# Multi-Class Prediction of Cirrhosis

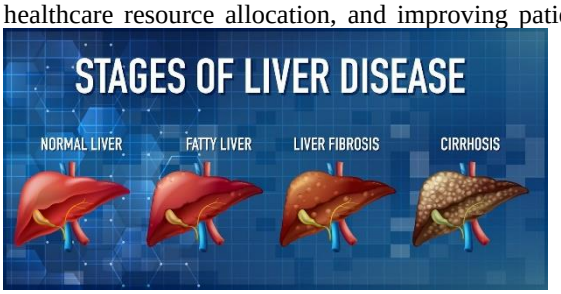
Introduction to Machine Learning
Systems and Biomedical Engineering Department,
Faculty of Engineering - Cairo University

***Abstract*—**Cirrhosis is a progressive liver disease marked by tissue scarring, leading to impaired liver function and life-threatening complications such as liver failure and portal hypertension. Accurate prognosis is critical for optimizing treatment strategies and improving patient outcomes. This study leverages machine learning to predict clinical outcomes for cirrhosis patients using the Cirrhosis Patient Survival Prediction dataset. The proposed model classifies patients into three distinct categories: Death, Censored, and Censored due to Liver Transplantation. By employing supervised learning techniques, this work aims to facilitate early identification of high-risk patients and support personalized clinical decision-making. The inclusion of censored cases, particularly post-transplantation outcomes, enhances the model's real-world applicability by addressing complex survival dynamics. Experimental results demonstrate the model's effectiveness in prognostic stratification, offering potential utility in resource allocation and patient management.

***Keywords— Cirrhosis, Machine Learning, Prognostic Model, Survival Prediction, Liver Transplantation***

## I. Introduction

Cirrhosis is a chronic, progressive liver disease characterized by the replacement of healthy liver tissue with scar tissue, leading to a gradual loss of liver function. This condition is often the result of long-term damage from factors such as chronic hepatitis, alcoholism, or non-alcoholic fatty liver disease. As cirrhosis advances, patients are at risk of life-threatening complications, including liver failure, portal hypertension, and an increased likelihood of developing liver cancer. The accurate prognosis of cirrhosis outcomes is crucial for informing treatment decisions, optimizing healthcare resource allocation, and improving patient care.




This project presents a comprehensive analysis aimed at predicting clinical outcomes for patients with cirrhosis, using data from the Cirrhosis Patient Survival Prediction dataset. The primary objective is to develop a robust machine learning model capable of classifying patients into one of three distinct outcome categories: **Death**, **Censored**, and **Censored due to Liver Transplantation**. By leveraging supervised learning techniques, this model seeks to support the early identification of high-risk patients and assist clinicians in tailoring individualized management strategies. The inclusion of censored cases, particularly those associated with liver transplantation, enhances the depth of the analysis by accounting for real-world treatment pathways and survival complexities.

## II. Related Work:

Recent advancements in machine learning have significantly contributed to the early detection and classification of liver cirrhosis. Several studies have explored multi-class prediction models to classify cirrhosis stages or outcomes, aiming to enhance clinical decision-making and patient management.

**1. Cirrhosis Prediction in Chronic Liver Disease Patients Using Machine Learning**

Athish V P and Sree Nandha S S (2023) presented a study at the 3rd International Conference on Pervasive Computing and Social Networking (ICPCSN), where they developed a multi-class classification model for cirrhosis prediction. The study employed algorithms such as Support Vector Machines (SVM), Gaussian Naïve Bayes, and K-Nearest Neighbors (KNN), achieving an impressive accuracy of 94.48%. The research emphasized the importance of data preprocessing and visualization in improving model performance.

**2. Liver Cirrhosis Stage Prediction Using Machine Learning**

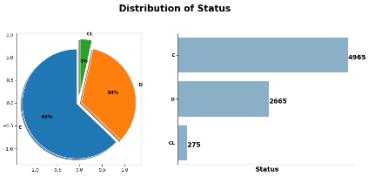
This study, published in a Springer book chapter, focuses on predicting the stage of liver cirrhosis using supervised learning techniques. The authors applied multi-class classification to categorize patients into one of four cirrhosis stages. The work highlights the potential of machine learning in supporting clinical diagnostics and staging of liver diseases.

## III. Dataset and Features

*1) Dataset Overview*

The dataset consists of 17 clinical features used to predict patient survival outcomes related to liver disease, each feature in the dataset provides crucial medical insights into the patient's condition. The target variable is categorized into three classes:

- **Class 0 (D – Death):** The patient did not survive.
- **Class 1 (C – Censored):** The patient was still alive at the end of the study period.
- **Class 2 (CL – Censored due to Liver Transplantation):** The patient did not experience death but was censored due to receiving a liver transplant.




*2) Data Preprocessing*

A. Handling Missing Values

To ensure data consistency, missing values dropped.

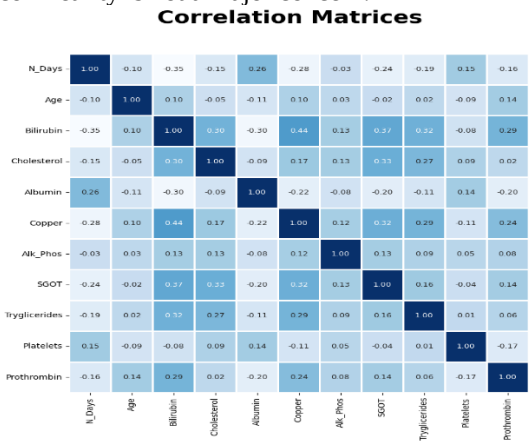
B. Outlier Detection & Removal

To improve model performance, extreme outliers were identified and removed. Observations with feature values greater than six standard deviations from the mean were considered outliers and excluded from the dataset.

- Before outlier removal: 7,905 observations
- After outlier removal: 7,739 observations

C. Feature Correlation

A correlation analysis revealed no significant high correlations between features, suggesting that multicollinearity is not a major concern.




D. Feature Encoding

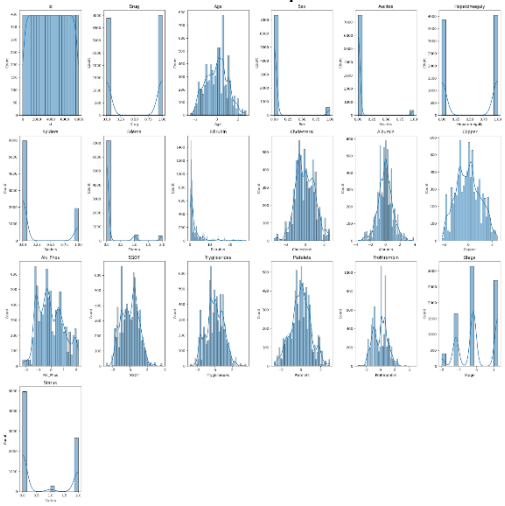
Categorical features were label-encoded to convert them into numerical representations. The following features were encoded:

Drug, Sex, Ascites, Hepatomegaly, Spiders, Edema, Stage

E. Feature Distributions and Normalization

The distribution of features was analyzed to assess skewness and normality:

- Right-skewed features: Bilirubin, Alk_Phos, Triglycerides
  → Consider applying log transformation for normalization.
- Normally distributed features: Albumin, Platelets, Prothrombin
  → No transformation required.



*3) Feature Engineering*

The following features were derived from the original dataset to enhance model interpretability and predictive power:

- ***date_of_diagnosis***:

$$date\_of\_diagnosis = Age - N\_Days$$

Represents the estimated date of diagnosis, calculated by subtracting the number of days since registration from the patient's age.

- ***diseases***:

$$diseases = Ascites + Hepatomegaly + Spiders + Edema$$

A cumulative score indicating the presence of liver-related complications.

- ***Albumin_Level***:A categorical variable based on albumin concentration:

$$\text{Albumin\_Level} = \begin{cases} 1, & \text{if Albumin} < 3.4 \\ 2, & \text{if Albumin} > 5.4 \\ 0, & \text{otherwise} \end{cases}$$

Identifies low, high, and normal albumin levels.

- ***Copper_Risk***:

$$\text{Copper\_Risk} = \begin{cases} 1, & \text{if Copper} > 140 \\ 0, & \text{otherwise} \end{cases}$$

Flag patients are at potential risk due to elevated copper levels.

- ***Platelets_Normal***:

$$\text{Platelets\_Normal} = \begin{cases} 1, & \text{if Platelets} < 150,000 \text{ or } > 450,000 \\ 0, & \text{otherwise} \end{cases}$$

Indicates abnormal platelet count outside the standard clinical range.

## IV. MEHODS

In this project, we employed several machine learning algorithms to classify the given dataset. Below is a brief description of each algorithm used:

**1. Random Forest (RF)**

- **Description**:
  Random Forest is an ensemble learning method that constructs multiple decision trees during training and outputs the mode of the classes (classification) of the individual trees. It reduces overfitting by averaging predictions from multiple trees.
  Random Forest Prediction

$$\hat{y} = \text{mode}\left(\{h_i(x)\}_{i=1}^{N}\right)$$

where:

$h_1(x) =$ prediction of the $i$-th decision tree

$N =$ number of trees (n_estimators)

**2.AdaBoost (Adaptive Boosting)**

2 **Description**:

AdaBoost is an ensemble method that combines weak learners (typically decision trees) sequentially, where each new learner corrects errors made by the previous ones by adjusting weights of misclassified samples.

AdaBoost Ensemble Output

$$F(x) = \sum_{t=1}^{T} \alpha_t h_t(x)$$

where:

$\alpha_t$ = weight of the $t$-th weak learner

$h_t(x)$ = prediction of the $t$-th weak learner

$T$ = n_estimators

### 3. XGBoost (Extreme Gradient Boosting)

- **Description**:

XGBoost is an optimized gradient boosting algorithm that improves efficiency and performance using regularization and parallel computing.

XGBoost Loss Function (Log Loss with Regularization)

$$\mathcal{L} = -\sum_{i=1}^{n} [y_i \log(p_i) + (1 - y_i)\log(1 - p_i)] + \Omega(f_k$$

where:

- $\Omega(f_k)$ = regularization term (controls model complexity)
- $p_i$ = predicted probability for sample $i$

### 4. Stacking Ensemble

- **Description**:

Stacking combines predictions from multiple models (meta-learners) using another model (meta-classifier) to improve generalization.

- **Implementation**:

We used predictions from Random Forest, AdaBoost, and XGBoost as inputs to a logistic regression meta-classifier.

### 5.Weighted Ensemble

- **Description**:
  A weighted average of predictions from base models, where weights are optimized based on validation performance.

## V. Experiments/ Results

### 1. Hyperparameter Optimization

#### 1.1 Individual Model Tuning

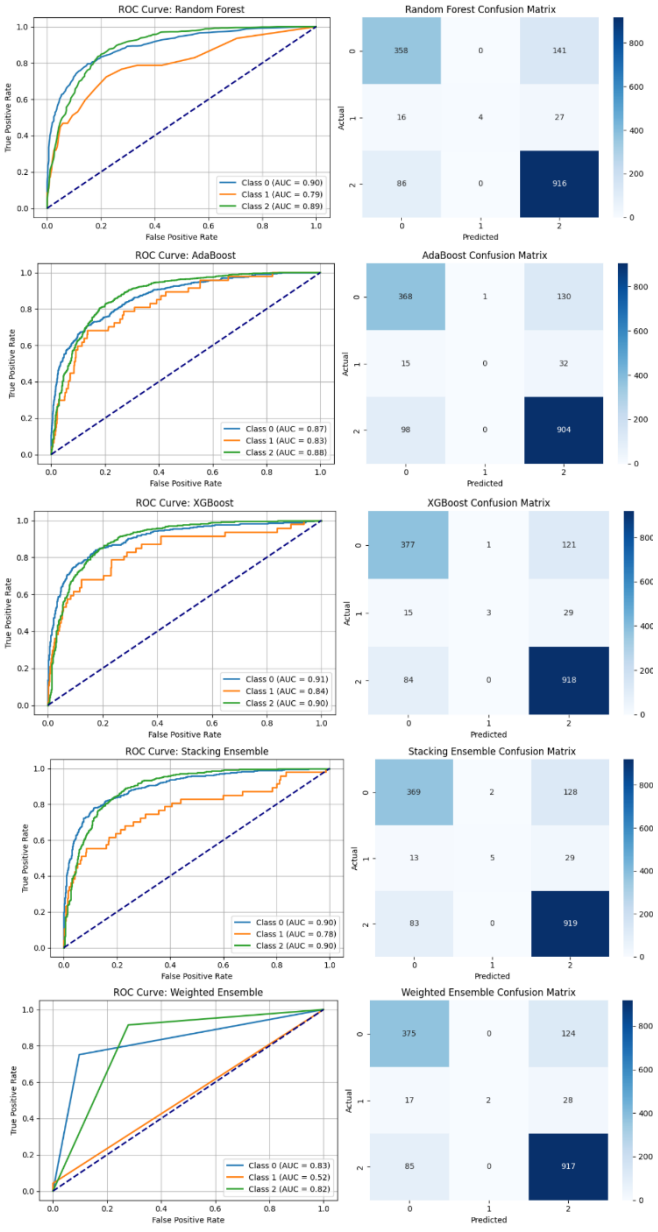
We conducted an extensive grid search with 5-fold cross-validation to optimize each model:

- **Random Forest** performed best with unlimited depth max depth=None and 100 trees, allowing comprehensive learning while preventing overfitting.
- **AdaBoost** achieved optimal results with a learning rate of 1.0 and 100 estimators, balancing convergence speed and accuracy.
- **XGBoost** required precise tuning, with our experiments showing best performance at:
  - Learning rate: 0.1
  - Max depth: 5
  - Number of estimators: 100

#### 1.2 Ensemble Methods

We implemented two advanced ensemble approaches:

1. **Stacking Ensemble**:
   a. Used logistic regression meta-learner
   b. Combined predictions from all base models
   c. Demonstrated superior performance
2. **Weighted Ensemble**:
   a. Assigned weights proportional to model accuracy
   b. Achieved 34% weight for Random Forest
   c. Balanced 33% weights for AdaBoost and XGBoost.



| Model | Accuracy | precision | Recall | F1_Score |
|---|---|---|---|---|
| Random Forest | 0.82558 | 0.82820 | 0.82558 | 0.81390 |
| AdaBoost | 0.82170 | 0.79554 | 0.82170 | 0.80799 |
| XGBoost | 0.83850 | 0.83445 | 0.83850 | 0.82697 |
| Stacking Ensemble | 0.83527 | 0.83032 | 0.83527 | 0.82493 |
| Weighted Ensemble | 0.83591 | 0.83903 | 0.83591 | 0.82340 |

## VI. Conclusion

### Summary of Findings & Key Points

- This study developed a machine learning-based prognostic model to predict clinical outcomes for cirrhosis patients, classifying them into three categories: Death, Censored, and Censored due to Liver Transplantation.
- Among the tested models (Random Forest, AdaBoost, XGBoost, Stacking Ensemble, and Weighted Ensemble), XGBoost emerged as the best-performing individual model, achieving the highest accuracy (0.8385) and recall (0.8385).
- The Weighted Ensemble had the highest precision (0.83903), making it valuable when minimizing false positives is critical.
- Ensemble methods (Stacking & Weighted Ensemble) generally outperformed single models, demonstrating the advantage of combining multiple algorithms for improved predictive power.

### Why Some Algorithms Performed Better

- XGBoost's superior performance can be attributed to its gradient-boosting framework, which effectively handles imbalanced data, optimizes loss functions, and reduces overfitting through regularization.
- Ensemble models (Stacking & Weighted Ensemble) improved accuracy by leveraging the strengths of multiple base models, compensating for individual weaknesses.
- AdaBoost underperformed slightly compared to XGBoost, likely due to its sensitivity to noisy data and outliers, whereas XGBoost's robustness gave it an edge.
- Random Forest performed well due to its ensemble of decision trees but lacked the fine-tuned optimization of boosting methods like XGBoost.

### Clinical & Practical Implications

- The model's ability to stratify patients into risk categories (Death, Censored, Transplant) can aid clinicians in early intervention strategies, improving survival outcomes.
- The inclusion of censored cases (post-transplantation outcomes) enhances real-world applicability, addressing complex survival dynamics in cirrhosis management.
- High recall (XGBoost) ensures fewer false negatives, critical for identifying at-risk patients, while high precision (Weighted Ensemble) reduces unnecessary interventions.

### Final Recommendation

For clinical deployment, XGBoost is recommended if the priority is maximizing detection of high-risk patients (high recall). Alternatively, the Weighted Ensemble is preferable when precision (minimizing false alarms) is more crucial. Future work could explore deep learning approaches and larger multi-center datasets to further enhance predictive performance.

## VII. Future Work

Building on our current results, we plan to enhance the model through improved preprocessing pipelines with smarter outlier handling and feature selection techniques. We will explore different advanced Gradient Boosting implementations, comparing different architectures and optimization approaches to better handle complex data patterns, and develop a production-ready application will focus on creating intuitive and interactive interfaces for end-users, incorporating explainability features, and enabling seamless system integration through robust APIs.

## References


[1] Athish, V. P., & Sree Nandha, S. S. (2023). *Cirrhosis Prediction in Chronic Liver Disease Patients Using Machine Learning Techniques*. In Proceedings of the 3rd International Conference on Pervasive Computing and Social Networking (ICPCSN). Retrieved from https://www.researchgate.net/publication/374467173

[2] Daniel, V. M. (n.d.). *Multi-Class Prediction of Cirrhosis Outcomes* [Computer software]. GitHub. Retrieved May 2025, from https://github.com/16danielvm/Multi-Class-Prediction-of-Cirrhosis-Outcomes

[3] Devarajan, R., & Others. (2023). *Liver Cirrhosis Stage Prediction Using Machine Learning*. In *Lecture Notes in Networks and Systems* (Vol. 857). Springer. DOI: 10.1007/978-3-031-64776-5_42

[4] B. Rieder, *Engines of Order: A Mechanology of Algorithmic Techniques*. Amsterdam, Netherlands: Amsterdam Univ. Press, 2020.

[5] I. Boglaev, "A numerical method for solving nonlinear integro-differential equations of Fredholm type," *J. Comput. Math.*, vol. 34, no. 3, pp. 262–284, May 2016, doi: 10.4208/jcm.1512-m2015-0241.


## Contribution

| | |
|---|---|
| **Somaia Ahmed** | **EDA & Preprocessing** |
| **Pavly Awad** | **Preprocessing** |
| **Abdullah Mahmoud** | **Modeling** |
| **Ahmed Amgad** | **Modeling** |