Intelligent Liver Disease Identification using Optimized Multilayer Perceptron using Whale Optimization Algorithm (WOA)

Panduranga Vital Terlapu, *Member, IAENG,* and Killi Bhumika

***Abstract*—Liver disease (LD) is a critical global health challenge, causing 1.5 million annual deaths (WHO 2021), with early detection essential for effective intervention. This study proposes a novel Whale Optimization Algorithm (WOA)-optimized Multilayer Perceptron (MLP) framework to enhance LD diagnosis accuracy using clinical, biochemical, and lifestyle data from 1,460 patients in AP, India's North Coastal region. Through exact feature selection (statistical + ML methods), key predictors were identified: Alkaline Phosphatase (AP, F11), Fever (F07), Aspartate Aminotransferase (AAT, F12), and Vomiting (F05). Comparative analysis showed the WOA+MLP model outperformed traditional classifiers with 92% accuracy, 0.91 AUC, and 0.92 F1-score—surpassing MLP (89% accuracy), Decision Trees (88%), and Naive Bayes (0.93 AUC but 80% accuracy). The optimized model reduced false positives by 31.6% and enhanced generalizability for early LD identification. The study shows significant potential for clinical deployment in early LD identification, reducing diagnostic burdens.**

***Index Terms*—Liver Disease, Machine Learning, Neural Networks, Feature Ranking, Optimization, WOA, Andhra Pradesh.**

I. Introduction

THE liver is an important and large organ in the human body. It weighs around 1.5 kg and makes up 2-3% of an adult's body weight. It is near several other organs, including the stomach, spleen, heart, right kidney, pancreas, and intestines. The gallbladder is located beneath the liver. LD is a global health concern with chronic diseases such as hepatitis, cirrhosis, and NAFLD. Reduces quality of life and burdens healthcare systems[1]. This research explores the prevalence and consequences of LD, highlighting the potential of machine learning for improved diagnosis and treatment strategies. According to WHO (2021) reports, LD is the 12th leading cause of death globally, with approximately 1.5 million deaths annually [2]. LD significantly affects individuals and families both personally and financially. Primary symptoms include fatigue, jaundice, and abdominal pain. Advanced conditions such as cirrhosis and liver cancer can result in organ failure, decreased life expectancy, and death. Chronic hepatitis C infections are the main risk factors for LD. Alcohol abuse also increases the risk [3]. Metabolic disorders such as obesity, diabetes, and dyslipidemia contribute to liver problems [4]. Chronic hepatitis (HC and HB) can cause cirrhosis [5]. They can also cause liver cancer.

The **Table I** shows that the stages of LD include treatment and diagnosis. It shows the stages of liver disease. It lists important symptoms, ways of diagnosing the disease, treatments, and prevention methods for each stage. This information helps us understand how liver disease progresses from inflammation to end-stage. LDs cause many deaths in people of all ages. We need better ways to predict these diseases in healthcare. Machine learning (ML) can help because it processes large amounts of data and acts as human intelligence. However, single ML classifiers often do not provide high accuracy. This is why ensemble methods, like Hybrid ML Models, are used. These models combine different classifiers to improve performance. Previous studies have looked at classifier ensembles for predicting diseases, but no single model works for various diseases. The WOA is a nature-inspired method. It helps optimize Multi-Layer Perceptrons (MLPs). WOA is based on the hunt for humpback whales. It effectively balances exploration and exploitation. The design makes it suitable for complex models like MLPs. Combining WOA with MLP improves accuracy and efficiency in identifying LD. It overcomes the limitations of manual tuning and traditional optimization techniques.

Objectives of the Study: i) The study's objective is to accurately identify individuals at risk of LD in the North Coastal region of AP, India through analysis of clinical, biochemical, and lifestyle data. ii) The primary objective is to develop a smart model that uses MLP and WOA to improve diagnostic accuracy in predicting LD. iii) Using statistical and ML-based feature selection (FS) techniques, the study aims to find the top features that affect LD. It will rank these features based on their influence. iv) To demonstrate Hybrid AI techniques for LD: Improves early detection. Supports clinical decision-making. Reduces healthcare burden.

This paper presents a comprehensive framework for the proposed approach, structured as follows: **Section 2** provides a detailed literature review, highlighting recent developments in LD prediction using machine learning (ML) and optimization techniques. **Section 3** describes the models and materials, including the dataset used, feature selection methods, baseline models, and the architecture of the WOA-optimized MLP model. **Section 4** outlines the results analysis, presenting the LD statistical analysis, the results of the feature selection, the model performance metrics, and comparative evaluations with traditional classifiers. **Section 5** provides discussions on the experimental findings,

Manuscript received June 14, 2025; revised August 28, 2025.

Panduranga Vital Terlapu is a professor of the Department of Computer Science and Engineering, Aditya Institute of Technology and Management, Tekkali, Srikakulam, Andhra Pradesh-532201, India (phone: +91 9491936489; fax: 08945 245266; e-mail: vital2927@gmail.com).

Killi Bhumika is a postgraduate student of the Department of Computer Science and Engineering, Aditya Institute of Technology and Management, Tekkali, Srikakulam, Andhra Pradesh, India (e-mail: bhumikakilli947@gmail.com).

TABLE I
STAGES OF LIVER DISEASE AND THEIR CONSEQUENCES LIKE SYMPTOMS AND TREATMENT

Stage	Symptoms	Diagnosis	Treatment	Prevention & Precautions
Inflammation (Stage 1)	Often asymptomatic; may include fatigue, mild abdominal discomfort.	Blood tests (e.g., liver function tests: AST, ALT, GGT), imaging (ultrasound).	Address underlying cause (e.g., stop alcohol, manage viral hepatitis); lifestyle changes (healthy diet, exercise).	Avoid alcohol, maintain healthy weight, vaccinate against hepatitis A and B, practice safe hygiene.
Fibrosis (Stage 2)	Fatigue, abdominal pain, itching (pruritus) due to bile salt buildup.	Blood tests (e.g., albumin, bilirubin levels), imaging (CT, MRI), elastography [1].	Treat underlying cause (e.g., antiviral drugs for hepatitis), monitor progression, lifestyle changes (weight loss if NAFLD).	Limit alcohol, control metabolic factors (blood sugar, cholesterol), avoid toxins (e.g., excessive acetaminophen) [2].
Cirrhosis (Stage 3)	Jaundice, ascites (abdominal swelling), easy bruising, confusion (hepatic encephalopathy).	Blood tests (e.g., prothrombin time), imaging, endoscopy (for varices), liver biopsy.	Manage complications (e.g., diuretics for ascites, beta-blockers for varices), lifestyle changes, consider transplant evaluation.	Regular health checkups, avoid alcohol/drugs, manage diet (low salt), screen for hepatocellular carcinoma (ultrasound every 6 months).
End-Stage LD (ESLD) / Liver Failure (Stage 4)	Severe jaundice, significant weight loss, internal bleeding, kidney/lung issues, coma [3].	Blood tests (e.g., low albumin, high bilirubin), imaging, clinical scoring (e.g., MELD score for transplant).	Liver transplant (if eligible), manage complications (e.g., lactulose for encephalopathy), palliative care if transplant not viable.	Early intervention, lifelong monitoring, avoid all liver stressors, adhere to medical advice, join support groups (e.g., British Liver Trust) [4].

addressing their clinical significance, model limitations, and potential enhancements. Finally, **Section 6** of the study provides a summary of key contributions and emphasizes the effectiveness of the research of the WOA-MLP framework and suggests directions for future work in the diagnosis of intelligent LD.

II. LITERATURE REVIEW

Improving LD classification models can increase their accuracy. It can also make them more efficient. Houssein et al. (2024) [5] introduced a new model KOA to enhance (FS) feature selection for LD classification. The new algorithm enhanced the selection of useful features from medical data sets, decreasing dimensionality and enhancing performance in classification tasks. The research demonstrated that the enhanced KOA performs better compared to conventional optimization techniques, with a more efficient solution for the early treatment with early diagnosis of LD. Raziani et al. (2022) [6] combined Modified WOA and MLP to develop a hybrid model for medical classification problems. Optimizing the parameters and structure of MLP, the WOA enhances its performance in fields like heart disease prediction and cancer diagnosis. Despite the improvement in the accuracy (ACC) of the model, problems such as overfitting, high computational cost, and interpretability persist. The conclusion of the research is that more research is needed to enhance such models for more generalization and practical pharmaceutical applications. Hybrid soft computing approaches can improve FS and classification. These methods combine different techniques to enhance efficiency. Varchagall et al. (2023) [7] targets early diagnosis of liver issues. It enhances the ACC of LD classification through methods like SVM and GA to select the most relevant information from medical data. The hybrid approach maximizes the classification model and eliminates redundant features in a bid to enhance diagnostic ACC. Compared to traditional techniques, the research reveals that using these methods can enhance liver disorder detection, offering a potential route to early treatment and diagnosis. Qiao et al.,'s (2024) [8] goal of their study was to combine deep learning with an enhanced WOA to create a liver tumor segmentation technique that is more accurate. To distinguish between different areas of the liver image, including background, healthy tissue, and tumor tissue, multi-threshold segmentation was employed. For feature (FE) extraction, CNN, a DL model, was utilized. By optimizing the multi-thresholding procedure, the IWOA increased the accuracy of segmentation. The suggested method demonstrated the potential to improve liver tumor treatment and diagnosis planning by outperforming conventional methods in terms of accuracy and computational efficiency, according to the results. The LD Literature Review **Table II** shows studies on using ML to predict LD. It shows recent studies on LD prediction from 2019 to 2024. It summarizes the use of ML models, datasets, and sample sizes. The table II also highlights the methodologies used and key findings of each study. The relevance of each study to LD research and diagnosis is evident.

A. Datset Description

Singh et al. (2024) [17] predicted the accuracy and performance of several ML classification algorithms in predicting LD. To find the best algorithm for precise LD prediction, it compared several models, such as DTs, SVM, and NNs. The outcomes emphasized each approach's advantages and disadvantages. To diagnose LD early, to enhance breast cancer diagnosis, Stephan et al. (2021) [18] suggested a hybrid (combined) model that combines the WOA and the ABC algorithm. WOA improved the optimization procedure, while the ABC algorithm was utilized for the best FS. By lowering computational complexity and improving classification ACC, this method outperformed conventional techniques in differentiating cancer tumors. Routray et al. (2023) [19] suggested using histopathological data to classify breast cancer and identify other organs at risk using an ensemble learning model in conjunction with the SOS optimization algorithm. While the ensemble method increased classification ACC, the SOS algorithm optimized FS. The hybrid model outperformed CNN. It is a promising tool for assessing organ risk. It can also aid in early cancer diagnosis. Kaur et al. (2022) [20] proposed WOA for liver cyst image segmentation. The segmentation procedure was optimized using the modified WOA, improving the effectiveness and efficiency of liver cyst identification in medical images. The proposed

TABLE II
SUMMARY OF RECENT STUDIES UTILIZING MLS AND DIAGNOSTIC TECHNIQUES FOR LIVER DISEASE (LD) PREDICTION AND ANALYSIS

Author Details	Study	Data Size	Methods	Key Findings and Result Analysis	Relevance to LD
Homsana et al. (2024) [9]	Cross-sectional	2826	Abdominal ultrasonography, interviews	SLD prevalence was 27.1%, higher in non-lean persons	Rising prevalence of SLD among females
Joloudari et al. (2019) [10]	Comparative study	583	ELTA and multiple classification models	PSO-SVM model achieved the highest ACC (94.42%) when compared to others	Demonstrates the effectiveness of data mining models
Islam (2024) [11]	Comparative study	583	ML algorithms tree structured	ETC with TPE achieved the highest ACC	Approach for early and accurate LD prediction
Ganie et al. (2024) [12]	Comparison	—	ML algo, hyperparameter tuning	GB achieved the highest ACC	Approach for early disease prediction
Mohamed et al. (2024) [13]	Comparative	583	Machine Learning	Ensemble Stacking achieved 93.88% without FS, 94.12% with it; Two-level stacking model reached 94% ACC	Improve disease prediction
Behera et al. (2023) [14]	Experiment	583	Hybrid model combining SVM and modified PSO	CCPSOSVM achieved the highest ACC of 92.59% for heart disease and 97.41% for LD	Provides a path in ML for innovation
Babatunde et al. (2024) [15]	Retrospective	1404	Neuro-Fuzzy System, Machine Learning	Achieved 97% classification	Advanced algorithm for early detection of LDs
van et al. (2023) [16]	Informatic	3000	AI-assisted data collection	Identified 38 distinct key events and 135 key event relationships	Understanding of cholestasis mechanisms
Singh et al. (2024) [17]	Experiment	1476	Data preprocessing, rough set theory for FS along with ML classifying algos	RF classifier achieved 88.66% ACC for hepatitis, 97.29% for dermatological conditions, 91.58% for hepatic disease	Approach for accurately identifying LDs

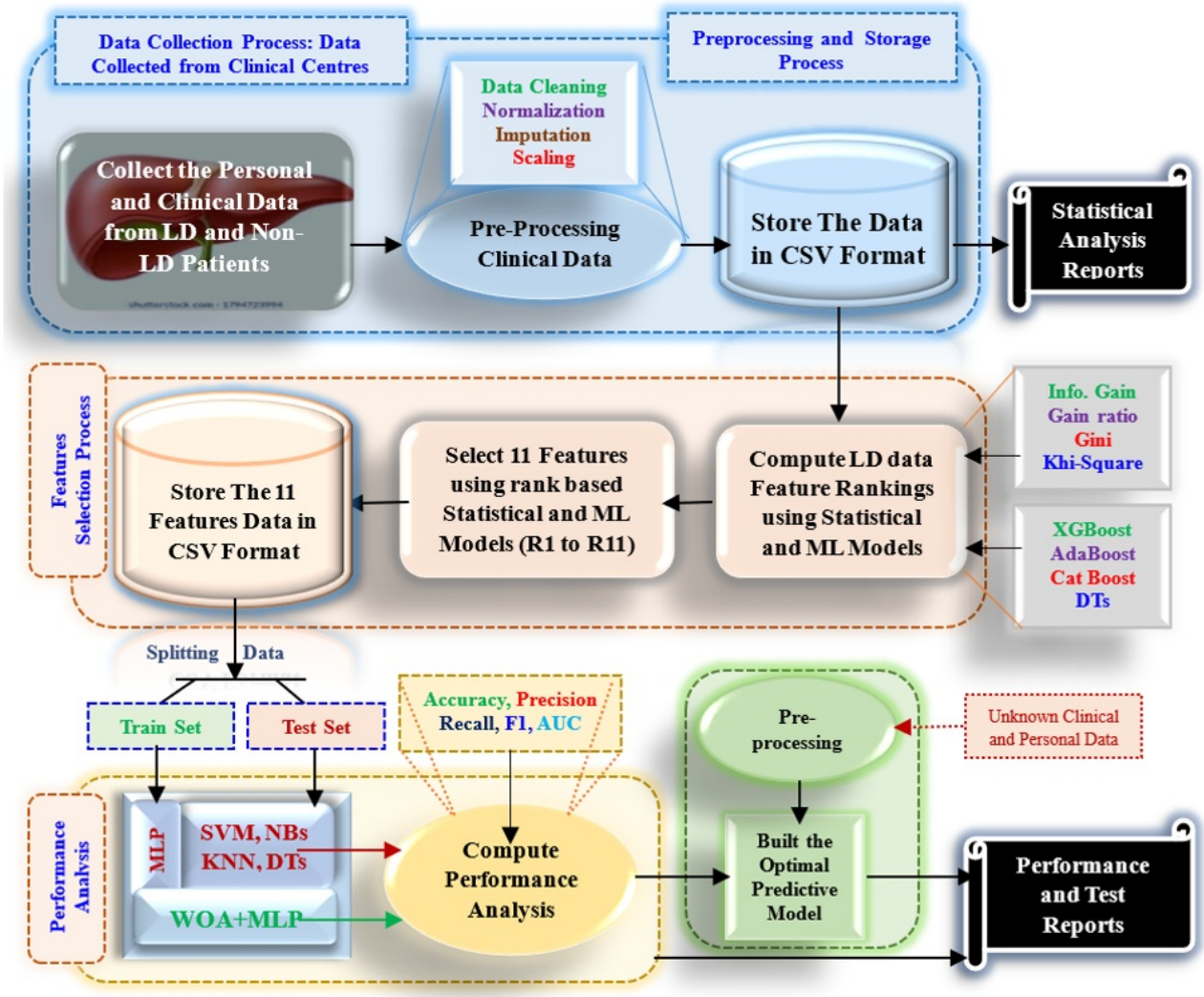

Fig. 1. Proposed Workflow for Intelligent Liver Disease Identification Using WOA-Optimized Multilayer Perceptron

method enhanced the algorithm's precision by 20%. It also increased the convergence speed by 15%. These improvements led to better segmentation results. According to experimental results, the modified WOA performed better than conventional techniques, offering a more useful instrument for precise liver cyst identification in medical imaging. Ekinci et al. (2023) [21] compared algorithms for FS and meta-heuristic optimization in ML-based heart disease classification, including GA, PSO, and ACO. It assessed each algorithm's performance according to its computational efficiency and classification ACC, demonstrating both its advantages and disadvantages. The study highlights how crucial FS is to enhancing heart disease prediction models, with the optimal algorithm relying on the demands of a given application.

III. Models and Meterials

This section outlines the models, materials, and methods for identifying LD. It focuses on a WOA-optimized MLP approach. Details of the workflow and dataset are provided. Additionally, it addresses baseline ML models and the MLP architecture. The discussion also includes the WOA-based optimization strategy.

1) Proposal Model: **Fig. 1** shows a workflow for identifying LD. It includes processing clinical data, feature engineering, and using WOA-optimized MLP modelling. The workflow has four main phases. It includes data collection, preprocessing, feature extraction, and classification. Clinical data is collected from patients at medical centres. This includes both personal and clinical information from patients with and without LD. After processing, the data is saved in CSV format for future use. Feature engineering involves selecting important features for a model. Eleven key features were identified using different ranking methods. These methods include statistical metrics like Information Gain and Gini Index and ML techniques like XGBoost and DTs. The chosen features are saved in a refined CSV format. The model detects LD using a MLP NN. It optimizes hyperparameters and weights through the WOA. This process enhances accuracy and reliability. The raw data is cleaned, normalized, and scaled to ensure consistency.

Performance Analysis using MLP and WOA: 1. Divide the dataset into two parts: 80% for training and 20% for testing. 2. Train four baseline classifiers: SVM, Naïve Bayes, KNN, and DTs. Check how well they perform. 3. Train an MLP model with several hidden layers using ReLU activations. 4. Use the WOA to optimize the MLP's hyperparameters. 5. Adjust the no.of neurons, learning rate, epochs, and batch size. Evaluate the model's performance metrics. Check the accuracy, F1 score, precision, AUC, and recall. Performance reports are created, including a confusion matrix and ROC curve. These reports help provide clinical insights and support diagnoses based on the model's predictions.

2) Dataset Description: For LD Dataset analysis (**Table III**), the North Coastal of AP, India LD (NCAPL) dataset contains one class attribute (C1) and sixteen feature attributes (F1–F16). Blood pressure (BP), biochemical markers like total bilirubin and alkaline phosphatase, clinical symptoms like vomiting and fever, lifestyle factors like smoking and drinking, and demographic information like age and gender are all included. Every attribute has distinct types and ranges, such as categorical, integral, and numerical. The target variable for predictive modelling is the diagnosis (C1), which divides patients into two groups: Non- LD (Class 0) and LD (Class 1).

TABLE III
North Coastal Region of AP, India Liver Disease Dataset Description

Feature Code	Range or Values	Data Type
F1.AG (Age)	6 – 99	Numeric
F2.GEN (Gender)	Female(0), Male(1)	Categorical
F3.SMK (Smoke)	No(0), Yes(1)	Categorical
F4.DRK (Drink)	No(0), Yes(1)	Categorical
F5.VMT (Vomiting)	Absent(0), Present(1)	Categorical
F6.BHA(BoneAche/ HeadAche)	Absent(0), Present(1)	Categorical
F7.FVR (Fever)	Absent(0), Present(1)	Categorical
F8.BPS(Blood Pressure)	Norm(0), Low(1), High(2)	Categorical
F9.TB(Total Bilirubin)	0.4 – 75	Integral
F10.DB(Direct Bilirubin)	0.1 – 19.7	Integral
F11.AP(Alkaline Phosphatase)	10 – 4929	Numeric
F12.AAT(Alanine Aminotransferase)	10 – 2000	Numeric
F13.ASAT(Aspartate Aminotransferase)	5 – 4929	Numeric
F14.TP(Total Proteins)	0.9 – 7.7	Integral
F15.ALB(Albumin)	0.9 – 7.7	Integral
F16.AGR(A-G Ratio)	0.3 – 4.0	Integral
C1.Class (Diagnosis)	NLD (0), LD (1)	Categorical

TABLE IV
Description of Parameters Used in the MLP Training Process

Parameters	Description
X	Input vector with n features [x_1,x_2,...,x_n]
H_l	Number of neurons in the l^{th} hidden layer
$W^{[l]}$	Matrix of weights between (layer $l-1$) and (layer l)
$b^{[l]}$	Bias vector for layer l (since $b^{[l]}$ is a vector)
$A^{[l]}$	Activation output at layer l
$Z^{[l]}$	Linear combination at layer l before activation
σ	ReLU, sigmoid, or softmax (Activation function)
η	Learning rate for gradient descent
$\hat{Y}$	Predicted output
Y	True output

3) ML Models: **Naïve Bayes** is a simple classification algorithm. It is based on Bayes' Theorem. This algorithm assumes that features are independent. It works well with categorical data. It is effective even with small amounts of training data. The assumption of conditional independence makes calculations easier. It reduces the model's complexity. A **Decision Tree** (DT) is an ML model that helps categorize and predict data by dividing it into branches based on various characteristics. It visually represents the decision-making process and is easy to understand, although it can sometimes be complex to explain. Rule-based splits utilize statistical methods to generate clear visualizations and effectively handle non-linear data without requiring feature adjustments. However, they may be influenced by high variance and slight changes in the data. **SVM** is a strong model [22] [23]. It is used for regression and classification. It is supervised learning. It is known for its accuracy. Its ability to handle high-dimensional data and its effectiveness in scenarios. SVM performs effectively with complex datasets that contain numerous features. It finds the best line or surface to separate data points. The **KNN** algorithm is a type of supervised

(SL) learning. It is non-parametric. KNN can be used for regression and classification. It predicts outcomes by looking at the 'k' nearest data points. Then, it chooses the most common class among them [24].

4) MLP Model architecture: The MLP model, as shown in the **Fig. 2**, comprises three primary layers: an input (I/P) layer, hidden (HL) layers, and an output (O/P)layer. It processes input features using weights, biases, and activation functions to find complex patterns. The O/P layer gives the final prediction. Training the model includes steps like forward propagation, backpropagation, and using gradient descent. The loss function checks how far the predictions are from the real results[25]. Activation functions add non-linearity, with ReLU for HL and sigmoid/softmax for output. The **table IV** shows the essential parameters involved in the MLP training process.

A) Forward Propagation

The MLP algorithm is a ML technique that processes inputs through each layer to generate outputs.

i. Input Layer(l = 0) for (n features):

$$A^{[0]} = X = [x_1, x_2, \ldots, x_n] \tag{1}$$

ii. Hidden Layers (l=1,2, ..., L)

$$Z^{[l]} = W^{[l]} A^{[l-1]} + b^{[l]} \tag{2}$$

$$A^{[l]} = \sigma(Z^{[l]}) \tag{3}$$

$$\sigma(z) = \max(0, z) \text{ if } \sigma \text{ is ReLU} \tag{4}$$

$$\sigma(z) = \frac{1}{1 + e^{-z}} \text{ if } \sigma \text{ is Sigmoid} \tag{5}$$

iii. Output Layer (l=L+1)

$$Z^{[L+1]} = W^{[L+1]} A^{[L]} + b^{[L+1]} \tag{6}$$

$$\hat{Y} = A^{[L+1]} = \sigma(Z^{[L+1]}) \tag{7}$$

For multi-class classification, Sigma is SoftMax.

$$\text{SoftMax}(z_i) = \frac{e^{z_i}}{\sum_{j=1}^{m} e^{z_i}} \tag{8}$$

iv. Loss Function The loss is the difference between the predicted and actual outputs. For binary classification (cross-entropy loss) as

$$\mathcal{L} = -\frac{1}{m} \sum_{i=1}^{m} [y_i \log(\hat{y}_i) + (1 - y_i) \log(1 - \hat{y}_i)] \tag{9}$$

$$\mathcal{L} = -\frac{1}{m} \sum_{i=1}^{m} \sum_{k=1}^{m} y_{i,k} \log(\hat{y}_{i,k}) \tag{10}$$

$$y_{i,k} \text{ is 1 } i^{\text{th}} \text{ sample } \in \text{ class } k \tag{11}$$

$$\text{else } y_{i,k} \ = 0 \tag{12}$$

B) Backpropagation (Error Reduction and Weights and Bias Updating Process)

Backpropagation is a method used in MLP. It calculates how much the loss changes when the parameters change. This helps improve the model's efficiency. The **Fig. 3** describes the iterative weight updating process in a multi-layer perceptron (MLP). It evaluates the error rate and adjusts the weights until the error attains a target level. When the goal is achieved, the process concludes. An Optimize-trained model is subsequently created for use.

i. Output Layer Gradient: Compute the gradient of the loss to $Z^{[L+1]}$ as

$$\delta^{[L+1]} = \frac{\partial \mathcal{L}}{\partial Z^{[L+1]}} = (\hat{Y} - Y) \odot \sigma'(Z^{[L+1]}) \tag{13}$$

where $\odot$ is the element-wise multiplication and σ' is the derivative of the activation function. Gradients for weights and biases.

$$\text{for Weights } \frac{\partial \mathcal{L}}{\partial W^{[L+1]}} = \delta^{[L+1]} (A^{[L]})^T \tag{14}$$

$$\text{for Biases } \frac{\partial \mathcal{L}}{\partial b^{[L+1]}} = \delta^{[L+1]} \tag{15}$$

ii. Hidden Layers Gradient

$$\text{for Weights } \frac{\partial \mathcal{L}}{\partial W^{[L+1]}} = \delta^{[L+1]} (A^{[L]})^T \tag{16}$$

$$\text{for Biases } \frac{\partial \mathcal{L}}{\partial b^{[L+1]}} = \delta^{[L+1]} \tag{17}$$

iii. Parameter Update (Gradient Descent) : Using Gradient Descent for weights and biases are likely to

$$\text{for Weights } \frac{\partial \mathcal{L}}{\partial W^{[L+1]}} = \delta^{[L+1]} (A^{[L]})^T \tag{18}$$

$$\text{for Biases } \frac{\partial \mathcal{L}}{\partial b^{[L+1]}} = \delta^{[L+1]} \tag{19}$$

Algorithm 1 describes the MLP training process. It includes three main steps: forward propagation, backpropagation, and updating weights and biases using gradient descent. The training continues for several epochs until the loss converges, or the epoch limit is reached.

5) WOA+MLP Optimization Models: The WOA+MLP model (**Fig. 4**) uses WOA to search for the best solutions. It helps improve the MLP's performance, especially in detecting CKD using the AP-CKD dataset. The model combines two parts. First, it uses the MLP's forward propagation and loss calculation. Second, it applies WOA's methods to update positions. It helps balance exploring new options and using known good ones.

WOA+MLP Model Workflow: MLP Structure: Feedforward NN with input, hidden, and output layers. WOA Optimization: Treats MLP's hyperparameters as a position vector in the search space. WOA Phases: Encircling Prey, Bubble-Net Attacking, and Search for Prey. Fitness Function: Minimizes MLP's error on a validation set. Stopping Criteria: Optimization stops after a maximum number of iterations or when the error converges to a satisfactory level. Output: Optimized MLP parameters used to train the final model on training data, evaluated on a test set. The WOA+MLP Optimization **Algorithm 2** combines the Whale Optimization Algorithm with MLP training. It starts by creating a group of whales and checking their fitness based on MLP validation loss. The algorithm updates the positions of the whales using specific strategies. It continues this process until it finds the best whale position with the lowest loss. This position is the optimal configuration for the MLP [26].

Algorithm 1 MLP Training Process

Initialization: Set up weights $W^{[l]}$ and biases $b^{[l]}$ for all layers (using small random values).
for each epoch **do**
    Perform forward propagation to compute $\hat{Y}$ and loss $\mathcal{L}$.
    Perform backpropagation to compute gradients $\frac{\partial \mathcal{L}}{\partial W^{[l]}}$ and $\frac{\partial \mathcal{L}}{\partial b^{[l]}}$.
    Update weights and biases using gradient descent:

$$W^{[l]} \leftarrow W^{[l]} - \eta \frac{\partial \mathcal{L}}{\partial W^{[l]}}, \quad b^{[l]} \leftarrow b^{[l]} - \eta \frac{\partial \mathcal{L}}{\partial b^{[l]}}$$

end for
Stop the process when the loss $\mathcal{L}$ converges or after a set number of epochs.

Algorithm 2 WOA+MLP Optimization

Initialization: Start by placing N whales at random positions $\vec{X}_i(0)$ in the search area, where i ranges from 1 to N. Define WOA parameters: a (linearly decreases from 2 to 0 over iterations), c_1, c_2 (random coefficients), and maximum iterations T.
Step 1: Fitness Evaluation:
for each whale $i = 1$ to N **do**
    Compute the fitness as the MLP validation loss:

$$f(\vec{X}_i(t)) = \mathcal{L}_{\text{val}}(\text{MLP}(\vec{X}_i(t))) \tag{20}$$

end for
Identify the best position $\vec{X}^*$ with the lowest fitness.
Step 2: Encircling Prey (Update Whale Positions):
for $t = 1$ to T **do**
    Update a:

$$\vec{a} = 2\left(1 - \frac{t}{T}\right) \tag{21}$$

    for each whale $i = 1$ to N **do**
        Generate random vectors:

$$\vec{A} = 2\vec{a} \cdot \vec{r} - \vec{a}, \quad \vec{C} = 2 \cdot \vec{r}, \quad p \in [0, 1] \tag{22}$$

where $\vec{r}$ is a random vector in $[0, 1]$.
        if $p < 0.5$ **then**
            if $|\vec{A}| < 1$ **then**
                Compute:

$$\vec{D} = |\vec{C} \cdot \vec{X}^*(t) - \vec{X}_i(t)| \tag{23}$$

$$\vec{X}_i(t+1) = \vec{X}^*(t) - \vec{A} \cdot \vec{D} \tag{24}$$

            else
                Select a random whale $\vec{X}_{\text{rand}}(t)$.
                Compute:

$$\vec{D} = |\vec{C} \cdot \vec{X}_{\text{rand}}(t) - \vec{X}_i(t)| \tag{25}$$

$$\vec{X}_i(t+1) = \vec{X}_{\text{rand}}(t) - \vec{A} \cdot \vec{D} \tag{26}$$

            end if
        else
            Compute:

$$\vec{D}' = |\vec{X}^*(t) - \vec{X}_i(t)| \tag{27}$$

$$\vec{X}_i(t+1) = \vec{X}^*(t) + \vec{D}' \cdot e^{bl} \cdot \cos(2\pi l) \tag{28}$$

where $l \in [-1, 1]$ is a random value and b is a constant defining the spiral shape.
        end if
    end for
    Re-evaluate fitness and update $\vec{X}^*$.
end for
Step 3: Return the Best Solution:
Output the optimized MLP parameters $\vec{X}^*$.

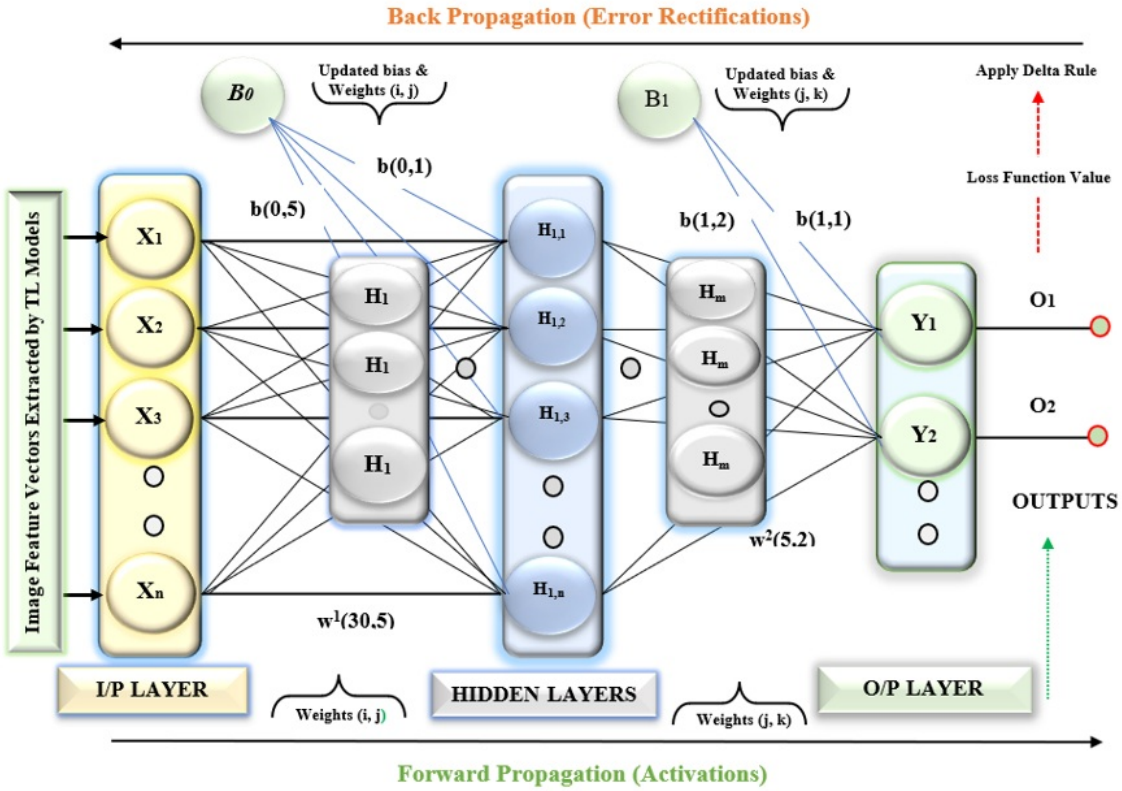

Fig. 2. MLP Architecture for detecting LD

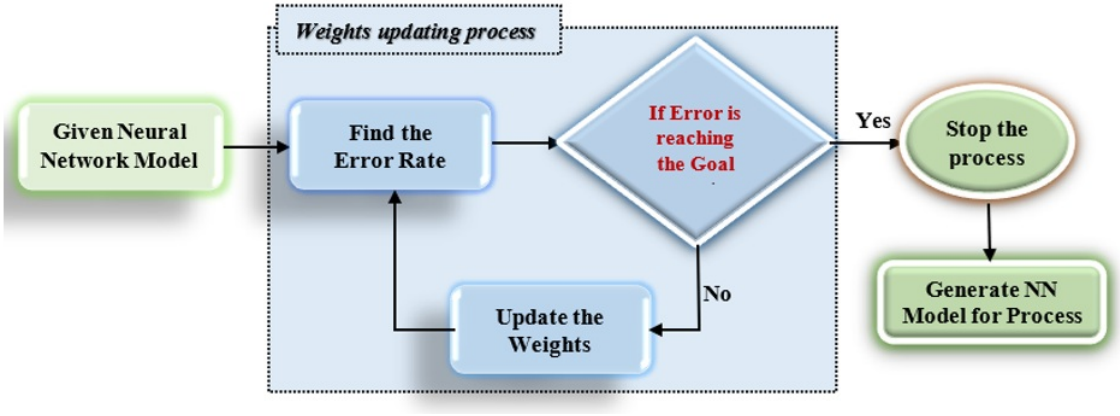

Fig. 3. Error Reduction and Weights and Bias Updating Process in the MLP

6) Confusion Matrix and Performance Parameters : The confusion matrix (**Fig. 5**) assesses a model's accuracy in predicting CKD and NCKD, encompassing TPs, False Negatives (FN), FPs, and TNs. The matrix also includes totals for predicted and actual cases of both diseases [27].

Performance matrices: ML metrics include accuracy (ACC), precision (PRE), recall (REC), and F1-score. Accuracy shows how many predictions were correct. Precision indicates the percentage of TPs among all positive predictions. The recall is a measure of how many TPs were recognized. The F1 score combines ACC and REC to provide a fair evaluation. Equations (5) to (14) show the whole metrics for classes LD(1) and NLD(0).

$$Accuracy = \frac{TruePos(LD) + TruePos(NLD)}{Total((LD) + (NLD))} \quad (29)$$

$$Precision(LD) = \frac{TruePos(LD)}{TruePos(LD) + FalsePos(LD)} \quad (30)$$

$$Precision(NLD) = \frac{TrueNeg(NLD)}{TrueNeg(NLD) + FalsePos(NLD)} \quad (31)$$

$$Precision = \frac{Precision(NLD) + Precision(LD)}{2} \quad (32)$$

$$Recall(LD) = \frac{TrueNeg(LD)}{TrueNeg(LD) + FalseNega(LD)} \quad (33)$$

$$Recall(NLD) = \frac{TruePos(NLD)}{TruePos(NLD) + FalseNega(NLD)} \quad (34)$$

$$Recall = \frac{Recall(LD) + Recall(NLD)}{2} \quad (35)$$

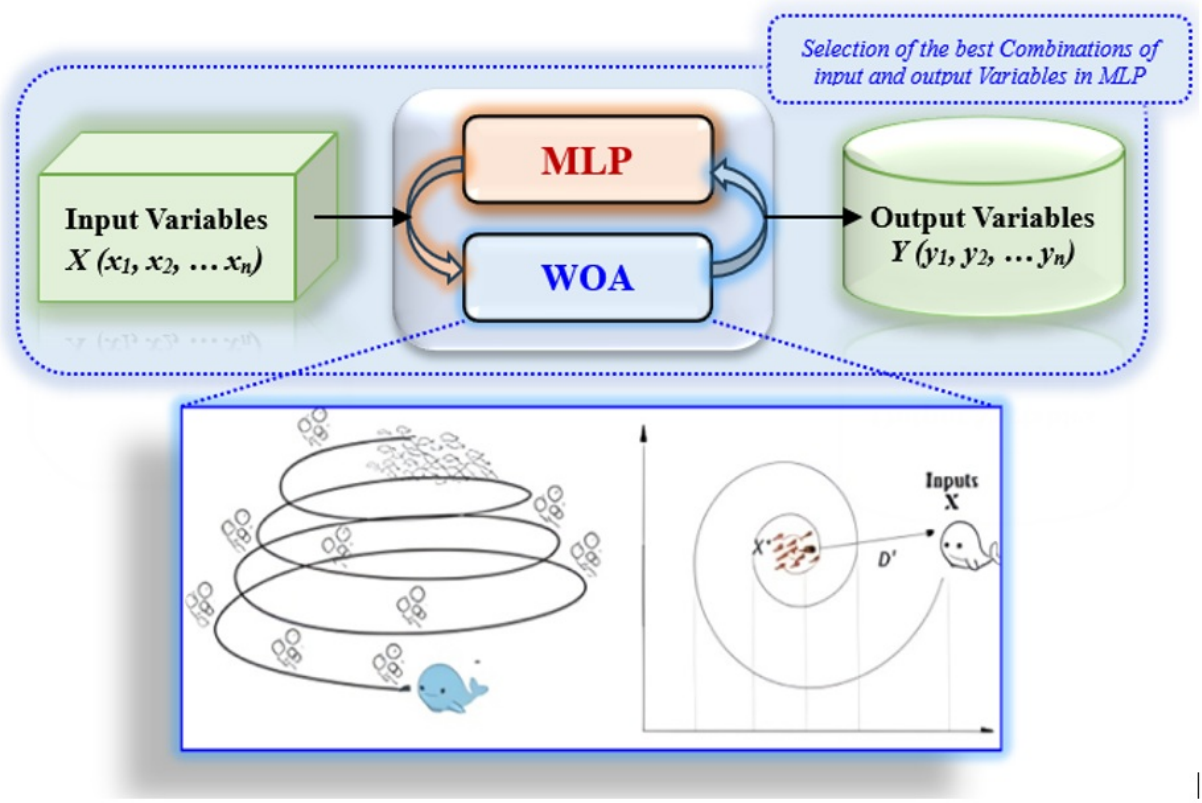

Fig. 4. MLP-WOA Training Detailed model

ConfusionMatrix		Predicted values		Total
Actual Values	Class	NLD(0)	LD(1)	
	NLD(0)	(0, 0)	(0,1)	T3
	LD(1)	(1,0)	(1,1)	T4
Total		T1	T2	T

Fig. 5. Confusion Matrix Structure for the LD Dataset Classification

$$F1Score(LD) = 2 \times \frac{Precision(LD) \times Recall(LD)}{Precision(LD) + Recall(LD)} \quad (36)$$

$$F1Score(LD) = 2 \times \frac{Precision(LD) \times Recall(LD)}{Precision(LD) + Recall(LD)} \quad (37)$$

$$F1Score = \frac{F1Score(LD) + F1Score(NLD)}{2} \quad (38)$$

TABLE V
STATISTICAL ANALYSIS ON CATEGORICAL FEATURES OF APLD DATASET

Features	Group (IDs)	NLD	LD	Total-Data
Gender	Male (1)	412	683	1095
	Female (0)	174	191	365
Smoke	No (0)	316	410	726
	Yes (1)	270	464	734
Drink	No (0)	496	418	914
	Yes (1)	88	458	546
Vomiting	Absent (0)	470	243	713
	Present (1)	116	631	747
Headache/ BoneAche	Absent (0)	443	290	733
	Present (1)	144	583	727
Fever	Absent (0)	468	208	676
	Present (1)	118	666	784
BP	Normal (0)	394	368	762
	Low BP (1)	91	182	273
	High BP (2)	98	327	425

IV. RESULTS ANALYSIS

This section analyses the results from the WOA-optimized MLP model for LD identification. It starts with statistics from the dataset. Then, it evaluates feature selection techniques using statistical and ML methods. The performance of different ML classifiers is compared using confusion matrices and evaluation metrics. The WOA+MLP framework shows better accuracy and predictive abilities.

1) Statistical Analysis of AP-LD Dataset: **Table V** presents a statistical analysis of categorical attributes in a dataset comparing individuals with NLD and LD. The analysis shows that males have a higher prevalence of LD (62.4%), while females have a more balanced distribution (174 in NLD and 191 in LD). Smoking is more prevalent among males (63.2%) than non-smokers (56.5%), suggesting a correlation between smoking and an increased likelihood of LD. Drinking is more prevalent among drinkers (83.9%), with a strong association between alcohol consumption and LD , aligning with known medical risk factors. The study shows that vomiting is a major symptom of LD. About 84.5% of people who vomit have LD. In contrast, only 34.1% of those who do not vomit have LD. Headaches and bone aches are also linked to LD , with 80.2% of affected individuals having the condition. Fever is another strong indicator, as 84.9% of those with fever have LD. Blood pressure is classified as Normal, Low, or High. Among individuals, 762 have normal BP, 273 have low BP, and 425

TABLE VI
STATISTICAL ANALYSIS ON CONTINUOUS FEATURES OF APLD DATASET

Continuous Attributes	Mean Values			Median Values		
	NLD (0)	LD (1)	Total Data	NLD (0)	LD (1)	Total Data
Age	41.75427	45.63616	44.07808	40	46	45
TB	1.219283	4.65389	3.275342	0.9	1.8	1.3
DB	0.384642	2.199199	1.47089	0.2	0.8	0.3
AP	96.87884	325.1705	233.5411	49	194	151
AA	36.3686	186.4611	126.2185	28	74	48
AAT	114.6672	213.7586	173.9863	93	168	126
TP	5.312799	6.237071	5.866096	5.5	6.4	6.2
Albumin	3.268601	3.071739	3.150753	3.2	3	3.1
AG-Ratio	1.074198	0.922197	0.983205	1	0.9	1

have high. The data shows a clear trend. The risk of LD goes up with abnormal blood pressure. High blood pressure has the strongest link to LD.

The **table VI** analyzed integral or continuous attributes to identify differences between NLD and LD cases. The mean age of LD patients was found to be higher (45.63) than non-LD patients (41.75), indicating a higher prevalence of LD in older individuals. Total Bilirubin (TB) levels were higher in LD cases. Direct Bilirubin (DB) levels were also higher in LD cases (Mean TB: 4.65, DB: 2.19), confirming bilirubin's role as a critical biomarker. Alkaline Phosphatase (AP) and Aminotransferases (AA, AAT) levels were also significantly elevated in LD cases (Mean AA: 186.46, Mean AAT: 213.75), emphasizing liver enzyme elevation as a strong indicator of disease presence. Total Protein (TP) was slightly higher in LD cases (Mean: 6.23) compared to NLD (Mean: 5.31), but the difference was not as pronounced. Albumin levels were lower in LD patients (Mean: 3.07) than non- LD cases (Mean: 3.26), suggesting that liver dysfunction may reduce albumin production. The Albumin-Globulin Ratio (AG-Ratio) was notably lower in LD patients (Mean: 0.92) compared to non- LD cases (Mean: 1.07), reinforcing its role as a diagnostic parameter for liver health.

2) Feature Selection of LD Dataset Through Statistical Models: The LD dataset is based on the North Coastal Districts of AP, India. It examines the relationship between 16 features and the diagnosis of LD. The correlation (**Fig. 6**) coefficients range from -1 to 1, and colours show the strength and direction of these correlations. The diagonal elements F1 and F2 have a self-correlation of 1.0. This sign indicates that the matrix structure is intact, as each feature correlates perfectly with itself. There are moderate to strong positive correlations among several biochemical markers. Total Bilirubin (TB) and Direct Bilirubin (DB) have a strong correlation of 0.99, indicating that higher total bilirubin levels are associated with higher direct bilirubin levels, often seen in liver dysfunction. Alkaline Phosphatase (AP) and Aspartate Aminotransferase (AST) have a significant connection of 0.54, meaning that when one liver enzyme level goes up, the other one might also rise in LD. Alanine Aminotransferase (AAT) and total proteins are somewhat related, with a score of 0.43, suggesting that changes in liver enzyme activity could be connected to how proteins are made.

The target variable T1 (Diagnosis: 0 = non-LD , 1 = LD) has weak to moderate negative correlations with several features. The A-G Ratio (F16) has a correlation of -0.26 with T1, indicating that a lower albumin-to-globulin ratio may be linked to LD. A15 feature (Albumin) had a -0.21 association with Target T1, indicating LD. This finding aligns with clinical predictions. LD usually leads to lower albumin levels and changes in A-G ratios because of impaired liver function. Lifestyle factors like smoking (F3) and drinking (F4) have weak correlations with T1, at 0.13 and 0.08, respectively, implying that while they may contribute to LD risk, their direct impact on diagnosis is limited. Symptoms such as Vomiting (F5), Headache/Bone Ache (F6), and Fever (F7) exhibit very weak correlations with T1 (ranging from 0.11 to 0.16). Blood pressure (BP, F8) has a moderate positive correlation of 0.33 with T1, indicating that higher BP may relate to LD , possibly due to conditions like cirrhosis-related hypertension. Age (F1) shows moderate positive correlations with total bilirubin (F9, 0.34) and direct bilirubin (F10, 0.33), suggesting older individuals may have higher bilirubin levels, a risk factor for liver issues. Gender (F2) has a weak positive correlation of 0.22 with Drink (F4), indicating possible gender-specific drinking habits, but its effect on T1 is minimal at 0.08.

The **table VII** shows the importance of features in the LD Dataset. It uses four metrics: Information Gain, Gain Ratio, Gini Index, and Chi-Square. The top five features are F11, F07, F12, F05, and F10. They have the highest Information Gain value of 0.262. It means they are important for classifying the dataset. The Gain Ratio and Gini Index also help measure feature significance. The top features have high scores, indicating they are crucial for the classification of LD. F11, F07, and F12 are the most important attributes. Middle-tier features can be used as extra inputs. Low-ranked features can be removed to simplify the model without losing much accuracy.

Table VIII shows the importance scores of features for the LD dataset. It uses ML models like XGBoost, AdaBoost, DT, and CatBoost. The table lists the most important attributes and their statistical weights, including the mean and SD in predicting liver disease. The LD Dataset shows important features using four methods: XGBoost, AdaBoost, DTs, and CAT Boost. The top features are F11, F07, F05, and F12. AdaBoost highlights F11 the most, with a mean of 0.1636. DTs also rank F11 highest, at 0.1468. However, CAT Boost ranks F07 first, at 0.0621. These features are important for predicting LD as they relate to key biochemical markers. The method used affects which features are prioritized. XGBoost, AdaBoost, and DTs fever F11, while CAT Boost focuses on F07.

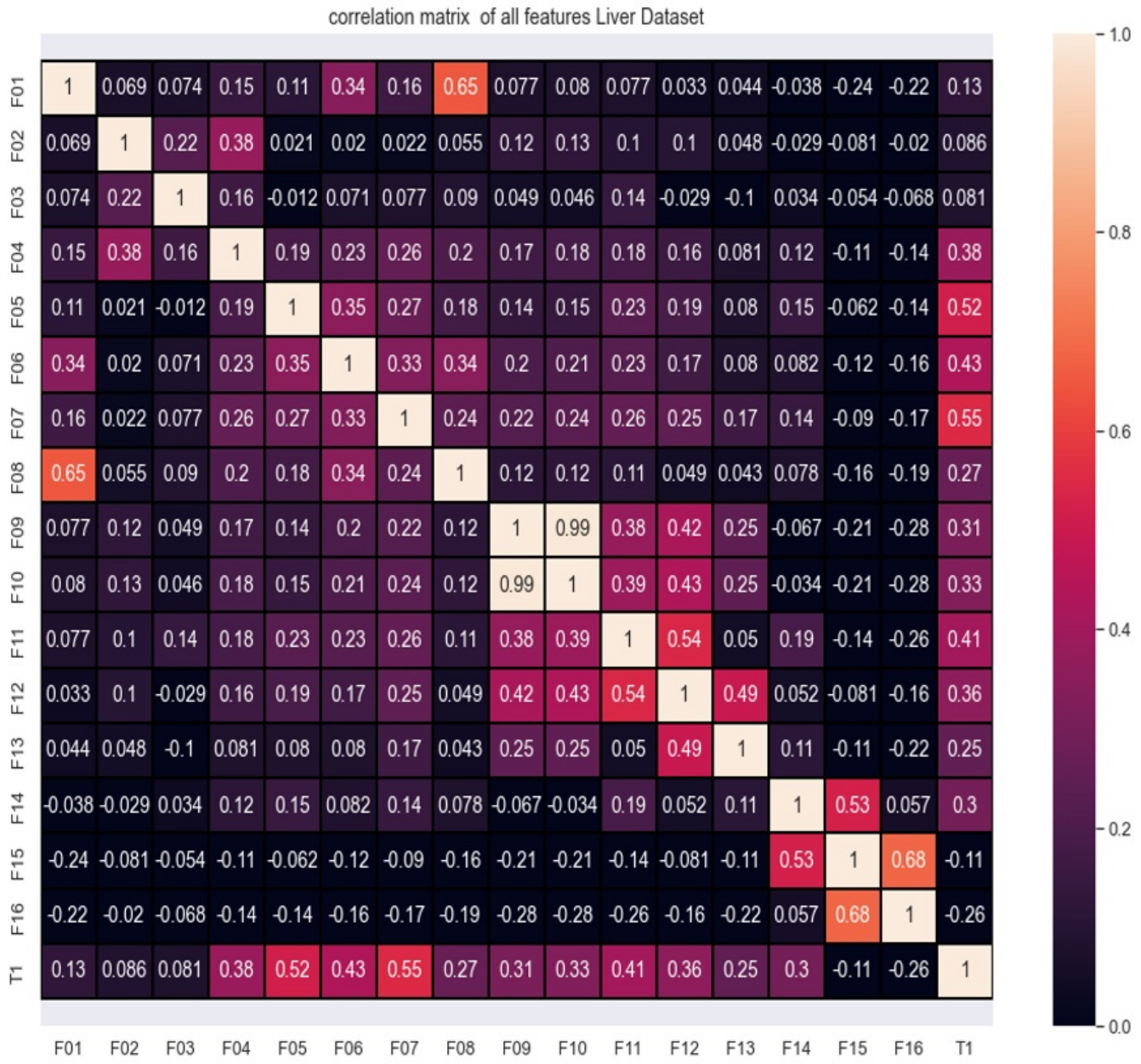

Fig. 6. Correlation Matrix analysis of AP LD Dataset

3) Feature Selection (FS) of AP-LD Dataset Through ML Models: The XGBoost Feature Ranking and Importance graph reveals F11, F05, and F07 as the most influential features in predicting LD. AdaBoost Feature Ranking and Importance figure 6 (A) shows how boosting enhances FS, with F11, F07, and F12 playing crucial roles in classification. Decision Trees (DTs) Feature Ranking and Importance plot shows F11, F07, and F05 as dominant factors in decision-making. CAT Boost Feature Ranking and Importance visualization highlights F07, F11, and F05 as top-ranking attributes. **Fig. 7** shows the top 10 attributes for predicting LD. These attributes were identified using ML methods. The data comes from the AP-LD dataset in the North Coastal region of Andhra Pradesh. Feature selection was done using four ML models: XGBoost, AdaBoost, Decision Trees, and CatBoost. Each model shows ranked attributes based on importance scores. This reveals both common and unique predictors important for diagnosing LD.

The target variable T1 (Diagnosis: 0 = non-LD , 1 = LD) has weak to moderate negative correlations with several features. The A-G Ratio (F16) has a correlation of -0.26 with T1, indicating that a lower albumin-to-globulin ratio may be linked to LD. A15 feature (Albumin) had a -0.21 association with Target T1, indicating LD. This finding aligns with clinical predictions. LD usually leads to lower albumin levels and changes in A-G ratios because of impaired liver function. Lifestyle factors like smoking (F3) and drinking (F4) have weak correlations with T1, at 0.13 and 0.08, respectively, implying that while they may contribute to LD risk, their direct impact on diagnosis is limited. Symptoms such as Vomiting (F5), Headache/Bone Ache (F6), and Fever (F7) exhibit very weak correlations with T1 (ranging from 0.11 to 0.16). Blood pressure (BP, F8) has a moderate positive correlation of 0.33 with T1, indicating that higher BP may relate to LD , possibly due to conditions like cirrhosis-related hypertension. Age (F1) shows moderate positive correlations

TABLE VII
ATTRIBUTE RANKS FOR THE LIVER DISEASE DATASET – RANK NUMBER AND FEATURE ATTRIBUTES

Information Gain		Gain Ratio		Gini Index		χ^2 (Chi-Square)	
Rank (F#)	Value	Rank (F#)	Value	Rank (F#)	Value	Rank (F#)	Value
R01 (F11)	0.262	R01 (F07)	0.232	R01 (F11)	0.162	R01 (F11)	301.294
R02 (F07)	0.231	R02 (F05)	0.205	R02 (F07)	0.148	R02 (F12)	301.278
R03 (F12)	0.227	R03 (F06)	0.139	R03 (F12)	0.137	R03 (F07)	198.957
R04 (F05)	0.205	R04 (F11)	0.131	R04 (F05)	0.132	R04 (F05)	184.719
R05 (F10)	0.143	R05 (F04)	0.120	R05 (F06)	0.091	R05 (F09)	160.350
R06 (F06)	0.139	R06 (F12)	0.114	R06 (F10)	0.090	R06 (F04)	128.059
R07 (F09)	0.130	R07 (F10)	0.076	R07 (F09)	0.080	R07 (F06)	127.791
R08 (F04)	0.113	R08 (F09)	0.065	R08 (F04)	0.072	R08 (F10)	118.743
R09 (F14)	0.089	R09 (F14)	0.045	R09 (F14)	0.060	R09 (F08)	96.540
R10 (F16)	0.079	R10 (F16)	0.041	R10 (F16)	0.051	R10 (F16)	77.355
R11 (F13)	0.075	R11 (F13)	0.037	R11 (F13)	0.049	R11 (F14)	76.589
R12 (F08)	0.053	R12 (F08)	0.037	R12 (F08)	0.035	R12 (F13)	57.236
R13 (F15)	0.041	R13 (F15)	0.020	R13 (F15)	0.027	R13 (F01)	23.139
R14 (F01)	0.017	R14 (F01)	0.008	R14 (F01)	0.011	R14 (F15)	9.980
R15 (F02)	0.005	R15 (F03)	0.008	R15 (F02)	0.004	R15 (F03)	7.535
R16 (F03)	0.005	R16 (F02)	0.007	R16 (F03)	0.003	R16 (F02)	2.560

TABLE VIII
ATTRIBUTE RANKS FOR THE LIVER DISEASE DATASET – FEATURE IMPORTANCE FROM ENSEMBLE AND TREE-BASED MODELS

XGBoost		AdaBoost		Decision Tree (DTs)		CatBoost	
Rank (F#)	Mean ± Std	Rank (F#)	Mean ± Std	Rank (F#)	Mean ± Std	Rank (F#)	Mean ± Std
R01 (F11)	0.0703 ± 0.0030	R01 (F11)	0.1636 ± 0.0062	R01 (F11)	0.1468 ± 0.0046	R01 (F07)	0.0621 ± 0.0062
R02 (F05)	0.0665 ± 0.0044	R02 (F07)	0.1284 ± 0.0033	R02 (F07)	0.0957 ± 0.0047	R02 (F11)	0.0572 ± 0.0046
R03 (F07)	0.0643 ± 0.0034	R03 (F12)	0.1049 ± 0.0068	R03 (F05)	0.0951 ± 0.0066	R03 (F05)	0.0528 ± 0.0058
R04 (F12)	0.0473 ± 0.0029	R04 (F05)	0.0933 ± 0.0055	R04 (F12)	0.0631 ± 0.0037	R04 (F12)	0.0478 ± 0.0014
R05 (F01)	0.0208 ± 0.0017	R05 (F14)	0.0529 ± 0.0032	R05 (F06)	0.0551 ± 0.0020	R05 (F01)	0.0273 ± 0.0032
R06 (F14)	0.0178 ± 0.0029	R06 (F13)	0.0463 ± 0.0042	R06 (F13)	0.0482 ± 0.0023	R06 (F14)	0.0258 ± 0.0035
R07 (F06)	0.0165 ± 0.0031	R07 (F01)	0.0395 ± 0.0034	R07 (F01)	0.0420 ± 0.0024	R07 (F13)	0.0197 ± 0.0024
R08 (F13)	0.0141 ± 0.0021	R08 (F02)	0.0348 ± 0.0016	R08 (F02)	0.0351 ± 0.0021	R08 (F02)	0.0181 ± 0.0034
R09 (F04)	0.0068 ± 0.0014	R09 (F06)	0.0343 ± 0.0026	R09 (F04)	0.0264 ± 0.0041	R09 (F16)	0.0156 ± 0.0033
R10 (F09)	0.0065 ± 0.0011	R10 (F16)	0.0276 ± 0.0039	R10 (F09)	0.0211 ± 0.0011	R10 (F06)	0.0127 ± 0.0026
R11 (F15)	0.0063 ± 0.0006	R11 (F04)	0.0267 ± 0.0042	R11 (F14)	0.0162 ± 0.0026	R11 (F09)	0.0122 ± 0.0012
R12 (F02)	0.0062 ± 0.0012	R12 (F09)	0.0199 ± 0.0011	R12 (F15)	0.0151 ± 0.0027	R12 (F04)	0.0077 ± 0.0018
R13 (F08)	0.0024 ± 0.0010	R13 (F15)	0.0164 ± 0.0017	R13 (F03)	0.0063 ± 0.0011	R13 (F15)	0.0049 ± 0.0012
R14 (F16)	0.0009 ± 0.0003	R14 (F08)	0.0101 ± 0.0011	R14 (F16)	0.0021 ± 0.0008	R14 (F08)	0.0034 ± 0.0011
R15 (F03)	0 ± 0	R15 (F10)	0.0063 ± 0.0018	R15 (F10)	0.0018 ± 0.0004	R15 (F03)	0.0003 ± 0.0004
R16 (F10)	0 ± 0	R16 (F03)	0 ± 0	R16 (F08)	0 ± 0	R16 (F10)	-0.0003 ± 0.0009

TABLE IX
PERFORMANCE PARAMETERS VALUES FOR THE EXPERIMENTAL ML MODELS ON AP-LIVER DATASET

ML Model	AUC	CA	Precision	Recall	F1-Score
SVM (RBF)	0.87	0.82	0.82	0.82	0.82
Naïve Bayes	0.93	0.80	0.81	0.81	0.81
KNN	0.85	0.79	0.82	0.82	0.82
Decision Tree	0.87	0.88	0.87	0.87	0.87

with total bilirubin (F9, 0.34) and direct bilirubin (F10, 0.33), suggesting older individuals may have higher bilirubin levels, a risk factor for liver issues. Gender (F2) has a weak positive correlation of 0.22 with Drink (F4), indicating possible gender-specific drinking habits, but its effect on T1 is minimal at 0.08.

4) Feature Selection using all Ranking of Statistical and ML Models: The AP-LD dataset analysis identified the top 11 attributes (R01 to R11) based on their consistently high rankings across various FS methods. These methods include Information Gain, Gain Ratio, Gini Index, Chi-Square (^{2}), XGBoost, AdaBoost, DTs, and CAT Boost. Attributes that frequently ranked high and maintained consistency across different methods were prioritized, with higher individual rankings receiving more importance. There are three main steps in the analysis methodology. Prioritizing frequently occurring attributes, ranking aggregation first extracts the top 11 ranks (R01 to R11) for each method. The second is feature importance scores, such as an Info. The ranking order is supported by the gain value of 0.262 for F11, which illustrates how strongly each feature contributes. Lastly, cross-method consistency identifies attributes that rank high in statistical methods (e.g., Info. Gain, ^{2}), and ML methods like XGBoost and CAT Boost are considered more reliable predictors. The items are ranked from R01 to R11 as follows: F11, F07, F12, F05, F06, F10, F09, F04, F14, F16,and F13.

5) ML Models Analysis on Selected Features on AP-LD Dataset: Confusion Matrices Analysis is shown in the **Fig. 8**. The image shows four confusion matrices for different ML models analysing a LD dataset. Each matrix is a 2x2 grid displaying TPs, TNs), FPs), and FNs for binary classification. Darker colours (like blue) represent higher values, while lighter colours (like white) indicate lower values. The Confusion Matrix for the SVM (Fig. 8 top left) shows 139 TPs, 19 FPs, 31 false negatives (FN), and 83 TNs. The NBs Confusion Matrix (Fig. 8 top right) shows the following values: TPs = 116, FPs = 42, False Negatives (FN) = 12, and TNs of 102. The KNN Confusion Matrix (Fig. 8 (bottom left)) shows 130 TPs, 28 FPs, 28 false negatives (FN), and 86 TNs. The Confusion Matrix (Fig. 8 (bottom right)) for the DTs shows 142 TPs, 16 FPs, 18 true negatives, and 96 false negatives. **Fig. 9** and **Table IX** evaluate the performance parameters values for experimental

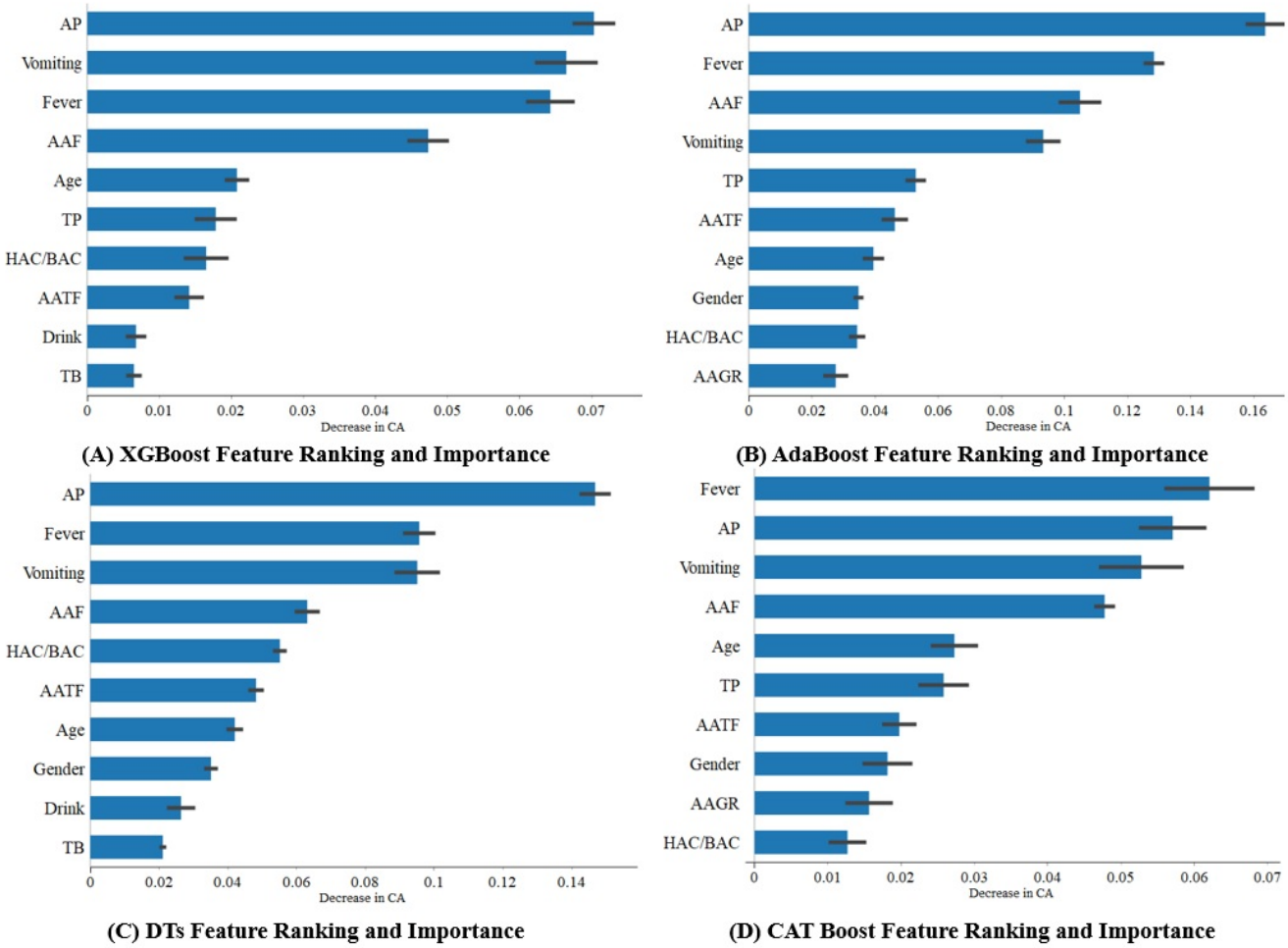

Fig. 7. Feature Selection Process (top 10 ranked features) though ML Models like XGBoost, AdaBoost, DTs and CAT Boost on AP-Liver Dataset

TABLE X
ROC-AUC ANALYSIS ON EXPERIMENTAL ML MODELS

Model	Colour	AUC Value	Performance
SVM	Blue	0.87	Very good. The SVM model can separate the classes well but not perfectly.
Naive Bayes	Orange	0.93	Excellent. NBs has the highest AUC, indicating it's the best among the compared models for this task.
KNN	Green	0.85	Good, but comparatively lower than SVM and Naive Bayes. Shows more confusion between classes.
Decision Tree	Red	0.87	Matches SVM in AUC score. Good discrimination ability but less smooth curve.

ML models on the AP-LD dataset. The models tested include SVM (RBF), NBs, KNN, and DTs. They were measured using important metrics like AUC, CA, F1-Score, etc.The ROC analysis (**Fig. 10**) shows that NBs is the best model for LD classification, achieving an AUC of 0.93. SVM and Decision Tree have an AUC of 0.87, but the Decision Tree may overfit. KNN has an AUC of 0.85, which is decent but needs tuning.

The **table X** shows the detailed comparison of ML model concerning ROC-AUC values. The SVM-RBF model showed a good separation between LD and NLD. It had an AUC of 0.87. The classification accuracy was 82%. The model kept steady precision and recall values. It reduced FPs and false negatives. The F1-score was 0.82, showing it worked effectively for the classification of LD and NLD. Naïve Bayes (NB) achieved an AUC value of 0.93, indicating excellent classification ability. The classification accuracy is 80%. Both Precision and Recall are 0.81, showing reliable performance in identifying positive cases. The F1-score is also 0.81, reflecting a good balance. The KNN model has an AUC of 0.85. Its classification accuracy is 79%, the lowest compared to other models. Both Precision and Recall are 0.82. The F1-score is also 0.82. The study indicates that KNN performs reasonably well. The Decision Tree (DT) model has a classification accuracy of 88%. It achieves a Precision, Recall, and F1-score of 0.87. The AUC score is also 0.87, like the SVM model. The DT model performs well. It can be effective with techniques like pruning.

TABLE XI
ROC-AUC ANALYSIS ON OPTIMIZED AND BASELINE MLP MODELS PERFORMANCE ON AP-LIVER DATASET

Model	Colour	AUC Value	Performance
MLP	Orange	0.89	Very good. The MLP model shows strong classification performance.
WOA + MLP	Green	0.91	Excellent. Optimization using WOA improves the model's discrimination power.

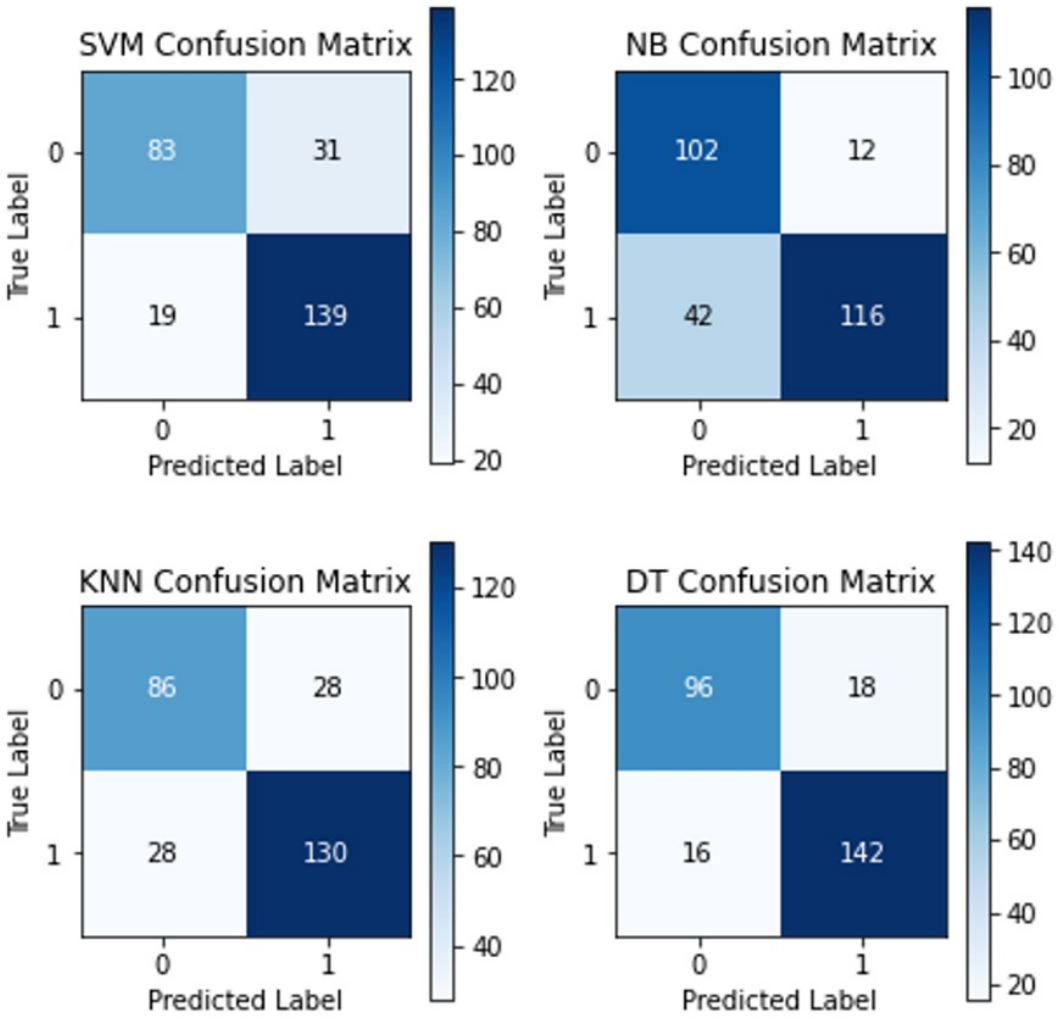

Fig. 8. Experimental ML Models Confusion Matrices Analysis

SVM:
Accuracy: 0.8162
Classification Report:

	precision	recall	f1-score	support
0	0.81	0.73	0.77	114
1	0.82	0.88	0.85	158
accuracy			0.82	272
macro avg	0.82	0.80	0.81	272
weighted avg	0.82	0.82	0.81	272

(A) SVM Performance Parameter Values

Naive Bayes:
Accuracy: 0.8015
Classification Report:

	precision	recall	f1-score	support
0	0.71	0.89	0.79	114
1	0.91	0.73	0.81	158
accuracy			0.80	272
macro avg	0.81	0.81	0.80	272
weighted avg	0.82	0.80	0.80	272

(B) Naïve Bayes Performance Parameter Values

KNN:
Accuracy: 0.7941
Classification Report:

	precision	recall	f1-score	support
0	0.75	0.75	0.75	114
1	0.82	0.82	0.82	158
accuracy			0.79	272
macro avg	0.79	0.79	0.79	272
weighted avg	0.79	0.79	0.79	272

(C) KNN Performance Parameters

Decision Tree:
Accuracy: 0.8750
Classification Report:

	precision	recall	f1-score	support
0	0.86	0.84	0.85	114
1	0.89	0.90	0.89	158
accuracy			0.88	272
macro avg	0.87	0.87	0.87	272
weighted avg	0.87	0.88	0.87	272

(D) DTs Performance Parameters

Fig. 9. Detailed Performance Parameters of each Experimental ML Models on AP-Liver Dataset

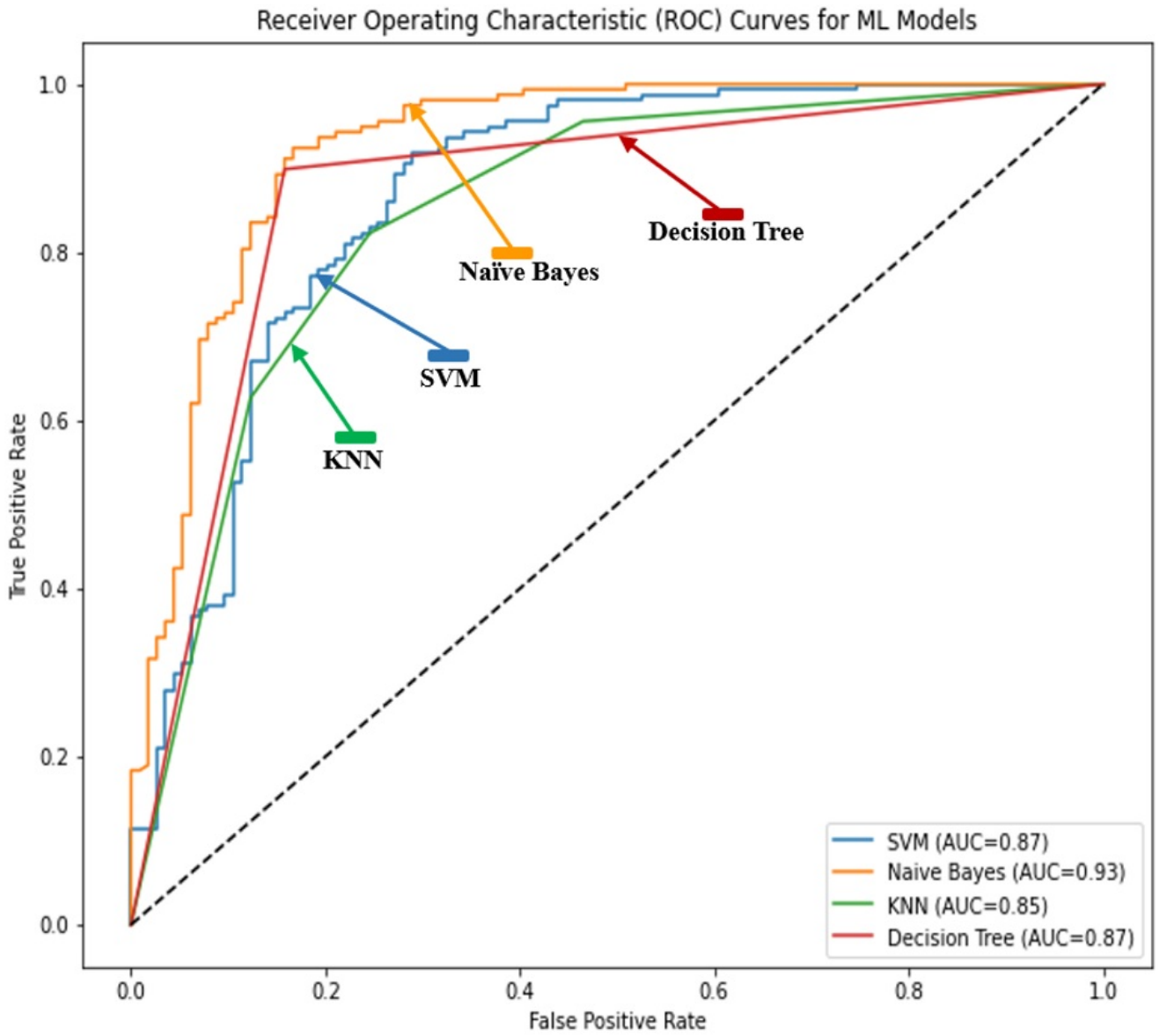

Fig. 10. ML ROC-AUC Analysis on AP-Liver Dataset

TABLE XII
PERFORMANCE PARAMETERS VALUES FOR THE MLP AND WOA+MLP ON AP-LIVER DATASET

MLP Models	AUC	CA	F1Score	Recall	Precision
MLP(10 10)	0.89	0.89	0.89	0.89	0.89
WOA+MLP(10 10)	0.91	0.92	0.92	0.92	0.92

TABLE XIII
COMPARATIVE PERFORMANCE ANALYSIS OF ML MODELS AND THE OPTIMIZED MLPS

ML Model	AUC	CA	Precision	Recall	F1-Score
SVM (RBF)	0.87	0.82	0.82	0.82	0.82
Naïve Bayes	0.93	0.80	0.81	0.81	0.81
KNN	0.85	0.79	0.82	0.82	0.82
Decision Tree	0.87	0.88	0.87	0.87	0.87
MLP(10 10)	0.89	0.89	0.89	0.89	0.89
WOA+MLP(10 10)	0.91	0.92	0.92	0.92	0.92

6) MLP and WOA+MLP Models Analysis: The confusion matrix (**Fig. 11**) compares two models for classifying LD. For the MLP (10,10) model: The model correctly identifies 148 cases of LD and 95 cases NLD. It misclassifies 19 non-disease cases as positive and misses 10 disease cases. This indicates The WOA + MLP model has the following results: (TP-148),(TN-10), (FP-13), and (FN-10). The model performs better in identifying non-disease cases, with 101 true negatives compared to 95. It also reduces FPs from 19 to 13. It leads to improved precision and overall accuracy.

The ROC analysis (**Fig. 12**) shows that the WOA-optimized MLP is better than the standalone MLP for classifying LD. The WOA + MLP model has an AUC of 0.91. The standard MLP model has an AUC of 0.89. The study indicates that the WOA + MLP model is more reliable. The ROC analysis shows that the WOA-optimized MLP is better than the standalone MLP for classifying LD. The WOA + MLP model has an AUC of 0.91. The standard MLP model has an AUC of 0.89. The study indicates that the WOA + MLP model is more reliable. The detailed analysis shows in the table XI.

The study (**Fig. 13** and **Table XII**) compares two models

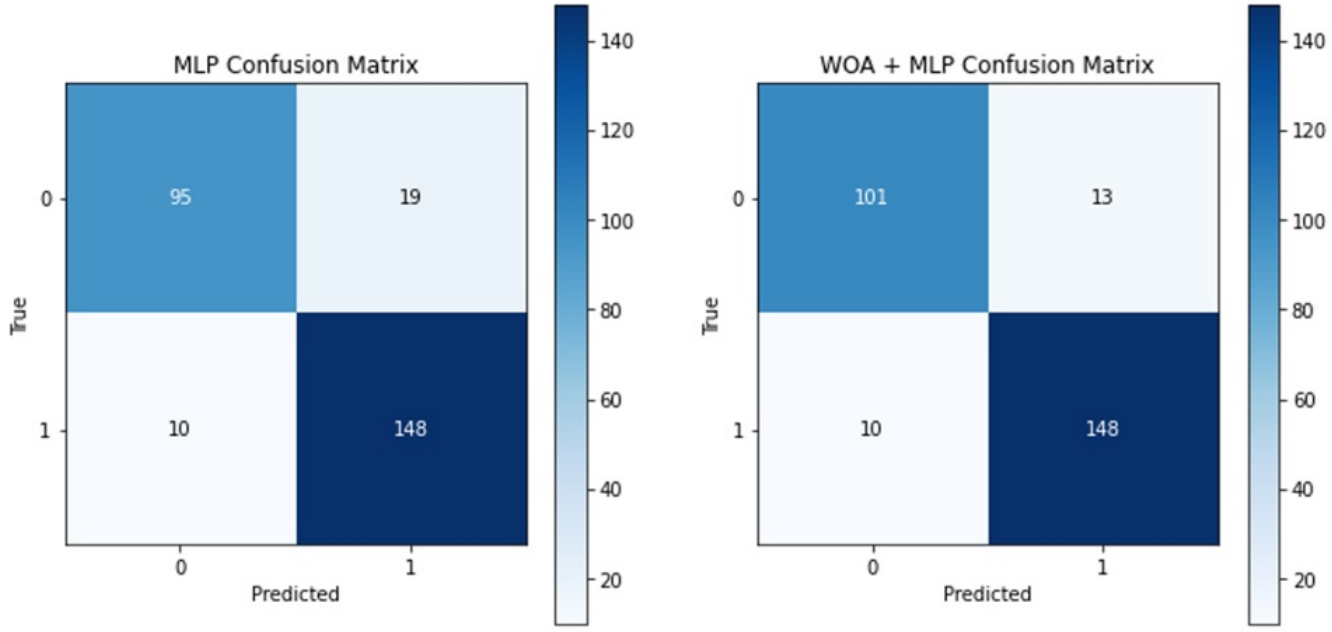

Fig. 11. MLP and WOA+MLP Models Confusion Matrices Analysis

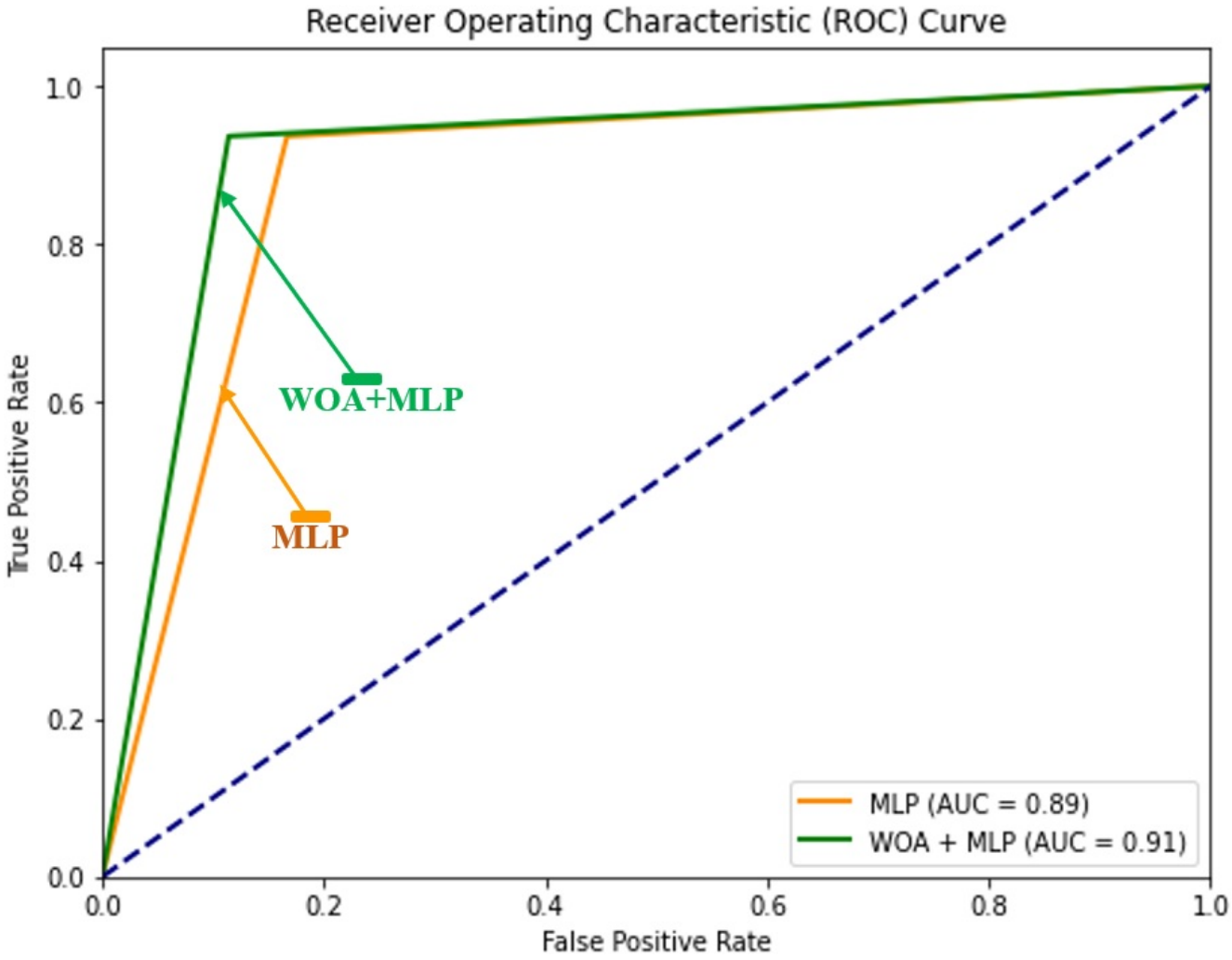

Fig. 12. MLP and WOA+MLP Models ROC-AUC Analysis

for LD classification. The MLP model is a standard structure with a (10,10) dimensions. It has an accuracy of 0.89 for both Class 0 (Non-Disease) and Class 1 (Disease). The second model is a WOA-optimized MLP, also with a structure

Accuracy: 0.8933823529411765
Classification Report:

	precision	recall	f1-score	support
0	0.90	0.83	0.87	114
1	0.89	0.94	0.91	158
accuracy			0.89	272
macro avg	0.90	0.89	0.89	272
weighted avg	0.89	0.89	0.89	272

(A) MLP Performance Parameter Values

Accuracy: 0.9154411764705882
Classification Report:

	precision	recall	f1-score	support
0	0.91	0.89	0.90	114
1	0.92	0.94	0.93	158
accuracy			0.92	272
macro avg	0.91	0.91	0.91	272
weighted avg	0.92	0.92	0.92	272

(B) WOA+MLP Performance Parameter Values

Fig. 13. Detailed Performance Parameters of each Experimental MLP and WOA+MLP Models Analysis on AP-Liver Dataset

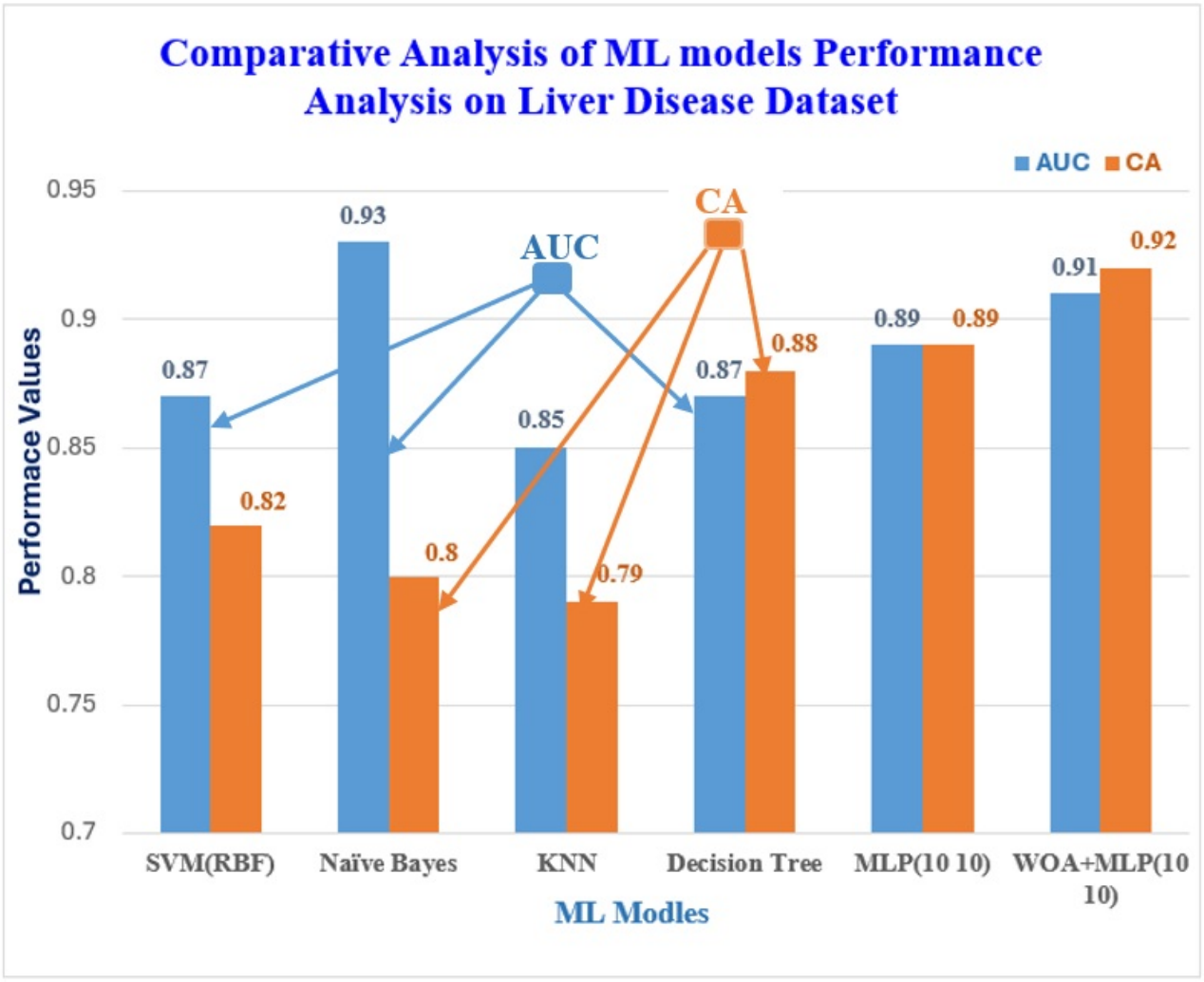

Fig. 14. Comparative Analysis of AUC and CA for Experimental ML, MLP, and WOA+MLP Models on the AP-LD Dataset.

of (10,10). This model achieves an accuracy of 0.92 for Class 0 (Non-LD Macro Average). The application of the WOA significantly improves performance, particularly in balancing Precision and recall, FPs reducing and enhancing generalization.

V. Discussions

This section evaluates the comparative performance of various ML models and an optimized MLP model on the AP-LD dataset. The study compares the effectiveness of WOA-based MLP tuning with existing studies, analysing evaluation metrics like AUC, CA, recall, precision, and F1-score.

1) Comparative Study on Experimental ML and Optimized MLP Models : This study (**Table XIII**) analyses the performance of different ML classifiers and an optimized Multilayer Perceptron model using WOA, considering metrics like AUC, CA, Precision, Recall, and F1-Score. NBs achieved the highest AUC of 0.93. This indicates strong discriminative capability. However, its overall accuracy was 0.80, and its F1-score was 0.81. This suggests some misclassification. SVM with RBF kernel had an AUC of

TABLE XIV
COMPARATIVE STUDY OF THE PROPOSED WORK WITH EXISTING LD PREDICTION MODELS USING ML TECHNIQUES

Author Details	Year	Data Size	Methods	Key Findings and Result Analysis	Relevance to LD
Dritsas et al. (2023) [28]	2023	828	Various ML models evaluated using SMOTE and 10-fold cross-validation. Performance measured by CA, AUC, F1-Value, PRE, and REC	The Voting classifier outperformed others with 80.1% accuracy, 80.4% precision, 80.1% recall, and AUC of 88.4%. AdaboostM1 and RF also showed good performance.	Findings support early detection and prediction of LD.
Amin et al. (2023) [29]	2023	583	Used the ILPD dataset. Methods like PCA, FA, LDA for feature extraction and ML algorithms for classification.	Achieved 88.10% accuracy, 85.33% recall, 88.68% F1 score, and 88.20% AUC.	Helped in early diagnosis of LD.
Rahman et al. (2019) [30]	2019	583	ML algorithms LR, KNN, DTs, SVM, NBs, RF are evaluated	Logistic Regression achieved the highest accuracy of 75%, while Naïve Bayes had the lowest at 53%.	This highlights the potential of ML to reduce costs and complexity.
Yang et al. (2024) [31]	2024	730	LASSO regression, multiple logistic regression used.	AUC: Training set: 0.877; Validation set: 0.871.	Personalized predictive model for NAFLD risk assessment.
Wu et al. (2019) [32]	2019	577	Classification models used.	RF model achieved an AUROC of 0.925 and accuracy of 87.48%, outperforming other models.	Effectiveness of RF model.
Current Study	–	1460	SVM(RBF), NBs, KNN, and DTs	SVM(RBF): AUC 87.00%, CA 82.00%; NBs: AUC 93.00%, CA 80.00%; KNN: AUC 85.00%, CA 79.00%; DTs: AUC 87.00%, CA 88.00%.	Proposed WOA+MLP model outperforms most existing models in accuracy and AUC, enhancing early LD detection.
Current Study	–	1460	MLP, WOA+MLP (Proposed Model)	MLP(10 10) achieved 89.00% accuracy, 89.00% recall, 89.00% F1-score, and 89.00% AUC. WOA+MLP(10 10) achieved 92.00% accuracy, 92.00% recall, 92.00% F1-score, and 91.00% AUC.	

0.87. DT classifiers also had an AUC of 0.87. The KNN model had a classification accuracy of 0.79. This indicates it is sensitive to noise and struggles in high-dimensional space. In contrast, the standard MLP (10-10) neural network architecture performed better. It achieved an AUC of 0.89 and an accuracy of 0.89. The WOA+MLP (10-10) model uses WOA for hyperparameter tuning. It achieved a classification accuracy of 0.92. The precision, recall, and F1-score were all 0.92. The model also improved its AUC to 0.91. The study shows that optimizing MLP weights and learning parameters enhances model performance. The study shows that combining optimization algorithms, such as WOA, with ML models, like MLP, improves performance. This combination enhances the reliability of conventional ML techniques. The **Fig. 14** compares various ML models, including SVM, NBs, KNN, DT, MLP, and WOA-optimized MLP, using their AUC and CA. NBs had the highest AUC of 0.93 but lower accuracy at 0.80. SVM and Decision Tree had similar AUCs of 0.87, with DT being more accurate (0.88) than SVM (0.82). KNN had the lowest AUC of 0.85 and accuracy of 0.79.

2) Comparative Study on Experimental ML and Optimized MLP Models: Zheng et al. (2025) [33] investigated the application of the WOA in conjunction with Kolmogorov-Arnold Networks (KAN) for FS in medical datasets. WOA, a bio-inspired algorithm that imitates whale hunting behavior, is combined with KAN, a type of NN recognized for its capacity to approximate complex functions. The proposed method aimed to efficiently select the most relevant features from medical datasets, enhancing the performance of ML models in tasks such as disease prediction. Shuaib et al. (2019) [34] utilized WOA to select the most pertinent features from the dataset for email spam classification. The Rotation Forest (RF) algorithm is used for classification after the best features have been chosen. By using WOA to optimize FS and RF's rotation-based transformation to improve generalization, this combination improves classification accuracy and creates an effective spam detection system. Chakraborty et al. (2023) [35] enhanced performance, a hybrid WOA for global optimization combines WOA with additional optimization methods like PSO or GA. This hybrid approach overcomes the drawbacks of WOA, such as premature convergence, and enhances speed, accuracy, and balance in exploiting and exploring. For complicated, multi-dimensional optimization problems, it provides a more reliable solution. **Table XIV** compares the ML and optimized models with existing LD prediction studies. It shows differences in datasets, methods, and performance metrics. The current approach is more effective in improving the accuracy and early diagnosis of LD. Wu et al. (2019)[32] used various ML algorithms to predict the likelihood of developing fatty LD entails by evaluating patient data, including age, BMI, liver enzymes, and other biomarkers. To find important variables and trends linked to fatty LD, algorithms such as DTs, RFs, SVM, or NN are trained on this data. Early predictions from the model can then help with prompt diagnosis and treatment and lower the chance of developing more serious LDs. Ghazal et al. (2023)[36]proposed a model for early LD prediction analyzes medical data using ML techniques to spot patterns linked to the disease early on. The model can predict the likelihood of LD and classify risk factors by utilizing algorithms such as NNs, DTs, and SVM. This allows for prompt interventions. This method increases the precision of diagnoses, and it helps in the early identification of LDs. Early management of LDs is also crucial. Shaban et al. (2024) [37] suggested combining ML techniques with an enhanced binary butterfly optimization algorithm. By

improving FS, it sought to increase prediction accuracy and produce more dependable and effective diagnosis models. When compared to conventional methods, the approach showed better performance.

VI. Conclusion

The study developed an intelligent LD identification framework by integrating an MLP with the WOA, addressing the need for accurate and early diagnosis in the North Coastal region of AP, India. The average total bilirubin level was 4.65 in LD patients and 1.22 in non-LD patients. The albumin-globulin ratio was 0.92 in LD patients compared to 1.07 in non-LD patients. There was a decrease in the albumin-globulin ratio (0.92 in LD compared to 1.07 in non-LD), and 83.9% of LD patients had alcohol consumption as a lifestyle factor. Feature selection techniques found important predictors. These include total bilirubin (F11), fever (F07), and vomiting (F05). Information Gain values reached 0.262. This value helps in analysing the most influential attributes. The WOA+MLP model has a (10,10) architecture. The model outperformed baseline models, including SVM, Naive Bayes, KNN, and Decision Tree. The model achieved an AUC of 0.91. Its classification precision was 0.92 and balanced precision, recall, and F1 score were 0.92. The standalone MLP model achieved an AUC of 0.89 and an accuracy of 0.89. NBs had the highest AUC but lower accuracy. The decision tree model showed signs of overfitting. The WOA+MLP model improved diagnostic reliability. Reduce FPs and increase true negative predictions. The findings demonstrate the value of integrating advanced ML methodologies in the medical field, highlighting their role in improving patient outcomes. Future research could build on these techniques, examining their use in other complex diseases to further improve diagnostic precision and treatment strategies.

Acknowledgment

The authors express their gratitude to the Aditya Institute of Technology and Management, Tekkali, Srikakulam, AP, India for their unwavering support throughout the research, and extend special thanks to Director Prof. VV Nageswara Rao, Principal Dr. A. S. Srinivasa Rao, and DST FIST Lab central coordinator Dr. Vital.

References

[1] N. Mistry and A. Yeoman, "Clinical and laboratory assessment of symptomatic and asymptomatic liver disease," *Medicine*, vol. 51, no. 5, pp. 305–310, 2023.

[2] P. Terlapu, R. P. Sadi, R. Pondreti, and C. Tippana, "Intelligent identification of liver diseases based on incremental hidden layer neurons ann model," *International Journal of Computing and Digital Systems*, vol. 11, no. 1, pp. 1027–1050, 2022.

[3] V. Lala, M. Zubair, and D. Minter, "Liver function tests," https://www.statpearls.com/point-of-care/20995/, 2023, statPearls.

[4] D. T. O'Gurek, "Diseases of the liver," in *Family Medicine: Principles and Practice*. Cham: Springer International Publishing, 2022, pp. 1265–1279.

[5] E. H. Houssein, N. Abdalkarim, N. A. Samee, M. Alabdulhafith, and E. Mohamed, "Improved kepler optimization algorithm for enhanced feature selection in liver disease classification," *Knowledge-Based Systems*, vol. 297, pp. 1–17(111 960), 2024.

[6] S. Raziani, S. Ahmadian, S. M. J. Jalali, and A. Chalechale, "An efficient hybrid model based on modified whale optimization algorithm and multilayer perceptron neural network for medical classification problems," *Journal of Bionic Engineering*, vol. 19, no. 5, pp. 1504–1521, 2022.

[7] M. Varchagall and P. AdaguruYogegowda, "Early detection of liver disorders using hybrid soft computing techniques for optimal feature selection and classification," *Concurrency and Computation: Practice and Experience*, vol. 35, no. 6, pp. 1–14, 2023.

[8] X. Qiao, X. Long, Y. Li, and T. Liu, "Multi-threshold segmentation of liver tumor images based on deep learning and improved whale optimization algorithm," in *Proceedings of the 2024 5th International Symposium on Artificial Intelligence for Medicine Science*, 2024, pp. 211–215.

[9] A. Homsana, P. Southisavath, K. Kling, J. Hattendorf, S. Vorasane, D. H. Paris, and N. Probst-Hensch, "Steatotic liver disease among lean and non-lean individuals in southern lao pdr: a cross-sectional study of risk factors," *Annals of Medicine*, vol. 56, no. 1, pp. 1–17(2 329 133), 2024.

[10] J. H. Joloudari, H. Saadatfar, A. Dehzangi, and S. Shamshirband, "Computer-aided decision-making for predicting liver disease using pso-based optimized svm with feature selection," *Informatics in Medicine Unlocked*, vol. 17, pp. 1–17(100 255), 2019.

[11] R. Islam, A. Sultana, and M. N. Tuhin, "A comparative analysis of machine learning algorithms with tree-structured parzen estimator for liver disease prediction," *Healthcare Analytics*, vol. 6, pp. 1–11(100 358), 2024.

[12] S. M. Ganie and P. K. D. Pramanik, "A comparative analysis of boosting algorithms for chronic liver disease prediction," *Healthcare Analytics*, vol. 5, pp. 1–12(100 313), 2024.

[13] M. H. Mohamed, B. H. Ali, A. I. Taloba, A. O. Aseeri, M. Abd Elaziz, S. El-sappgah, and N. El-Rashidy, "Towards an accurate liver disease prediction based on two-level ensemble stacking model," *IEEE Access*, vol. 12, pp. 180 210–180 237, 2024.

[14] M. P. Behera, A. Sarangi, D. Mishra, and S. K. Sarangi, "A hybrid machine learning algorithm for heart and liver disease prediction using modified particle swarm optimization with support vector machine," *Procedia Computer Science*, vol. 218, pp. 818–827, 2023.

[15] R. S. Babatunde, A. N. Babatunde, B. F. Balogun, A. T. Abdulahi, E. Umar, S. B. Mohammed, and A. A. Oke, "A neuro-fuzzy-based approach to detect liver diseases," *Pakistan Journal of Engineering and Technology*, vol. 7, no. 2, pp. 50–58, 2024.

[16] J. van Ertvelde, A. Verhoeven, A. Maerten, A. Cooreman, B. dos Santos Rodrigues, J. Sanz-Serrano, and M. Vinken, "Optimization of an adverse outcome pathway network on chemical-induced cholestasis using an artificial intelligence-assisted data collection and confidence level quantification approach," *Journal of Biomedical Informatics*, vol. 145, pp. 1–9(104 465), 2023.

[17] K. N. Singh and J. K. Mantri, "A clinical decision support system using rough set theory and machine learning for disease prediction," *Intelligent Medicine*, vol. 4, no. 3, pp. 200–208, 2024.

[18] P. Stephan, T. Stephan, R. Kannan, and A. Abraham, "A hybrid artificial bee colony with whale optimization algorithm for improved breast cancer diagnosis," *Neural Computing and Applications*, vol. 33, no. 20, pp. 13 667–13 691, 2021.

[19] N. Routray, S. K. Rout, B. Sahu, S. K. Panda, and D. Godavarthi, "Ensemble learning with symbiotic organism search optimization algorithm for breast cancer classification and risk identification of other organs on histopathological images," *IEEE Access*, vol. 11, pp. 110 544–110 557, 2023.

[20] R. Kaur and B. S. Khehra, "Modified whale optimization for liver cyst image segmentation," in *2022 5th International Conference on Contemporary Computing and Informatics (IC3I)*. IEEE, 2022, pp. 1329–1335.

[21] E. Ekinci and Z. Garip, "A comparative analysis of meta-heuristic optimization algorithms for feature selection on ml-based classification of heart-related diseases," *The Journal of Supercomputing*, vol. 79, no. 11, pp. 11 797–11 826, 2023.

[22] P. V. Terlapu, U. D. Prasan, T. R. Kumar, V. Bendalam, S. R. Pappu, M. J. Rao, and M. R. Mohitha, "Rice category identification through deep transfer learning features and machine learning classifiers: An intelligent approach," *IAENG International Journal of Computer Science*, vol. 51, no. 7, pp. 765–784, 2024.

[23] A. Alemu, A. Girma, M. Abebe, and R. Srinivasagan, "Genetic algorithm-optimized k-nearest neighbors and support vector machines for breast cancer detection in resource-constrained environments," *IAENG International Journal of Computer Science*, vol. 52, no. 3, pp. 848–861, 2025.

[24] A. D. Praveen, T. P. Vital, D. Jayaram, and L. V. Satyanarayana, "Intelligent liver disease prediction (ildp) system using machine learning models," in *Intelligent Computing in Control and Communication*. Springer Singapore, 2021, pp. 609–625.

[25] P. V. Terlapu, M. Yugandhar, B. Ramesh, B. V. Kumar, and R. Pemula, "Student cognitive learning capability (sclc) prediction system using pca-ann based model," in *2022 International Conference*

on Computing, Communication and Power Technology (IC3P). IEEE, 2022, pp. 11–18.
[26] T. P. Vital, "Intellectual gestational diabetes diagnosis system using mlp-whale optimization algorithm including statistical analysis," *International Journal of Computing and Digital Systems*, vol. 14, no. 1, pp. 485–505, 2023.
[27] P. V. Terlapu, K. K. Raju, G. K. Kumar, G. J. Rao, K. Kavitha, and S. Samreen, "Improved software effort estimation through machine learning: Challenges, applications, and feature importance analysis," *IEEE Access*, vol. 12, pp. 138 663–138 701, 2024.
[28] E. Dritsas and M. Trigka, "Supervised machine learning models for liver disease risk prediction," *Computers*, vol. 12, no. 1, pp. 1–15, 2023.
[29] 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, pp. 1–11, 2023.
[30] A. S. Rahman, F. J. M. Shamrat, Z. Tasnim, J. Roy, and S. A. Hossain, "A comparative study on liver disease prediction using supervised machine learning algorithms," *International Journal of Scientific & Technology Research*, vol. 8, no. 11, pp. 419–422, 2019.
[31] C. Yang, T. Du, Y. Zhao, Y. Qian, J. Tang, X. Li, and L. Ma, "Development and validation of a risk prediction model for nafld: A study based on a physical exdritsasation population," *Diabetes, Metabolic Syndrome and Obesity*, pp. 143–155, 2024.
[32] C. C. Wu, W. C. Yeh, W. D. Hsu, M. M. Islam, P. A. A. Nguyen, T. N. Poly, and Y. C. J. Li, "Prediction of fatty liver disease using machine learning algorithms," *Computer Methods and Programs in Biomedicine*, vol. 170, pp. 23–29, 2019.
[33] B. Zheng, Y. Chen, C. Wang, A. A. Heidari, L. Liu, H. Chen, and G. Liang, "Kolmogorov-arnold networks guided whale optimization algorithm for feature selection in medical datasets," *Journal of Big Data*, vol. 12, no. 1, pp. 1–36, 2025.
[34] M. Shuaib, S. I. M. Abdulhamid, O. S. Adebayo, O. Osho, I. Idris, J. K. Alhassan, and N. Rana, "Whale optimization algorithm-based email spam feature selection method using rotation forest algorithm for classification," *SN Applied Sciences*, vol. 1, pp. 1–17, 2019.
[35] S. Chakraborty, A. K. Saha, S. Sharma, R. Chakraborty, and S. Debnath, "A hybrid whale optimization algorithm for global optimization," *Journal of Ambient Intelligence and Humanized Computing*, vol. 14, no. 1, pp. 431–467, 2023.
[36] T. M. Ghazal, A. U. Rehman, M. Saleem, M. Ahmad, S. Ahmad, and F. Mehmood, "Intelligent model to predict early liver disease using machine learning technique," in *2022 International Conference on Business Analytics for Technology and Security (ICBATS)*. IEEE, 2022, pp. 1–5.
[37] W. M. Shaban, "Early diagnosis of liver disease using improved binary butterfly optimization and machine learning algorithms," *Multimedia Tools and Applications*, vol. 83, no. 10, pp. 30 867–30 895, 2024.

Dr. Panduranga Vital Terlapu received the B.Sc. and M.C.A. degrees in computer science from Andhra University, Andhra Pradesh, India, in 1995 and 1998, respectively, the M.Tech. degree in computer science and engineering from Acharya Nagarjuna University, Andhra Pradesh, and the Ph.D. degree in computer science and engineering from GITAMUniversity, Andhra Pradesh.With an impressive career spanning 24 years in teaching and 18 years in research, he is currently a Professor with the Department of Computer Science and Engineering, Aditya Institute of Technology and Management (AITAM), India. His scholarly contributions are significant, boasting over 81 research papers published in distinguished international journals, including those indexed in SCI and SCOPUS, along with conferences hosted by IEEE, Springer, and Elsevier, among others, and accessible online. He is also actively reviewing submissions for prominent journals affiliated with Springer, Elsevier, and IEEE databases. He is a renowned expert in knowledge dissemination, has published two books, and filed three patents, including one granted by South Africa, showcasing his innovative thinking and significant impact in the field. His research interests include a broad spectrum within the field of computer science, including machine learning, image processing, deep learning, data mining, big data analytics, the IoT and computational intelligence, voice analysis, and software engineering. His dedication and expertise contribute substantially to the advancement of these areas of study. He holds an esteemed membership in professional organizations, such as the Association for Computing Machinery (ACM), and lifetime memberships from the International Computer Science and Engineering Society (ICSES), USA, and Indian Society for Technical Education (ISTE), New Delhi, India. In addition, he maintains affiliations with five other reputable associations.

Killi Bhumika is pursuing his M.Tech in Computer Science and Engineering (CSE) at Aditya Institute of Technology and Management (AITAM), Tekkali, Andhra Pradesh. She completed his B.Tech from Jawaharlal Nehru Technological University (JNTU), Kakinada. She has two research articles and the successfully executed five notable projects. Her projects span diverse domains, including web design, data mining, machine learning (ML), deep learning (DL), and computer networks and security. Her commitment to exploring innovative solutions and utilizing advanced technologies demonstrates his dedication to advancing these fields.