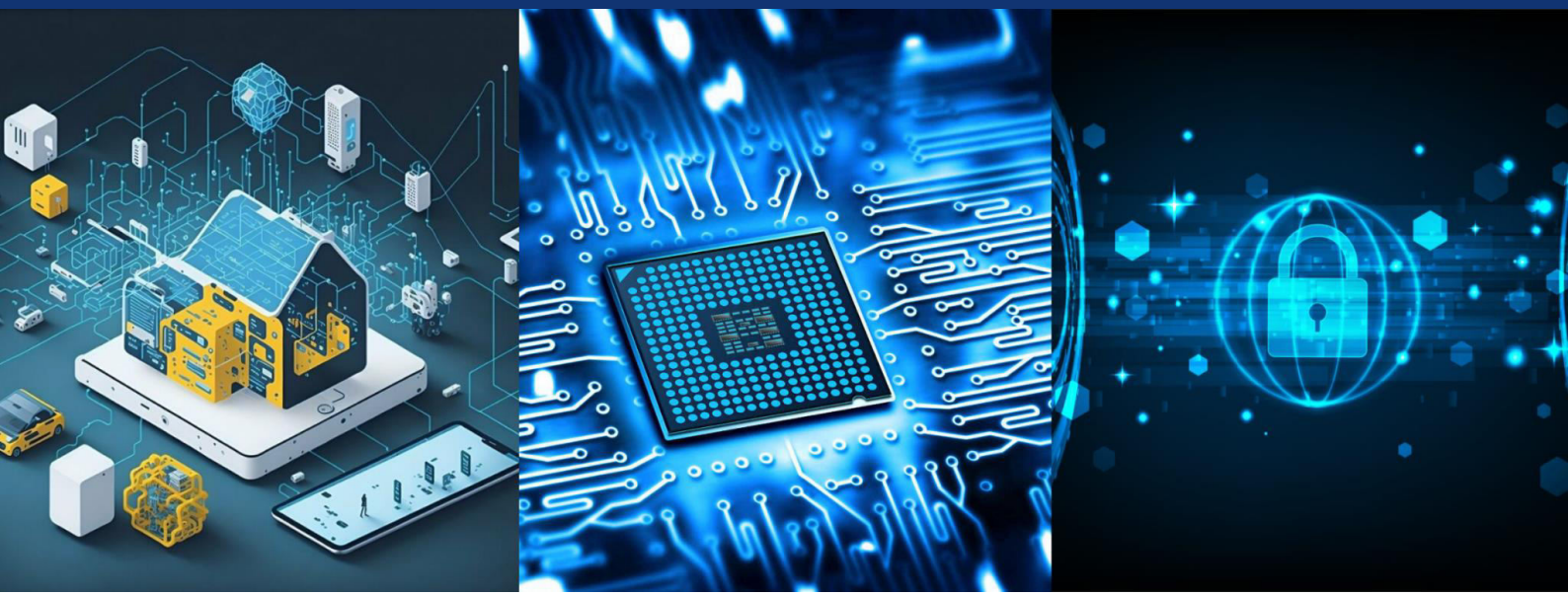
IJIRCCE

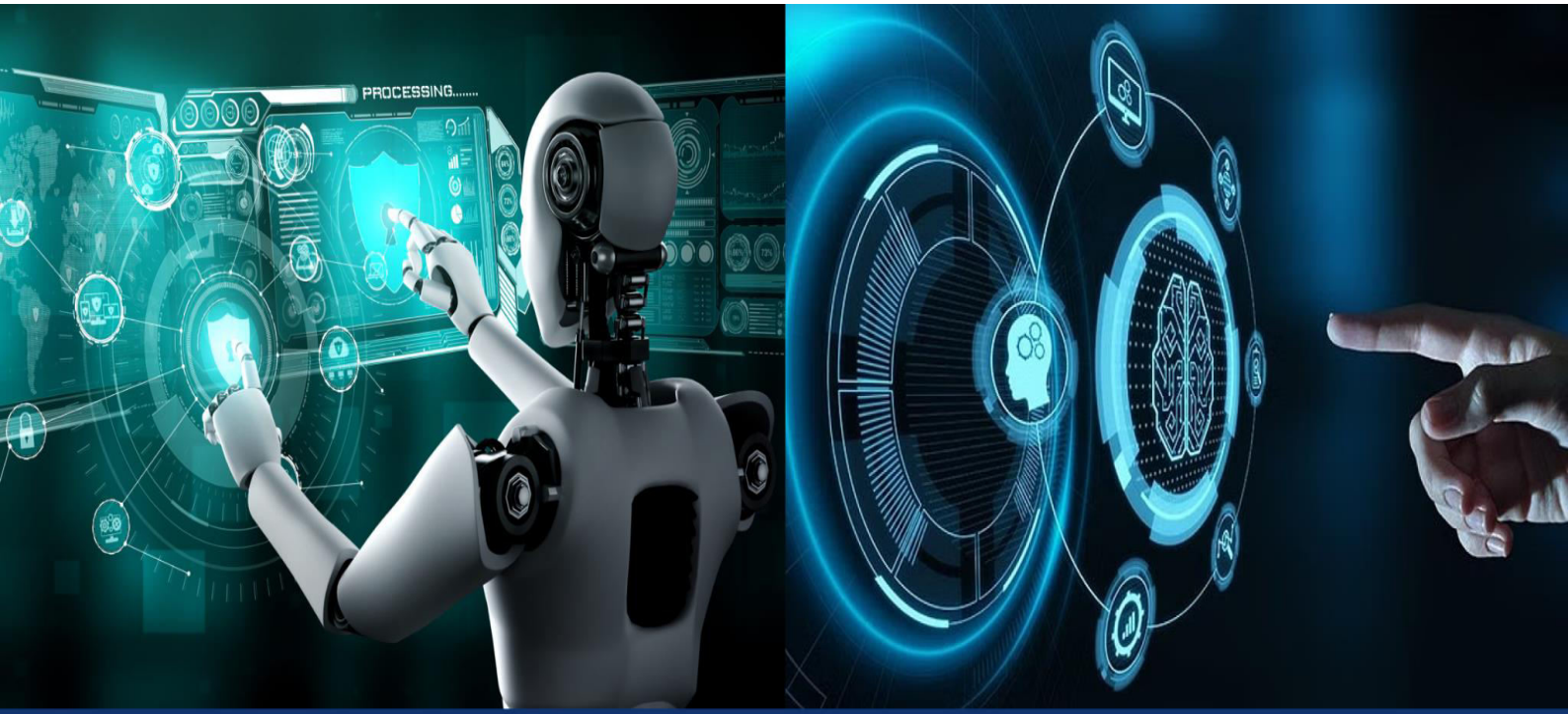
ISSN(O): 2320-9801

ISSN(P): 2320-9798



International Journal of Innovative Research in Computer and Communication Engineering

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)



Impact Factor: 8.771

Volume 14, Issue 4, April 2026

© 2026 IJIRCCE | Volume 14, Issue 4, April 2026 DOI: 10.15680/IJIRCCE.2026.1404144

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.771| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

An Explainable Machine Learning Framework for Liver Disease Prediction Using Dual Statistical Feature Optimization and Ensemble Learning

Polemreddy Venkata Alekhya, Dr. Deepak Nedunuri

M. Tech Scholar, Department of CSE, Sir C R Reddy College of Engineering, Eluru, India

Professor, Department of CSE, Sir C R Reddy College of Engineering, Eluru, India

ABSTRACT: Liver disease continues to be a major global health concern due to its silent progression and delayed diagnosis. This study proposes an efficient and interpretable machine learning framework for early liver disease prediction by integrating advanced preprocessing, Dual Statistical Feature Optimization (DSFO), and ensemble learning. Missing values are handled using iterative imputation techniques, while class imbalance is addressed using SMOTE to ensure balanced data distribution. The proposed DSFO method combines Chi-Square dependency analysis and ANOVA-based variance evaluation to identify the most significant clinical features. Unlike conventional stacking approaches, the model employs a simplified ensemble strategy using LightGBM and Extra Trees Classifier combined through soft voting, improving generalization while reducing model complexity. The framework achieved an accuracy of **97.96%** and ROC-AUC of **0.9894**, demonstrating strong predictive capability. To enhance interpretability, SHAP is used to provide insights into feature contributions. Results reveal that biochemical attributes such as bilirubin and liver enzyme levels play a dominant role in prediction. The proposed system offers a reliable and transparent solution suitable for real-world clinical applications.

KEYWORDS: Liver Disease Prediction, Feature Optimization, Ensemble Learning, Explainable AI, SHAP

I. INTRODUCTION

Liver disease represents a significant global health burden, contributing substantially to morbidity and mortality across diverse populations. Conditions such as cirrhosis, hepatitis, and non-alcoholic fatty liver disease often progress silently, with symptoms appearing only at advanced stages. This delayed manifestation makes early diagnosis both challenging and critical for improving patient outcomes. Traditional diagnostic practices rely heavily on biochemical tests and clinical expertise; however, these approaches are often limited in their ability to efficiently analyze complex and large-scale patient data. As healthcare systems increasingly generate vast amounts of structured and unstructured data, there is a growing need for intelligent computational methods capable of supporting accurate and timely diagnosis.

In recent years, the application of machine learning (ML) techniques in healthcare has gained considerable attention due to their ability to uncover hidden patterns and relationships within clinical datasets. Various models, including decision trees, support vector machines, and ensemble methods, have been employed for liver disease prediction, often achieving promising accuracy levels. Despite these advancements, several critical challenges persist in the existing literature. Many studies rely on datasets with missing values and imbalanced class distributions, which can bias model performance. Additionally, traditional feature selection techniques often fail to capture both statistical dependency and variance characteristics simultaneously, leading to suboptimal feature subsets. Another major limitation is the lack of interpretability in many high-performing models, particularly ensemble and boosting techniques, which are often considered “black-box” systems and therefore less suitable for clinical decision-making.

These gaps highlight the need for a comprehensive framework that not only improves predictive accuracy but also ensures data quality, feature relevance, and model transparency. Accordingly, the primary objective of this study is to develop an

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.771| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

interpretable and efficient machine learning framework for liver disease prediction. Specifically, this research seeks to address the following questions:

- How can data preprocessing techniques improve the reliability of clinical datasets for liver disease prediction?
- Can a hybrid statistical feature selection approach effectively identify the most relevant clinical attributes?
- How does a simplified ensemble learning strategy compare with more complex stacking models in terms of performance and generalization?
- In what ways can explainable AI techniques enhance the interpretability and trustworthiness of predictive models in healthcare?

To address these questions, the study proposes a novel framework that integrates advanced preprocessing methods, a Dual Statistical Feature Optimization (DSFO) technique for feature selection, and an ensemble learning model combining LightGBM and Extra Trees Classifier. Furthermore, SHapley Additive exPlanations (SHAP) are employed to provide both global and local interpretability of model predictions. The rationale behind this approach lies in balancing predictive performance with model simplicity and transparency, thereby making the system more suitable for real-world clinical applications.

The contributions of this research are threefold. First, it introduces a hybrid feature selection strategy that captures both dependency and variance characteristics of clinical features. Second, it demonstrates the effectiveness of a simplified ensemble model in achieving high accuracy while reducing computational complexity. Third, it integrates explainable AI techniques to enhance model interpretability, thereby addressing one of the major limitations of existing approaches.

The remainder of this paper is organized as follows: Section 2 reviews related work in liver disease prediction using machine learning. Section 3 describes the proposed methodology, including preprocessing, feature selection, and model design. Section 4 presents the experimental results and discussion. Finally, Section 5 concludes the paper and outlines potential directions for future research.

II. LITERATURE SURVEY

A study conducted by S. M. Ganie and P. K. Dutta Pramanik (2024) [5] presents a comprehensive comparative evaluation of boosting-based ensemble algorithms for chronic liver disease prediction, emphasizing the importance of early diagnosis in improving clinical outcomes and reducing disease burden. The authors note that conventional machine learning models often struggle with complex and heterogeneous clinical data, whereas boosting techniques provide enhanced predictive performance. The study investigates seven widely used algorithms, including Gradient Boosting, AdaBoost, LogitBoost, SGBoost, XGBoost, LightGBM, and CatBoost, using two publicly available datasets, namely the Liver Disease Patient Dataset and the Indian Liver Patient Dataset, which differ in size and demographic characteristics. A systematic preprocessing framework is applied, incorporating exploratory data analysis, normalization, class balancing through upsampling, and hyperparameter optimization to improve model effectiveness. Additionally, the contribution of individual features is analyzed to understand their relevance in disease prediction. Model performance is assessed using k-fold cross-validation, multiple evaluation metrics, and runtime analysis. The results indicate that Gradient Boosting achieves the highest overall performance, with accuracy values of 98.80 percent and 98.29 percent on the respective datasets, outperforming other boosting algorithms across most metrics, although it requires higher computational time. This work highlights the effectiveness of boosting-based ensemble methods in medical prediction tasks and underscores the need for further research on efficient and interpretable machine learning frameworks for liver disease detection.

A study by F. Mostafa, E. Hasan, M. Williamson, and H. Khan (2021) [6] investigates the application of statistical machine learning techniques for liver disease prediction, emphasizing the importance of accurate medical diagnosis in enhancing patient care, research, and healthcare policy. The authors highlight that traditional diagnostic approaches rely on pathological assessments, while recent advancements integrate artificial intelligence and machine learning with clinical findings to improve disease detection. The study utilizes a dataset of 615 patient records and applies data visualization techniques to identify patterns and address missing values. Missing data are handled using multiple imputation by chained equations, while principal component analysis is employed to reduce dimensionality and extract significant features. Feature importance is further validated using the Gini index. The dataset is divided into training and testing subsets to evaluate model performance using binary classification algorithms, including artificial neural networks, random forest, and support vector machines. Additionally, the synthetic minority oversampling technique is applied to

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.771| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

address class imbalance and reduce overfitting. The results indicate that the random forest model achieves the highest performance, with an accuracy of 98.14 percent and statistically significant improvement over other methods. The findings demonstrate that machine learning approaches can effectively identify key risk factors and enhance inference-based diagnosis, supporting more accurate and reliable clinical decision-making in liver disease prediction.

A study by G. S. Harshpreet Kaur (2021) [7] examines the application of machine learning techniques for the diagnosis of chronic liver disease, highlighting the growing global burden of liver-related mortality, which accounts for approximately two million deaths annually and represents about 3.5 percent of total deaths worldwide. Chronic liver disease is recognized as a life-threatening condition, where early detection and timely treatment are essential for improving patient outcomes. The study emphasizes the role of artificial intelligence, particularly machine learning algorithms such as support vector machines, k-means clustering, k-nearest neighbors, random forest, and logistic regression, in enhancing predictive accuracy and supporting early-stage diagnosis. With the increasing availability of large-scale data driven by advancements in data collection technologies and automation, machine learning models can effectively analyze complex clinical datasets. The research proposes a hybrid classification approach for liver disease prediction using the Indian Liver Patient Dataset obtained from Kaggle. The methodology includes data preprocessing, feature extraction, and classification stages. The model is implemented using Python in the Spyder environment, and performance is evaluated using metrics such as accuracy, precision, and recall. The proposed approach achieves an accuracy of 77.58 percent, demonstrating the potential of machine learning techniques in supporting clinical diagnosis while indicating the need for further improvements in predictive performance.

A study by R. Amin, R. Yasmin, S. Ruhi, M. H. Rahman, and M. S. Reza (2023) [8] explores the application of machine learning techniques for the prediction of chronic liver disease, emphasizing the critical role of the liver in performing numerous physiological functions and the importance of early diagnosis in improving patient survival. The authors highlight that conventional machine learning pipelines typically involve data preprocessing, feature extraction, and classification; however, projection-based feature extraction methods often fail to eliminate redundancy effectively and may not capture the true significance of features. To address this limitation, the study proposes an integrated feature extraction approach using the Indian Liver Patient Dataset obtained from the UCI repository, which contains 583 patient records with an imbalanced distribution of disease cases. The methodology begins with preprocessing steps, including the handling of missing values and outliers, followed by the application of a combined feature extraction strategy to identify the most relevant attributes. Multiple machine learning algorithms, such as logistic regression, random forest, k-nearest neighbors, support vector machine, multilayer perceptron, and an ensemble voting classifier, are employed for classification. The proposed system achieves an accuracy of 88.10 percent, along with strong precision, recall, F1 score, and AUC values, outperforming several existing methods. The findings indicate that the integrated approach enhances predictive performance and can serve as a supportive tool for clinicians in diagnosing liver disease more effectively.

III. METHODOLOGY

The proposed methodology is designed as a structured pipeline that integrates data preprocessing, feature optimization, ensemble learning, and model interpretability to achieve accurate and reliable liver disease prediction. Initially, data preprocessing is performed to enhance dataset quality by addressing common issues such as missing values, class imbalance, and inconsistent feature scales. Missing data are handled using iterative imputation techniques to preserve statistical relationships, while class imbalance is mitigated through the Synthetic Minority Over-sampling Technique, ensuring balanced representation of both classes. Additionally, feature scaling and categorical encoding are applied to standardize the dataset and make it suitable for machine learning algorithms. Following preprocessing, a Dual Statistical Feature Optimization (DSFO) approach is introduced to identify the most relevant clinical features. This method integrates Chi-Square analysis to measure the dependency between features and the target variable with ANOVA to evaluate variance across different classes. By selecting only those features that demonstrate significance in both statistical measures, the approach ensures improved feature relevance, reduced dimensionality, and enhanced model performance. For prediction, a simplified ensemble learning strategy is adopted instead of complex stacking architectures. The model combines LightGBM and Extra Trees Classifier using a soft voting mechanism, where prediction probabilities from both models are aggregated. LightGBM effectively captures complex patterns in the data through gradient boosting, while Extra Trees introduces randomness to improve model diversity and robustness. This combination helps in reducing overfitting while maintaining strong generalization capability. Finally, to address the critical requirement of interpretability in healthcare applications, SHapley Additive exPlanations are incorporated. This technique provides both

© 2026 IJIRCCE | Volume 14, Issue 4, April 2026 DOI: 10.15680/IJIRCCE.2026.1404144

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.771| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

global insights into feature importance and local explanations for individual predictions, enabling transparency and trust in the model's decision-making process. Overall, the methodology ensures a balanced trade-off between accuracy, efficiency, and interpretability, making it suitable for real-world clinical applications.

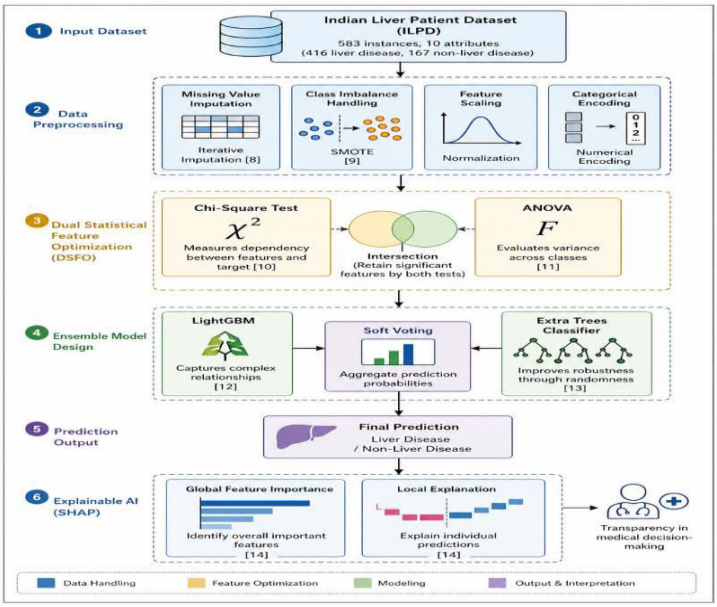


Fig.1 Research Workflow

3.1 Dataset Description

The study utilizes the Indian Liver Patient Dataset (ILPD), a widely used benchmark dataset for liver disease prediction obtained from the UCI Machine Learning Repository. This dataset contains clinical and biochemical records of patients collected from the northeastern region of India, making it highly relevant for real-world medical analysis. The dataset consists of 583 instances, among which 416 correspond to patients diagnosed with liver disease and 167 represent healthy individuals, indicating a clear class imbalance that must be addressed during preprocessing. Each record in the dataset includes a combination of demographic and laboratory-based attributes. The demographic features include age and gender, which provide basic patient information. The biochemical attributes represent key indicators of liver function and include Total Bilirubin, Direct Bilirubin, Alkaline Phosphotase, Alamine Aminotransferase (SGPT), Aspartate Aminotransferase (SGOT), Total Proteins, Albumin, and the Albumin-to-Globulin Ratio. These parameters are clinically significant, as abnormalities in these values are strongly associated with liver dysfunction. The dataset also contains missing values in certain attributes, particularly in the Albumin-to-Globulin Ratio, which necessitates appropriate imputation techniques to maintain data integrity. Additionally, the imbalance between liver disease and non-disease cases can lead to biased model predictions if not properly handled. Therefore, preprocessing steps such as missing value treatment, normalization, and class balancing are essential before model training. Overall, the ILPD dataset provides a comprehensive representation of liver-related clinical features, enabling the development and evaluation of machine learning models for accurate and reliable liver disease prediction.

3.2 Data Preprocessing

Data preprocessing is a critical step in the proposed framework, aimed at enhancing the quality, consistency, and reliability of the dataset before model training. Clinical datasets often contain missing values, imbalanced class distributions, and variations in feature scales, which can negatively impact model performance if not properly addressed. In this study, missing values are handled using iterative imputation techniques, which estimate and replace incomplete data by preserving underlying statistical relationships among features. To address the issue of class imbalance, the Synthetic Minority Over-sampling Technique is applied to generate synthetic samples for the minority class, thereby ensuring a balanced distribution and reducing model bias. Additionally, feature scaling is performed to normalize the range of numerical attributes, preventing features with larger magnitudes from dominating the learning process. Categorical variables, such as gender, are converted into numerical representations to make them compatible with

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.771| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

machine learning algorithms. Overall, these preprocessing steps ensure that the dataset is clean, balanced, and properly formatted, enabling the development of a robust and accurate predictive model.

3.3 Dual Statistical Feature Optimization (DSFO)

The Dual Statistical Feature Optimization (DSFO) method is proposed to enhance the feature selection process by integrating two complementary statistical techniques. In many medical datasets, the presence of redundant and irrelevant features can negatively affect model performance and increase computational complexity. To address this, the Chi-Square test is employed to measure the statistical dependency between each feature and the target variable. This helps in identifying features that have a strong association with liver disease outcomes. Alongside this, Analysis of Variance (ANOVA) is utilized to evaluate the variation of feature values across different classes, enabling the identification of attributes that provide strong discriminatory power between liver disease and non-disease cases.

By combining the strengths of both methods, DSFO ensures a more robust and reliable feature selection process. Only those features that demonstrate significance in both dependency-based and variance-based evaluations are retained for further modeling. This dual filtering approach reduces dimensionality, removes noise, and improves the overall efficiency of the learning process. As a result, the selected feature subset not only enhances predictive accuracy but also contributes to better interpretability, as the retained features are statistically meaningful and clinically relevant.

3.4 Ensemble Model Design

The proposed framework adopts a simplified ensemble learning strategy to improve predictive performance while avoiding the complexity associated with stacking architectures. Ensemble learning is widely recognized for its ability to enhance model accuracy and generalization by combining multiple learning algorithms. In this study, two efficient and complementary models, namely LightGBM and Extra Trees Classifier, are selected as base learners due to their strong performance on structured medical datasets.

LightGBM is a gradient boosting algorithm that is particularly effective in handling large-scale data and capturing complex nonlinear relationships between features. It employs a leaf-wise tree growth strategy, which enables faster training and improved accuracy. On the other hand, the Extra Trees Classifier introduces randomness in feature selection and split generation, which helps in reducing variance and improving model robustness. The diversity between these models ensures that different aspects of the data are effectively learned.

To combine the strengths of both models, a soft voting mechanism is employed, where the predicted probabilities from each model are aggregated to produce the final output. This probabilistic combination allows for more balanced and stable predictions compared to hard voting or single-model approaches. The overall ensemble design reduces the risk of overfitting while maintaining strong predictive performance, making it suitable for real-world clinical applications.

3.5 Explainable AI (SHAP)

In healthcare applications, model interpretability is as important as predictive accuracy, as clinical decisions require transparency and trust. To address this need, the proposed framework incorporates SHapley Additive exPlanations (SHAP), a widely used explainable AI technique based on cooperative game theory. SHAP assigns contribution values to each feature, indicating its impact on the model's prediction. At the global level, SHAP provides an overall view of feature importance across the entire dataset. This helps in identifying which clinical attributes, such as bilirubin levels or enzyme measurements, have the most significant influence on liver disease prediction. Such insights are valuable for understanding the behavior of the model and validating its decisions against medical knowledge. At the local level, SHAP explains individual predictions by showing how each feature contributes positively or negatively to a specific outcome. This level of explanation is particularly useful for clinicians, as it allows them to interpret the reasoning behind each prediction for a given patient. By offering both global and local interpretability, SHAP enhances transparency, builds trust in the model, and supports informed medical decision-making.

IV. RESULTS AND DISCUSSION

4.1 Performance Metrics

The performance of the proposed model is evaluated using standard classification metrics to ensure a comprehensive assessment of its predictive capability. The model achieves an accuracy of 97.96 percent, indicating a high proportion of

© 2026 IJIRCCE | Volume 14, Issue 4, April 2026 DOI: 10.15680/IJIRCCE.2026.1404144

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.771| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

correctly classified instances. Precision and recall values of 97.72 percent and 97.48 percent, respectively, demonstrate the model's effectiveness in identifying both positive and negative cases with minimal misclassification. The F1-score of 97.60 percent reflects a balanced trade-off between precision and recall, confirming the model's reliability in handling class imbalance.

In addition to these metrics, the ROC-AUC score of 0.9894 highlights the model's strong discriminative ability across different classification thresholds. A high ROC-AUC value indicates that the model can effectively distinguish between liver disease and non-disease cases. This is particularly important in medical applications where accurate classification is critical for early diagnosis and treatment planning. The combination of these metrics confirms the robustness and consistency of the proposed framework.

Overall, the performance results demonstrate that the integration of preprocessing, feature optimization, and ensemble learning contributes significantly to improved predictive accuracy. The use of multiple evaluation metrics ensures that the model's performance is not biased toward a single criterion, thereby providing a more reliable assessment suitable for clinical applications.

Table 1. Performance Metrics

Metric	Value
Accuracy	97.96%
Precision	97.72%
Recall	97.48%
F1-Score	97.60%
ROC-AUC	0.9894

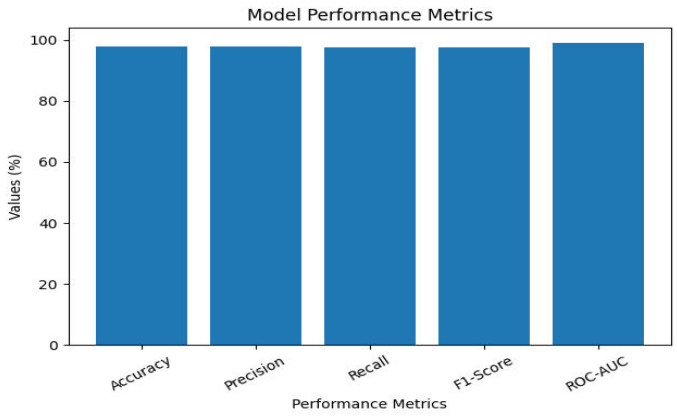


Fig.2 Model performance

4.2 Analysis

The experimental results reveal that the proposed Dual Statistical Feature Optimization method plays a significant role in enhancing model efficiency by eliminating redundant and irrelevant features. By selecting only statistically significant attributes, the model benefits from reduced dimensionality, which in turn improves training speed and prediction accuracy. This optimized feature set allows the model to focus on the most informative clinical indicators, thereby strengthening its predictive capability.

The use of ensemble learning further improves the stability and generalization of the model. By combining LightGBM and Extra Trees Classifier through a soft voting mechanism, the framework effectively captures complex patterns while reducing variance. Unlike stacking-based models, the simplified ensemble architecture lowers computational complexity without compromising performance, making it more practical for real-world deployment. Additionally, the incorporation

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.771| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

of SHAP enhances model transparency by providing clear explanations of feature contributions, addressing one of the major limitations of traditional machine learning approaches.

Analysis of feature importance indicates that biochemical attributes such as Total Bilirubin, Direct Bilirubin, SGPT, SGOT, and Albumin levels have the highest influence on prediction outcomes. These findings are consistent with clinical knowledge, reinforcing the validity of the model. Compared to existing studies, the proposed framework achieves competitive accuracy while maintaining interpretability and reduced complexity. This balance between performance and transparency makes the model a strong candidate for clinical decision support systems.

V. CONCLUSION

This study presents an effective and interpretable machine learning framework for liver disease prediction by integrating advanced preprocessing, Dual Statistical Feature Optimization (DSFO), and a simplified ensemble learning approach. The proposed methodology addresses key challenges in medical data analysis, including missing values, class imbalance, and irrelevant feature selection, thereby improving the overall quality and reliability of the dataset. By combining Chi-Square and ANOVA techniques, the DSFO method ensures the selection of statistically significant and discriminative features, contributing to enhanced model performance. The ensemble model, built using LightGBM and Extra Trees Classifier with a soft voting mechanism, demonstrates strong predictive capability while maintaining reduced computational complexity compared to traditional stacking approaches. The achieved performance metrics, including high accuracy and ROC-AUC values, indicate the robustness and generalization ability of the model. Furthermore, the integration of SHAP enhances interpretability by providing both global and local explanations, making the model more transparent and suitable for clinical applications. Overall, the proposed framework offers a balanced solution that combines accuracy, efficiency, and interpretability, which are essential requirements in healthcare systems. The findings suggest that the model can serve as a reliable decision-support tool for early detection of liver disease, ultimately contributing to improved patient outcomes. Future work may focus on validating the model on larger and more diverse datasets, incorporating real-time clinical data, and exploring advanced optimization techniques to further enhance performance.

REFERENCES

[1] S. Hashem et al., “Machine Learning Prediction Models for Diagnosing Hepatocellular Carcinoma with HCV-related Chronic Liver Disease,” Computer Methods and Programs in Biomedicine, vol. 196, p. 105551, Nov. 2020, doi: 10.1016/j.cmpb.2020.105551.
[2] E. Dritsas and M. Trigka, “Supervised Machine Learning Models for Liver Disease Risk Prediction,” Computers, vol. 12, no. 1, p. 19, Jan. 2023, doi: 10.3390/computers12010019.
[3] S. Hashem *et al.*, "Comparison of Machine Learning Approaches for Prediction of Advanced Liver Fibrosis in Chronic Hepatitis C Patients," in *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, vol. 15, no. 3, pp. 861-868, 1 May-June 2018, doi: 10.1109/TCBB.2017.2690848.
[4] W. El Atifi, O. El Rhazouani, F. M. Khan, and H. Sekkat, “Optimizing ensemble machine learning models for accurate liver disease prediction in healthcare,” PLoS One, vol. 20, no. 8, p. e0330899, Aug. 2025, doi: 10.1371/journal.pone.0330899.
[5] S. M. Ganie and P. K. Dutta Pramanik, “A comparative analysis of boosting algorithms for chronic liver disease prediction,” Healthcare Analytics, vol. 5, p. 100313, Jun. 2024, doi: 10.1016/j.health.2024.100313.
[6] F. Mostafa, E. Hasan, M. Williamson, and H. Khan, “Statistical Machine Learning Approaches to Liver Disease Prediction,” Livers, vol. 1, no. 4, pp. 294–312, Dec. 2021, doi: 10.3390/livers1040023.
[7] G. S. Harshpreet Kaur, “The Diagnosis of Chronic Liver Disease using Machine Learning Techniques,” ITII, vol. 9, no. 2, pp. 554–564, Mar. 2021, doi: 10.17762/itii.v9i2.382.
[8] 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, p. 101155, 2023, doi: 10.1016/j.imu.2022.101155.

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

INNO SPACE SJIF Scientific Journal Impact Factor

doi crossref

निस्केयर NISCAIR

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

9940 572 462 6381 907 438 ijircce@gmail.com

www.ijircce.com

Scan to save the contact details