Scalable Computing: Practice and Experience, ISSN 1895-1767, http://www.scpe.org
© 2025 SCPE. Volume 26, Issues 3, pp. 1104–1115, DOI 10.12694/scpe.v26i3.4254

ENHANCED EARLY DIAGNOSIS OF LIVER DISEASES USING FEATURE SELECTION AND MACHINE LEARNING TECHNIQUES ON THE INDIAN LIVER PATIENT DATASET

ARUN GANJI *, D. USHA †, AND P.S. RAJAKUMAR ‡

Abstract. Liver diseases are a significant global health concern, with timely diagnosis crucial for effective treatment and prevention of further damage. This study addresses the challenge of early liver disease detection using machine learning techniques applied to the Indian Liver Patient Dataset (ILPD). Our proposed method comprises a four-phase approach: (1) initial model training using five machine learning algorithms - Multilayer Perceptron (MLP), Support Vector Machine (SVM), Decision Trees (DT-CART), Light Gradient Boosting Machine (LGBM), and Logistic Regression (LR) - on the original dataset; (2) feature selection using Forward Selection (FS) to identify the most relevant attributes; (3) model retraining with the selected features; and (4) model optimization to enhance prediction accuracy. The dataset was split into 80% training and 20% testing sets, with 10-fold cross-validation applied throughout. Our findings demonstrate the significant impact of feature selection and model optimization on algorithm performance. The Light Gradient Boosting Machine (LGBM) emerged as the top-performing model, achieving an accuracy of 82.12% after optimization, compared to its initial 76.21%. LGBM also showed balanced performance across specificity, sensitivity, precision, and F1-score metrics. This study contributes to the field by presenting a comprehensive approach to liver disease prediction, emphasizing the importance of feature selection and model optimization in improving diagnostic accuracy.

Key words: Indian Liver Patient Dataset (ILPD), Machine learning (ML), Forward Selection (FL), Classification Algorithms.

1. Introduction.

The liver, the largest internal organ in the human body, plays a pivotal role in numerous physiological processes, performing more functions than any other organ. Its significance in maintaining overall health cannot be overstated, making the early detection and treatment of liver diseases a critical medical priority. Liver enzymes, particularly Aspartate Aminotransferase (SGOT) and Alanine Aminotransferase (SGPT), are crucial indicators in diagnosing liver diseases. Various factors, including lifestyle habits such as smoking and alcohol consumption, can trigger liver diseases and elevate these enzyme levels [1]. Moreover, conditions like diabetes, hepatitis B, and hepatitis C can lead to liver damage and, if left untreated, progress to liver failure. The consequences of severe liver damage are dire, often necessitating organ transplantation or resulting in mortality [2]. These facts underscore the urgent need for early and accurate diagnosis of liver diseases.

In recent years, the rapid advancement of technology has ushered in a new era of medical diagnostics, with machine learning emerging as a powerful tool in various healthcare domains. Machine learning methods can analyze vast amounts of data, both structured and unstructured, to create predictive models using statistical and mathematical techniques. The efficacy of these predictions is intrinsically linked to the quality of the underlying model.

In the realm of classification algorithms, the selection of relevant features plays a crucial role in model performance. Not all attributes in a dataset contribute equally to the model's predictive power. To enhance model accuracy and efficiency, researchers often employ feature selection methods to identify the most relevant attributes. This process is vital in creating effective machine learning classification models for medical diagnostics.

The present study addresses the critical need for improved liver disease diagnosis by proposing a hybrid model that combines feature selection with advanced machine learning techniques. We utilize the Indian Liver Patient Dataset (ILPD) and employ the forward selection method, a wrapper approach, for feature selection. The study then applies various machine learning algorithms, including Support Vector Machine

*Dr. M.G.R. Educational and Research Institute, Chennai-95, Tamilnadu, India (arun.ganji@gmail.com).
†Dr. M.G.R. Educational and Research Institute, Chennai-95, Tamilnadu, India (usha.cse@drmgredu.ac.in).
‡Dr. M.G.R. Educational and Research Institute, Chennai-95, Tamilnadu, India (rajakumar.subramanian@drmgredu.ac.in).

(SVM), Multilayer Perceptron (MLP), Decision Trees (DT), Logistic Regression (LR), and Light Gradient Boosting Machine (LGBM), to diagnose liver failure using the most important features identified through feature selection. Furthermore, we enhance the diagnostic accuracy by optimizing the hyperparameters of these machine learning models.

Our approach comprises four key stages:

1. Initial application of classification algorithms on the complete dataset.
2. Feature selection to create a subset of important attributes, followed by the application of classification algorithms on this subset.
3. Model improvement through hyperparameter optimization and comparison of prediction accuracies across the first three stages.
4. Comparative analysis of our optimized models' performance with existing literature using the same dataset.

This comprehensive approach aims to significantly improve the accuracy of liver disease prediction, potentially leading to earlier diagnoses and more effective treatment strategies. By combining feature selection, advanced machine learning techniques, and model optimization, we address the pressing need for more accurate and efficient diagnostic tools in hepatology.

The subsequent sections of this paper provide a detailed review of related studies, a comprehensive description of our methodology, including the dataset and machine learning methods employed, and a thorough comparison of our results with existing literature. Through this research, we aim to contribute to the ongoing efforts to enhance liver disease diagnosis and, ultimately, improve patient outcomes.

2. Related Study. Numerous studies have been carried out on various datasets related to liver diseases. This section provides a summary of some of the key studies that are prominent in the literature on this topic. Literature [3] modelled the SVM method in the diagnosis of diabetes and chronic liver disease on Diabetes, BUPA and ILPD datasets in the MATLAB study environment. They achieved 63% success in the diagnosis made with the first 4 attributes of the BUPA data set, 70% with the first 6 attributes, and 70% with the first 8 attributes. On the ILPD data set, they stated that the diagnosis made with the first 4 attributes of the data set was 71%, 73% with the first 6 attributes, and 73.2% with the first 8 attributes.

In the literature [4] Gulia et al. used J48, Random Forest (RF), MLP, SVM, and Bayesian Net machine learning methods towards diagnose liver diseases on the ILPD dataset. In the first stage, they made diagnostic success measurements with the original version of the data set without selecting attributes on the data set. In the first stage, they achieved the highest diagnostic success rate with 71.35% from the SVM classification method. They stated that the other algorithms were successful in RF 70.31%, MLP 68.25%, J48 68.70%, and Bayes Net 67.23%, respectively. In the second stage, they applied the Greedy Stepwise (Greedy algorithm) feature selection on the dataset to identify significant features. They found that the crucial attributes are Total Bilirubin and Direct Bilirubin, Total Proteins, Albumin and A/G ratio. In the second stage, the accuracy of the RF algorithm was determined with a rate of 71.86%. Other methods were successful in SVM 71.34%, J48 70.65%, MLP 70.82%, and Bayeses Net 68.11%, respectively.

Literature [5] used ANN, RF, Functional Tree and Radial Based Functional machine learning methods for the diagnosis of liver failure on BUPA and ILPD datasets. 86.95% of BUPA data were used as training (300 samples) and 13.05% (45 samples) as tests. For the ILPD data set, 87.48% (510 samples) are utilised for training and 12.52% (73 samples) are used for testing. They tested ANN using the MATLAB environment along with other methods in the WEKA tool. In the analyses, they used a 10-fold cross-validation test technique for the methods they used in WEKA and 10 hidden layers and a network of forward-fed neurons for ANN. As a result of this study, the highest diagnostic success was 76% for the BUPA dataset with ANN and 78% for the ILPD dataset.

In the research [6] author evaluated the ILPD dataset in two different stages. They applied a feature model also comparative study to enhance forecast accuracy. In the first stage, they applied a minimum maximum normalization filter to the original data set. In the second stage, they identified the attributes containing important features by using PSO (Particle Swarm Optimization) feature selection on data set. They identified the important attributes were Direct Bilirubin, Total Bilirubin, Total Proteins, A/G ratio, Albumin, SGOT, SGPT, and Alkphos. In the first stage, Bayes Net was the most successful method with an accuracy rate of

74.25%. In other methods, J.48 achieved 73.32%, MLP 69.22%, SVM 71.44%, and RF 68.43%, respectively. In the second stage, the J48 method was the most successful algorithm with a rate of 95.04%. In the relevant study, Bayes' Net was 90.33%, RF was 80.22%, MLP was 77.54% and SVM was 73.44%.

Alice [7] applied DT, RF, Naive Bayes, AMM and SVM machine learning methods to identify liver failure using ILPD dataset. R programming language was used in the study. The DT algorithm achieved the highest diagnostic success rate with 81%. Among the additional methods, RF was 77%, ANN was 71%, SVM was 77%, and Naive Bayes was 37%.

Literature [8] examined the success rates of machine learning methods in medical datasets. The designated medical datasets are Breast Cancer Data, Cryotherapy, Chronic Kidney Disease, Hepatitis, ILPD and Immunotherapy. They analyzed the datasets with Naive Bayes, J48, MLP, JRip, IBk and Bagging machine learning methods. In their study, they used 10-fold cross-validation. In the ILPD data set, they obtained the highest diagnostic success rate from the Bagging method with a rate of 69.30%. In other methods, they stated that they achieved 68.95% success in MLP, 68.78% in J48, 66.38% in JRIP, 64.49% in IBk and 55.75% in Native Bayes, respectively.

In the literature [9], the author focused on predicting liver disease built on a software engineering methodology using the trait selection and classification technique. Intelligent liver disease prediction software (ILDPS) was developed using feature selection and classification prediction techniques based on a software engineering model. They applied LR, RF, SMO, Naive Bayes, IBk and J48 machine learning approaches to find accuracy on the ILPD dataset. They evaluated the dataset in two stages. In the first stage, they made diagnostic success measurements with the original version of the data set without selecting attributes upon data set. In second stage, they utilized the Greedy Stepwise feature selection algorithm to identify significant features within the dataset. They determined that the key attributes are Total Bilirubin, Alkphos, Direct Bilirubin, SGOT, and SGPT. Their accuracy rate was tested using 10x cross-validation. In the first stage, the highest diagnostic success rate of 72.53% was achieved using the RF machine learning method. Among the other methods, LR was 72.50%, SMO was 71.35%, J48 was 68.78%, IBk was 64.15% and Naive Bayes was 55.74%. In the second stage, the LR machine learning method achieved the highest diagnostic success rate at 73.36%. In further methods, RF was 71.87%, SMO was 71.36%, J48 was 70.67%, IBk was 67.41% and Naive Bayes was 55.90%.

In the research [10] author used LR, SVM, K-Nearest Neighbour (K-NN) and on ILDP datasets, ANN machine learning approaches used to diagnose liver failure. They presented the model with the maximum accuracy as a Graphical User Interface (GUI) using the Tkinter package in Python. The utmost diagnostic success rate has been obtained from the ANN machine learning technique by a rate of 92.80%. Of the other methods, SVM was 75.04%, LR was 73.23% and C-NN was 72.05%. In ANN, number of inputs is set to 10, number of hidden layers is set to 2, The first hidden layer is configured with 400 neurons, and the second hidden layer is also configured with 400 neurons.

This study we developed and presented a hybrid model and a comparative analysis to improve the prediction accuracy of liver failure patients in four phases.

3. Materials and Methods. In this study, the ILPD dataset shared for research in the UCI machine learning pool has been used for liver failure disease diagnosis. There are 583 specimens in the ILPD dataset. The data were split into 80% training set (466 samples) and 20% test set (117 samples). A random state value of 42 was used to use the same training and test set in each data set separation process. Gender attribute was not taken into account in the data set. This study presents a hybrid model aims to enhance the prediction accuracy for liver failure patients across four phases. In first stage, SVM, MLP, DT, LR and LGBM machine learning methods were applied to the original dataset with the attributes "Total Bilirubin, Age, Direct Bilirubin, Alanine Aminotransferase, Alkaline Phosphatase, Total Protein, Albumin, Albumin/Globulin Ratio" and diagnostic prediction successes were measured. In the second stage, forward selection, which is one of the Wrapper methods, was used for feature selection on the data set. After the feature selection, the predictive success of the diagnosis of liver failure disease was measured by SVM, MLP, DT, LR and LGBM machine learning methods using the attributes "Direct Bilirubin, Age, Alanine Aminotransferase, Alkaline Phosphatase". In the third stage, the models were re-established by applying model improvement to the SVM, MLP, DT, LR and LGBM machine learning methods applied in the second stage, and their accuracy performance was increased. The forecast accuracy performances of the first three stages were compared. In the fourth stage, the accuracy

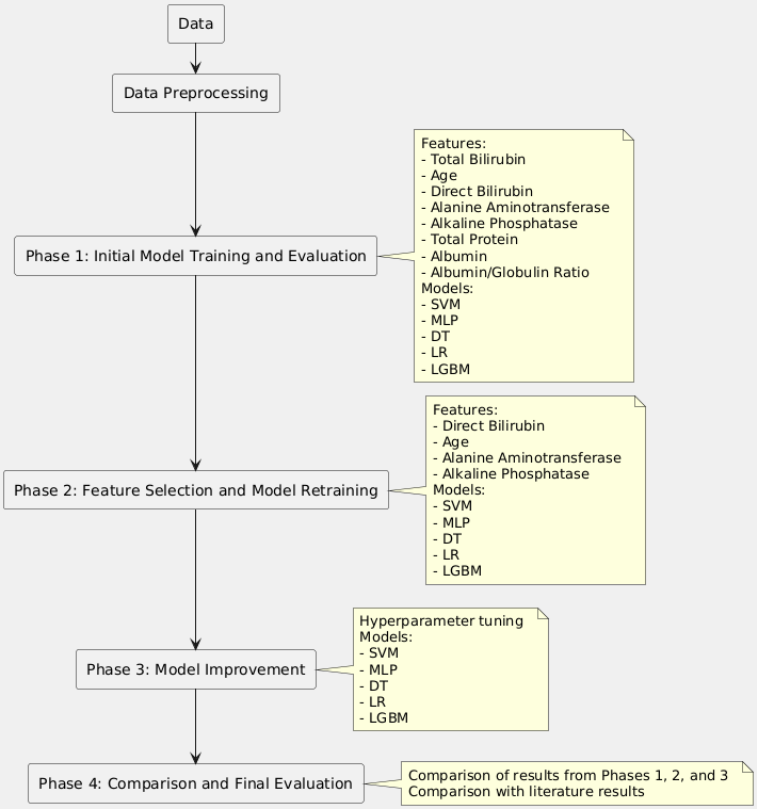

Fig. 3.1: Proposed Architecture.

performance results of the classification methods with model improvement were compared with the data set used in this study in the literature. The modelling of machine learning methods was carried out in the JupyterLab environment by means of the Python programming language. While measuring the success of machine learning methods, cross-validation method (cv) was used as 10 times.

3.1. Proposed Architecture. The proposed architecture Figure 3.1 and the algorithm (Algorithm 1) for liver disease prediction consists of four main phases:

Data Preparation and Initial Model Training: The ILPD dataset is preprocessed, with the gender attribute removed. The data is then split into 80% training (466 samples) and 20% test (117 samples) sets, using a random state of 42 for reproducibility. Five machine learning algorithms (SVM, MLP, DT, LR, and LGBM) are trained on the original dataset using all eight attributes: Total Bilirubin, Age, Direct Bilirubin, Alanine Aminotransferase, Alkaline Phosphatase, Total Protein, Albumin, and Albumin/Globulin Ratio.

Feature Selection: The Forward Selection method, a wrapper approach, is applied to identify the most important features in the dataset. This process reduces the feature set to four key attributes: Direct Bilirubin, Age, Alanine Aminotransferase, and Alkaline Phosphatase. Model Retraining with Selected Features: The five machine learning algorithms are retrained using only the four selected features. This step aims to improve model performance by focusing on the most relevant attributes.

Model Optimization and Evaluation: Each model undergoes hyperparameter tuning to further enhance its performance. The optimized models are then evaluated using 10-fold cross-validation. Performance metrics such as accuracy, specificity, sensitivity, precision, and F1-score are calculated and compared across all phases.

Table 3.1: Dataset properties

Sl#	Properties	Descrption	Data Type	Property Range Value
1	Age	Age	Numeric	[4-90]
2	Gender	Gender	Norminal	”male” or ”female”
4	TB	Total Bilirubin	Numeric	[0.4-75]
5	DB	Direct Bilirubin	Numeric	[0.1-19.7]
6	AlkPhos	Alkalen fosphatase	Numeric	[63-2110]
7	Sggt	Alanine Aminotransferase	Numeric	[10-2000]
8	Sgot	Aspartate Aminotransferase	Numeric	[10-4029]
9	TP	Total Protein	Numeric	[2.7-9.6]
10	ALB	Albumen	Numeric	[0.9-5.5]
11	A/G Ratio	Albumin/Globulin Ratio	Numeric	[0.3-2.8]
12	Selector Field	Selective Domsin Information	Numeric (12)	1-Sick,2-Not Sick

3.2. Data Set.

Dataset Selection. For this study, we chose the Indian Liver Patient Dataset (ILPD) [11]for our analysis and model development. The ILPD was selected for several key reasons:

Relevance and Specificity. The ILPD focuses specifically on liver patients, aligning perfectly with our research goal of improving liver disease diagnosis. It contains crucial liver function tests and patient attributes essential for identifying liver disorders.

Diversity and Representation. This dataset represents a diverse population from India, a country with a high prevalence of liver diseases, enhancing the generalizability of our findings. Dataset Quality: The ILPD is a well-curated dataset collected by reputable medical institutions, ensuring high data quality and reliability.

Balanced Representation. It includes both liver patients and non-liver patients, providing a balanced dataset crucial for developing accurate classification models.

Feature Richness. The ILPD contains various features including age, gender, and multiple blood tests, allowing for comprehensive analysis and feature selection experiments.

Benchmark Status. As a widely used dataset in liver disease prediction research, the ILPD enables direct comparison of our results with other state-of-the-art methods in the literature [16].

Public Availability. Being part of the UCI Machine Learning Repository, the ILPD is freely available, promoting reproducibility and further research in the field.

Challenging Nature. The dataset presents a complex classification problem, making it an excellent testbed for evaluating our proposed hybrid model and feature selection approach [17].

By utilizing the ILPD, we aim to contribute meaningfully to liver disease diagnosis research while ensuring our results are comparable, relevant, and potentially impactful for a significant patient population. This choice aligns well with our research objectives and the broader goal of improving early detection of liver diseases [18].

The total 583 sample patient records in the repository. Of these records, 441 were male and 142 were female. Of the patient records, 416 had liver disease and 167 did not have liver disease. The total men with liver disease is 324 and total women with liver disease is 92. There are 11 features in the dataset. While 10 of them can be used as attributes, the eleventh feature is the field where the presence of the disease is shown. SGPT and SGOT, which are found in attributes, are used under different names today. In the new terminology, SGOT is called ALT, and SGPT is called AST. Detailed information about the dataset characteristics is presented in Table 3.1.

3.2.1. Normalization Filter. The high number of changes between data affects the classification methods learning accuracy in some way. Normalization aims to eliminate discrepancies between mathematical operations and data, facilitating easier data comparison. In this study, the data were normalized using the Standard Scaler from the sklearn.preprocessing library in Python, applied to both MLP and SVM algorithms.

3.2.2. Attribute Selection. In this study, among wrapper methods (SFS) Step Forward Selection is one of them, was used to select features in the data set. In the SFS method, the cycle process is initiated

Table 3.2: Important attributes obtained after the SFS method.

Attribute Name	**Abbreviation**
Direct Bilirubin	DB
Age	Age
Alanine Aminotransferase	Sgpt
Alkaline Phosphatase	AlkFos

Table 3.3: Hyperparameters used in MLP.

Activation	**relu**
alpha	0.1
hidden_layer_size	(10,10,10,10,10)
random_state	42
max_iter	215

with the attribute that contributes the most to the performance of the established model, that is, it has the highest correlation with the dependent variable. Other attributes are checked in order according to the specified severity level ( ). Attributes that satisfy the condition based on the level of importance are added to the model. Attributes that don't meet the initial severity level requirement are not added to the model. Once all variables have been checked, the cycle ends. In this study, the attribute was selected with a significance level of 95%. The attributes determined by the SFS method has been presented in Table 3.2.

3.3. Support Vector Machine. Vapnik and his developed the Support Vector Machine algorithm in the 1990s. The Support Vector Machine, which is one of the supervised machine learning algorithms, is used to distinguish between binary base class data by applying statistical processing on it [12].

Since a classification belonging to two classes will be used in our data set, linear support vector classification is preferred. Reducing the value of the maximum number of iterations during the model setup phase while model optimization increased the accuracy value. Therefore, it max_iter the best performance in the accuracy rate, the hyperparameter is obtained as 23. While creating the improved model, the random state value was used as 42 and the maximum number of iters was used as 23. In this study, the Linear SVM machine learning method called LinearSVC, which is available in the Python programming language sklearn.svm library, was used.

3.4. Multilayer Sensors. Multilayer Perceptrons consist of input, intermediate and output layers. Unlike a single-layer sensor, the intermediate layer acts as a bridge between the output layer and the input layer. The middleware evaluates the inputs from the input layer against the problem before sending them to the output layer. As a result of the evaluation, a better decision is made according to the problem. The number of interlayers can be increased according to the condition of the problem [13].

In MLP, a five-layered model consisting of 10 neurons in each hidden layer was used in the improved model studies. The corrected linear unit function is preferred for activation function in hidden layers. A value of 0.1 as the L2 penalty parameter increased the accuracy. By adding the maximum number of iterations to 215, the best performance in accuracy was achieved. The hyperparameters that are changed when building an optimized model are shown in Table 3.3. In this study, the MLPClassifier classification algorithm in the Python programming language sklearn.neural_network library was used.

3.5. Logistic Regression. It is a statistical method that solves problems. It specifies two possible (0 or 1) outcomes by regression analysis of the dataset [13]. The outcome dependent variable used in this study is kept numerically (1-liver patient, 2-not liver disease).

In the improved model studies in LR, it has been observed that the use of liblinear as the preferred algorithm for optimization problems increases the success because the size of our data set is small.

Table 3.4: Hyperparameters used in logistic regression.

random_state	42
multi_class	ovr

Table 3.5: Hyperparameters used for the decision tree.

max_features	auto
max_depth	5
random_state	42

Table 3.6: Hyperparameters used in LGBM.

n_estimators	150
learning_rate	0.2
random_state	42

Since our result variable belongs to 2 classes (1-liver patient, 2-not liver disease), the multi-class feature was used as ovr. The hyperparameters that are changed when creating an optimized model are shown in Table 3.4. In this study, the LogisticRegression classification method in the Python programming language sklearn.linear_model library was used.

3.6. Decision Trees. Algorithms in supervised learning most widely used one is Decision Trees. The primary goal is to convert complex structures in the dataset into simpler ones. With decision trees, both numerical and categorical data can be processed. It is one of the commonly used approaches in classification because the test and training are fast and the results can be interpreted visually easily. It consists of roots, nodes and branching. In decision trees, operations are concluded in two steps. In step one, creation of the tree. In step two, classification rules are created from the created tree. Depending on the algorithms used to create the decision tree, the construction of the decision tree may vary [14]. As given in Algorithm 2, the tree is built using a top-down, recursive approach known as recursive partitioning. At each step, the algorithm chooses the attribute that best splits the set of items, typically using measures like Gini impurity or information gain.

In this study, the Decision Tree Classifier machine learning method in the Python programming language sklearn.tree library was used. The decision tree was created with the Classification and Regression Tree (CART) algorithm. Entropy algorithm was used for branching. Limiting the maximum depth of the tree to 5 increased the accuracy rate. For the best division, the max_features parameter auto corresponding to max_features=sqrt(n_features) was used. The hyperparameters that are changed when building an optimized model are shown in Table 3.5.

3.7. Light Gradient Reinforcement Machine Classifier. The Light Gradient Boosting Machine Classifier is a library developed by Microsoft in 2017. LGBM is one of the Gradient Boosting Machine (GBM) types developed to increase the training time performance of XGBoost. XGBoost makes the initial search transversely, while LGBM performs the first search in depth [15]. Increasing the learning rate from default parameters and limiting the number of trees to 150 in improved model studies in LGBM increased the prediction success.

The hyperparameters that are changed when creating an optimized model are shown in Table 3.6. In this study, the LGBMClassifier classification algorithm in the Python programming language lightgbm library was used.

4. Results and Evaluation. There are some criteria to evaluate the classification success of machine learning approaches used on data set. In the calculation of these criteria, the confusion matrix specified in Table 4.1 is used. The confusion matrix indicates the prediction accuracy success in the created model in a 2x2

Table 4.1: Confusion matrix.

	Class	Negative	Positive
Real Class	Negative	TN	FP
	Positive	FN	TP

matrix. It allows the comparison of actual values with the estimates made.

The terms mentioned in Table 4.1 can be explained as follows:

1. True Positive (TP): Specifies the accurately predicted liver disease class value.
2. False Positive (FP): Specifies the inaccurately predicted liver disease class value.
3. False Negative (FN): Specifies the inaccurately predicted non-liver disease class value.
4. True Negative (TN): Specifies the accurately predicted non-liver disease class value.

There are many metrics in the literature that are used to evaluate classification achievements. These metrics and their formulas, which are also used within the scope of this study, are explained in this section.

Accuracy. This metric represents the ratio of accurately predicted liver patients to all values Equ. 4.1.

$$\text{Accuracy } = \frac{TP + TN}{TP + FN + TN + FP} \tag{4.1}$$

Sensitivity. Also known as recall, this metric shows the ratio of accurately predicted liver patients to the sum of correctly predicted liver patients and those incorrectly predicted as not having liver disease. It demonstrates the model's ability to detect true positive cases Equ. 4.2.

$$\text{Sensitivity } = \frac{TP}{TP + FN} \tag{4.2}$$

Specificity. This metric represents the ratio of accurately predicted non-liver disease cases to the sum of all predicted non-liver disease cases. It shows the model's ability to correctly identify true negative cases Equ. 4.3

$$\text{Sensitivity } = \frac{TN}{TN + FP} \tag{4.3}$$

Precision. Ratio of accurately predicted liver patient values to the sum of accurately predicted liver patient values and correctly predicted liver patient values. It shows the ratio of liver patient class values to which liver disease is estimated shown in Equ. 4.4.

$$\text{Precision } = \frac{TP}{TP + FP} \tag{4.4}$$

F-Criterion. Also known as F1-score, this metric is the harmonic mean of Precision and Sensitivity. It provides a single score that balances both precision and recall Equ. 4.5.

$$F1 - \text{ Criterion} = 2 \times \frac{\text{Precision } \times \text{ Recall}}{\text{Precision } + \text{R}ecall} \frac{2 \times TP}{2 \times TP + FP + FN} \tag{4.5}$$

MAE (Mean Absolute Error). It is the amount of the difference between two continuous variables. Values close to zero perform better. The result values of the machine learning methods applied after the original version and attribute extraction from the data set are given in Table 4.2.

Many datasets can contain many related or unrelated data. The 11 attributes in the dataset were reduced to 4 attributes after the feature selection process. When Table 4.2 is examined, it is seen that the accuracy rates of MLP, SVM, DT, LGBM machine learning algorithms increase after feature extraction. A comparison of the accuracy values after the model refinement is applied after feature selection is given in Table 4.3 and Figure 4.1.

Table 4.2: Comparison of truth values before and after feature extraction.

Classification Algorithm	(%) Accuracy	
	Before Feature Selection	After Feature Selection
SVM	75.21	76.13
MLP	74.32	78.63
DT-CART	75.91	77.04
LR	78.09	77.81
LGBM	76.20	77.95

Table 4.3: Comparison of accuracy values before and after feature selection.

Classification Algorithm	(%) Accuracy		
	Before Feature Selection	After Feature Selection	After Model Optimization
SVM	75.22	76.13	77.87
MLP	74.31	78.63	81.13
DT-CART	75.90	77.04	81.13
LR	78.09	77.80	77.80
LGBM	76.21	77.95	82.12

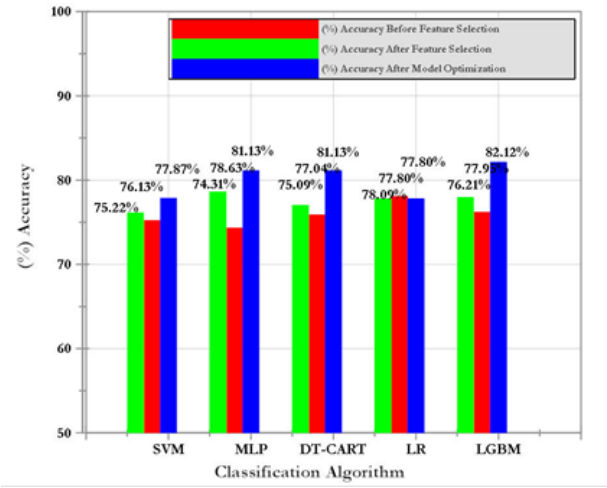

Fig. 4.1: Comparison of accuracy values before, after and model optimization

When Table 4.3 is examined, it is seen that the accuracy rates of all machine learning methods as in the Figure 4.1 increase after feature selection. The performance values of the machine learning methods used after the model improvement application stage after the feature extraction are given in Table 4.4. Figure 4.2 illustrated the performance evaluation of different ML algorithms, in terms of accuracy, sensitivity, specificity, precision and F-criterion.

When Table 4.4 is examined, the maximum diagnostic success rate was got from LGBM machine learning technique with the rate of 82.12%. With a success rate of 81.13%, which is close to the LGBM algorithm,

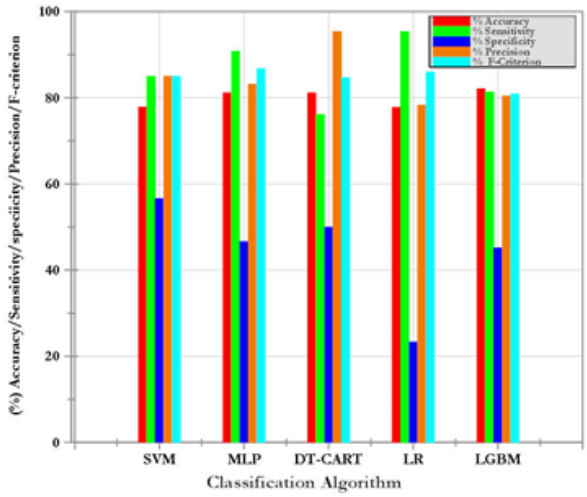

Fig. 4.2: Performance evaluation of different ML algorithms.

Table 4.4: Machine learning methods performance values.

Algorithm	% Accuracy	% Sensitivity	% Specificity	% Precision	% F-Criterion	Mean Absolute Error
MLP	81.13	90.80	46.66	83.15	86.81	0.2
SVM	77.87	85.05	56.66	85.05	85.05	0.22
LR	77.80	95.40	23.33	78.30	86.01	0.23
DT-CART	81.13	76.14	50.00	95.40	84.69	0.25
LGBM	82.12	81.39	45.16	80.45	80.92	0.28

DT-CART and MLP are the second high-performance methods. According to the sensitivity criterion, the highest rate was obtained from the LR method with 95.40%. When the results were evaluated according to the measure of precision, DT-CART was the most successful method with a rate of 95.40%. According to the criterion, MLP is the method that gives the best results with a rate of 86.81%. When the relevant studies in this field were examined, it was seen that LGBM machine learning method was not applied on the ILPD data set before.

Figure 4.3, when the studies evaluated according to MAE were considered, the lowest MAE error rate was obtained. Machine learning is also frequently used in the field of healthcare. With the development of technology, it is aimed to diagnose diseases early by using smart applications as well as traditional diagnostic methods in disease diagnosis. Machine learning algorithms are often utilised in the development of intelligent applications. In this study, the focus is on the early diagnosis of liver failure, which has been seen frequently in recent years and leads to loss of life if not diagnosed in the early stages, with high accuracy. With the feature selection method, it was reduced from 11 attributes to 4 attributes and higher prediction successes were obtained. Models were created on five different machine learning methods used. Model improvements were made to these models and the disease prediction success of the models was increased. In the selection of the model, highest accuracy, precision, sensitivity, and F-Criterion ratios were taken into account. The most successful result was obtained from the LGBM method with an accuracy rate of 82.12%.

5. Conclusion. In this study, we employed Forward Selection (FS) for feature selection and applied various machine learning methods, including Multilayer Perceptron (MLP), Support Vector Machine (SVM), Decision Trees (DT-CART), Light Gradient Boosting Machine (LGBM), and Logistic Regression (LR), to the Indian

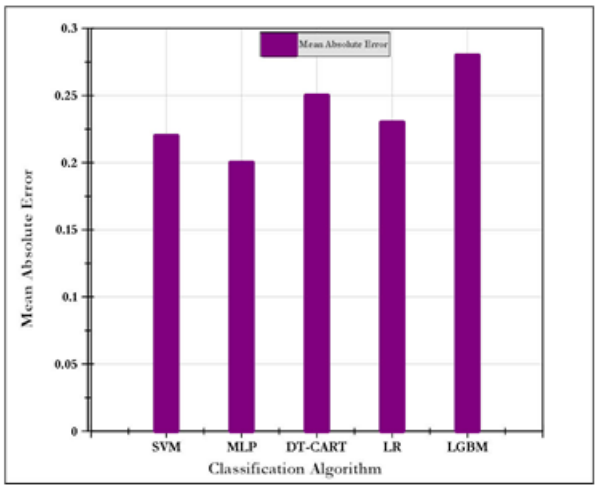

Fig. 4.3: Mean Absolute Error comparison for different Classification Algorithms.

Liver Patient Dataset (ILPD) for the early diagnosis of liver diseases. The diagnostic success of these methods was evaluated before and after feature selection, as well as after model optimization. Among the evaluated methods, the Light Gradient Boosting Machine (LGBM) demonstrated the best overall performance. After feature selection and model optimization, the LGBM achieved the highest accuracy of 82.12%, significantly improving from its initial accuracy of 76.21%. Additionally, the LGBM model showed a balanced performance across various metrics, with a sensitivity of 81.39%, specificity of 45.16%, precision of 80.45%, and F1-score of 80.92%, alongside a mean absolute error of 0.28.

Comparatively, the Multilayer Perceptron (MLP) also showed strong performance, with an accuracy of 81.13% after model optimization, improving from 74.31% before feature selection. However, LGBM's higher accuracy and balanced performance across different metrics make it the preferred choice in this study for early diagnosis of liver diseases.

The results underscore the effectiveness of feature selection and model optimization in enhancing performance of machine learning methods for medicinal diagnostics. Future research could focus on incorporating additional clinical data and exploring more advanced machine learning techniques to further improve diagnostic accuracy and reliability.

REFERENCES

[1] I. Arshad, C. Dutta, T. Choudhury and A. Thakral, , *"Liver disease detection due to excessive alcoholism using data mining techniques",* 2018 International Conference on Advances in Computing and Communication Engineering (ICACCE), pp. 163-168, 2018.

[2] H. W. Lee, J. J. Sung, and S. H. Ahn, , *"Artificial intelligence in liver disease,"* Journal of Gastroenterology and Hepatology, Vol.36, Iss.3, pp.539-542, 2021.

[3] Mamdouh E. ve Mabrouk M. , *"A Study of Support Vector Machine Algorithm For Liver Disease Diagnosis",* American Journal of Intelligent Systems, 4 (1), 9-14, 2021.

[4] Gulia A. Vohra R. Rani P. , *"Liver Patient Classification Using Intelligent Techniques",* International Journal of Computer Science and Information Technologies), 5 (4): 5110-5115, 2014.

[5] Naeem, S., Ali, A., Qadri, S., Khan Mashwani, W., Tairan, N., Shah, H., ... & Anam, S. , *Machine-learning based hybrid-feature analysis for liver cancer classification using fused (MR and CT) images.* Applied Sciences, 10(9), 3134, 2020.

[6] Priya M.B. Juliet P. L. Tamilselvi P.R." , *Performance Analysis of Liver Disease Prediction Using Machine Learning Algorithms"*, International Research Journal of Engineering and Technology (IRJET), 206-211, 2018.

[7] Chaurasia, V., Pal, S., , . *Machine learning algorithms using binary classification and multi model ensemble techniques for skin diseases prediction.* International Journal of Biomedical Engineering and Technology, 34(1), 57-74, 2020.

[8] Gertz, M.; Adams, D.; Ando, Y.; Beirão, J.M.; Bokhari, S.; Coelho, T.; Comenzo, R.L.; Damy, T.; Dorbala, S.; Drachman, B.M.; et al. , *Avoiding misdiagnosis: Expert consensus recommendations for the suspicion and diagnosis of transthyretin amyloidosis for the general practitioner.* BMC Fam. Pract., 21, 198, 2020.

[9] Ahmed, S. T., Kumar, V. V., & Kim, J. , *AITel: eHealth augmented-intelligence-based telemedicine resource recommendation framework for IoT devices in smart cities.* IEEE Internet of Things Journal, 10(21), 18461-184685, 2023.

[10] Eltanashi, S., & Atasoy, F. , *Proposed speaker recognition model using optimized feed forward neural network and hybrid time-mel speech feature.* ICATCES 2020 Proceeding Book, 130-140.

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

[12] Koike, H.; Katsuno, M. , *Transthyretin Amyloidosis: Update on the Clinical Spectrum, Pathogenesis, and Disease-Modifying Therapies.* Neurol. Ther. 9, 317–333, 2020.

[13] Tozza, S.; Severi, D.; Spina, E.; Di Paolantonio, A.; Iovino, A.; Guglielmino, V.; Aruta, F.; Nolano, M.; Sabatelli, M.; Santoro, L.; et al. , *A compound score to screen patients with hereditary transthyretin amyloidosis.* J. Neurol. 269, 4281–4287, 2022.

[14] Manjunatha, S., Swetha, M. D., Rashmi, S., & Subramanian, A. K. , *Convolutional Neural Network-based image tamper detection with Error Level Analysis.*, 2024.

[15] Di Stefano, V.; Thomas, E.; Alonge, P.; Giustino, V.; Pillitteri, G.; Leale, I.; Torrente, A.; Pignolo, A.; Norata, D.; Iacono, S.; et al. , *Patisiran Enhances Muscle Mass after Nine Months of Treatment in ATTRv Amyloidosis: A Study with Bioelectrical Impedance Analysis and Handgrip Strength.* Biomedicines, 11, 62, 2022.

[16] Subashchandrabose, U., John, R., Anbazhagu, U. V., Venkatesan, V. K., & Thyluru Ramakrishna, M. , *Ensemble Federated learning approach for diagnostics of multi-order lung cancer. Diagnostics* 13(19), 3053, 2023.

[17] Luhadia, G., Joshi, A. D., Vijayarajan, V., & Vinoth Kumar, V. (2023, April)., *Fusion of Variational Autoencoder-Generative Adversarial Networks and Siamese Neural Networks for Face Matching. In International Conference on Frontiers of Intelligent Computing: Theory and Applications, Singapore: Springer Nature Singapore.*,231-241, 2023.

[18] Dhiman, G., Vinoth Kumar, V., Kaur, A., & Sharma, A. , *Don: deep learning and optimization-based framework for detection of novel coronavirus disease using x-ray images. Interdisciplinary Sciences: Computational Life Sciences*, 13, 260-272, 2021.

Edited by: Dhilip Kumar V
Special issue on: Unleashing the power of Edge AI for Scalable Image and Video Processing
Received: Jul 7, 2024
Accepted: Sep 18, 2024