# Liver Disease Prediction Using a Hybrid Machine Learning Approach

**Sanjit Kumar Dash**
Odisha University of Technology and Research, Bhubaneswar, Odisha, India
skdash@outr.ac.in

**Nitish Agrawal**
Odisha University of Technology and Research, Bhubaneswar, Odisha, India
nitishagrawal2022@gmail.com

**Rahul Agarwalla**
Odisha University of Technology and Research, Bhubaneswar, Odisha, India
rahulagarwalla1234@gmail.com

**Mohammed Altaf Ahmed**
Department of Computer Engineering, College of Computer Engineering and Sciences, Prince Sattam bin Abdulaziz University, Al-Kharj, Saudi Arabia
m.altaf@psau.edu.sa (corresponding author)

**Suleman Alnatheer**
Department of Computer Engineering, College of Computer Engineering and Sciences, Prince Sattam bin Abdulaziz University, Al-Kharj, Saudi Arabia
s.alnatheer@psau.edu.sa

**Qutubuddin Mohammed**
Department of Electrical Engineering, College of Engineering Wadi Addawasir, Prince Sattam bin Abdulaziz University, Saudi Arabia
q.mohammed@psau.edu.sa

*Received: 4 November 2025 | Revised: 17 December 2025 and 29 December 2025 | Accepted: 6 January 2026*

 *DOI: https://doi.org/10.48084/etasr.15959*

**ABSTRACT**

**Liver disease poses a severe threat to human health if not detected early. Existing diagnostic methods are usually time-consuming, expensive, and require expertise, which is often unavailable in healthcare facilities. This study introduces a hybrid AI-based diagnostic framework that integrates both Deep Learning (DL) and Machine Learning (ML) techniques to support the early and accurate detection of liver disease. The proposed hybrid model integrates a MultiLayer Perceptron Neural Network (MLPNN) with a soft Voting Classifier, which includes Extreme Gradient Boosting (XGB) and Light Gradient Boosting Machine (LGBM). To enhance the predictive performance of the model, advanced feature engineering techniques were employed, including formulating medically pertinent ratios and balancing the data using SMOTE-Tomek resampling. The proposed hybrid model achieved an accuracy of 95.49%, demonstrating remarkable generalization capabilities across the dataset. The proposed model is strong and reliable, as demonstrated by the confusion matrix, classification report, and ROC-AUC curve results.**

***Keywords-liver disease; neural network; deep learning; voting classifier; boosting; SMOTE analysis; sustainable healthcare***

## I. INTRODUCTION

Liver disease is a significant health issue worldwide, accounting for millions of deaths every year due to cirrhosis, hepatitis, and liver cancer [1]. Since the liver plays a central function in metabolism, detoxification, and protein production, proper liver function is essential for health. Early diagnosis and intervention are paramount to prevent progression of the disease, but conventional diagnostics such as biopsy and imaging scans are usually time-consuming, expensive, and invasive [2]. In addition, symptoms may not be visible until late, and therefore, it is difficult for clinicians to diagnose the disease promptly. Consequently, there is a growing need for affordable and effective diagnostic techniques that can provide early and accurate identification of liver disease.

Machine Learning (ML) and Deep Learning (DL) have emerged as influential techniques in medical diagnostics, facilitating the creation of automated and highly precise disease prediction models. Hybrid approaches have integrated DL with ensemble ML techniques, demonstrating significant potential for classifying liver disease, paving the way for enhanced early detection and better patient outcomes [3, 4]. In addition, some studies have employed techniques such as feature selection and class balancing methods, such as the Synthetic Minority Over-sampling Technique (SMOTE), to enhance model performance [5]. Recent advances in ML for liver disease prediction have greatly improved diagnostic capabilities, outpacing traditional methods and providing essential support to healthcare professionals. Moreover, their ability to understand intricate patterns renders them a crucial component in contemporary healthcare frameworks for anticipatory disease management and intervention [6, 7].

With the development of smart healthcare, the integration of predictive models with IoT devices and intelligent systems is necessary to bring about a revolution in patient care. An automated and accurate diagnostic tool for liver disease is a building block in the development of connected health environments. Such a model could become the heart of many practical applications, such as IoT-based point-of-care diagnostic devices in remote clinics or decision support modules in telemedicine. By putting this intelligence on the edge, healthcare organizations can realize constant risk assessments in real time to facilitate proactive individual patient management. This work is not just about the accuracy of algorithms in a single model, but how models can be embedded in a larger ecosystem of intelligent health monitoring.

The main objective of this research was to develop a hybrid diagnostic model that combines Deep Learning (MLPNN) with ensemble ML techniques (XGB and LGBM) for the timely and precise diagnosis of liver disease. This hybrid approach aimed to combine the strengths of both DL and ensemble ML techniques to capture complex, non-linear patterns in the data while improving accuracy and robustness through model diversification. The key novelty lies in the integration of MLPNN with ensemble classifiers, which allows the model to effectively handle the intricacies of liver disease classification. Additionally, model performance is improved using enhanced feature engineering techniques and applying SMOTE-Tomek for class balancing. The novelty of this work lies not only in proposing a new classification algorithm but in the practical integration of DL and ensemble boosting through a dual-path fusion framework, enhanced by medically meaningful feature ratios and effective class imbalance handling. In addition, this method prioritizes computational efficiency and suitability for deployment in resource-constrained healthcare settings.

## II. RELATED WORKS

A significant development in the application of ML for liver disease diagnosis is the use of hybrid and ensemble models that combine multiple classifiers to improve predictive performance. Early foundational work in this area proposed an intelligent hybrid diagnosis model [8]. This two-phase approach first used an Artificial Neural Network (ANN) to classify patients as healthy or diseased. Then, it employed a combination of Case-Based Reasoning (CBR) and Analytic Hierarchy Process (AHP) to identify the specific type of liver disease the patient was suffering from, demonstrating the potential of multi-stage models.

Recent studies have focused on improving these models, frequently using benchmark datasets such as the Indian Liver Patient Dataset (ILPD). In [9], several ML models were investigated, showing that an ensemble voting classifier outperformed individual models. This study underscored the paramount importance of preprocessing, as optimal model performance was attained after the application of SMOTE to rectify the dataset's class imbalance. In [10], a comprehensive adaptive analysis of nine different classifiers across 12 unique medical datasets further highlighted the challenge of algorithm selection. This study demonstrated that no single algorithm is universally superior, as for the ILPD, SVM and ZeroR performed well, while for a hepatitis dataset, SVM and Voting Frequency Intervals (VFI) were optimal.

In addition to ensemble methods, researchers have investigated DL models to determine whether they can capture complex dependencies. In [11], a Bidirectional Long Short-Term Memory (BiLSTM) model on the ILPD dataset demonstrated its ability to monitor long-term relationships both in forward and backward directions within patient data, improving predictive accuracy. Moving further into hybrid DL, in [12], a Convolutional Neural Network (CNN) was used to process histopathological images and fused with an LSTM to analyze structured clinical data, creating a comprehensive set of features for the detection of cirrhosis. In [13], a hybrid CNN-Transformer model extracted both local and global spatio-temporal features from Contrast-Enhanced Ultrasound (CEUS) videos to classify focal liver lesions [13]. Similarly, in [14], a multi-class ensemble DL framework was employed for real-time liver tumor detection from CT images, specifically designed to address variability in tumor shapes and noise.

However, the most promising results have often been achieved with advanced ensemble stacking techniques. In [15], a sophisticated two-level ensemble stacking model on ILPD included comprehensive preprocessing and feature selection, leading to a highly accurate model. In [16], a multiclass liver disease prediction framework incorporated adaptive data

preprocessing steps such as class-specific imputation, outlier handling, feature normalization, and ensemble modeling. This study demonstrated that strong adaptive preprocessing combined with ensemble learning significantly improves predictive accuracy across multiple liver disease stages. Further validating this direction, another stacked ensemble framework used Extra Trees (ET) and Random Forest (RF) as base classifiers and an SVM as meta-classifier, achieving remarkable predictive accuracy [17]. This hybrid philosophy was also applied in [18], combining SVM, for its optimal hyperplanes, with XGBoost, for its robust gradient boosting, to improve the classification accuracy for cirrhosis.

Beyond the model architecture, feature selection and hyperparameter tuning are critical. In [19], a hybrid framework also focused on XGBoost, combining a dual feature selection strategy (Mutual Information and RFE) with a genetically-optimized XGBoost classifier. This focus on preprocessing and optimization achieved strong predictive capabilities, providing an explainable model using SHAP.

This trend is confirmed by comparative studies, such as [20], in which analysis on ILPD showed that a stacking classifier achieved a high level of performance compared to other standalone models. Collectively, these studies show a clear trend: from single classifiers to advanced, stacked ensembles on both tabular and imaging data, combined with rigorous preprocessing, to build highly accurate and reliable diagnostic models for liver disease.

## III. METHODOLOGY

Figure 1 shows an ML framework that starts with collecting data and then moves to preprocessing, splitting, and scaling the data. This model then deals with class imbalance by using over- or undersampling techniques. This framework combined DL and ensemble-based models to classify liver disease so that predictions are strong, accurate, and easy to understand. The proposed model combines an MLPNN with a Voting Classifier that uses two strong gradient boosting methods, XGB and LGBM. The process concludes by visualizing performance metrics and comparing the models to identify the most effective.

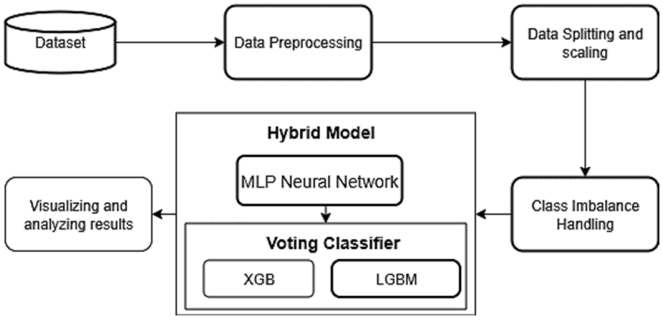


Fig. 1. Model framework representation.

### A. Dataset Description and Analysis

This study used the Indian Liver Patient Database (ILPD), which is publicly available on the UCI machine learning repository [21]. The dataset provides data of 583 unique patients having 11 attributes, 1 target variable, and 10 features. The 10 features used in prediction are age, gender, total bilirubin, direct bilirubin, total proteins, albumin, A/G ratio, SGPT, SGOT, and alk-phos. The target variable showcases two outcomes: patients without liver disease and patients with liver disease. The dataset consists of 416 patient records with liver disease and 167 without.

Figure 2 visualizes the correlations between the features using a heatmap, demonstrating relationships that affect model performance. High correlations between Total Bilirubin and Direct Bilirubin (0.87) and Albumin and A/G Ratio (0.85) suggest redundancy, which was met by excluding highly correlated features to reduce multicollinearity. Low correlations among the majority of other features imply that they add independent information, providing heterogeneous predictive inputs to the model.

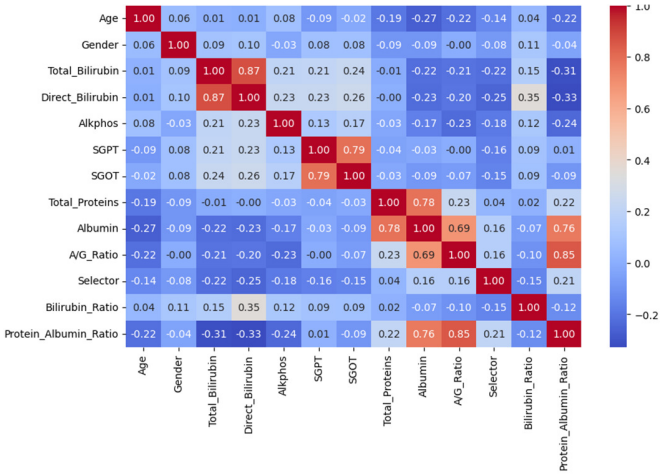


Fig. 2. Feature correlation heatmap.

### B. Data Preprocessing

Through various steps, the data was cleaned to be uniform and suitable for modelling. This procedure also addressed missing values, interrelated features, and imbalanced class sizes. First, the dataset was loaded, and the columns were labelled based on relevant properties that relate to the subject of interest. The categorical feature Gender was converted into a numerical format, using label encoding, to make it compatible with machine learning algorithms. Missing values were handled using median-based statistical imputation, which is robust to outliers and well-suited for skewed medical attributes such as liver enzyme measurements. This approach preserves the statistical integrity of the dataset while ensuring completeness before feature engineering and model training. Two medically meaningful features, Bilirubin_Ratio (Direct Bilirubin to Total Bilirubin) and Protein_Albumin_Ratio (Albumin to Total Proteins), were engineered to capture essential biological relationships. As shown in Figure 2 high correlation between 2 pairs of features, Total Bilirubin and Direct Bilirubin (0.87) and Albumin and A/G Ratio (0.85), were identified. Hence, features with an absolute correlation above 0.8 were removed, minimizing redundancy and preventing model overfitting due to collinearity.

### *C. Data Splitting and Scaling*

The dataset was divided into 80% for training and 20% for testing, ensuring that the proportions of the categories remained the same. This is important to obtain unbiased and correct results while testing for model performance. After splitting the data, standardization was performed using the StandardScaler. This technique rescales features to have a mean of zero and a standard deviation of one. Thus, all the factors in the data contribute equally during learning, which results in greater accuracy and more rapid learning of the model.

### *D. Class Imbalance Handling*

Class imbalance is a common problem in medical datasets that can result in biased model predictions, where the majority class is favored and the minority class is ignored. To counteract this, SMOTE with Tomek links (SMOTE-Tomek) was used. SMOTE creates synthetic samples for the minority class to decrease the imbalance, and Tomek links detect and eliminate borderline majority class samples, sharpening decision boundaries and enhancing classification performance. SMOTE-Tomek ensures that the classifier is given a better-balanced training set and thus reduces bias and improves model generalizability across classes. This hybrid process is especially effective in medical databases, where misclassification of the minority class can result in serious harm, preventing overfitting without losing discrimination between classes [9]. The resulting balanced dataset improves model training, offering enhanced sensitivity and predictive performance overall.

### *E. Proposed Hybrid Model*

Figure 3 illustrates the proposed hybrid model. MLPNN consists of multiple layers of interconnected neurons, each applying non-linear transformations through activation functions to predict outputs. XGB merges regularization and advanced optimization techniques for accurate and generalized learning [22], while LGBM enhances speed and scalability using histogram-based decision tree learning. First, the MLPNN is trained to learn complex nonlinear relationships from the dataset. The Voting Classifier combines two powerful tree-based models, which individually predict the class labels, and the final decision is made by aggregating their predictions either by majority vote (hard voting) or by averaging the class probabilities (soft voting).

The proposed architecture consists of two modules that function in parallel: MPLNN and Boosting. The MLPNN was configured with two hidden layers consisting of 100 and 50 neurons, respectively, using the ReLU activation function. The Adam optimizer was employed for training with a maximum of 500 iterations and binary cross-entropy as the loss function. For the boosting models, XGBoost was configured with 100 estimators, a maximum tree depth of 3, and a learning rate of 0.05. Similarly, LightGBM was implemented with 100 estimators, a maximum depth of 3, and a learning rate of 0.05. These hyperparameters were fixed empirically and kept constant across all experiments to ensure reproducibility and fair comparison.

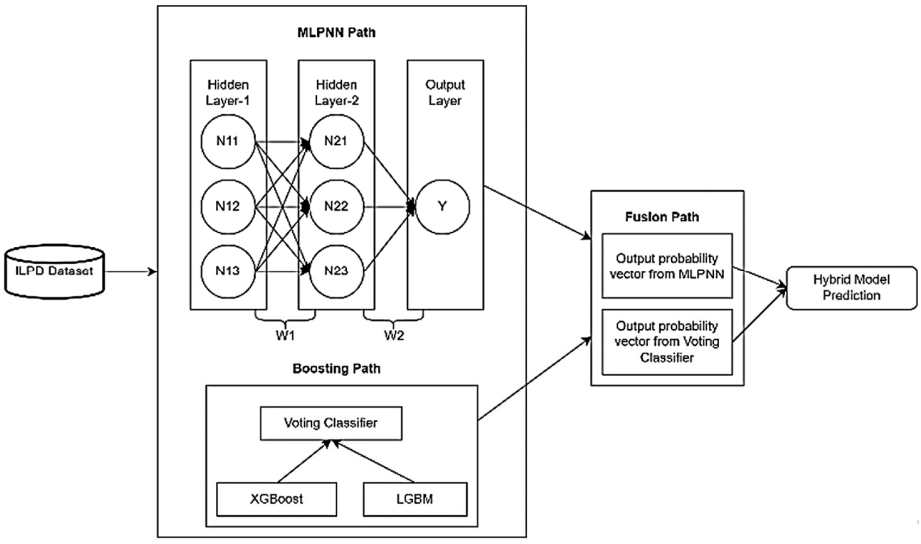


Fig. 3. Hybrid model architecture using MLPNN and Boosting.

#### *1) MLPNN*

MLP is a deep neural network trained to learn hierarchical non-linear representations from the input features, consisting of an input layer, two hidden layers, and an output layer. Each layer performs a transformation using weight matrices and bias vectors, followed by an activation function. The first hidden layer computes a weighted sum of the previous layer's output ($x$) and applies a (usually nonlinear) transformation, learning fundamental shapes and patterns from raw data (1). The second hidden layer takes the output from the first and passes it through weights, biases, and activation once again by stacking together features of level 1 to capture higher-level patterns (2). The output layer takes the final hidden representation and uses a sigmoid activation function to compute the class probability, returning a value between 0 and 1 (3).

$$h^{(1)} = \sigma(W^{(1)}x + b^{(1)}) \quad (1)$$

$$h^{(2)} = \sigma(W^{(2)}h^{(1)} + b^{(2)}) \quad (2)$$

$$\hat{y} = sigmoid(W^{(3)}h^{(2)} + b^{(3)}) \quad (3)$$

where $x$ represents the input, $W^{(1)}$ and $b^{(1)}$ are weights and biases, and $\sigma$ represents a non-linear activation function. $h^{(1)}$ and $h^{(2)}$ represent the output of hidden layers 1 and 2, respectively, and $\hat{y}$ is the predicted probability.

To train the MLPNN effectively, the Binary Cross-Entropy Loss Function ($\mathcal{L}$) determines the extent to which predicted probabilities diverge from the true binary labels (4). This loss guides the learning process, helping the model to adjust its internal parameters (weights and biases) through backpropagation and optimize the weights of the MLPNN.

$$\mathcal{L} = -\frac{1}{N}\sum_{i=1}^{N}(y_i \log(\hat{y}_i) + (1 - y_i)\log(1 - \hat{y}_i)) \quad (4)$$

where $N$ is the number of observations, $\hat{y}_i$ is the predicted probability, and $y_i$ is the actual probability.

*2) Boosting*

Two independent classifiers—XGBoost and LightGBM—were trained on the same input, with a Voting Classifier integrating their individual predictive capabilities. This meta-model takes the output probability vectors from both XGBoost and LightGBM and averages them using soft voting (5). This process enhances generalization by reducing the variance and bias of individual models. Both XGBoost and LightGBM are trained separately to output class probabilities, which are averaged to produce the Voting Classifier output:

$$f_{\text{Voting}}(X) = \frac{f_{\text{XGB}}(X) + f_{\text{LGBM}}(X)}{2} \quad (5)$$

where $X$ represents the input features, $f_{\text{XGB}}(X)$ represents the output probability vector from XGBoost, and $f_{\text{LGBM}}(X)$ is the output probability vector from LGBM.

*3) Fusion Layer: Combining Deep and Boosted Learning*

After generating prediction probability vectors from both the MLPNN (deep learning path) and the Voting Classifier (boosting path), a crucial step is to combine these outputs to produce a single, unified prediction. This approach employed the Simple Average Fusion strategy,

$$f_{\text{Hybrid}}(X) = \frac{f_{\text{MLP}}(X) + f_{\text{Voting}}(X)}{2} \quad (6)$$

where the final class probability is calculated as the direct average of output probabilities generated from both paths. This simple yet effective method strongly combines the predictive power of both the constituent models.

*4) Decision Layer*

Once the fusion layer calculates the combined output probabilities from the MLPNN and the Voting Classifier, the model makes a final decision regarding the classification label. This is achieved using a decision rule that selects the class with the highest fused probability. This final step transforms the continuous probabilistic output into a discrete label (0 or 1), representing whether a patient is likely to have liver disease:

$$\hat{y} = \arg\max\left(f_{\text{Hybrid}}(X)\right) \quad (7)$$

## IV. RESULTS AND DISCUSSION

### A. Experimental Setup

The model was developed and evaluated in a Google Colab environment using Python 3. The key libraries for the implementation included Scikit-learn, Pandas, and NumPy, with Keras (TensorFlow backend) used for the MLPNN and XGBoost and LightGBM for the boosting classifiers. All experiments were conducted on a standard CPU runtime, confirming the model's computational efficiency without requiring specialized hardware accelerators. This standardized setup ensures the complete reproducibility of the results.

To evaluate the model performance, the dataset was processed using a unified preprocessing pipeline, followed by performance assessment using 10-fold cross-validation to ensure model reliability.

### B. Hybrid Model Performance

Figure 4 shows the confusion matrix obtained from the proposed hybrid model, providing a breakdown of correct and incorrect predictions across the two classes: "Disease" and "No Disease". Figure 5 presents the AUC-ROC generated for the model, plotting the True Positive Rate (sensitivity) against the False Positive Rate, illustrating the diagnostic ability of the classifier at various threshold settings.

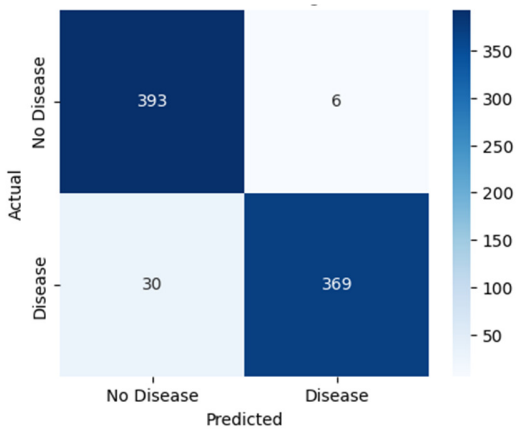


Fig. 4. Confusion matrix for the proposed hybrid model.

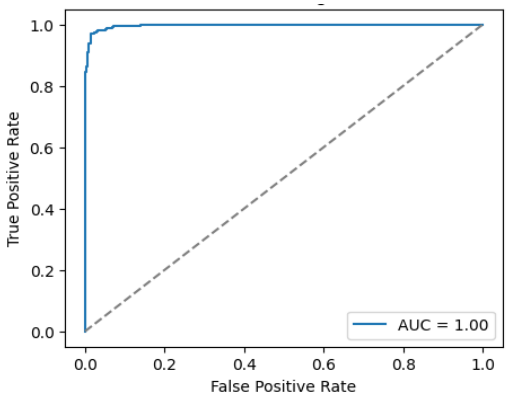


Fig. 5. AUC-ROC for the hybrid model.

Of all test cases, 393 were correctly identified as "No Disease", while 369 were accurately predicted as "Disease". Only 6 cases were misclassified as having the disease when they did not, and 30 diseased cases were incorrectly predicted as healthy. With an AUC-ROC score of 0.9970, the proposed hybrid model demonstrated near-perfect discrimination between classes. Table I presents a comprehensive performance evaluation of the proposed model, reflecting the effectiveness of the model in handling both positive and negative class predictions, confirming its robustness and reliability.

TABLE I. MODEL PERFORMANCE SCORES

| Evaluation metrics | Proposed model score (%) |
|---|---|
| Accuracy | 95.49 |
| Precision | 98.40 |
| Recall (Sensitivity) | 92.48 |
| F1 score | 95 |
| Specificity | 98.50 |
| AUC-ROC | 99.67 |

### C. Comparison Analysis

Table II presents a comparative evaluation of the proposed hybrid model (MLPNN+Boosting) against various existing studies in the domain of liver disease prediction. The proposed model outperforms all previously reported performances, achieving the highest accuracy of 95.49%, precision of 98.40%, and a high specificity of 98.50%, highlighting the effectiveness of integrating MLPNN with boosting techniques such as XGBoost and LGBM in enhancing classification performance for liver disease prediction tasks.

TABLE II. COMPARISON OF PROPOSED MODEL AGAINST EXISTING STUDIES

| Study | Algorithm used | Model | Accuracy (%) |
|---|---|---|---|
| [15] | Two-level stacking | Ensemble | 94.01 |
| [12] | CNN+LSTM | Hybrid | 92.0 |
| [16] | Adaptive preprocessing + Ensemble modeling | Ensemble | 99.0 |
| [17] | Autoencoder+SVM | Hybrid | 91.3 |
| This | MLPNN+Boosting | Hybrid | 95.49 |

## V. CONCLUSION & FUTURE WORKS

This study presented a hybrid ensemble framework that combines MLPNN with advanced boosting methods (XGBoost and LightGBM) to make better predictions about liver disease. The model has been thoroughly trained and tested using a range of performance metrics, such as accuracy, precision, recall, specificity, and AUC-ROC. This approach involved preprocessing the data, creating new features, and addressing the data imbalance with SMOTE Tomek. The ensemble method takes advantage of neural networks' ability to identify nonlinear relationships in data, while the boosting algorithms improve both predictive performance and generalization. The experimental results demonstrate that the proposed model achieved an exceptional accuracy of 95.49%, surpassing existing models. Cross-validation and comparative analysis further demonstrated the model's strength.

A noteworthy aspect of this research is the computational efficiency and real-time suitability of the model, which contributes to the Sustainable Development Goal in healthcare. All reported results were achieved in a standard CPU-only environment, verifying that the proposed model is computationally efficient and does not require special hardware acceleration. This efficiency results in a short inference time, leading to near-real-time risk assessment of new patient data for a clinical decision support application. Moreover, the model is highly scalable, while its base algorithms (MLPNN, XGBoost, and LightGBM) are known for good performance on large-scale datasets. This allows the model to be retrained on increasing datasets in order to improve its accuracy and degree of generalization at a cost that grows only moderately with the size of the training set.

Although the proposed model presents encouraging results, its performance could be truly tested within real-world clinical frameworks and intelligent healthcare systems. Future work will focus on deploying this high-accuracy model for practical application and integrating it with Internet of Things (IoT) systems. Furthermore, deploying the model on edge computing devices to process data locally (reduced latency), without leaking out patient privacy information, will work in regions with limited internet connectivity. Another important immediate step is to connect the model to telemedicine repositories and hospital EHR systems, offering a powerful clinical decision support tool: A physician could enter a patient's liver function test results and immediately receive an AI-derived risk assessment in the course of that consultation. Through exploring such avenues, this work can transform from a predictor to an integral part of an intelligent and proactive healthcare system for liver disease care.

## ACKNOWLEDGMENT

The authors extend their appreciation to Prince Sattam bin Abdulaziz University for funding this research work through project number PSAU/ 2025/01/32835.

## DATA AVAILABILITY

The dataset used to train and test the model is accessible in [21].

## REFERENCES

[1] H. Devarbhavi, S. K. Asrani, J. P. Arab, Y. A. Nartey, E. Pose, and P. S. Kamath, "Global burden of liver disease: 2023 update," *Journal of Hepatology*, vol. 79, no. 2, pp. 516–537, Aug. 2023, https://doi.org/10.1016/j.jhep.2023.03.017.

[2] J. C. T. Lai, L. Y. Liang, and G. L. H. Wong, "Noninvasive tests for liver fibrosis in 2024: are there different scales for different diseases?," *Gastroenterology Report*, vol. 12, Jan. 2024, Art. no. goae024, https://doi.org/10.1093/gastro/goae024.

[3] A. Quadir *et al.*, "Enhanced Preprocessing Approach Using Ensemble Machine Learning Algorithms for Detecting Liver Disease," *Biomedicines*, vol. 11, no. 2, Feb. 2023, https://doi.org/10.3390/biomedicines11020581.

[4] R. Amin, R. Yasmin, S. Ruhi, M. H. Rahman, and M. S. Reza, "Prediction of chronic liver disease patients using integrated projection based statistical feature extraction with machine learning algorithms," *Informatics in Medicine Unlocked*, vol. 36, Jan. 2023, Art. no. 101155, https://doi.org/10.1016/j.imu.2022.101155.

[5] A. Dhyani *et al.*, "Comparative Analysis of Supervised Machine Learning Algorithms for Liver Disease Prediction with SMOTE Enhancement," in *2023 3rd Asian Conference on Innovation in Technology (ASIANCON)*, Ravet IN, India, Aug. 2023, pp. 1–6, https://doi.org/10.1109/ASIANCON58793.2023.10270381.

[6] R. Ashtagi, V. Musale, V. S. Rajput, S. Chinchmalatpure, S. Mohite, and R. V. Bidwe, "Revolutionizing Early Liver Disease Detection: Exploring Machine Learning and Ensemble Models," *International Journal of Intelligent Systems and Applications in Engineering*, vol. 12, no. 13s, pp. 528–541, Jan. 2024.

[7] M. Ghosh *et al.*, "A Comparative Analysis of Machine Learning Algorithms to Predict Liver Disease," *Intelligent Automation & Soft Computing*, vol. 30, no. 3, pp. 917–928, 2021, https://doi.org/10.32604/iasc.2021.017989.

[8] R. H. Lin and C. L. Chuang, "A hybrid diagnosis model for determining the types of the liver disease," *Computers in Biology and Medicine*, vol. 40, no. 7, pp. 665–670, July 2010, https://doi.org/10.1016/j.compbiomed.2010.06.002.

[9] E. Dritsas, M. Trigka, E. Dritsas, and M. Trigka, "Supervised Machine Learning Models for Liver Disease Risk Prediction," *Computers*, vol. 12, no. 1, Jan. 2023, https://doi.org/10.3390/computers12010019.

[10] S. Elsayed, M. M. Ismail, A. F. Abdel-Gawad, and I. Mohamed, "Adaptive Analysis of Machine Learning Algorithms on Medical Datasets for Disease Prediction," *Information Sciences with Applications*, vol. 3, pp. 1–21, Apr. 2024, https://doi.org/10.61356/j.iswa.2024.3232.

[11] N. Jillani, A. M. Khattak, M. Z. Asghar, and H. Ullah, "Efficient Diagnosis of Liver Disease using Deep Learning Technique," in *2023 IEEE International Symposium on Medical Measurements and Applications (MeMeA)*, Jeju, Republic of Korea, June 2023, pp. 1–6, https://doi.org/10.1109/MeMeA57477.2023.10171906.

[12] B. A. Julian and T. A. Annie, "Advanced Detection of Liver Cirrhosis Utilizing LSTM and Neural Networks," in *2025 11th International Conference on Communication and Signal Processing (ICCSP)*, Melmaruvathur, India, June 2025, pp. 1685–1690, https://doi.org/10.1109/ICCSP64183.2025.11088752.

[13] L. Zhao *et al.*, "A Hybrid CNN-Transformer for Focal Liver Lesion Classification," in *ICASSP 2024 - 2024 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Seoul, Republic of Korea, Apr. 2024, pp. 13001–13005, https://doi.org/10.1109/ICASSP48485.2024.10447765.

[14] N. P. Nelaturi, V. Rajesh, and I. Syed, "Real-Time Liver Tumor Detection with a Multi-Class Ensemble Deep Learning Framework," *Engineering, Technology & Applied Science Research*, vol. 14, no. 5, pp. 16103–16108, Oct. 2024, https://doi.org/10.48084/etasr.8106.

[15] H. M. Marghany *et al.*, "Toward an Accurate Liver Disease Prediction Based on Two-Level Ensemble Stacking Model," *IEEE Access*, vol. 12, pp. 180210–180237, 2024, https://doi.org/10.1109/ACCESS.2024.3459429.

[16] A. A. Ahad, B. Das, M. R. Khan, N. Saha, A. Zahid, and M. Ahmad, "Multiclass liver disease prediction with adaptive data preprocessing and ensemble modeling," *Results in Engineering*, vol. 22, June 2024, Art. no. 102059, https://doi.org/10.1016/j.rineng.2024.102059.

[17] A. A. Almelibari, M. I. Labib, and Y. Ramadan, "Enhancing Liver Disease Classification Based on a Stacked Machine Learning Model," *Engineering, Technology & Applied Science Research*, vol. 15, no. 5, pp. 26403–26409, Oct. 2025, https://doi.org/10.48084/etasr.11526.

[18] D. Logeshwari, R. Ramkumar, and R. Sanjay, "Hybrid Models for Classifying Liver Cirrhosis with SVM and Xgboost Algorithms," in *2024 International Conference on Innovative Computing, Intelligent Communication and Smart Electrical Systems (ICSES)*, Chennai, India, Dec. 2024, pp. 1–6, https://doi.org/10.1109/ICSES63760.2024.10910344.

[19] C. Kate, Rajkumari, B. C. Anil, S. C. Geerlapally, P. Sudheer, and W. M. Stalin, "A Hybrid Feature Selection and Hyperparameter-Tuned XGBoost Approach for Liver Disease Prediction," in *2025 5th International Conference on Soft Computing for Security Applications (ICSCSA)*, Salem, India, Aug. 2025, pp. 1144–1152, https://doi.org/10.1109/ICSCSA66339.2025.11171205.

[20] P. Sharma and S. Kamboj, "Machine Learning Approaches for Early Detection of Liver Ailments: Comprehensive Analysis and Comparative Study," in *2024 15th International Conference on Computing Communication and Networking Technologies (ICCCNT)*, Kamand, India, June 2024, pp. 1–6, https://doi.org/10.1109/ICCCNT61001.2024.10724600.

[21] B. Ramana and N. Venkateswarlu, "ILPD (Indian Liver Patient Dataset)." UCI Machine Learning Repository, 2022, https://doi.org/10.24432/C5D02C.

[22] T. Chen and C. Guestrin, "XGBoost: A Scalable Tree Boosting System," in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, San Francisco, CA, USA, Aug. 2016, pp. 785–794, https://doi.org/10.1145/2939672.2939785.