

Application of Biochemical Tests and Machine Learning Techniques to Diagnose and Evaluate Liver Disease

Shamima Akter¹, Hossain Uddin Shekhar^{2*}, Sharif Akhteruzzaman³

¹Department of Bioinformatics and Computational Biology, George Mason University, Fairfax, VA, USA

²Department Biochemistry & Molecular Biology, University of Dhaka, Dhaka, Bangladesh

³Department Genetic Engineering & Biotechnology, University of Dhaka, Dhaka, Bangladesh

Email: *hossainshekhar@du.ac.bd

How to cite this paper: Akter, S., Shekhar, H.U. and Akhteruzzaman, S. (2021) Application of Biochemical Tests and Machine Learning Techniques to Diagnose and Evaluate Liver Disease. *Advances in Bioscience and Biotechnology*, 12, 154-172. <https://doi.org/10.4236/abb.2021.126011>

Received: May 11, 2021

Accepted: June 22, 2021

Published: June 25, 2021

Copyright © 2021 by author(s) and Scientific Research Publishing Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

Abstract

Background: The liver function tests (LFTs) remain one of the most commonly employed clinical measures for the diagnosis of hepatobiliary disease. LFTs sometimes referred to as hepatic panel help to determine the health of liver, monitor the progression of a disease and measure the severity of a disease particularly scarring or cirrhosis of the liver. **Aims:** In this study, we present a new approach to evaluate the natural progression of liver disease through the assessment of eight biochemical parameters: serum total bilirubin (TB), alanine aminotransferase (ALT), aspartate aminotransferase (AST), Alkaline phosphatase (ALP), total protein (TP), albumin (ALB), albumin/globulin (A/G) ratio, and alpha-fetoprotein (AFP) as well as two machine learning (ML) tools—Random Forest and CART to substantiate the outcome. **Methods:** The study was carried out in a total of 100 subjects which included healthy controls (group I-25 patients), patients with acute hepatitis (group II-25 patients), chronic hepatitis (group III-25 patients) and hepatocellular carcinoma (group IV-25 patients) applying both biochemical and Machine Learning methods. **Results:** Of the eight parameters tested, all except ALP ($p = 0.426$), showed an overall discriminatory ability as judged by one-factor analysis of variance ($p < 0.0001$). We also assessed the differences among group means by least significance difference (LSD). The analysis showed that TB remained significantly elevated in groups II, III, and IV as compared to controls ($p < 0.05$). ALP did not have any discriminatory power among the four groups tested. ALT and AST were good discriminators only between the control groups and groups II and III. TP, ALB, and A/G ratio were decreased significantly in groups III and IV as compared to controls. Group III and IV were almost indistinguishable using these biochemical parameters except for AFP, which was found to be elevated only in group IV.

The accuracy of classification into different liver patient groups using random Forest and CART was 94% and 95% respectively. **Conclusion:** Acute hepatitis (group II) shows a higher level of AST, ALT and ALP compared to chronic hepatitis (group III) and hepatocellular carcinoma (group IV). Two machine learning algorithms also predicted and supported the same biochemical results by correctly classifying liver disease patients. We also recommend that the AFP test can be performed if hepatocellular carcinoma is suspected.

Keywords

Hepatitis, Hepatocellular Carcinoma, Bilirubin, Transaminases, Cholestasis, Machine Learning, Random Forest, CART

1. Introduction

As a body's largest organ, the liver plays an important role in our whole body's blood transfer and controls the concentrations of most of the chemical and waste products in our blood. Thus, it is important to keep the liver healthy. Parasites, viruses cause inflammation and decrease function by infecting the liver, subsequently causing Liver Disease (LD) [1].

LD is a common clinical disorder; it is also associated with high morbidity and mortality [2]. Additionally, LD has been increasing in parallel with the prevalence of diabetes, metabolic syndrome, alcohol and obesity [3]. Higher prevalence of LD has appeared as a greater economic burden. Therefore, accurate identification of individuals at risk and early recognition of LD could offer immense benefits for diagnosis, prevention, or even proper treatment. Subsequently, reliance on a single diagnostics test is not sufficient to evaluate liver function [4]. A wide variety of biochemical measures are therefore used to determine the general condition of the liver.

Different biochemical tests commonly referred to as Liver Function Tests (LFT) provide secondary evidence for hepatic diseases [5]. Metrical record, physical examination along with diagnostic test (LFTs) results entail to 1) recognize patients with liver disease; 2) diagnosis of differential jaundice; 3) monitor the severity (*i.e.*, course and response of the disease); and 4) detect hepatotoxicity caused by various agents [6]. In addition, commonly used LFTs are mainly used to determine liver damage instead of monitoring hepatic functions which can make the identification of disease complicated [7]. Certainly, these biochemical tests can also detect problems such as hemolysis (high bilirubin), higher alkaline phosphatase level (bone disease). Abnormal LFTs often suggest that the liver may not function properly and indicate the severity of the problem. But still, the correctness and accuracy to predict liver disease remain uncertain [6].

In this essence, computer-based diagnosis methods/tools such as Machine Learning (ML) can help to predict liver diseases correctly with precision. The knowledge discovery of ML has made it possible to handle valuable data to en-

hance decision making both in medical diagnosis and prognosis. Researchers show potential interest in ML to support data mining, classification techniques (based on features or characteristics) towards medical diagnosis and prediction of liver diseases [8]. In medical settings, groups of patients can be diagnosed into different classes with respect to types and/or subtypes of diseases. In ML, classification is defined as a supervised method where training data is fitted into the model, then the model is trained with the dataset(s) with a known class of sample based on the features [9] [10]. Later the model predicts the test sample class which is unknown [9]. The classifier performance is evaluated by measuring the accuracy of classification that describes the percentage of correctly classified occurrences. ML provides the promise to improve the diagnosis and predict diseases that are of concern in liver diseases [8].

Several liver functions tests are carried out for the estimation and medication of hepatic dysfunction in patients. The biochemical markers: serum bilirubin, alanine aminotransferase (ALT), aspartate aminotransferase (AST), alkaline phosphatase (ALP), α -fetoprotein, 5' nucleosidase, ceruloplasmin are liver function tests [5] [11] [12]. Few other studies have indicated that ALT, AST, ALP, GGT, Bilirubin, prothrombin time, serum Albumin are the tests that are commonly performed in liver disease patients [13] [14]. In these studies, researchers highlighted the importance of these tests to reflect liver functionality. For instance, bilirubin presents the excretions of anions, transaminases explain the hepatocellular integrity, bilirubin, ALP describes the creation of bile and subsequent flow of bile freely, and albumin denotes the protein synthesis. Overall, researchers comprehensively presented all these biomarkers along with AFP, serum proteins to screen the liver functions [4]. Also, they found that AFP was increased in the case of hepatocellular carcinoma. In addition, in the case of asymptomatic patients, a mild increase of serum ALT levels are observed and around one-third of the patients show normal liver enzyme function persistently [6] [15].

Recently, authors proposed to build intelligent medical decision support systems by using ML through classification of liver disease and clustering to patterns that can benefit physicians for the treatment [16]. Other studies have recommended the data categorization based on liver ailment and used different algorithms (J-48, SVM, Random Forest tree, etc.) for classifying these liver disease conditions [17] [18]. For example, researchers use six ML methods; LR, KNN, DT, SVM, NB, RF to classify the patients with liver disease. They estimated accuracy, recall (specificity) and precision (sensitivity) of the applied methods while classifying the patients into groups [19]. In another study, CART (Classification and Regression Tree) was represented to detect liver disease patients and obtained 92.94% accuracy [20]. Authors applied classification algorithms; naïve bayes, C4.5 decision tree, back propagation, SVM and KNN. They compared the performance of classifiers to classify the patients on the basis of accuracy, precision, sensitivity and specificity [21]. Very recently in one study, authors pre-

dicted the risk factor of chronic kidney disease using different machine learning algorithms; Random forest, Decision Stump, Linear regression, Naïve Bayes, Simple logistic regression while classifying the CKD patients [22].

In this study, considering the importance and utilities of both of these methods, we strived to find two key research questions: 1) which biochemical parameter is significantly associated with different liver disease(s), and 2) assess/validate the outcome by using ML tools. To get the answers, we performed a cross-sectional study to a) evaluate liver patients using different biochemical markers (TB, ALT, AST, ALP, TP, ALB, A/G ratio, and AFP, and ALT, b) employ two prediction models (Random Forest and CART) of ML to support the findings. We aim to establish the relationships between conventional LFTs of liver with ML tools to verify the accuracy. We believe that the study helps clinicians to correctly identify and make an actionable decision of prevention, early diagnosis, and targeted intervention. The study provides a fresh insight about utilizing both the biochemical test and ML tools to predict liver disease. To our knowledge, the study is the first initiative to connect traditional biochemical tests with computational analysis and subsequently establishes the reliance of ML to support the outcome from biochemical tests.

2. Materials and Methods

Four different groups (Control, AH, CH, AH) of patient data provided by Shekhar Lab, Biochemistry and Molecular Biology, University of Dhaka, were analyzed in two different ways; 1) Testing biochemical parameters-serum total bilirubin (TB), alanine aminotransferase (ALT), aspartate aminotransferase (AST), Alkaline phosphatase (ALP), total protein (TP), albumin (ALB), albumin/globulin (A/G) ratio, and alpha-fetoprotein (AFP), followed by statistical analysis to obtain the distribution of these parameters in different groups. 2) Using ML methods-RF and CART to classify the samples into 4 different groups and obtain the accuracy of the classification model followed by finding the important features as biochemical important parameter for the liver disease. The process of the study has been shown in **Figure 1**.

2.1. Biochemical Tests

2.1.1. Data Collection

Patients were selected from a pathological laboratory of a diagnostic clinic, based on case records for a three years period. Samples were collected and subsequently analysed in collaboration with Shekhar Lab. Principal investigator Dr. Hussain Uddin Shekhar had the authorization to collect the samples. The written informed consent was taken from all the patients. The patient criteria were as follows: 1) The patients with acute viral hepatitis (AH) were diagnosed clinical features, serological investigations, and ultrasound scanning; 2) Chronic hepatitis (CH) Patients had persisted jaundice for more than six months without HC; 3) Diagnosis of HC had been confirmed by histological examination and biopsy of

a liver specimen. Hepatitis B as detected by HbsAg enzyme immunoassay kit (DiaSorin, Vercelli, Italy) was present in 67% with AVH, 75% in CH and 80% patient in HC. The controls were healthy volunteers and included students and university staff.

2.1.2. Blood Samples

A sample of venous blood was collected and allowed to clot. It was then centrifuged at 4°C for 10 minutes at 4000 rpm. The serum was then separated and stored at -20°C until analyzed. The process flow has been shown in **Figure 2**.

2.1.3. Laboratory Tests

The following routine liver function tests were executed by standard methods. The reference number—Serum bilirubin [23], serum alanine aminotransferase [24] and aspartate aminotransferase [24], serum alkaline phosphatase [25], serum total protein [26], serum albumin [27]—are followed to prepare samples, and to measure the quantity of each parameter as LFTs in those patients. The value of albumin was subtracted from the total protein and serum globulin concentration was calculated as the total protein value consists of both albumin and globulin values.

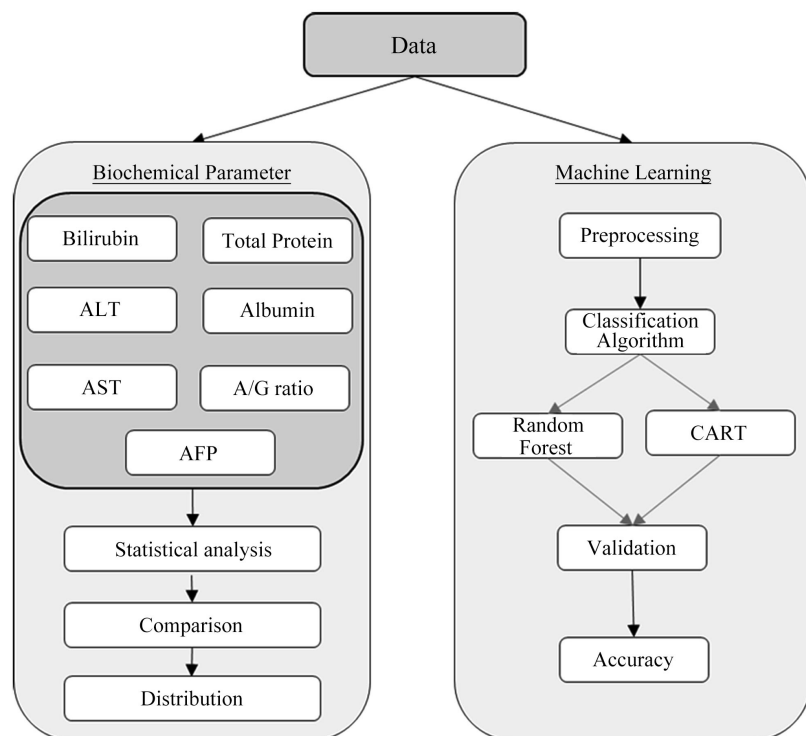


Figure 1. Data analyzing pipeline of different liver disease patients.

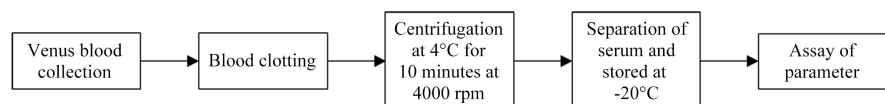


Figure 2. Process flow of the blood serum collection.

Bilirubin

Total and direct Bilirubin was measured by the modified Jendrassik-Grof Analysis method with centrifugal analyzer method [23].

Alpha-fetoprotein (AFP)

Serum alpha-fetoprotein levels were determined by microparticle enzyme immunoassay (MEIA) technology in IMx systems (Abbott Laboratories, IL 60064, USA) [28].

AST and ALT

To measure both of these two enzymes correctly, conditions were optimized with basic variables (buffer kind, conc of buffer, ion, pH) and kinetic parameters (inhibitor and Michaelis constant were determined) [24].

ALP

Serum ALP was measured using a test kit named “Iso-ALP”, Boehringer Mannheim. The principle of the assay is based on Rosalki and Foo (Clin Chem 1984; 30: 1182-6) [25].

Total protein

TP was determined based on the Biuret method [26].

Albumin

Serum albumin was measured using a prompt and consistent method with Bromocresol green. In this method, when albumin was added to 0.075M succinate buffer in pH4.20 with the Bromocresol solution, absorbance increased at 628 nm [27].

2.1.4. Statistical Analysis

All statistical analyses were done using the SPSS statistical package (Ver. 10.0). One-way analysis of variance (ANOVA) following a logarithmic transformation of the data was initially used to detect the overall difference in group means of the eight biochemical parameters. Differences among group means were assessed using the least significance (LSD) [29].

2.2. Machine Learning (ML)**2.2.1. Random Forest**

RF is an aggregate classifier which works efficiently on a large dataset. It is a regression classification method that performs through generating decision trees from randomly selected subset of training data and gives output class (*i.e.*, which is the output from individual trees [30]). There is no need to reduce variables during the analysis because RF can maintain thousands of input attributes easily. RF provides the estimation of variables which are important in the classification. During the generation of multiple trees, input vectors are put down to each of the decision trees in the forest to classify a new liver object. Each tree gets a vote for new data classification for accuracy. Most votes containing classes are chosen by the forest. The Following equations are employed to classify the liver data with RF model [31].

$$ni_j = w_j C_j - w_{\text{left}(j)} C_{\text{left}(j)} - w_{\text{right}(j)} C_{\text{right}(j)} \quad (1)$$

- ni_j = the importance of node j
- w_j = weighted number of samples reaching node j
- C_j = the impurity value of node j
- left(j) = child node from left split on node j
- right(j) = child node from right split on node j

$$f\hat{i}_i = \frac{\sum_{j:\text{node } j \text{ splits on feature } i} ni_j}{\sum_{k \in \text{all nodes}} ni_k} \tag{2}$$

- $f\hat{i}_i$ = the importance of feature i
- ni_j = the importance of node j

$$\text{norm}f\hat{i}_i = \frac{f\hat{i}_i}{\sum_{j \in \text{all features}} f\hat{i}_j} \tag{3}$$

$$\text{RF}f\hat{i}_i = \frac{\sum_{j \in \text{all trees}} \text{norm}f\hat{i}_{ij}}{T} \tag{4}$$

- $\text{RF}f\hat{i}_i$ = the importance of feature i calculated from all trees in the Random Forest model
- $\text{norm}f\hat{i}_{ij}$ = the normalized feature importance for i in tree j
- T = total number of trees

$$\text{Accuracy} = (\text{TP} + \text{TN}) / (\text{TP} + \text{TN} + \text{FP} + \text{FN}) \tag{5}$$

2.2.2. CART

CART (Classification and regression trees) is a ML classification model which helps to obtain a variable depending on other variables which are labeled and then predict the class through asking a set of if-else questions [32]. There are two advantages to using these models; 1) Nonlinear dataset is handled by this model, 2) data normalization or Standardization is not needed as distance or other quantitative parameters between data points are not calculated. In the construction of the tree model, there are three types of nodes (Root, Internal, and Leaf) involved in the tree. Each node has its own if-else question for variables that can direct to a specific leaf-node for the final prediction of class using decision boundaries [33]. Information Gain (IG) is a criterion to detect the purity of a node and can be measured depending on the split of items by a node. Corresponding impurity criterion is used to split the features (Table 1).

Table 1. CART classification based on impurity criterion described with equations by Scikit-learn and spark [31].

Impurity	Task	Formula	Description
Gini impurity	Classification	$\sum_{i=1}^C f_i(1 - f_i)$	f_i is the frequency of label i at node and C is the number of unique labels
Entropy	Classification	$\sum_{i=1}^C -f_i \log(f_i)$	f_i is the frequency of label i at node and C is the number of unique labels
Variance/Mean Square Error(MSE)	Regression	$\frac{1}{N} \sum_{i=1}^N (y_i - \mu)^2$	y_i is the label for an instance, N is the number of instance and μ is the mean given by $\frac{1}{N} \sum_{i=1}^N y_i$
Variance/Mean Absoluter Error (MAE) (Scikit-learn only)	Regression	$\frac{1}{N} \sum_{i=1}^N y_i - \mu $	y_i is the label for an instance, N is the number of instance and μ is the mean given by $\frac{1}{N} \sum_{i=1}^N y_i$

$$\text{Information Gain-Gain}(T, X) = \text{Entropy}(T) - \text{Entropy}(T, X) \quad (6)$$

- T = target variable
- X = Feature to be split on
- $\text{Entropy}(T, X)$ = The entropy calculated after the data is split on feature X

3. Results

3.1. Patient Characteristics

The summary data of the number and age of the various groups of patients and controls used in this study is represented in **Table 2**. The mean age did not differ significantly among groups. Inspection of the data showed clear evidence of non-Gaussian distributions. However, a transformation to log10 was successful in normalizing the data. All subsequent analyses were based on the logarithm of the raw values.

3.2. Biochemical

3.2.1. Comparison of Biomedical Parameters

There were large and statistically highly significant differences in means among the four groups of all eight biochemical parameters, as expected. These are summarized in **Table 3**, which also shows how the groups differed for each parameter. For example, the mean of the controls differed from that of all other groups for bilirubin, ALT and AST, but did not differ from AH and CH for AFP. **Table 3** shows the mean value (\pm SD) of each biochemical liver function tests for the four groups: Healthy individual (Control), AH, CH, and HC. In addition, p values were derived from ANOVA. Within each group, arithmetic means showing the same letter (*i.e.*, a or b subscript) are not significantly different from each other.

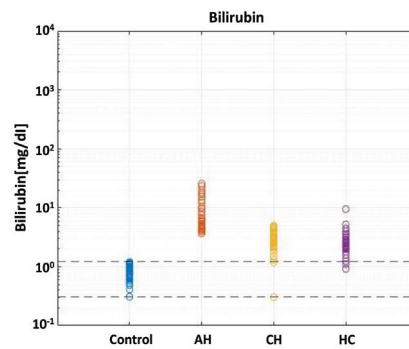
Although group means differed, there were considerable overlaps among individual patients for each biochemical parameter. The result is illustrated in **Figure 3(a)-(h)**. The figures show that many of the liver disease patients were clearly in pathological range, while others were within normal range. Therefore, it would not be possible to draw a line on any of these plots which would separate all the liver disease patients from controls, or which could distinguish between the liver disease groups.

Table 2. Summary data of patients and controls.

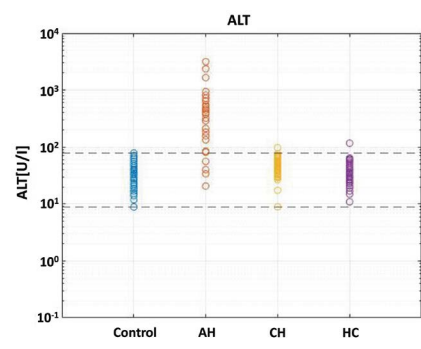
Group	Male	Female	Total	Age
	(n)	(n)	(n)	(years)
Control (Gr-I)	15	10	25	39.4 \pm 11.8
Acute Hepatitis (Gr-II)	13	12	25	42.3 \pm 11.5
Chronic Hepatitis (Gr-III)	11	14	25	40.7 \pm 11.0
Hepatocellular carcinoma (Gr-IV)	16	9	25	45.4 \pm 9.7

Table 3. Geometric Mean \pm SD values of biochemical parameters in four group of subjects.

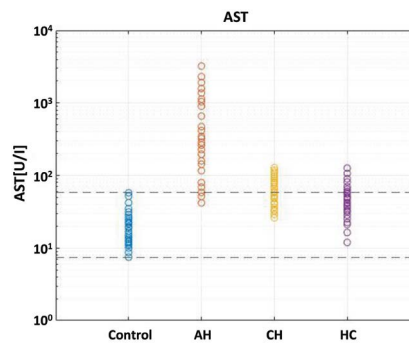
Value	Group				p value
	I (Control)	II (AVH)	III (CH)	IV (HC)	
AFP (IU/mL)	1.85 \pm 1.52 ^a	1.07 \pm 1.73 ^a	1.75 \pm 1.51 ^a	161 \pm 2.22	<0.0001
A/G ratio	1.36 \pm 1.25 ^a	1.36 \pm 1.44 ^a	1.10 \pm 1A7 ^b	0.92 \pm 1.46 ^b	<0.0001
Albumin (g/dL)	4.31 \pm 1.21 ^a	4.03 \pm 1.20 ^a	3.37 \pm 1.37 ^b	2.83 \pm 1.32 ^b	<0.0001
ALP (UL)	181 \pm 1 43 ^a	228 \pm 2.45 ^a	181 \pm 1.45 ^a	187 \pm 1.58 ^a	>0.05
ALT (UL)	26.0 \pm 1.65 ^a	264 \pm 3.64	42.1 \pm 1.58 ^b	34.8 \pm 1.68 ^{a/b}	<0.0001
AST (UL)	20.9 \pm 1 52	290 \pm 3.53	52.8 \pm 1.36 ^a	41.5 \pm 1.64 ^a	<0.0001
Bilirubin (mg/dL)	0.60 \pm 1.43	8 44 \pm 1.75	2.30 \pm 1.78 ^a	2.46 \pm 1.62 ^a	<0.0001
Total Protein (g/dL)	7.19 \pm 1.10 ^a	7.09 \pm 1.09 ^a	6.55 \pm 1.17 ^b	6.96 \pm 1.15 ^b	<0.0001



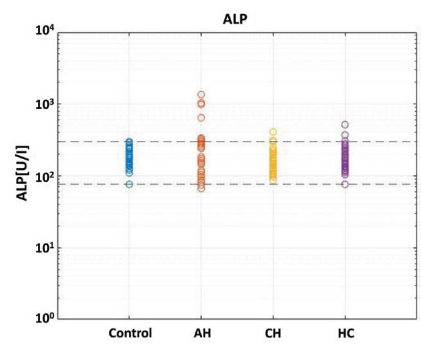
(a)



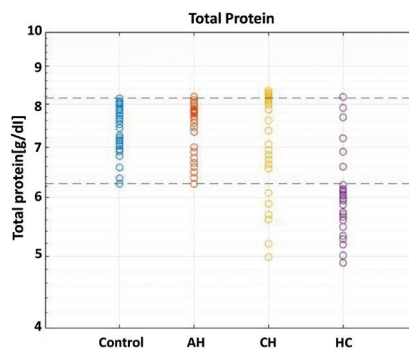
(b)



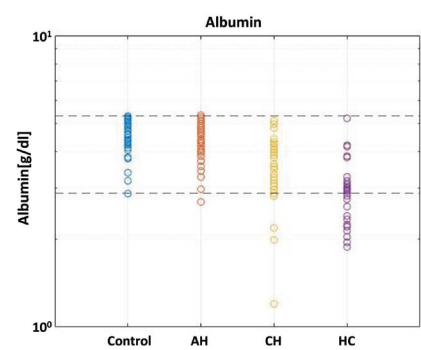
(c)



(d)



(e)



(f)

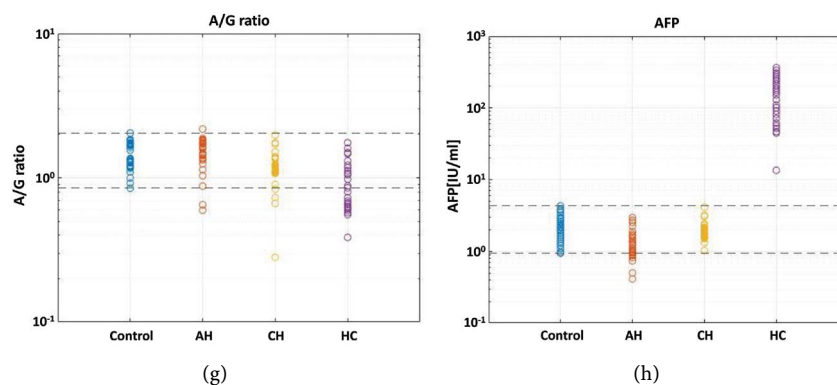


Figure 3. (a)-(h): Individual values of each of the eight biochemical parameters for control subjects and patients in each of the three diagnostic groups. Based lines show extremes of values in control subjects. The abbreviations are AH: Acute viral hepatitis, CH: Chronic hepatitis HC: Hepatocellular carcinoma.

3.2.2. Examination of Biochemical Parameters (with Respect to Bilirubin)

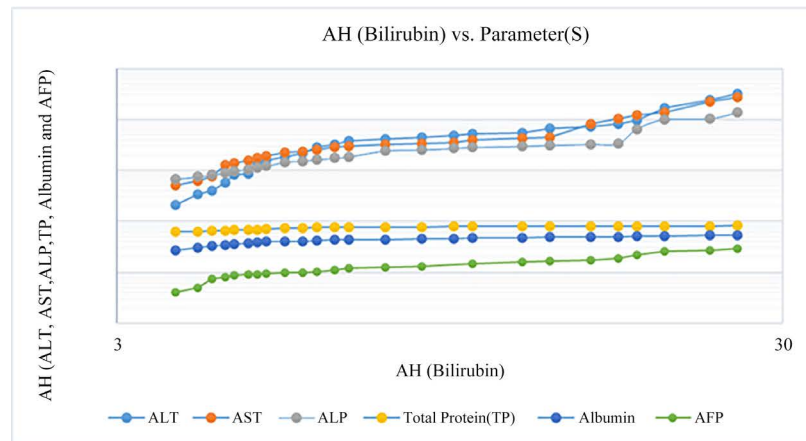
From **Figure 4**, it has been observed that in the AH group (**Figure 4(a)**), when the patient's bilirubin was low, then the concentrations of AST, ALT, and ALP were low and patients with high bilirubin values showed a high concentration of AST, ALT, and ALP enzymes that describe the severity of the disease in acute conditions. Other parameters such as total Protein, Albumin, and AFP of all the patients observed constant regarding elevated bilirubin level. As apparent from **Figure 4(b)**, in the case of the low and high value of bilirubin for CH, the corresponding liver enzymes *i.e.*, AST, ALT, and ALP showed higher values, respectively. The other parameters showed consistency with bilirubin value that elaborates insignificant change of the value of AFP, TP, or albumin with the higher value of bilirubin. We observed a similar (CH) outcome while comparing HC Bilirubin with other liver enzymes (**Figure 4(c)**), In the case of HC, Liver enzyme levels did not increase substantially and patients with high-value bilirubin did not show high value of ALT, AST. AFP concentration showed substantially higher than the other two enzymes. However, the values of other parameters showed a similar outcome with AH and CH. All these findings are consistent with the results shown in **Figure 3**.

3.3. Machine Learning

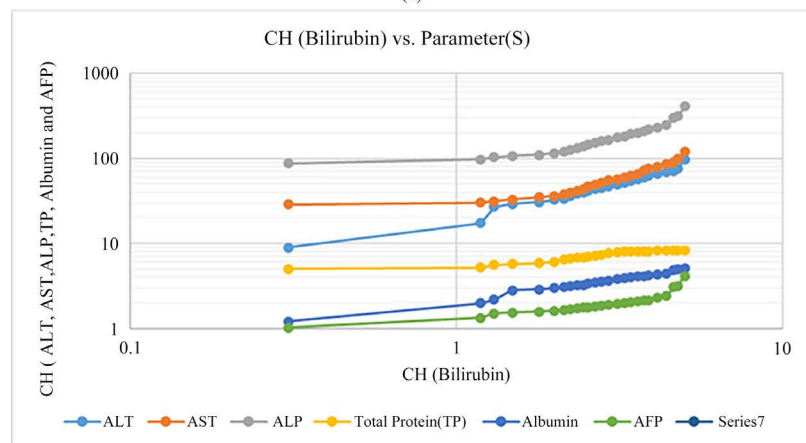
3.3.1. Random Forest

RF Classifier model have classified 100 patients into 4 groups (Control, AH, CH, HC). Out of 100 patients, the model has classified 95 patients correctly with whole data set (**Table 4**). When data was split into train (70%) and test (30%) samples, then model have classified 30 patients as test samples and 28 patients were classified correctly (**Table 5**). The model also has estimated the features and ranked them 1 to 7 based on feature importance (**Table 6**). From this estimation, it has been identified that Bilirubin is the most important parameter in diagnosing all 4 liver disease patients. The 2nd important parameter is the AFP;

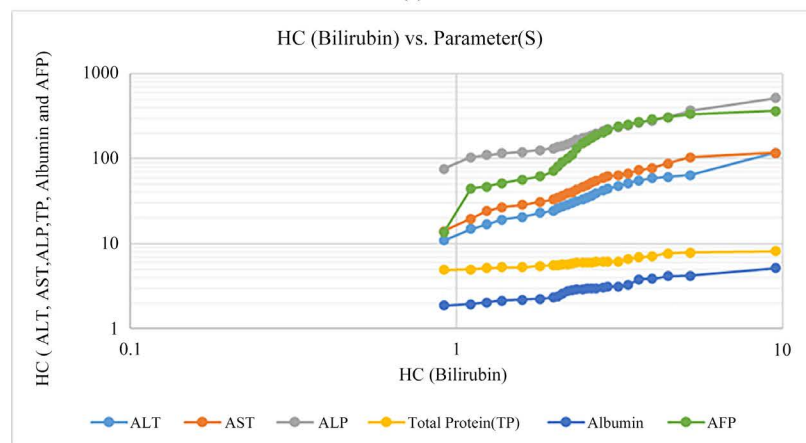
3rd one is AST and 4th one is ALT (Figure 5). Therefore, the model classified the patients correctly into different groups with 94% accuracy.



(a)



(b)



(c)

Figure 4. Shows the relationship of Bilirubin level with other parameters in three different groups *i.e.*, AH, CH, HC except control. For each group, x-axis denotes the conc of Bilirubin and y axis denotes the values of all other parameters in logarithmic scale. Liver enzymes *i.e.*, AST and ALT show high value with the high value of Bilirubin. AH, CH, HC groups are represented by the (a), (b), (c), respectively. Each dot represents the values of the corresponding parameter of each patient.

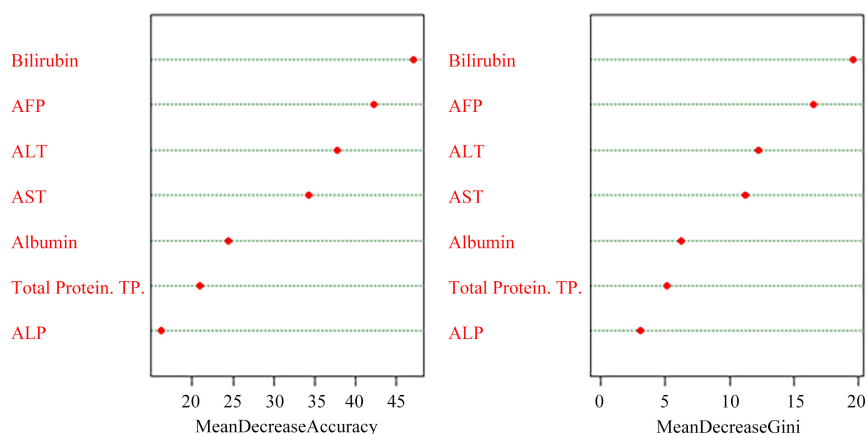


Figure 5. Variable important plot to rank the feature or parameter using RF model.

Table 4. Confusion matrix.

Liver Patient Class	AH	CH	Control	HC
AH	25	4	0	0
CH	0	19	0	0
Control	0	2	25	0
HC	0	0	0	25

Table 5. Confusion matrix for test data set for RF.

Liver Patients Group	AH	CH	Control	HC
AH	9	0	0	0
CH	0	6	2	0
Control	0	0	7	0
HC	0	0	0	6

Table 6. Feature ranking with score.

Ranking	Feature	score
1	Bilirubin	0.2623
7	AFP	0.2311
3	AST	0.1754
2	ALT	0.1359
6	Albumin	0.08
5	Total.Protein.TP.	0.0718
4	ALP	0.0434

3.3.2. CART

The CART classifier model correctly classified 94 patients correctly into 4 different groups out of 100 patients. The whole data is trained and predicts the class. The results are shown in **Table 7** as a confusion matrix. One of the impor-

tant findings of the classification tree is that out of 7.4 features; AFP, Bilirubin, ALT, AST were important in the classification process. These 4 (four) features are required to perform the classification. The results show that only AFP is required to classify the 25 HC patients correctly, Bilirubin is required for the Control group, and all patients were classified correctly. Both ALT and AST were required for AH and CH patients classification. All 25 AH and 21 CH patients were classified correctly, 6 patients were wrongly classified for HC patients (Figure 6).

Table 7. Confusion matrix.

Liver Patients Group	AH	CH	Control	HC
AH	23	2	0	0
CH	1	23	1	0
Control	0	0	25	0
HC	0	0	1	24

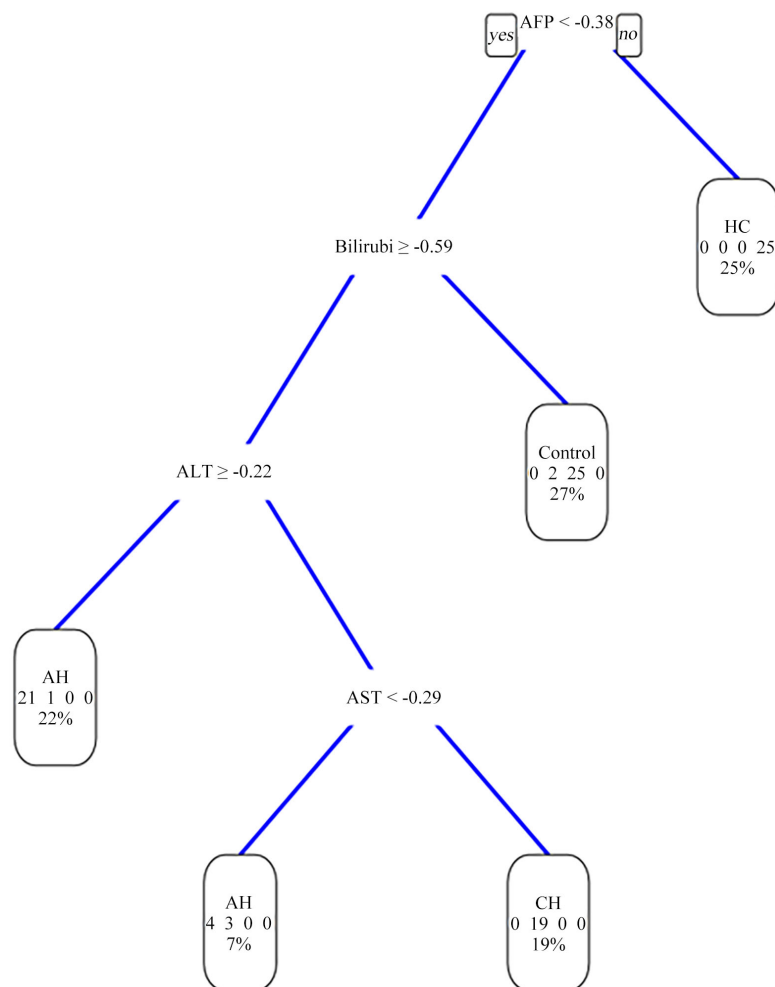


Figure 6. Classification Tree (CART) of the four liver diseases groups. Here Bilirubi indicates Bilirubin.

4. Discussion

4.1. Biochemical Tests

In this study, the diagnostic effectiveness of eight biochemical parameters were evaluated among four subject groups including controls and patients with three types of liver disease. The biochemical parameters were chosen to measure a range of known biochemical functions of the organ. Since the liver performs a wide variety of tasks, hence relying on a single test is not sufficient to evaluate liver function. Therefore, a wide variety of diagnostic tests is imperative for the indication of hepatobiliary disease along with ML classification scheme with two models *i.e.* RF & CART.

4.1.1. Bilirubin

Although a small amount of unconjugated (indirect) bilirubin is present in healthy people, there is virtually no observable conjugation (direct) bilirubin in the blood. This is due to the rapid secretion of conjugated bilirubin into the bile. Serum bilirubin levels will not increase no less than half of the liver's excretion potential is lost. In this study, we found that serum total bilirubin was increased significantly in the patients with acute hepatitis, and it showed no overlapping with the control group (**Figure 3(a)**). In groups III (chronic hepatitis) and IV (hepatocellular carcinoma) bilirubin levels, though remained elevated compared to the controls, were not as high as in acute hepatitis. Serum bilirubin help to determine abnormalities in hepatic uptake, conjugation and secretion [34].

4.1.2. ALT and AST

The levels of ALT and AST describe the most used markers of liver impairment. This study demonstrated a significant elevation of serum ALT and AST in the patients of acute hepatitis compared to controls (**Figure 3(b)**, **Figure 3(c)**). In group IV, *i.e.*, in hepatocellular carcinoma patients, ALT levels did not differ significantly from the controls. In hepatocellular carcinoma, apoptosis induced dead hepatocytes wither away and likely synthesize less of the enzymes. This is likely to clarify why most patients with hepatocellular carcinoma have persistently normal liver enzymes though having inflammation in the liver biopsy [35] [36]. The mean AST value in group IV, though was significantly higher than it was in the control group, nevertheless had considerable overlap with the control groups. The extrahepatic origin of this enzyme might explain this difference from ALT. Moreover, pyridoxine deficiency might be another reason behind this phenomenon. Though ALT formation is inhibited more strongly than that of AST by pyridoxine deficiency, both ALT and AST use pyridoxine as coenzymes [37]. In group III, *i.e.*, in chronic hepatitis patients, both ALT and AST values significantly higher than control groups but not as high as it was in case of group were acute hepatitis patients. The reason behind this is not well understood. It is observed from the recent studies that histological evidence can be found for chronic Hepatitis C Virus (HCV infection in patients, while normal or slightly elevated serum transaminases are frequently present [38].

4.1.3. ALP

Serum ALP values did not show any substantial divergence between controls and the three other patient groups (**Figure 3(d)**). The results obtained also suggest that ALP is not a good marker to identify hepatocellular injury or intrahepatic problems strikingly high ALP levels indicate the risk of extrahepatic biliary obstruction, primary liver cirrhosis and cholestasis triggered by drugs. Since this study did not recruit any such patient, the serum ALP value in all the four groups studied, were almost comparable to each other.

4.1.4. Serum Albumin

Serum total protein and A/G ratio are also indirect measures for the synthetic capacity of the liver, as most plasma proteins are synthesized in the liver. We also found in this study that total protein, albumin, and A/G ratio are decreased significantly in groups III (CH) and IV (HC) but not in group II (AH) (**Figure 3(e)**, **Figure 3(f)**, **Figure 3(g)**). The likely explanations are as follows: 1) Approximately, three weeks is the half-life of albumin, and it is a large life cycle, 2) the reduced synthetic ability of liver is compensated by the double production of albumin compared to normal synthesis rate. So, serum albumin concentration changes slowly in response to alterations in protein synthesis. This is possibly why serum albumin, total protein and A/G ratio in acute hepatitis are within the normal range. Overall, these three biochemical tests are positive indicators of chronic liver disease, but chronic renal failure, urinary protein loss or loss of gastrointestinal properties may affect levels [39] [40].

4.1.5. Alpha-Fetoprotein

In this study, AFP was assessed to discriminate hepatocellular carcinoma from other liver diseases. This study clearly demonstrates that AFP is elevated significantly only in group IV (HC) compared to controls (**Figure 3(h)**). Benign hepatic disease; for instance acute chronic active hepatitis, viral hepatitis, and liver cirrhosis occasionally show elevated AFP levels [41]. However, this study did not find any such association. Another important finding of this study is that groups III and IV are not distinguishable by any of the parameters used in this study except AFP.

4.2. Machine Learning (ML)

4.2.1. Random Forest

ML study supports the finding from the biochemical analysis. RF model estimates feature importance when predicting the class or groups of patients (**Figure 5**). This finding resulted Bilirubin as most important features that denotes the marker in diagnosing all 4 liver diseases. AFP turns out to be second important features that correlates with finding and discussion for biochemical findings as AFP can differentiate HC groups from other groups. ALT and AST are predicted another crucial feature in ML study correlating with biochemical results as these two parameters values highly increased in AH and CH patients. Feature ranking (**Figure 5**) shows that the first four variables (Bilirubin, AFP,

ALT, AST) turn out to be required to diagnose liver patients, other features have insignificant importance for classification that are also discussed in the biochemical part and evident from the result.

4.2.2. CART

From this ML classification using the CART model, it has been shown that HC patients are classified from the data of AFP that correlates with the experimental results (Section 3.2). AH and CH patients are classified from the data of AST and ALT, meaning that these enzymes are important markers for the diagnosis of AH, CH. This finding also correlates with the experimental results. From the tree, the 4 features (*i.e.* Bilirubin, ALT, AST, AFP) are shown and predicted to be important for classification and overall, from biochemical analysis also provided the same outcome.

5. Conclusion

In conclusion, the determination of abnormal liver tests requires close attention to the relevant data from case records as well as physical evaluation. It is usually helpful that liver tests are divided into three groups: evaluate synthetic function (Albumin, Total Protein and A/G ratio), evaluate hepatocellular injury or inflammation (ALT and AST), evaluate cholestasis (ALP and Glutamyl transferase). AFP can only be employed if hepatocellular malignancy is suspected. The clinical conditions and the specific pattern of liver disorders will not only minimize various diagnoses but also offer a cost-effective method to evaluate patients and recognize individuals who require a liver biopsy. The models have classified the patients correctly into different groups with 94% and 95% accuracy with respective models (RF and CART) and these results establish and validate the identification of important parameters as features.

Informed Consent

Written informed consent was obtained from the patients who participated in this study.

Author Contributions

Concept—S.A.; S.A.Z.; Design—S.A.; S.A.Z.; Supervision—S.A.Z.; H.U.S.; Resource—S.I.; Materials—S.A.; S.A.Z.; Data Collection and/or Processing—S.I.; S.A.; Analysis and/or Interpretation—S.A.; S.A.Z.; Manuscripts—S.A.; S.A.Z.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

- [1] Singh, A. and Pandey, B. (2014) Intelligent Techniques and Applications in Liver Disorders: A Survey. *International Journal of Biomedical Engineering and Tech-*

- nology, **16**, 27-70. <https://doi.org/10.1504/IJBET.2014.065638>
- [2] Hay, J.E. (2004) Acute Liver Failure. *Current Treatment Options in Gastroenterology*, **7**, 459-468. <https://doi.org/10.1007/s11938-004-0005-5>
- [3] Khan, B., Naseem, R., Ali, M., Arshad, M. and Jan, N. (2019) Machine Learning Approaches for Liver Disease Diagnosing. *International Journal of Data Science and Advanced Analytics*, **1**, 27-31.
- [4] Thapa, B.R. and Walia, A. (2007) Liver Function Tests and Their Interpretation. *The Indian Journal of Pediatrics*, **74**, 663-671. <https://doi.org/10.1007/s12098-007-0118-7>
- [5] Agrawal, S., Dhiman, R.K. and Limdi, J.K. (2016) Evaluation of Abnormal Liver Function Tests. *Postgraduate Medical Journal*, **92**, 223-234. <https://doi.org/10.1136/postgradmedj-2015-133715>
- [6] Kamath, P.S. (1996) Clinical Approach to the Patient with Abnormal Liver Test Results. *Mayo Clinic Proceedings*, **71**, 1089-1095. <https://doi.org/10.4065/71.11.1089>
- [7] Lala, V., Goyal, A., Bansal, P. and Minter, D.A. (2020) Liver Function Tests. *Stat-Pearls* [Internet].
- [8] Takkar, S., Singh, A. and Pandey, B. (2017) Application of Machine Learning Algorithms to a Well Defined Clinical Problem: Liver Disease. *International Journal of E-Health and Medical Communications*, **8**, 38-60. <https://doi.org/10.4018/IJEHMC.2017100103>
- [9] Dankerl, P., Cavallaro, A., Tsymbal, A., Costa, M.J., Suehling, M., Janka, R., et al. (2013) A Retrieval-Based Computer-Aided Diagnosis System for the Characterization of Liver Lesions in CT Scans. *Academic Radiology*, **20**, 1526-1534. <https://doi.org/10.1016/j.acra.2013.09.001>
- [10] Kumar, S.S., Moni, R.S. and Rajeesh, J. (2013) An Automatic Computer-Aided Diagnosis System for Liver Tumours on Computed Tomography Images. *Computers & Electrical Engineering*, **39**, 1516-1526. <https://doi.org/10.1016/j.compeleceng.2013.02.008>
- [11] Snell, A.M. (1958) Liver Function Tests and Their Interpretation. *Gastroenterology*, **34**, 675-682. [https://doi.org/10.1016/S0016-5085\(58\)80052-9](https://doi.org/10.1016/S0016-5085(58)80052-9)
- [12] Harris, E.H. (2005) Elevated Liver Function Tests in Type 2 Diabetes. *Clinical Diabetes*, **23**, 115-119. <https://doi.org/10.2337/diaclin.23.3.115>
- [13] Rochling, F.A. (2001) Evaluation of Abnormal Liver Tests. *Clinical Cornerstone*, **3**, 1-12. [https://doi.org/10.1016/S1098-3597\(01\)90074-2](https://doi.org/10.1016/S1098-3597(01)90074-2)
- [14] Gowda, S., Desai, P.B., Hull, V.V., Math, A.A.K., Vernekar, S.N. and Kulkarni, S.S. (2009) A Review on Laboratory Liver Function Tests. *The Pan African Medical Journal*, **3**, Article No. 17.
- [15] Theal, R.M. and Scott, K. (1996) Evaluating Asymptomatic Patients with Abnormal Liver Function Test Results. *American Family Physician*, **53**, 2111-2119.
- [16] Vijayarani, S. and Dhayanand, S. (2015) Liver Disease Prediction Using SVM and Naïve Bayes Algorithms. *International Journal of Science, Engineering and Technology Research*, **4**, 816-820.
- [17] Rajeswari, P. and Reena, G. S. (2010) Analysis of Liver Disorder Using Data Mining Algorithm. *Global Journal of Computer Science and Technology*, **10**, 48-52.
- [18] Baitharu, T.R. and Pani, S.K. (2016) Analysis of Data Mining Techniques for Healthcare Decision Support System Using Liver Disorder Dataset. *Procedia Computer Science*, **85**, 862-870. <https://doi.org/10.1016/j.procs.2016.05.276>

- [19] Rahman, A.K.M.S., Javed Mehedi Shamrat, F.M., Tasnim, Z., Roy, J. and Hossain, S.A. (2019) A Comparative Study on Liver Disease Prediction Using Supervised Machine Learning Algorithms. *International Journal of Scientific and Technology Research*, **8**, 419-422.
- [20] Lin, R.-H. (2009) An Intelligent Model for Liver Disease Diagnosis. *Artificial Intelligence in Medicine*, **47**, 53-62. <https://doi.org/10.1016/j.artmed.2009.05.005>
- [21] Venkata Ramana, B., Babu, M.S.P. and Venkateswarlu, N. (2011) A Critical Study of Selected Classification Algorithms for Liver Disease Diagnosis. *International Journal of Database Management Systems*, **3**, 101-114. <https://doi.org/10.5121/ijdms.2011.3207>
- [22] Islam, M.A., Akter, S., Hossen, M.S., Keya, S.A., Tisha, S.A. and Hossain, S. (2020) Risk Factor Prediction of Chronic Kidney Disease based on Machine Learning Algorithms. *2020 3rd International Conference on Intelligent Sustainable Systems*, Thoothukudi, 3-5 December 2020, 952-957. <https://doi.org/10.1109/ICISS49785.2020.9315878>
- [23] Koch, T.R. and Doumas, B.T. (1982) Bilirubin: Total and Conjugated, Modified Jendrassik-Grof Method. *American Association for Clinical Chemistry*, **9**, 113.
- [24] Bergmeyer, H.U., Scheibe, P. and Wahlefeld, A.W. (1978) Optimization of Methods for Aspartate Aminotransferase and Alanine Aminotransferase. *Clinical Chemistry*, **24**, 58-73. <https://doi.org/10.1093/clinchem/24.1.58>
- [25] Rosalki, S.B. et al. (1993) Multicenter Evaluation of Iso-ALP Test Kit for Measurement of Bone Alkaline Phosphatase Activity in Serum and Plasma. *Clinical Chemistry*, **39**, 648-652. <https://doi.org/10.1093/clinchem/39.4.648>
- [26] Peters, T., Biamote, G.T. and Doumas, B.T. (1982) Total Protein in Serum, Urine, and Cerebrospinal Fluid, Albumin, in Serum. In: Faulkner, W.R. and Meites, S., Eds., *Selected Methods of Clinical Chemistry*, American Association for Clinical Chemistry, Volume 9, Washington DC.
- [27] Doumas, B.T., Watson, W.A. and Biggs, H.G. (1971) Albumin Standards and the Measurement of Serum Albumin with Bromocresol Green. *Clinica Chimica Acta*, **31**, 87-96. [https://doi.org/10.1016/0009-8981\(71\)90365-2](https://doi.org/10.1016/0009-8981(71)90365-2)
- [28] Fiore, M., Mitchell, J., Doan, T., Nelson, R., Winter, G., Grandone, C., et al. (1988) The Abbott IMx Automated Benchtop Immunochemistry Analyzer System. *Clinical Chemistry*, **34**, 1726-1732. <https://doi.org/10.1093/clinchem/34.9.1720>
- [29] Norušis, M.J. (2006) SPSS 14.0 Guide to Data Analysis. Prentice Hall Upper Saddle River, NJ.
- [30] Breiman, L. (2001) Random Forests. *Machine Learning*, **45**, 5-32. <https://doi.org/10.1023/A:1010933404324>
- [31] Ronaghan, S. (2018) Trees, Random Forest and Feature Importance in Scikit-Learn and Spark. *Towards Data Science*. <https://towardsdatascience.com/the-mathematics-of-decision-trees-random-forest-and-feature-importance-in-scikit-learn-and-spark-f2861df67e3>
- [32] Sharma, M. (2019) Fundamentals of Classification and Regression Trees (CART). *Analytics Vidhya*. <https://mathanrajsharma.medium.com/fundamentals-of-classification-and-regression-trees-cart-e9af0b152503>.
- [33] Loh, W.Y. (2011) Classification and Regression Trees. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, **1**, 14-23. <https://doi.org/10.1002/widm.8>
- [34] Erlinger, S., Arias, I.M. and Dhumeaux, D. (2014) Inherited Disorders of Bilirubin

- Transport and Conjugation: New Insights into Molecular Mechanisms and Consequences. *Gastroenterology*, **146**, 1625-1638.
<https://doi.org/10.1053/j.gastro.2014.03.047>
- [35] Healey, C.J., Chapman, R.W. and Fleming, K.A. (1995) Liver Histology in Hepatitis C Infection: A Comparison between Patients with Persistently Normal or Abnormal Transaminases. *Gut*, **37**, 274-278. <https://doi.org/10.1136/gut.37.2.274>
- [36] Haber, M.M., West, A.B., Haber, A.D. and Reuben, A. (1995) Relationship of Aminotransferases to Liver Histological Status in Chronic Hepatitis C 1. *American Journal of Gastroenterology (Springer Nature)*, **90**, 1250-1257.
- [37] Diehl, A.M., Potter, J., Boitnott, J., Van Duyn, M.A., Herlong, H.F. and Mezey, E. (1984) Relationship between Pyridoxal 5'-Phosphate Deficiency and Aminotransferase Levels in Alcoholic Hepatitis. *Gastroenterology*, **86**, 632-636.
[https://doi.org/10.1016/S0016-5085\(84\)80110-9](https://doi.org/10.1016/S0016-5085(84)80110-9)
- [38] Friedman, L.S., Martin, P. and Munoz, S.J. (1996) Liver Function Tests and the Objective Evaluation of the Patient with Liver Disease. In: Zakim, D. and Boyer, T.D., Eds., *Hepatology: A Textbook of Liver Disease*, Volume 1, WB Saunders, Philadelphia, 791-833.
- [39] Martin, P. and Friedman, L.S. (2018) Assessment of Liver Function and Diagnostic Studies. In: Friedman, L.S. and Martin, P., Eds., *Handbook of Liver Disease*, Elsevier, Amsterdam, 1-17. <https://doi.org/10.1016/B978-0-323-47874-8.00001-8>
- [40] Zolk, M., Cordiani, M.R., Marchesini, G., Iervese, T., Labate, A.M., Bonazzi, C., et al. (1991) Prognostic Indicators in Compensated Cirrhosis. *American Journal of Gastroenterology*, **86**, 1508-1513.
- [41] Chen, D. and Sung, J. (1979) Relationship of Hepatitis b Surface Antigen to Serum Alpha-Fetoprotein in Nonmalignant Diseases of the Liver. *Cancer*, **44**, 984-992.
[https://doi.org/10.1002/1097-0142\(197909\)44:3%3C984::AID-CNCR2820440328%3E3.0.CO;2-6](https://doi.org/10.1002/1097-0142(197909)44:3%3C984::AID-CNCR2820440328%3E3.0.CO;2-6)