.5 U/L [29]. The procedure of calibrating and then validating the calibration of the spectrophotometers for low-level readings was extremely sensitive to the user's methodology as well as the environment in which the measurements were taken. Interferences generated by bubbles, particle pollution, and stray light became substantial problems as observed activity levels decreased below 1.0 nephelometric turbidity units (NTU). One-point and two-point calibrations are the most frequent approaches, although there are various other strategies to reduce the impact of these mistakes. The criteria that we used to measure the variables for GGT resulted in an increase in the test's sensitivity to 53%, but it resulted in a decline in the test's specificity to 75%. [30] Research found that a positive predictive value of 90% for NAFLD was achieved with an increased transaminase level. Even though measuring transaminases as a population-based screening test is not without its flaws, it is nonetheless commonly employed in clinical practise to stratify individuals who have risk factors for nonalcoholic fatty liver disease (NAFLD) [31].

A trans-abdominal-sonography-calibrated sonography machine was used to make the diagnosis of NAFLD based on increased echogenicity of the liver parenchyma as well as attenuation of the portal vein or echogenicity of the diaphragmatic area. Before the study began, the sonographers went through training [32]. The study was assessed by the ethics committee of Shiraz University of Medical Sciences, and it was given the green light (Code: 92-6869, Date:May 11, 2012). Everyone who took part in the research gave their informed consent after reading and understanding the study's goals.

A. Methods

CTs are a non-parametric classification method that builds decision trees in a hierarchical structure. Root, internal, and leaf nodes are the components that make up a CT structure. Therefore, CT can be used to model the relationships among the predictors and their interactions in influencing the outcome [33-36]. Each branch node represents a choice that can be made between a number of possibilities, and each leaf node represents a classification or judgement. The creation of a CT is dependent on the following three primary factors: 1) Selecting a sampling-splitting rule that will define the tree branch that will be connected to the

classification nodes; 2) Evaluating the Classification that is Produced by the Splitting Rule at Each Node; and 3) Selecting an Optimal or Final Tree by Making Use of the Criteria for Classification Purposes [37].

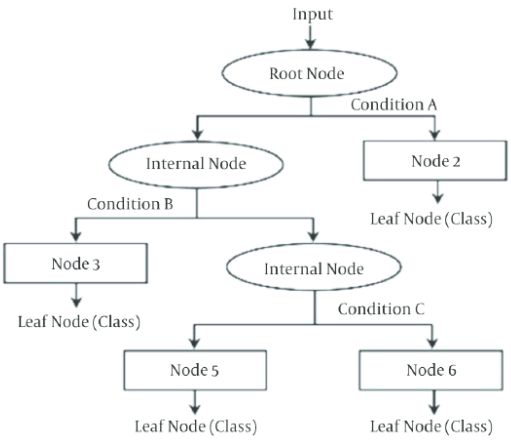

Fig 1. CT Structure

IV. Machine Learning and Liver Disease

Classification algorithms are frequently defined to be used for forecasting liver disease because they can help determine whether a patient has the disease or not depending on specific qualities or characteristics. This is one of the reasons why classification algorithms are frequently defined to be used for forecasting liver disease. It was observed that the F-Tree method has the highest accuracy among the algorithms that were examined, making it a suitable alternative for forecasting liver illness. This was determined based on the solutions that were already available. The classification of liver illnesses frequently makes use of feature selection in conjunction with fuzzy K-means classification algorithms. These techniques can be helpful in identifying crucial characteristics that can be utilised to differentiate between the many forms of liver problems. Using a fuzzy-based classification system might help improve the performance of the classification process since it takes into account the degree to which examples are similar to one another. This is because the same attribute values may be present in several liver illnesses. [38], researchers classified liver illnesses using a number of different classification algorithms. These algorithms included J48, SVM, RF, and MLP. The performance of these state-of-the-art algorithms was examined and contrasted in the study, with measures such as data correctness, data effectiveness, and correction rate serving as the basis for the evaluation. According to the results of the study, the accuracy produced by the multilayer perceptron algorithm was superior to that of the other algorithms that were investigated during the course of the study. The researchers in this study [39] utilised Bayesian classification in order to differentiate between a variety of liver disorders, such as cirrhosis, hepatitis, and diseases that are not related to the liver. They categorised the patients in the liver patient dataset into various subtypes of sickness using both the Naive Bayes and FT tree algorithms and then examined the effectiveness of these methods in terms of their accuracy and the amount of time it took for them to execute. According to the findings of their research, the Naive Bayes algorithm fared better than the FT tree method when it came to how long it took to execute.

The authors of this paper [40] investigated the efficacy of several classification approaches in terms of making accurate diagnoses of liver disorders. The Naive Bayes approach, the Decision Tree method, the Multilayer Perceptron method, the Random Forest method, and the Logistic Regression method were among these methods. Metrics like as precision, recall, sensitivity, and specificity were utilised by the writers in order to carry out the performance evaluation. When compared to the other algorithms that were investigated, the results demonstrated that Naive Bayes attained the highest level of precision. In addition, when recall was taken into account, it was discovered that the Logistic Regression and Random Forest algorithms both produced satisfactory results.

The WEKA tool was used in the study that was cited [41] in order to construct a model for evaluating liver disease. The Naive Bayes, Decision Tree (DT), and J48 algorithms were utilised in order to develop the model that was suggested by them. The accuracy of the algorithms as well as the amount of time it took to execute them were measured, and the findings were compared to those of the many options that were available. The research showed that the J48 and DT algorithms were more accurate than the Naive Bayes method. These algorithms outperformed the Naive Bayes algorithm.

Using a dataset consisting of Indian individuals suffering from liver illness, [42] executes a number of different classification algorithms on it. These methods include LR, K-NN, and SVM. In addition to that, confusion matrices were utilised in order to compare and contrast the various algorithms. The experimental examination of these algorithms was looked at and analysed in terms of how correct they were. According to the findings, the LR approach and the KNN technique are both capable of attaining a significant degree of accuracy; however, the LR technique has a higher sensitivity. Given this information, it is reasonable to draw the conclusion that LR is an effective method for disease prediction. It was suggested in [43], [44], and [45] that a novel feature model may be created using the classification techniques of Random Forest, SVM, J48, Bayesian Net, and MLP.

In order to investigate and carry out a comparison study for the purpose of predicting liver disease, three separate approaches are utilised. During the preliminary phase, the normalisation technique known as min-max was investigated and put through its paces using the dataset that

contained records of liver disorders. In the second stage, the PSO feature selection approach is utilised in order to choose the essential components of the dataset for the purpose of forecasting liver sickness. During the third phase, an evaluation of the methods' performance based on their accuracy was carried out, and the implementation of the classification algorithms was carried out. Following the results of the experimental study, one may draw the inference that the J48 technique performed better than PSO feature selection did.

By employing a variety of sample and oversampling strategies, the authors of this study [46] sought to achieve data homogeneity in order to make reliable projections on liver disorders. The J48 method, Multilayer perceptron, Random Forest, Multiple linear regression, Support Vector Machines, and Genetic programming were among the classification algorithms that were utilised. When it came to forecasting liver illnesses, the Random Forest algorithm that included oversampling at greater rates displayed the best performance. In a similar manner, this work [47] classified liver illnesses using a number of different classification algorithms. These classification algorithms included Naive Bayes, FT tree, J48, SVM, RF, MLP, K-NN, and LR. The performance of these algorithms was improved using a variety of methods, including feature selection, fuzzy K-means classification, normalisation, PSO feature selection, oversampling, and undersampling, among others.

The effectiveness of these algorithms was analysed using a variety of metrics, including accuracy, execution time, precision, recall, sensitivity, specificity, and F-measures, among others. According to the findings of these investigations, the performance of various algorithms can vary significantly depending on the dataset and the application [48]. As a result, it is essential to analyse the performance of a number of different algorithms and methods in order to determine which strategy is the most effective for a particular dataset and application. Although it has been shown that the results can vary depending on the dataset that was used and the particular algorithm that was applied, it has also been discovered that some algorithms, such as Multilayer perceptron, Random Forest, Support Vector Machines, and Artificial Neural Network, generally do better than others when it comes to predicting liver illnesses.

In the present day, ML has developed into a widely general routine. The application of machine learning (ML) as a tool to assist in therapeutic and pharmacological diagnostics is growing. Nevertheless, one thing that is consistent is the steady expansion of access to massive volumes of data.

The information that is presented in the following content provides a comprehensive assessment of the most significant issues and the most common solutions to ML challenges [49] in medicine. The goal of feature selection is to lessen the difficulty of the problem by selecting a subset of the valuable features contained in the input and getting rid of the features that are left over. The preceding statement makes it abundantly evident that several studies have been carried out utilising a wide variety of machine learning algorithms to make predictions about liver illnesses and assess how well these algorithms function.

In these investigations, a variety of methods, including Naive Bayes, Decision Trees, Random Forest, Logistic Regression, K-NN, SVM, Artificial Neural Networks, and C4.5, were utilised. The effectiveness of these algorithms has been assessed using a variety of metrics, including accuracy, precision, recall, F1-score, specificity, and execution time, among others. It should also be mentioned that different research come to different conclusions regarding the algorithm that has the best performance; this is because the dataset and assessment criteria that are employed can have a significant impact.

A. Logistic Regression

The previous model has been rebranded as the Logit model [50]. It is used to mimic the possibility of a particular class or event prevailing, such as passing or failing an exam, being well or ill, being alive or dead, or winning or losing a game. It is a model based on mathematics. In its simplest version, it makes use of a logistic method to advance a binary dependent variable, but it may also be expanded in a number of different, more intricate ways. A variable that is dependent and has two alternative values, such as pass/fail, which represents an indicator variable, will be included in a binary logistic model [51], and these two values will be treated as 0 and 1, respectively.

Support Vector Machine, Option B

It seeks to identify a suitable hyperplane that can partition the available data into a variety of categories. Utilising the scikit-learn package is one method for putting the Support Vector Machine (SVM) [52] algorithm into practise in Python. The information is cleaned up and separated into two distinct sets: the test data and the training sets. Twenty out of a total of one hundred samples will be used for the test. The number one hundred sixty out of every two hundred is always used as the set for the development of the algorithm. Either a set of hyperplanes or an SVM can be formed using an SVM. It is constructed in a place with an extension that is immeasurable. When a hyperplane achieves a decent separation, also known as a functional margin, it does so by having the greatest distance possible between itself and the nearest data point of any class. To boil it all down, a larger margin indicates that the classifier has a lower level of simplification defect.

Convolutional Neural Networks is the third option.

[53] CNN is a division of DNN. Its primary function is to facilitate the investigation of various imageries. It is also possible to refer to them as space invariant ANN or shift invariant ANN. They share a bulks architecture between themselves. Convolution is the name of the mathematical process that the network creates at this point. They are purely neural networks that perform convolution in place of generic matrix multiplication in at least one of their layers. This occurs in at least one of their layers.

MLP Classifier, Letter D

A section of (ANN) is denoted by the notation MLP [54]. Additionally known as feed-forward ANN. An MLP includes the nodes on the bottom of three sheets. The first input layer, the mid-hidden layer, and the final output layer are all contained within the three sheets. In addition, with the exception of the node that serves as the input, every other node in the network functions as a neuron that employs a nonlinear activation. When it comes to training, MLP uses a method that is referred to as backpropagation, which is an administered learning system. MLP [55] is distinguishable from linear perceptron due to the manifold layers it possesses as well as the non-linear inducement it employs. E. Random Forest The random forest algorithm [56] uses data samples to generate decision trees, and then it derives its forecast from all of the decision trees that it has constructed. In the end, it selects the explanation that received the most votes based on its clarity.

The information is cleaned up and separated into two distinct sets: the test data and the training sets. Twenty out of a total of one hundred samples will be used for the test. The number one hundred sixty out of every two hundred is always used as the set for the development of the algorithm. The data are broken up into a great deal of different groups and subgroups by the programme. In the event that a single person creates lines between the data points that belong to a subgroup, lines that connect the subgroups that make up the group, and so on. The erection resembles a tree in its appearance. The hyperplane that provides a consistent separation across classes is the one that minimises the distance to the next closest data point in the training set and maximises it.

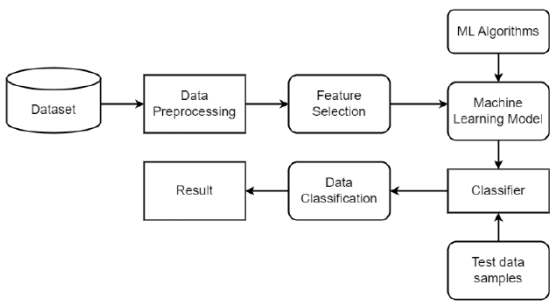

Fig 2. Processing of ML algorithms

By analysing photos of liver biopsies, the author intended to develop a method that does not require patient participation in order to diagnose liver illness. The author considered using images obtained from a biopsy since these images have a reasonable possibility of distinguishing between a liver that is unwell and one that is healthy. Using image study and deep learning to select an effective CNN [57] architecture and then carrying out the rest of the plan is the strategy that is currently being considered. NAFLD is very prevalent. According to the findings of an investigation, NAFLD is the root cause of about forty percent of all liver diseases that occur around the world. The presence of nonalcoholic fatty liver disease (NAFLD) in the liver can be identified by the sign of hepatic steatosis, in addition to other causes of fat accumulation, such as heavy alcohol intake, long-term use of drugs that cause steatosis, and inherited conditions. In the procedure that has been suggested, biopsy images of the liver are obtained, and the hepatic structures that are visible in these images are analysed with the assistance of two CNN [58] systems that share the same architecture. They intend to create a four-class detection system that can initially identify sinusoids, inflated hepatocytes, veins, and fat droplets. In the last stage of the process, all of the components of the detecting system are brought together and integrated with one another. After that, it determines the ratio between fat and ballooning. The ratio that was discovered will help in reaching a conclusion regarding the patient's condition. They used seven hundred and twenty different photos from the liver biopsy. Six hundred and twenty of these photographs are utilised in the validation process, sixty of these images are utilised in the testing process, and the remaining forty images are utilised in the testing process. These photographs were created with a resolution of 10,000 by 10,000 pixels or higher from the beginning. The portion of each image in which the tissue is still present is extracted, and the resulting images have a resolution of 64 by 64 pixels. B. Assessing the Accuracy of Liver Disease Prediction Using Machine Learning The utilisation of and modelling of medical data sets is currently being addressed by specialists all around the world. The primary objective of this approach is to cut down on the amount of time that passes between performing the liver test and creating the report and final result. They implemented several Machine Learning algorithms [59] such as decision trees, naïve Bayes trees, artificial neural networks, and random forests. After that, the Pearson correlation [60] is utilised in order to identify the irregularities, which include TP (true positives), FN (false negatives), FP (false positives), and TN (true negatives). This is done in order to determine the level of precision, specificity, and affectability possessed by the employed algorithms. Using previously established formulae, the sensitivity, affectability, specificity, and accuracy of the test are computed based on the words that are created. The author's intention was to design a user interface that would allow the

patient's report to be entered by the user as an input. After that, the algorithms are given some skill training by making use of the data set that has been allotted, and the output of the user input information is determined.

The result will consist of a single number, which will either be a zero or a one. A binary one will indicate that the patient's liver is sickly, while a binary zero will indicate that the patient's liver is healthy. The data that is given as input by the user will be logged, and then that data will be used to train the algorithm once more. The algorithm will be trained and skilled to progress with greater precision when additional tabular values are incorporated. The data set was retrieved by the authors from the UCI computer, which is the site. These individual algorithms' final outputs are charted here for your viewing pleasure. These grids illustrate the outcomes, which provide cause for optimism. Accuracy not cited. When developing the model, they took use of machine learning algorithms such as ANN, Naive Bayes, and Decision tree. Despite the fact that there is a wide variety of liver conditions, the authors focused on the most common forms in order to simplify the process and arrive at accurate findings.

C. The Division of the Liver Utilising CT Scan and the Localization of Disease

The authors sparked an idea by demonstrating that liver disease can be detected by the utilisation of abdominal CT. Standard X-ray technology is unable to provide a clear image of many inside organs. These are the circumstances that reinforce the incentives to use CT scans rather than X-rays because CT scans are able to display the structures more clearly. The resolution of these images, which were created via a CT scan, will be correct. They suggested employing WTA in order to segment the scan image, determine the location of the liver, and separate it from the background. In the final stage of the process, the percentage of the region that was impacted is computed. One needs to make use of a method that is quite accurate, and that method is a CT scan, in order to gain an understanding of the severity of the progression of liver disease. In the field of medicine, the CT scan is put to wide use in order to obtain information about humanoid structures. In the first step of the process, the imageries are analysed in order to identify the various components.

The Erode and Dilate algorithm is applied to the photos so they can be processed in the subsequent stage. This is where the viewing point values are adjusted. In addition, they go through a process with WTA that divides the liver into sections. The WTA provides two outputs, which are the liver area and the non-liver area respectively. The term "cropping" refers to the product that is obtained as a result. The collected data is then converted into a binary format, with this organ represented by the colour white and the remaining data represented by black. Next, a median filter is applied in order to lessen the amount of noise and make the texture more even. After that, a formula is used to determine how much of the liver has been destroyed. According to the authors, the average level of precision for image breakdown is approximately 81%, while the average level of precision for liver breakdown is approximately 92%. The WTA was utilised by the author in order to differentiate the liver. Another strategy that has been suggested is to make use of the binary threshold in order to separate the healthy areas of the liver from the sick areas.

D. The Prediction of Liver Disease Through the Use of Classification Algorithms

The procedure that was suggested by the authors demonstrates that machine learning is capable of not only predicting the disease but also recognising hidden patterns for the purposes of diagnostics and decision-making. The steadily expanding number of reported cases of liver problems is seen as a widespread health concern around the globe. The purpose of this thesis is to present findings that are reliable in diagnosing liver illness by making use of classification algorithms [61]. The algorithms known as Logistic Regression, K-Nearest Neighbour, and Support Vector Machines are the ones that are utilised for work of this nature. These are the algorithms that go by the name Classification. The box plot is a representation of the outliers that are present in the attribute algorithms [62], which are largely utilised in the process of disease prediction using machine learning algorithms. In the field of medicine, machine learning techniques are proving to be quite beneficial in the prediction of diseases based on medical databases. There are researchers and scientists working on virtually every continent who are diligently applying machine learning models [63] in conjunction with classification algorithms in order to strategically improve medical diagnostics and are achieving improved results as a result of their efforts. In this particular piece of research, the prediction of liver illness is accomplished by the utilisation of Logistic Regression, K-Nearest Neighbour, and Support Vector Machines. It is well knowledge that the liver is the largest internal organ in the body. It is responsible for a variety of important processes, including the synthesis of glycogen and bile, the generation of triglycerides and cholesterol, and the manufacture of blood clotting factors and proteins. In most instances, significantly more than 75% of the liver tissue needs to be affected in order for there to be a reduction in function. Because of this, it is of the utmost importance to recognise the condition at an early stage, when it is still treatable, before it reaches a more severe state.

V. Conclusion

Fatty liver disease, often known as FLD, is a prevalent clinical issue that is not only associated with a high morbidity but also a high mortality rate. FLD ultimately

results in cirrhosis that does not involve cholestasis as well as hepatocellular cancer. In addition, the incidence of FLD has been on the rise concurrently with the rise in the prevalence of diabetes, metabolic syndrome, and obesity. A more widespread occurrence of FLD has been identified as a bigger economic burden. Therefore, accurate identification of individuals who are at risk and early diagnosis of Fatty Liver Disease (FLD) could offer enormous benefits to stratify patients and be considered as a diagnostic reference standard for the evaluation of fatty infiltration of the liver. However, this procedure is not only extremely intrusive and expensive, but it also carries the risk of producing unwanted side effects and errors in sampling when it is put into practise. Despite the fact that ultrasonography is currently being used as a diagnostic tool for FLD with a greater level of accuracy, the identification accuracy is heavily dependent on the practitioner. Machine learning, commonly known as ML, is a subfield of computer science that makes use of computer algorithms to recognise patterns in vast amounts of data. These patterns can then be used to assist in making predictions about various outcomes based on the data. In a wide variety of fields, machine learning techniques have recently emerged as a promising instrument for predictive analysis and decision-making. As a result of the availability of clinical data, machine learning has also been playing an important part in the process of making medical decisions. The creation of a model based on machine learning would be a useful aid in determining the presence of disease and arriving at an efficient clinical decision in real time. Additionally, it would make it possible to optimise the use of hospital resources by early categorising the appropriate patients as having significant numbers of risk indicators. In this article, we have included a brief assessment and discussion of the various methodologies for the diagnosis and classification of fatty liver.

VI. Reference

[1]. A. Arjmand, C. T. Angelis, A. T. Tzallas, M. G. Tsipouras, E. Glavas,R. Forlano, P. Manousou, and N. Giannakeas, “Deep learning in liverbiopsies using convolutional neural networks,” in 2019 42nd InternationalConference on Telecommunications and Signal Processing (TSP).IEEE, 2019, pp. 496–499.

[2]. L. A. Auxilia, “Accuracy prediction using machine learning techniquesfor indian patient liver disease,” in 2018 2nd International Conferenceon Trends in Electronics and Informatics (ICOEI). IEEE, 2018, pp.45–50.

[3]. A. Spann, A. Yasodhara, J. Kang, K. Watt, B. Wang, A. Goldenberg,and M. Bhat, “Applying machine learning in liver disease and transplantation:a comprehensive review,” Hepatology, vol. 71, no. 3, pp.1093–1105, 2020.

[4]. S. Sontakke, J. Lohokare, and R. Dani, “Diagnosis of liver diseasesusing machine learning,” in 2017 International Conference on EmergingTrends & Innovation in ICT (ICEI). IEEE, 2017, pp. 129–133.

[5]. J. C. Cohen, J. D. Horton, and H. H. Hobbs, “Human fatty liver disease:old questions and new insights,” Science, vol. 332, no. 6037, pp. 1519–1523, 2011.

[6]. F. Himmah, R. Sigit, and T. Harsono, “Segmentation of liver usingabdominal ct scan to detection liver desease area,” in 2018 InternationalElectronics Symposium on Knowledge Creation and IntelligentComputing (IES-KCIC). IEEE, 2018, pp. 225–228.

[7]. M. B. Priya, P. L. Juliet, and P. Tamilselvi, “Performance analysis ofliver disease prediction using machine learning algorithms,” InternationalResearch Journal of Engineering and Technology (IRJET), vol. 5,no. 1, pp. 206–211, 2018.

[8]. Bellentani S, Marino M. Epidemiology and natural history of nonalcoholicfatty liver disease (NAFLD). *Ann Hepatol.* 2009;**8 Suppl 1**:S4–8. [PubMed: 19381118].

[9]. Dowman JK, Tomlinson JW, Newsome PN. Systematic review: the diagnosisand staging of non-alcoholic fatty liver disease and nonalcoholicsteatohepatitis. *Aliment Pharmacol Ther.* 2011;**33**(5):525–40.doi: 10.1111/j.1365-2036.2010.04556.x. [PubMed: 21198708].

[10]. Kelishadi R, Poursafa P. Obesity and air pollution: global riskfactors for pediatric non-alcoholic fatty liver disease. *Hepat Mon.*2011;**11**(10):794–802. doi: 10.5812/kowsar.1735143X.746. [PubMed:22224077].

[11]. Browning JD, Szczepaniak LS, Dobbins R, Nuremberg P, Horton JD, CohenJC, et al. Prevalence of hepatic steatosis in an urban population inthe United States: impact of ethnicity. *Hepatology.* 2004;**40**(6):1387–95. doi: 10.1002/hep.20466. [PubMed: 15565570].

[12]. Kelishadi R, Farajian S, Mirlohi M. Probiotics as a novel treatment fornon-alcoholic Fatty liver disease; a systematic review on the currentevidences. *Hepat Mon.* 2013;**13**(4):ee7233. doi: 10.5812/hepatmon.7233.[PubMed: 23885277].

[13]. Bjornsson E, Angulo P. Non-alcoholic fatty liver disease. *Scand JGastroenterol.*2007;**42**(9):1023–30.doi: 10.1080/00365520701514529.[PubMed: 17710666].

[14]. Bedogni G, Miglioli L, Masutti F, Tiribelli C, Marchesini G, Bellentani S.Prevalence of and risk factors for nonalcoholic fatty liver disease: theDionysos nutrition and liver study. *Hepatology.* 2005;**42**(1):44–52. doi:10.1002/hep.20734. [PubMed: 15895401].

[15]. Chitturi S,Wong VW, Farrell G. Nonalcoholic fatty liver in Asia: Firmlyentrenched and rapidly gaining ground. *J Gastroenterol Hepatol.*2011;**26 Suppl 1**:163–72. doi: 10.1111/j.1440-1746.2010.06548.x. [PubMed:21199528].

[16]. Browning JD. Statins and hepatic steatosis: perspectives fromthe Dallas Heart Study. *Hepatology.* 2006;**44**(2):466–71. doi:10.1002/hep.21248. [PubMed: 16871575].

[17]. Alavian SM, Mohammad-Alizadeh AH, Esna-Ashari F, Ardalan G, HajarizadehB. Non-alcoholic fatty liver disease prevalence amongschool-aged children and adolescents in Iran and its association withbiochemical and anthropometric measures. *Liver Int.* 2009;**29**(2):159–63. doi: 10.1111/j.1478-3231.2008.01790.x. [PubMed: 18492015].

[18]. Jamali R, Khonsari M, Merat S, Khoshnia M, Jafari E, Bahram Kalhori A,et al. Persistent alanine aminotransferase elevation among the generalIranian population: prevalence and causes.*World J Gastroenterol.*2008;**14**(18):2867–71. [PubMed: 18473412].

[19]. Lankarani KB, Ghaffarpasand F, Mahmoodi M, Lotfi M, Zamiri N, HeydariST. Non alcoholic fatty liver disease in southern Iran: a populationbased study. *Hepatitis Mon.* 2013;**13**(5).

[20]. Eshraghian A, Dabbaghmanesh MH, Eshraghian H, Fattahi MR, OmraniGR. Nonalcoholic fatty liver disease in a cluster of Iranian population:thyroid status and metabolic risk factors. *Arch Iran Med.*2013;**16**(10):584–9. [PubMed: 24093139].

[21]. Breiman L. Classification and regression trees. USA: Chapman & Hall;1984.

[22]. Marchesini G, Moscatiello S, Di Domizio S, Forlani G. Obesityassociatedliver disease. *J Clin Endocrinol Metab.* 2008;**93**(11 Suppl1):S74–80. doi: 10.1210/jc.2008-1399. [PubMed: 18987273].

[23]. Targher G, Arcaro G. Non-alcoholic fatty liver disease and increasedrisk of cardiovascular disease. *Atherosclerosis.* 2007;**191**(2):235–40. doi:10.1016/j.atherosclerosis.2006.08.021. [PubMed: 16970951].

[24]. Targher G, Marra F, Marchesini G. Increased risk of cardiovasculardisease in non-alcoholic fatty liver disease: causal effect or epiphenomenon?.*Diabetologia.* 2008;**51**(11):1947–53. doi: 10.1007/s00125-008-1135-4. [PubMed: 18762907].

[25]. Ballestri S, Lonardo A, Bonapace S, Byrne CD, Loria P, Targher G.Risk of cardiovascular, cardiac and arrhythmic complications in patientswith non-alcoholic fatty liver disease. *World J Gastroenterol.*2014;**20**(7):1724–45. doi: 10.3748/wjg.v20.i7.1724. [PubMed: 24587651].

[26]. Piper ME, Loh WY, Smith SS, Japuntich SJ, Baker TB. Using decision treeanalysis to identify risk factors for relapse to smoking. *Subst Use Misuse.*2011;**46**(4):492–510. doi: 10.3109/10826081003682222. [PubMed:20397871].

[27]. Rokach L. Data Mining with Decision Trees. Incorporated; 2008.

[28]. Gorunescu F. Data Mining: Concepts, Models and Techniques. USA:Springer; 2011.

[29]. Anuurad E, Shiwaku K, Nogi A, Kitajima K, Enkhmaa B, Shimono K, etal. The new BMI criteria for asians by the regional office for the westernpacific region of WHO are suitable for screening of overweightto prevent metabolic syndrome in elder Japanese workers. *J OccupHealth.* 2003;**45**(6):335–43. [PubMed: 14676412].

[30]. Torres DM, Harrison SA. Diagnosis and therapy of nonalcoholicsteatohepatitis. *Gastroenterology.* 2008;**134**(6):1682–98. doi:10.1053/j.gastro.2008.02.077. [PubMed: 18471547].

[31]. Ruhl CE, Everhart JE. Elevated serum alanine aminotransferase andglutamyltransferase and mortality in the United States population.*Gastroenterology.* 2009;**136**(2):477–85.

[32]. Tahan V, Canbakan B, Balci H, Dane F, Akin H, Can G. Serum gammaglutamyltranspeptidasedistinguishes non-alcoholic fatty liver diseaseat high risk. *Hepato-Gastroenterol.* 2007;**55**(85):1433–8.

[33]. Choudhury J, Sanyal AJ. Clinical aspects of fatty liver disease. Seminarsin liver disease. .

[34]. Clark JM, Brancati FL, Diehl AM. Nonalcoholic fatty liver disease. *Gastroenterology.*2002;**122**(6):1649–57. [PubMed: 12016429].

[35]. Saverymuttu S, Joseph A, Maxwell J. Ultrasound scanning in the detectionof hepatic fibrosis and steatosis. *British Med J.* 1986;**292**(13).

[36]. Anyanwu MN, Shiva SG. Comparative analysis of serial decisiontree classification algorithms. *Int J Computer Science Security.*2009;**3**(3):230–40.

[37]. Maroco J, Silva D, Rodrigues A, Guerreiro M, Santana I, de MendoncaData mining methods in the prediction of Dementia: A real-datacomparison of the accuracy, sensitivity and specificity of linear discriminantanalysis, logistic regression, neural networks, support vectormachines, classification trees and random forests. *BMC Res Notes.*2011;**4**:299. doi: 10.1186/1756-0500-4-299. [PubMed: 21849043].

[38]. T. R. Baitharu and S. K. Pani, “Analysis of data mining techniquesfor healthcare decision support system using liver disorder dataset,”Procedia Computer Science, vol. 85, pp. 862–870, 2016.

[39]. U. R. Acharya, S. V. Sree, R. Ribeiro, G. Krishnamurthi, R. T. Marinho,J. Sanches, and J. S. Suri, “Data mining framework for fatty liver diseaseclassification in ultrasound: a hybrid feature extraction paradigm,”Medical physics, vol. 39, no. 7Part1, pp. 4255–4264, 2012.

[40]. N. Nahar and F. Ara, “Liver disease prediction by using differentdecision tree techniques,” International Journal of Data Mining &Knowledge Management Process, vol. 8, no. 2, pp. 01–09, 2018.

[41]. A. Naik and L. Samant, “Correlation review of classification algorithmusing data mining tool: Weka, rapidminer, tanagra, orange and knime,”Procedia Computer Science, vol. 85, pp. 662–668, 2016.

[42]. A. N. Arbain and B. Y. P. Balakrishnan, “A comparison of datamining algorithms for liver disease prediction on imbalanced data,”International Journal of

Data Science and Advanced Analytics (ISSN2563-4429), vol. 1, no. 1, pp. 1–11, 2019.
[43]. M. A. Kuzhippallil, C. Joseph, and A. Kannan, "Comparative analysisof machine learning techniques for indian liver disease patients,"in 2020 6th International Conference on Advanced Computing andCommunication Systems (ICACCS). IEEE, 2020, pp. 778–782.
[44]. K. R. Asish, A. Gupta, A. Kumar, A. Mason, M. K. Enduri, andS. Anamalamudi, "A tool for fake news detection using machine learningtechniques," in 2022 2nd International Conference on IntelligentTechnologies (CONIT). IEEE, 2022, pp. 1–6.
[45]. M. K. Enduri, A. R. Sangi, S. Anamalamudi, R. C. B. Manikanta,K. Y. Reddy, P. L. Yeswanth, S. K. S. Reddy, and G. A. Karthikeya,"Comparative study on sentimental analysis using machine learningtechniques," Mehran University Research Journal of Engineering andTechnology, vol. 42, no. 1, pp. 207–215, 2023.
[46]. M. Islam, C.-C. Wu, T. N. Poly, H.-C. Yang, Y.-C. J. Li et al.,"Applications of machine learning in fatty live disease prediction," inBuilding Continents of Knowledge in Oceans of Data: The Future ofCo-Created eHealth. IOS Press, 2018, pp. 166–170.
[47]. S. Mohanty, P. K. Gantayat, S. Dash, B. P. Mishra, and S. C. Barik,"Liver disease prediction using machine learning algorithm," in DataEngineering and Intelligent Computing: Proceedings of ICICC 2020.Springer, 2021, pp. 589–596.
[48]. C. Liang and L. Peng, "An automated diagnosis system of liverdisease using artificial immune and genetic algorithms," JOURNAL OFMEDICAL SYSTEMS, vol. 37, no. 2, 2013.
[49]. R. A. Khan, Y. Luo, and F.-X. Wu, "Machine learning based liverdisease diagnosis: A systematic review," Neurocomputing, vol. 468, pp.492–509, 2022.
[50]. A. S. Abdalrada, O. H. Yahya, A. H. M. Alaidi, N. A. Hussein,H. T. Alrikabi, and T. A.-Q. Al-Quraishi, "A predictive model for liverdisease progression based on logistic regression algorithm," Periodicalsof Engineering and Natural Sciences (PEN), vol. 7, no. 3, pp. 1255–1264, 2019.
[51]. F. E. Harrell, Jr and F. E. Harrell, "Binary logistic regression," Regressionmodeling strategies: With applications to linear models, logisticand ordinal regression, and survival analysis, pp. 219–274, 2015.
[52]. E. M. Hashem and M. S. Mabrouk, "A study of support vector machinealgorithm for liver disease diagnosis," American Journal of IntelligentSystems, vol. 4, no. 1, pp. 9–14, 2014.
[53]. Z. Yao, J. Li, Z. Guan, Y. Ye, and Y. Chen, "Liver disease screeningbased on densely connected deep neural networks," Neural Networks,vol. 123, pp. 299–304, 2020.
[54]. M. Abdar, N. Y. Yen, and J. C.-S. Hung, "Improving the diagnosis ofliver disease using multilayer perceptron neural network and boosteddecision trees," Journal of Medical and Biological Engineering, vol. 38,no. 6, pp. 953–965, 2018.
[55]. T. Bikku, "Multi-layered deep learning perceptron approach for healthrisk prediction," Journal of Big Data, vol. 7, no. 1, pp. 1–14, 2020.
[56]. T. A. Assegie, R. Subhashni, N. K. Kumar, J. P. Manivannan, P. Duraisamy,and M. F. Engidaye, "Random forest and support vectormachine based hybrid liver disease detection," Bulletin of ElectricalEngineering and Informatics, vol. 11, no. 3, pp. 1650–1656, 2022.
[57]. J. Gu, Z. Wang, J. Kuen, L. Ma, A. Shahroudy, B. Shuai, T. Liu,X. Wang, G. Wang, J. Cai et al., "Recent advances in convolutionalneural networks," Pattern recognition, vol. 77, pp. 354–377, 2018.
[58]. J. Murphy, "An overview of convolutional neural network architecturesfor deep learning," Microway Inc, pp. 1–22, 2016.
[59]. G. Bonaccorso, Machine learning algorithms. Packt Publishing Ltd,2017.
[60]. I. Cohen, Y. Huang, J. Chen, J. Benesty, J. Benesty, J. Chen, Y. Huang,and I. Cohen, "Pearson correlation coefficient," Noise reduction inspeech processing, pp. 1–4, 2009.
[61]. M. Ghosh, M. Raihan, M. Sarker, M. Raihan, L. Akter, A. K. Bairagi,S. S. Alshamrani, and M. Masud, "A comparative analysis of machinelearning algorithms to predict liver disease." Intelligent Automation &Soft Computing, vol. 30, no. 3, 2021.
[62]. F. Osisanwo, J. Akinsola, O. Awodele, J. Hinmikaiye, O. Olakanmi, andJ. Akinjobi, "Supervised machine learning algorithms: classification andcomparison," International Journal of Computer Trends and Technology(IJCTT), vol. 48, no. 3, pp. 128–138, 2017.
[63]. P. C. Sen, M. Hajra, and M. Ghosh, "Supervised classification algorithmsin machine learning: A survey and review," in Emergingtechnology in modelling and graphics. Springer, 2020, pp. 99–111.