Tuijin Jishu/Journal of Propulsion Technology
ISSN: 1001-4055
Vol. 45 No. 4 (2024)

A Web Application to predict the Liver Functionality using Ensemble Learning Technique

Mrs. MD Reshma[1*], Dr. S. Thulasi Prasad[2], Dr. M. Rajendra prasad[3], Mr. G. Nagendra[4]

[1]M. Tech Student, Department of Electronics and Communication Engineering, Vidya Jyothi Institute of Technology, Aziz Nagar, Hyderabad- 500075, India.*

[2]Professor, Department of Electronics and Communication Engineering, Vidya Jyothi Institute of Technology, Aziz Nagar, Hyderabad- 500075, India.

[3]Professor, Department of Electronics and Communication Engineering, Vidya Jyothi Institute of Technology, Aziz Nagar, Hyderabad- 500075, India.

[4] Associate Professor, Department of Electronics and Communication Engineering, Vidya Jyothi Institute of Technology, Aziz Nagar, Hyderabad- 500075, India.

Abstract- Liver diseases pose a significant global public health challenge due to various causes like viral infections and metabolic issues. Early and precise diagnosis is crucial for effective management. While traditional diagnostic methods have limitations, the application of machine learning (ML) in medical diagnosis shows promise, enhancing predictive capabilities. By integrating Adaboost, Catboost, LGBM, and MLP classifiers into our approach, we aim to leverage their collective predictive power with a voting classifier. The goal is to improve the accuracy and reliability of liver disease prediction. This method achieved an 85% accuracy rate in determining whether a patient has a normal or pathological liver condition, generating two distinct outcomes. These findings suggest that the proposed system can complement a doctor's diagnosis of liver disease.

Keywords: *Liver Disease, Ensemble Learning, Machine Learning, Classification*

1. Introduction

Over a million people worldwide receive a liver disease diagnosis each year. 3.5% of deaths this century were connected to disorders of the liver. The main causes of liver impairment are drug misuse, obesity, diabetes, and chronic alcoholism [1]. Common liver illnesses include liver cancer, hepatitis A, B, and C, and liver cirrhosis. Liver cirrhosis claimed the lives of 1.32 million more people worldwide in 2017 than it did in 1990, with 33.3% of the victims being women and 66.7% of men. Even so, overall mortality is decreasing due to advances in cutting-edge care and leading healthy lives [2]. Prompt diagnosis of illnesses is crucial in reducing the risk to patients' lives. Timely treatment can effectively eliminate such risks. The medical field follows various stages to evaluate a disease including prevention, diagnosis, and treatment. Each phase involves unique methods and criteria. Liver ailments manifest with distinct symptoms and diagnostic procedures. When the liver reaches irreparable damage, the only viable option is undergoing a liver transplant. Liver disease advances through four stages: cirrhosis, fibrosis, healthy, and cancer. Symptoms such as internal bleeding, dry mouth, constipation, and stomach discomfort can impact the digestive system sig- nificantly [3].Several symptoms of liver disease might include memory loss, numbness, and fainting, alongside skin issues like yellow skin, spider veins, and redness in the feet. To ward off liver problems, regular doctor visits are crucial. Im- munizations, limiting alcohol and soda consumption, staying active, and keeping weight under control are all beneficial. A notable advancement in the healthcare field involves machine learning. Utilizing classification and regression techniques aids in disease prediction. Machine learning contributes to unbiased decision-making and ensures precise disease diagnosis and

forecasting. These advancements have sparked interest within the biomedical sector [4]. Machine learning methods have the potential to speed up and lower the expenses of addressing medical issues. The primary goals of this approach are to cut down on diagnostic expenses in the healthcare sector and improve prognostic accuracy. The document is structured into several parts. The literature review segment outlines the assessment and techniques utilized in various studies. Within the dataset description section, there is an explanation of the ILPD dataset [5] along with its characteristics. The proposed methodology section details the various stages involved in disease prediction and interface design. The results segment delves into the findings. Finally, the conclusion and future prospects section brings everything to a close, wrapping up the analysis.

2. Literature Survey

Different classification techniques have been applied to medical problems individually or have been ensemble together to get accurate results. Various machine learning modes exist, such as support vector machines(SVM), Naive Bayes, K- Nearest Neighbor, Artificial neural network, Random Forest etc. to perform classification for different tasks. However, most liver disorders don't show any noticeable signs when they're first developing. With today's sophisticated databases, extracting data and gaining insights to help treat any ill- ness is a straightforward process [6].The most correlated feature space can be chosen and extracted using a variety of techniques in order to anticipate any illness or object. Several researchers have attempted to apply machine learning algorithms in different ways in an effort to accurately and early classify liver patients. The paper by Sreejith et al. evaluated the performance of classification using datasets from Pima Indian Diabetes (PID), Thoracic Surgery (TSD), and ILPD. The main goal of their study is to compare the performance and without feature selection. They applied the Chaotic Multi- Verse Optimization (CMVO) evolutionary feature selection method and the Synthetic Minority Over-sampling Technique (SMOTE) for class-balancing. When using a random forest classifier without OSMOTE and CMVO-based feature selec- tion, they achieved an accuracy of 69.43% on the ILPD dataset. With OSMOTE and CMVO-based feature selection, the ac- curacy improved to 82.620%. Additionally, with OSMOTE and CMVO-based feature selection, they obtained an accuracy of 82.46% [7]. Javad Hassannataj (2018) conducted a study comparing various data mining models, including Support Vector Machine (SVM), Particle Swarm Optimization (PSO)- SVM, Random Forest, Multi-Layer Perceptron (MLP) Neural Network, Bayesian Networks, using the ELTA technique. The results showed that the PSO-SVM model outperformed other models in terms of specificity, sensitivity, accuracy, Area under the Curve (AUC), F-measure, precision, and False Positive Rate (FPR). In addition to this, a 10-fold cross-validation technique was utilized to assess the models on a liver disease- related dataset. The average estimated accuracy scores for Ran- dom Forest, MLP Neural Network, Bayesian Network, SVM, and PSO-SVM were 87.35%, 78.91%, 66.78%, 76.51%, and 95.17% respectively [8]. Yugal et al. put forward a rule-based approach to categorize data across different liver conditions employing machine learning methods [9]. They utilized k- fold cross-validation in their model, making use of SVM, rule induction (RI), decision trees, naive Bayes, and ANN. The decision tree with a rule-based classification system showed superior performance among the various models examined. They compiled a dataset with 12 features and 583 entries [9]. Han Ma et al. evaluated the most precise predictive model for detecting non-alcoholic fatty liver disease (NAFLD) Data stemming from individuals undergoing health assessments at Zhejiang University's First Affiliated Hospital was utilized for this model. Among the eleven models considered, the Bayesian network emerged as the top performer [10]. In the suggested study of A. Gulia et al. [11], researchers used algorithms such Random Forest, Bayesian Network, Support Vector Machine, J48, Multi-Layer Perceptron, and Random Forest to categorise liver patient data. The Centre of Machine Learning and Intelligent Systems has made use of the data found in the UCI repository. With an accuracy of 71.87%, the Random Forest Algorithm is determined to be the best when their three-phase analysis is complete.

3. Dataset Description

The ILPD, short for the Indian Liver Patient Dataset, com- prises liver disease diagnoses for 416 patients in CSV format along with liver-free data for 167 individuals. A liver function- ality test is carried out on individuals from Andhra Pradesh, focusing mainly on Srikakulam and Vijayawada areas. This dataset includes

ten distinct attributes: age, gender, total Biliru- bin, direct Bilirubin total proteins, Alkaline Phosphotase, Alamine Aminotransferase, Aspartate Aminotransferase, Al- bumin, Albumin and Globulin Ratio, and dataset itself. There exists a total of 583 patient records with 441 being male and 152 females. Several researchers have employed this dataset to explore chronic liver disorders' viability. Table 1 describes the dataset attributes and their range of values.

TABLE 1: DATASET DESCRIPTION [12]

Attribute	Type	Description	Range
Age	Numeric	Age of the patient.	[4-90]
Gender	String	Gender of the patient 'Male' or 'Female'	NA
TB	Numeric	Total Bilirubin	[0.5-75]
DB	Numeric	Direct Bilirubin	[0.1-19.7]
Alkphos	Numeric	Alkaline Phosphotase	[63-2110]
Sgpt	Numeric	Alamine Aminotransferase	[10-2000]
Sgot	Numeric	Aspartate Aminotransferase	[10-4929]
TP	Numeric	Total Proteins	[2.7-9.6]
ALB	Numeric	Albumin	[0.9-5.5]
A/G Ratio	Numeric	Albumin and Globulin Ratio	[0.3-2.8]
Dataset	Numeric	Target field used to split the data into two classes (labeled by the experts)	1(Positive) 2(Negative)

The dataset consists of 11 columns and 583 rows. Different data types like int, float and object are present in the dataset with all non-null values. The default value of person with a normal liver is 1 in the dataset and for abnormal liver is 2. The dataset needs a bit modification according to our proposed method.

3.1 Pre-Processing of the Dataset

The dataset initially consists of 583 rows and 11 columns. First, we need to check for any null values present in the dataset. Subsequently, we will replace these null values with the mean of the entire column values. Additionally, a standard scaler can come in handy when dealing with input datasets that vary significantly in their ranges. In cases where the mean is zero, the scaler will adjust the values to fit within the range of -1 to +1.Moving on to the next step of pre-processing involves eliminating any duplicate entries in the dataset. We also need to encode the gender values such that males are represented as 0 and females as 1. Moreover, since one of the dataset columns indicates whether a disease is present, we adjust its values to represent the presence (1) or absence (0) of the disease. In order to assess the relationship between variables within the dataset, we have generated a correlation matrix and plotted the correlation coefficients which is depicted in fig. 1. This matrix provides insights into how strongly different variables are related to each other. Based on these correlation values, we have decided to drop three columns: Total Bilirubin, Alamine Aminotransferase, and Total protein due to their low coefficient values.After this process, our dataset now consists of 570 rows and 7 columns. Finally, to address any imbalances present in the data, a random over sampler is applied by introducing a few dummy values. With these preparatory steps completed, our dataset is now ready for training machine learning models.

4. Proposed Methodology

There are different steps in which the proposed methodology is executed as shown in fig.2. The predicted value of our proposed method would be if a person has normal or abnormal liver.

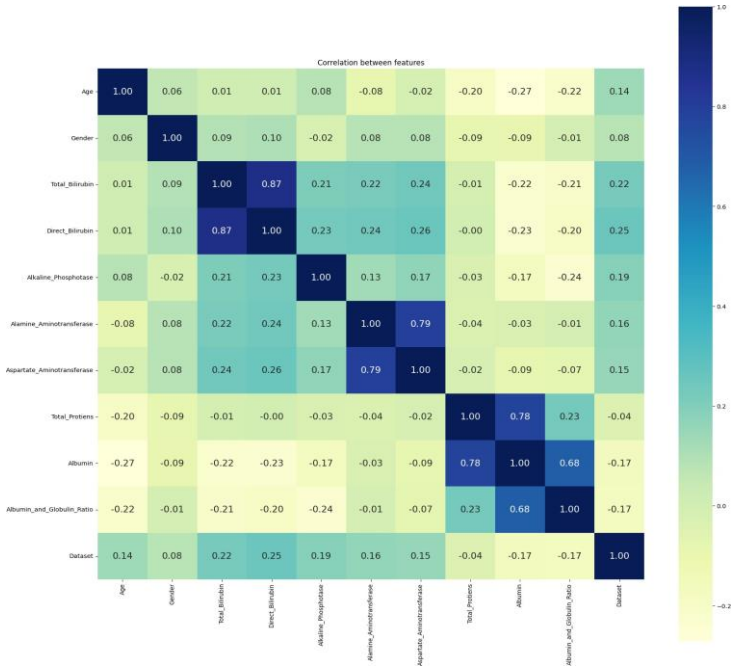


Fig. 1. correlation between the attributes

4.1 Indian Liver Patient Dataset

The medical records of individuals with chronic liver dis- ease from the Andhra Pradesh region make up this dataset. There are 583 rows and 11 columns in this dataset. Consider- ations include age, gender, total protein, alamine aminotrans- ferase, bilrubin, and other traits. As your body degrades old red blood cells, it naturally creates the yellowish protein known as bilirubin. The fluid called bile, which is made by your liver and includes bilirubin, aids in the digestion of meals. If your liver is in good condition, it will remove the majority of the bilirubin from your body. If your liver is damaged, bilirubin may leak into your circulation. To assess your liver's health, a bilirubin blood test is conducted. So out of all the features bilirubin is considered as the important one.

4.2 Data Pre-Processing

It describes the steps used to prepare data for analysis by cleaning, transforming, and integrating it. Enhancing the quality of the data and tailoring it to the particular data mining task are the objectives of data preprocessing. The steps followed in this procedure have been clearly explained in the third section. In this process we will fill the null values, eliminate the duplicates using some standard procedure like random sampling and standard scaler. After the altering the dataset we are left with 570 rows and 7 columns.

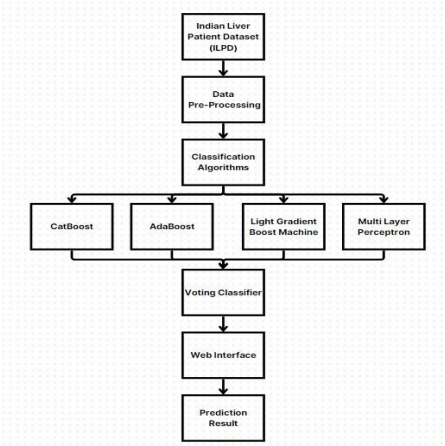


Fig. 2. correlation between the attributes

4.3 Classification Algorithms

In our problem statement we have the binary classification technique so we have used four different algorithms. These al- gorithms have different functionalities to solve a classification problem.

AdaBoost, also known as adaptive boosting, is an ensem- ble learning method for machine learning that is used to regression and classification problems. The fundamental idea of AdaBoost is to repeatedly train the weak classifier on the training dataset, increasing the weight of the erroneously classified data points in each new classifier. To create the final AdaBoost model, all of the weak classifiers used in the training process are integrated with the weights given to the models according to their accuracies. The weakest model with the most accuracy is given the highest weight, while the model with the lowest accuracy is given the lowest weight. AdaBoost methods combine a number of weak models from machine learning to produce an output with a strong classification model.

CatBoost, sometimes known as Categorical Boosting is intended for use with very large numbers of independent features in regression and classification tasks. Catboost is a variant of gradient boosting that handles both category and numerical variables. To translate category information into numerical features, no feature encoding technology such as One-Hot Encoder or Label Encoder is needed. In order to lessen over fitting and enhance overall efficiency, it also makes use of a method known as the symmetric weighted quantilesketch (SWQS), which automatically manages the dataset's missing values.

LightGBM is a gradient-boosting framework that uses de- cision trees to lower memory utilisation and improve model efficiency. It employs two cutting-edge methods:Exclusive Feature Bundling (EFB) and Gradient-based One Side Sam- pling (GOSS). The histogram-based approach, which is the main component of all GBDT (Gradient Boosting Decision Tree) systems, has constraints that these methods address.

A multi layer perceptron comprises one or more hidden layers with several neurons stacked on top of each other, as well as input and output layers. In a multilayer perceptron, neurons can use any kind of activation function; however, in a perceptron, the neuron needs to have an activation function (like sigmoid or ReLU) that imposes a threshold.Similar to the Perceptron, the multi layer Perceptron belongs to the feed forward algorithm family because, prior to being exposed to the activation function, inputs are combined with the starting weights in a weighted sum. Each layer feeds the next layer by using its internal representation of the data or the result of its calculation.The output from these individual algorithms is fed into the next step where the final output and accuracy is decided.

4.4 Voting Classifier

A voting classifier forecasts an output depending on which model has the highest chance of providing the target class after being trained on a wide ensemble of models. All it does is aggregate all the classifier outputs that are given into the voting classifier, which determines the output class based on which voting majority is greatest. The idea is not to construct separate dedicated models and measure each one's accuracy, but to train a single model that predicts output based on the cumulative majority of votes from each output class. Thus, the output of the model with the best accuracy is chosen based on accuracy, and the outcomes are displayed.

4.5 Web Interface

An interactive user interface is created using Flask micro framework architecture. Then the machine learning model is integrated with the flask to create a working model. The interface consists of different boxes and buttons that requires some input values provided by the user and output will be predicted.

4.6 Prediction Result

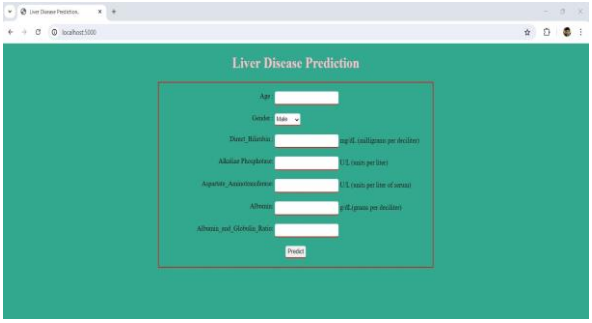
For getting the output user has to give details to the model which includes the patient's age, gender, Direct bilirubin, Al- kaline Phosphotase, Asparate Aminotransferase,Albumin and Albumin and Gobulin ratio values. With the help of these values the output is predicted as normal or abnormal.

5. Results And Discussions

5.1 Prediction through the Web Application

Our predicted result is displayed through the interface which shows only two values normal or abnormal liver based upon the input values given to it. Our classifier classifies the patient into these two categories.Our Application interface home page is shown in fig. 3.

Fig. 3. Liver Disease Prediction Home Page



The next two screenshots fig. 4 and fig. 5 shows input and output of a patient having abnormality in liver functioning.

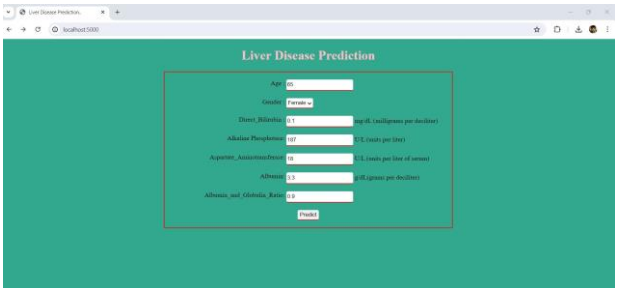


Fig. 4. Liver Disease Prediction input of a abnormal patient

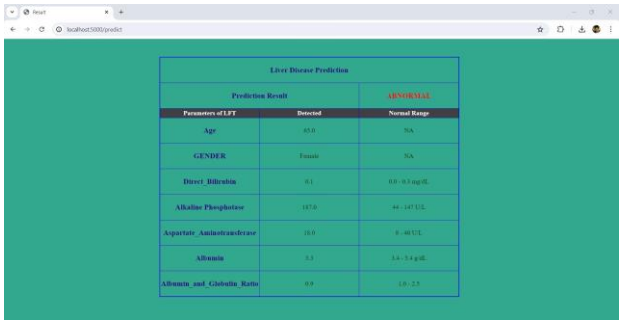


Fig. 5. Liver Disease Prediction output of a abnormal patient

The following two screenshots fig. 6 and fig. 7 shows input and output of a patient having normal liver functioning.

5.2 Performance Evaluation

The models are initially taught separately on the training data, and the accuracies of their testing and training are then calculated. Twenty percent of the dataset is used for testing, while the remaining eighty percent is for training. After the models have been trained, the necessary testing is carried out. The table II contains the

individual accuracy data. After training on individual models, a voting classifier is used to consider the highest accuracy and predict the output. In this work we used hard voting strategy where, after individual

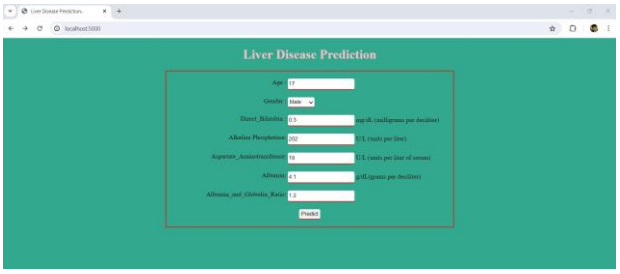


Fig. 6. Liver Disease Prediction input of a normal patient

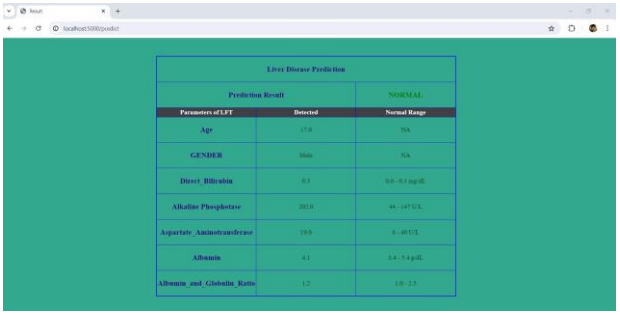


Fig. 7. Liver Disease Prediction output of a normal patient

model training majority of predicted class by individual mod- els is taken as final prediction. A normal liver has an output value of 0, and an aberrant liver has a value of 1. Following the voting classifier's application, 0.9661 for training accuracy and 0.8567 for testing accuracy are recorded. In addition to this, during training and testing, the values for precision, recall, f-1 score, and support are computed for both 0 and 1 outputs. These values are listed in tables 3 and 4 below.

An overview of a machine learning model's performance on a set of test data is displayed in a confusion matrix. It serves as a visual representation of the percentage of the model's predictions that were accurate or inaccurate. This approach is frequently used to assess the effectiveness of categorization models, which seek to assign a category label to each input occurrence. The number of model instances that were created on the test set is shown in the matrix.

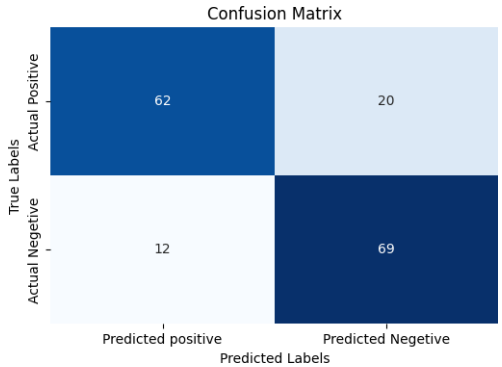


Fig. 8. Confusion matrix on test data prediction

TABLE 2: MODEL ACCURACY TABLE

S.No	Model Name	Training Accuracy	Testing Accuracy
1.	LGBM	1.000000	0.846626

2.	AdaBoost	0.853621	0.742331
3.	CatBoost	0.781202	0.723926
4.	MLP	0.959938	0.791411

TABLE 3: TRAINING CLASSIFICATION REPORT

Output	Precision	Recall	F1-Score	Support
0	0.98	0.95	0.97	325
1	0.95	0.98	0.97	324

TABLE 4: TESTING CLASSIFICATION REPORT

Output	Precision	Recall	F1-Score	Support
0	0.78	0.85	0.81	81
1	0.84	0.76	0.79	82

When a positive data point is correctly predicted by the model, a true positive (TP) happens. When the model correctly forecasts a negative data point, a true negative (TN) happens. A model produces false positives (FP) when it predicts a positive data point incorrectly, and false negatives (FN) when it predicts a negative piece of data. Confusion matrix based on the actual label and predicted label of the test data with 163 instances belonging to 82 abnormal and 81 normal clinical records is plotted in fig 8. with the results of voting classifier.

6. Conclusion And Future Scope

In the proposed work we have ensemble four machine learning classification algorithms Adaboost, Catboost, LGBM and MLP together with a voting classifier to get the maxi- mum accuracy. The ILPD dataset in CSV format contains 11 different features to predict if the patient is suffering from a liver disease or not. We have achieved the testing accuracy of 85%. The web interface is also very user friendly and easy to operate, By giving few input values to the system we can predict the liver condition of the patient. The voting classifier is used to select the algorithm with the highest accuracy to predict the patients condition precisely. In the future we can create an interface that would directly take the reports and collect the values from them instead of giving them manually. Ensembling other algorithms also might improve the accuracy.

References

[1] GS Harshpreet Kaur. The diagnosis of chronic liver disease using machine learning techniques. *Inf. Technol. Ind*.9(2):554–64. https://doi.org/10.17762/itii. v9i2.382. 2021.

[2] C Collaborators. Articles the global, regional, and national burden of cirrhosis by cause in 195 countries and territories, 1990 – 2017: *a systematic analysis for the Global Burden of Disease Study 2017*. p. 245–66. https://doi.org/10.1016/ S2468-1253(19)30349-8. 2017.

[3] G Shaheamlung, H Kaur, M Kaur. A Survey on machine learning techniques for the diagnosis of liver disease. *In Proceedings of the 2020 International Conference on Intelligent Engineering and Management (ICIEM)*, London, UK, 17–19 June 2020.

[4] S. M. Mahmud, et al. Machine Learning Based Unified Framework for Diabetes Prediction. *Proceedings of the 2018 International Conference on Big Data Engineering and Technology*. ACM , 2018.

[5] B. V. Ramana, P. M. Surendra, P. Babu, and P. N. B. Venkateswarlu, A Critical Comparative Study of Liver Patients from USAand INDIA: An Exploratory Analysis, *Int. J. Comput. Sci. Issues*, vol. 9, no. 3, pp. 506–516, 2012.

[6] J Hassannataj, H Saadatfar, A Dehzangi, S Shamshirband. Infor- matics in Medicine Unlocked Computer-aided decision-making for predicting liver disease using PSObased optimized SVM with fea- ture selection. *Inform Med Unlocked 2019*;17 (August):100255. https://doi.org/10.1016/j.imu.2019.100255. 2019.

[7] S Sreejith, H Khanna Nehemiah, A Kannan. Clinical data classification using an enhanced SMOTE and chaotic evolution- ary feature selection. *Comput Biol Med 126*(February):103991. https://doi.org/10.1016/j. compbiomed.2020.103991. 2020.

[8] Joloudari, Javad Hassannataj, et al. Computer-aided decision-making for predicting liver disease using PSO-based optimized SVM with feature selection. *Informatics in medicine unlocked* 17 100255. 2019.

[9] Y Kumar. Sahoo, G. Prediction of different types of liver diseases using rule based classification model. *Technol*. Health Care 21, 417–432. [CrossRef] [PubMed] 2013.

[10] J Jacob, J.C Mathew, J Mathew, E Issac. Diagnosis of liver disease using machine learning techniques. *Int. Res. J. Eng. Technol*. 5, 412–423. 2018.

[11] R Gulia. P Vohra, Rani. Liver Patient Classification Using Intelligent Techniques, *International Journal of Computer Science and Information Technologies* (IJCSIT), vol. 5, no. 4, pp. 5110-5115. 2014.

[12] Singh, Jagdeep, Sachin Bagga, and Ranjodh Kaur. Software-based prediction of liver disease with feature selection and classification techniques. *Procedia Computer Science* 167 1970-1980. 2020.