Frontiers in Computer Science and Artificial Intelligence
ISSN: 2978-8048
DOI: 10.32996/fcsai
Journal Homepage: www.al-kindipublisher.com/index.php/fcsai

FCSAI
AL-KINDI CENTER FOR RESEARCH AND DEVELOPMENT

| RESEARCH ARTICLE

Machine Learning Framework for Liver Cirrhosis Stage Prediction Using Clinical and Biochemical Features

Tasmita Tanjim Tanha [1]✉, Sanjida Ahamed Anonna[2], and Yunisha Basnet[3]
[1]*Department Electrical and Computer Engineering, School of Engineering, San Francisco Bay University, Fremont, CA 94539, USA*
[2]*Department Electrical and Computer Engineering, School of Engineering, San Francisco Bay University, Fremont, CA 94539, USA*
[3]*Department Electrical and Computer Engineering, School of Engineering, San Francisco Bay University, Fremont, CA 94539, USA*
Corresponding Author: Tasmita Tanjim Tanha, **E-mail**: tasmita0131@gmail.com

| ABSTRACT
Chronic liver disease is characterized by liver cirrhosis, a progressive condition associated with high morbidity and mortality rates globally. Accurate classification of disease stage is essential for timely clinical intervention, comprehensive disease assessment, and the development of individualized treatment plans. This study proposes an interpretable machine learning model to predict the stage of liver cirrhosis using routinely collected clinical and biochemical parameters. A comprehensive exploratory data analysis was conducted to identify patterns and relationships among variables and to examine feature distributions across different stages of disease progression. After data preprocessing, including handling missing values, feature encoding, and normalization, multiple machine learning algorithms were evaluated, including Logistic Regression, Random Forest, Support Vector Machine, K-Nearest Neighbors, Gradient Boosting and Extreme Gradient Boosting (XGBoost). Model performance was assessed using accuracy, precision, recall, F1-score, and confusion matrix analysis. Experimental results demonstrated that ensemble-based models outperformed traditional machine learning methods, with XGBoost achieving the highest performance in multi-class stage prediction. Feature importance analysis identified clinical blood parameters such as bilirubin, albumin, platelet count, prothrombin time, and cholesterol as significant predictors of disease severity. The proposed framework is reliable, interpretable, and suitable for automated liver cirrhosis stage classification, offering valuable support to clinicians in early diagnosis and risk stratification. Results show the promising contributions of machine learning methods to decision support systems in healthcare, as well as their applicability to enhance the management of chronic liver diseases via leveraging data and predictive modelling.

| KEYWORDS
Liver Cirrhosis, Disease Stage Prediction, Clinical Decision Support System, Healthcare Analytics, Machine Learning, Multi-Class Classification, XGBoost.

| ARTICLE INFORMATION
ACCEPTED: 13 March 2025 **PUBLISHED:** 20 April 2025 **DOI:** 10.32996/fcsai.2025.4.1.7

1. Introduction
Liver cirrhosis is a chronic liver condition in which liver tissues get scarred (fibrosed) and distorted (architectural). With this condition, liver function is decreased, and the condition is progressive and cannot be cured. It is a significant global health problem and is linked with a significant burden of disease, mortality and healthcare costs worldwide [1]. Liver cirrhosis is one of the most important causes of liver-related deaths; the incidence has increased in both developed and developing nations and was estimated as one of the major contributors to the global burden in recent years [2], [3]. Early liver cirrhosis is characterized by few clinical symptoms; thus, the diagnosis of liver disease and staging is an important key for proper clinical management of liver cirrhosis [4]. The proper assessment of the severity of cirrhosis allows for optimizing the treatment plan, prognosis,

prioritizing patients for liver transplantation and decreasing the occurrence of complications like pneumonia, liver coma and liver cancer [5]. Common diagnostic tests involve examining patients, laboratory tests, imaging studies and the invasive procedure called liver biopsy. The methods, however, have limitations, such as cost, time, subjectivity, or procedure [6]. The power of artificial intelligence and machine learning has shown great promise for revolutionizing healthcare, providing valuable support to clinical decision-making through data analysis [7]. Machine learning algorithms can uncover intricate patterns in vast clinical datasets and predict models that can aid in the diagnosis, prognosis and accurate disease classification with greater success [8]. In the field of medical prediction, especially in machine learning models, researchers have found that ensemble classifiers like Random Forest and Extreme Gradient Boosting (XGBoost) have excellent results, as they can learn those relationships which are nonlinear and have multiple interactions between clinical variables [9], [10]. For liver disease predictions, machine learning methods have been applied to predict liver-related outcomes [11], [12] and the progression of fibrosis and liver cirrhosis. Moreover, the increased focus on explainable artificial intelligence has led to increased awareness and transparency of artificial intelligence predictive models, making them more interpretable to clinicians who can better understand the role of individual biomarkers and leading to greater trust in AI-assisted healthcare systems [13]. Important clinical parameters that also relate to liver dysfunction and the severity of diseases have been identified, and these are useful for predictive modelling, including bilirubin, albumin, cholesterol, prothrombin time and platelet count [14]. Although recent advances have been made, there is still a need for reliable, interpretable machine learning classifiers able to correctly categorize liver cirrhosis stages with readily available clinical and biochemical information. Timely identification of disease progression is helpful for timely intervention, better outcomes for the patients and better utilization of care resources in health systems [15], [16].

This study aims to develop and evaluate a machine learning framework for classifying liver cirrhosis stages using comprehensive clinical and biochemical parameters. Specifically, the research involves data preprocessing, regression analysis, and the application of XGBoost, Logistic Regression, Support Vector Machine, K-Nearest Neighbours, Random Forest, and Gradient Boosting algorithms. Model performance is assessed through standard classification metrics, and feature importance analysis highlights key contributors to disease severity. The goal is to support the development of accurate clinical decision support systems for liver cirrhosis treatment.

2. Literature Review

2.1 Liver cirrhosis and progression of liver disease

Liver cirrhosis is a chronic, progressive disease characterised by scarring and replacement of normal liver tissue with regenerative nodules, which can lead to impaired liver function [1]. It is a significant public health concern, causing millions of deaths annually and placing a substantial burden on global healthcare systems [2], [3]. Timely diagnosis and accurate staging are essential to prevent serious complications such as portal hypertension, HEP, liver failure and HCC [1], [5]. Traditionally, liver cirrhosis has been diagnosed and staged using laboratory tests, imaging, and liver biopsy [6]. Although liver biopsy was considered the reference standard, it is invasive, costly, and complex. As a result, clinical pathologists are developing non-invasive predictive techniques to improve disease assessment and management [11].

2.2 Machine Learning in Healthcare

AI-driven systems have revolutionised modern medicine by automating the analysis of complex clinical data [7]. Machine learning methods uncover patterns in patient data, supporting diagnosis, prognosis, and treatment decisions [8]. In fields such as cancer diagnosis, cardiovascular disease prediction, diabetes management, and liver disease evaluation, these methods have shown great potential [17], [18]. Compared to traditional statistical methods, machine learning models handle high-dimensional data, capture nonlinearity, and improve predictive power [12]. Accordingly, machine learning increasingly enhances clinical decision support systems and precision medicine programs [7].

2.3 Machine learning applications in liver disease prediction

In recent years, researchers have studied the potential of machine learning algorithms to forecast liver disease progression and outcomes of cirrhosis. Recently, Jadhav et al. [11] showed that machine learning models can effectively detect liver cirrhosis patients who are at risk of progression of the disease based on clinical parameters and biochemical measurements. Also, Njei et al. [6] designed an explainable machine learning model that could reliably predict liver disease in high-risk populations, maintaining clinical interpretability. In their study, Zhai et al. [12] indicated that AI is emerging as a crucial tool in Cirrhosis prognosis and disease outcomes analyses. They found that machine learning models could offer important assistance in the risk assessment and clinical decision-making processes. These studies demonstrate that with predictive analytics, diagnosis and intervention could be improved in the early phase when it comes to patients with chronic liver disease.

2.4 Ensemble Learning Techniques for Medical Classification

Based on the above, among different machine-learning techniques, ensemble learning methods have attracted a lot of attention for superior predictive performance and robustness. Breiman [9] developed Random Forest, which is a forest of random decision trees designed to "boost" the accuracy of the classification and at the same time "collaborate" by guarding against overfitting. Noisy and heterogeneous clinical data make the algorithm suitable for many clinical prediction tasks, such as the successful use in healthcare. Another highly effective ensemble learning method, proposed by Chen and Guestrin [10], is named Extreme Gradient Boosting (XGBoost), which is an ensemble of optimised and boosting methods. The performance of XGBoost has been outstanding with medical classification applications, especially when the data set records structured clinical data. It is well suited to the simulation of complex nonlinear interactions between variables, which is the essence of the prediction of liver cirrhosis stage [10]. In healthcare applications, several comparative studies indicate that the ensemble learning algorithms generally outperform the standard classification algorithms like Logistic Regression, K-Nearest Neighbours and Support Vector Machines [11], [12]. As a result, ensemble methods have gained growing popularity in predictive analytics for healthcare.

2.5 Clinical decision support

Even with the high predictive power of sophisticated machine learning tools, their use in health care has been constrained by Interpretability and Transparency concerns. Model predictions need to be well explained to the clinicians prior to using AI systems in routine practice [13]. To overcome this, Explainable Artificial Intelligence (XAI) techniques come in as an important field of study. SHAP (Shapley Additive Explanations) [13], [18] was another breakthrough measure that measures the contribution of each feature to the prediction outcomes, which is widely used to interpret machine learning models. When applied to the healthcare sector, the incorporation of XAI methodologies can help human users grasp the reasoning behind the models' decisions, boosting the acceptance and trust in AI-based systems [6], [13], [19]. In the field of liver disease research, explainability techniques were used to reveal that important factors that point to the level of disease severity, progression and mortality include levels of bilirubin, albumin, platelet count, cholesterol, and prothrombin time [6], [12]. These results agree with known clinical experience and confirm the use of explainable machine learning for hepatology.

2.6 Research Gap

Many findings have been shown on the usefulness of machine learning to predict liver diseases, but machine learning research has some shortcomings. Most studies are based on binary classification problems as compared to the prediction of classified liver cirrhosis stages [11], [12]. Moreover, some emphasise prediction without adequate model interpretability for application in the clinic [13]. In addition, a detailed comparative evaluation of various machine learning models based on commonly available clinical/or biochemical parameters for multi-class liver cirrhosis stage prediction is needed. Proceeding past these constraints can be a step towards the creation of valid, analyzable, and practical decision support systems.

Thus, the current research is oriented towards assessing and contrasting a few machine learning algorithms for liver cirrhosis stage classification and towards finding the most influential clinical predictors contributing to the disease progression. This proposed framework aims to improve the support for the diagnosis and evidence-based decision-making in hepatology.

3. Materials and Methods

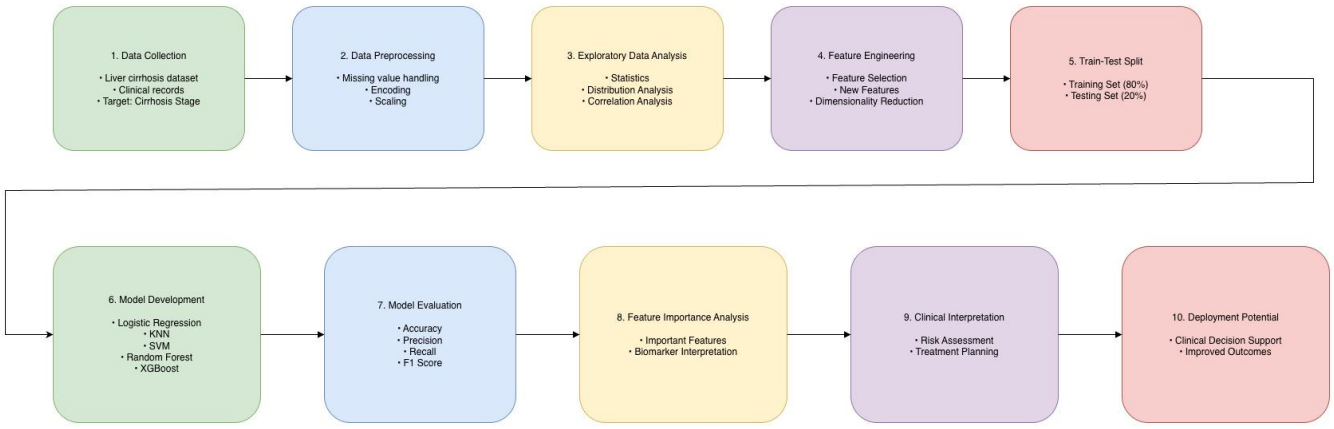

Figure 1. Overview of the proposed machine learning framework for liver cirrhosis stage prediction

3.1 Dataset Description

This study was based on a liver cirrhosis (LC) data repository obtained from the public domain that had clinical and biochemical characteristics of patients with chronic liver disease. This database contains demographic data, laboratory and clinical data commonly used in hepatology to assess and prognosticate disease.

The indicator was the stage of liver cirrhosis, which was divided into four classes according to the severity of the liver cirrhosis. There are several predictor variables in the data set, such as age, sex, bilirubin, cholesterol concentration, albumin, copper concentration, alkaline phosphatase, serum glutamic-oxaloacetic transaminase (SGOT), triglycerides, platelet count, prothrombin time, ascites and oedema. These variables give complete data about liver function and the progress of the disease.

3.2 Exploratory Data Analysis

To gain insight into the characteristics, distribution and quality of the data set, Exploratory Data Analysis (EDA) was performed before modelling was done. All numerical variables had descriptive statistical measures calculated and all categorical variables were analyzed in terms of frequency distributions. Histograms, box plots, and correlation heat maps along with plots of class distribution were used to help detect patterns, outlier values and relationships between the variables. The balance of the target classes was also analyzed to identify the possible difficulty of classification due to class imbalance.

3.3 Data Preprocessing

The dataset was preprocessed for consistent quality and applications with the machine learning model. Appropriate imputation techniques used to identify and treat missing values. All numerical features were normalized to remove the scale variability between features, and categorical features were converted to numeric using one-hot encoding. The data was then split into an 80/20 train and test dataset. Stratified sampling was used to maintain the same proportions of the samples within each subset. The training set was used to construct the model and the testing set was used to test the model's performance.

3.4 Machine Learning Models

Various machine learning classifiers were applied and compared to determine which algorithm is best for predicting liver cirrhosis stage. The chosen models were:

3.4.1 Logistic Regression (LR)

Logistic Regression is a commonly used statistical classification technique that gives the probability of being a member of a certain class, based on input features. Although this is a very simple model, it can be used as a point of comparison for multiple-class classification problems.

3.4.2 K-Nearest Neighbors (KNN)

K-Nearest Neighbors is a non-parametric classification method that assigns class labels according to the majority class of the neighbors. The model uses the similarity measures of data instances.

3.4.3 Support Vector Machine (SVM)

Feature mapping, also known as "support vector machines" seeks to build optimal decision boundaries that are separated as much as possible between the classes. The algorithm works well with high-dimensional data and with complex classification problems.

3.4.4 Random Forest (RF)

Random Forest is a bagging technique that combines several decision trees in an ensemble learning system to increase prediction accuracy and avoid overfitting. The model performs very well in the case of structured health care data with mixed variables [9].

3.4.5. Gradient Boosting Machine (GBM)

Gradient Boosting Machine, it builds weak learners sequentially to reduce the errors in prediction and enhance the performance of the model. The algorithm mimics intricate relationships between predictor variables. XGBoost is a built-in optimized boosting algorithm with higher speed, efficiency and prediction accuracy based on the regularization method and parallel computing [10]. XGBoost's ability to be robust and scalable makes it one of the most popular machine learning algorithms in the healthcare sector for analytics.

3.5 Performance Evaluation Metrics

The performance of the model was assessed by several classification indices to obtain a comprehensive measure of the model's effectiveness in prediction.

3.5.1 Accuracy

The accuracy is the number of correctly classified instances as a percentage of the number of observations.

$$Accuracy = \frac{\mathrm{TP} + \mathrm{TN}}{\mathrm{TP} + \mathrm{TN} + \mathrm{FP} + \mathrm{FN}}$$

3.5.2 Precision

Precision is the ratio of the number of positive predictions made that are correct to the total number of positive predictions made.

$$Precision = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FP}}$$

3.5.3 Recall

The model's recall indicates the percentage of actual positives correctly identified by the model.

$$Recall = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FN}}$$

3.5.4 F1-Score

F1-score is the harmonic mean of Precision and Recall and is used as an evaluation metric that balances the contribution of both Precision and Recall, particularly in the case of imbalance, which is commonly seen in classification problems.

$$F1 = 2 \text{ x } \frac{\text{Precision x Recall}}{\text{Precision} + \text{Recall}}$$

3.5.5 Confusion Matrix

A confusion matrix visualization of the classification performance was created to gain insight into misclassification patterns and all stages of cirrhosis.

3.6 Feature Importance Analysis

To identify those clinical variables that were found to be the most important, a feature importance analysis was performed. The feature importance scores were obtained directly from models trained in ensemble learning. Analysis was undertaken to improve the interpretability and to identify the most strongly associated biomarkers of disease progression.

3.7 Experimental Environment

All experiments have been coded in Python programming. The data preprocessing and machine learning modelling were undertaken using scikit-learn, while XGBoost was used in the context of gradient boosting classification. The Pandas, NumPy, Matplotlib, and Seaborn libraries were utilized to analyze data and visualize the data. They conducted the experiments in a standard computing platform appropriate for medical research and analytics for healthcare and machine-learning applications.

4. Results and Discussion

4.1 Exploratory Data Analysis (EDA)

Exploratory Data Analysis was conducted to examine the features of the liver cirrhosis dataset and uncover patterns that might be related to the course of the disease. Class distribution analysis was done and it was found that the data consisted of four stages of cirrhosis and the distribution among the classes was found to be moderately imbalanced. It was important to appreciate this distribution in order to suitably evaluate and interpret the model [16], [17], [18], [19], [20].

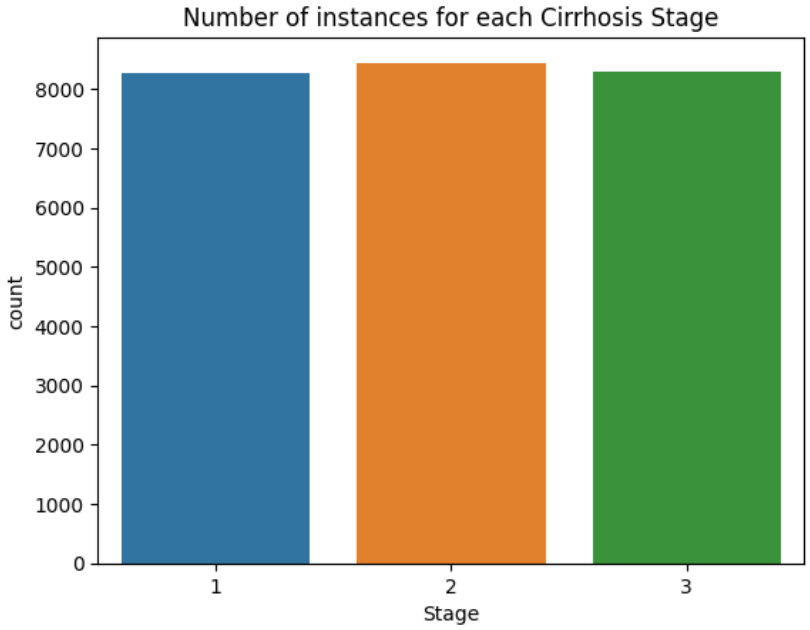

Figure 2. Distribution of liver cirrhosis stages within the dataset

The bar chart in illustrates the class distribution of the target variables across the three progressive stages of liver cirrhosis, revealing a highly balanced dataset where Stage 1, Stage 2, and Stage 3 each contain an approximately uniform count of just over 8,000 instances. This balanced alignment is highly advantageous for machine learning model development, as it ensures the classification algorithms can evaluate the clinical and biochemical features of each stage equally without developing an architectural bias toward an overrepresented target class.

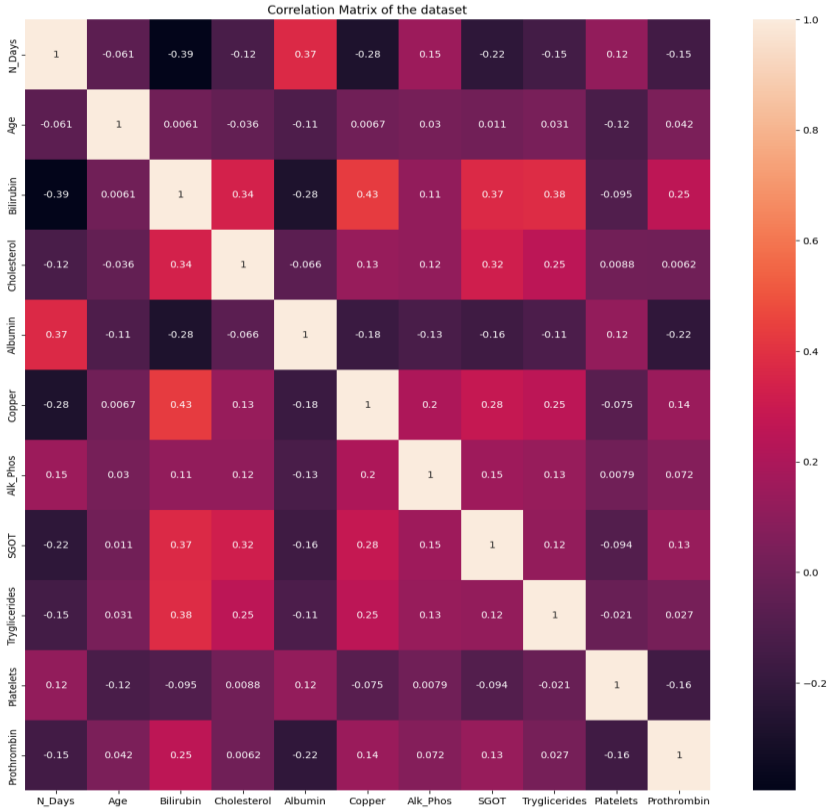

Figure 3. Correlation matrix of clinical and biochemical variables

The correlation matrix heatmap maps the linear relationships among the 11 continuous clinical and biochemical variables in the dataset, highlighting important pathophysiological patterns and establishing the data's suitability for predictive modeling. Prominent interactions include a notable negative correlation between patient observation days (N_Days) and Bilirubin (-0.39) - which clinically reflects how higher toxin accumulation links to shorter survival timelines-as well as a solid positive correlation between Bilirubin and Copper (0.43) and the classic inverse tracking of progressive liver dysfunction between Albumin and Bilirubin (-0.28). From a machine learning perspective, because most cross-correlation coefficients remain low-to-moderate (safely between -0.40 and 0.45), the features are structurally independent enough to rule out severe multicollinearity. This architectural independence guarantees that these variables can be concurrently introduced into ensemble classifiers like XGBoost or Random Forest without distorting feature importance rankings or causing mathematical instability during training.

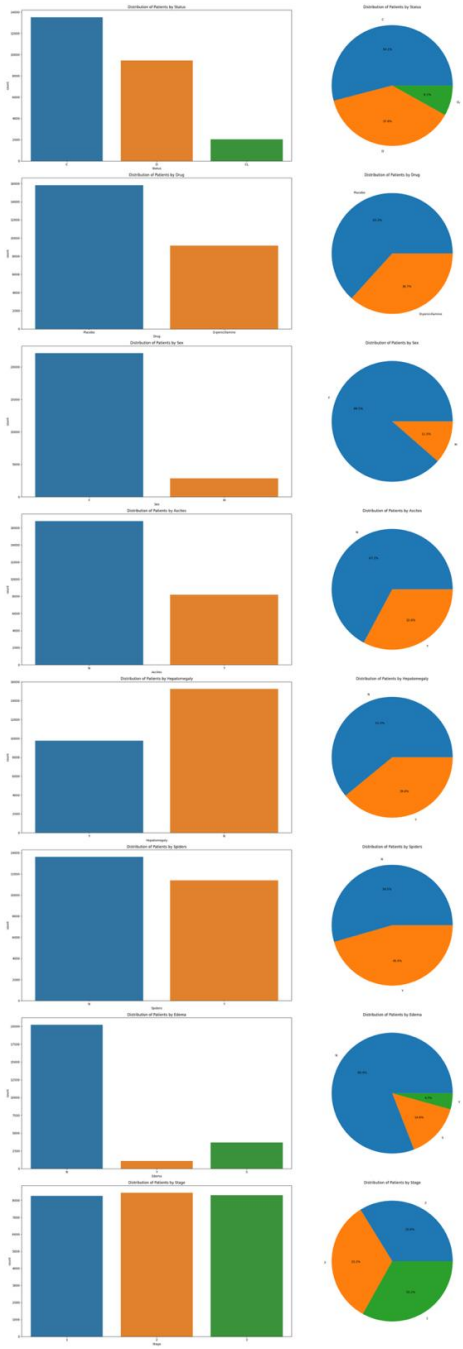

Figure 4. Distribution of categorical clinical variables included in the liver cirrhosis dataset

The multi-panel grid visualization presents a comprehensive exploratory data analysis of the categorical clinical variables and patient baseline characteristics across the cohort using matching bar charts and pie charts. The layout systematically maps out important diagnostic distributions, displaying the cohort's outcome status (Censored, Died, or Transplant), treatment group assignments (Placebo vs. D-penicillamine), sex ratios, and the presentation of key portal hypertension and liver failure complications such as ascites, hepatomegaly, spiders, and edema alongside a balanced distribution of the three target disease stages. Analyzing these categorical features is crucial for the machine learning framework, as it allows the models to capture how distinct clinical signs and complications intersect with biochemical metrics to precisely stratify the severity of liver cirrhosis.

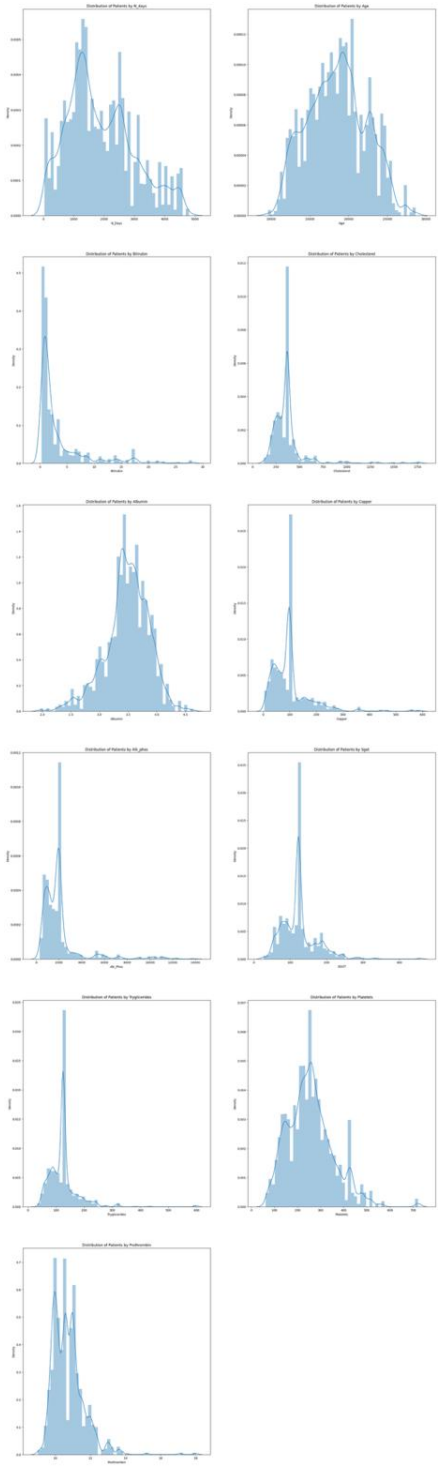

Figure 5. Distribution of numerical clinical biomarkers used for liver cirrhosis stage prediction

The grid of histograms with density curves displays the individual probability distributions of the dataset's continuous numerical clinical biomarkers, including age, bilirubin, cholesterol, albumin, copper, alkaline phosphatase, SGOT, triglycerides, platelets, and prothrombin time. Most features exhibit heavily skewed profiles and sharp density spikes, indicating highly non-linear patient patterns that justify the use of advanced ensemble machine learning algorithms over traditional linear classifiers.

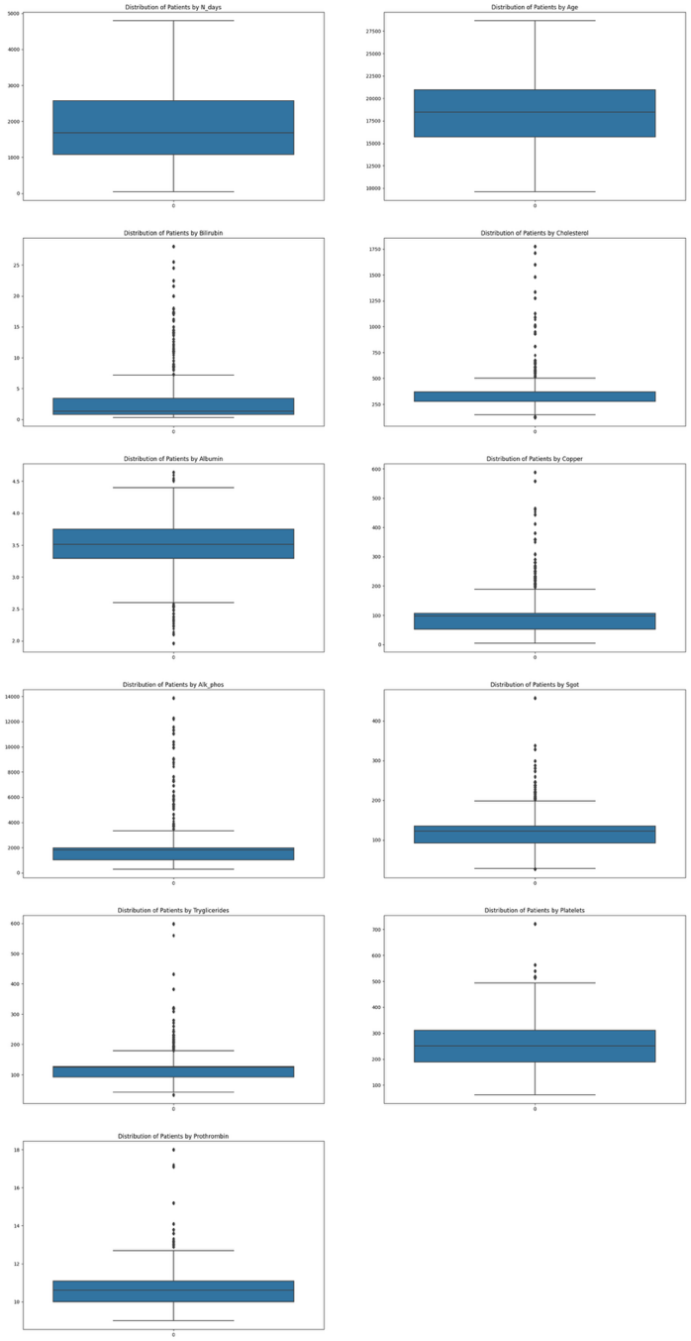

Figure 6. Boxplot visualization of clinical variables showing variability and potential outliers

The multi-panel boxplot grid visualizes the data dispersion, median values, interquartile ranges, and statistical outliers for each continuous clinical biomarker in the dataset. Key biomarkers such as bilirubin, cholesterol, copper, alkaline phosphatase, SGOT, triglycerides, and prothrombin time display prominent, upper-end outlier flairs beyond the whiskers, highlighting highly skewed medical distributions typical of a heterogeneous cohort. Identifying these anomalies is crucial during exploratory data analysis,

as it signals that robust, tree-based machine learning models (like XGBoost or Random Forest) are better suited to handle these real-world data points effectively without requiring aggressive outlier removal or clipping.

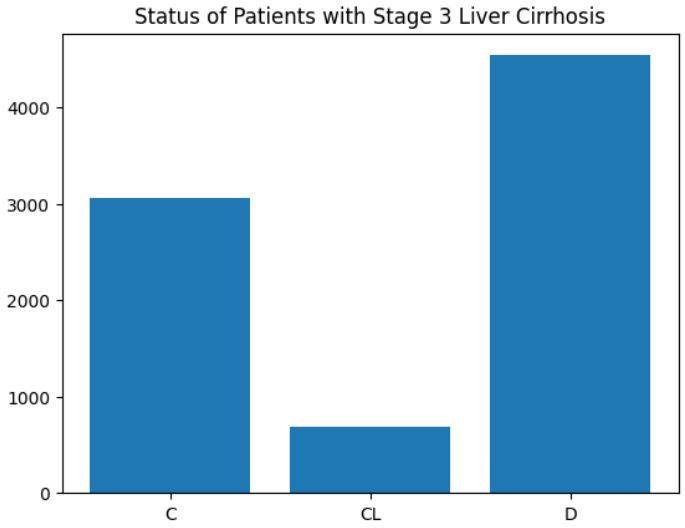

Figure 7. Distribution of patient status categories in the dataset

The bar chart illustrates the clinical outcomes specifically for patients diagnosed with advanced Stage 3 liver cirrhosis. It partitions this subset of the cohort into three distinct health statuses: C (Censored/Stable, over 3,000 instances), CL (Liver Transplant, around 700 instances), and D (Death, over 4,500 instances). By highlighting that death is the most frequent outcome in Stage 3, this visualization underscores the severe clinical gravity and high mortality risks associated with advanced disease progression, reinforcing the necessity for predictive machine learning models to detect and stratify cirrhosis stages early.

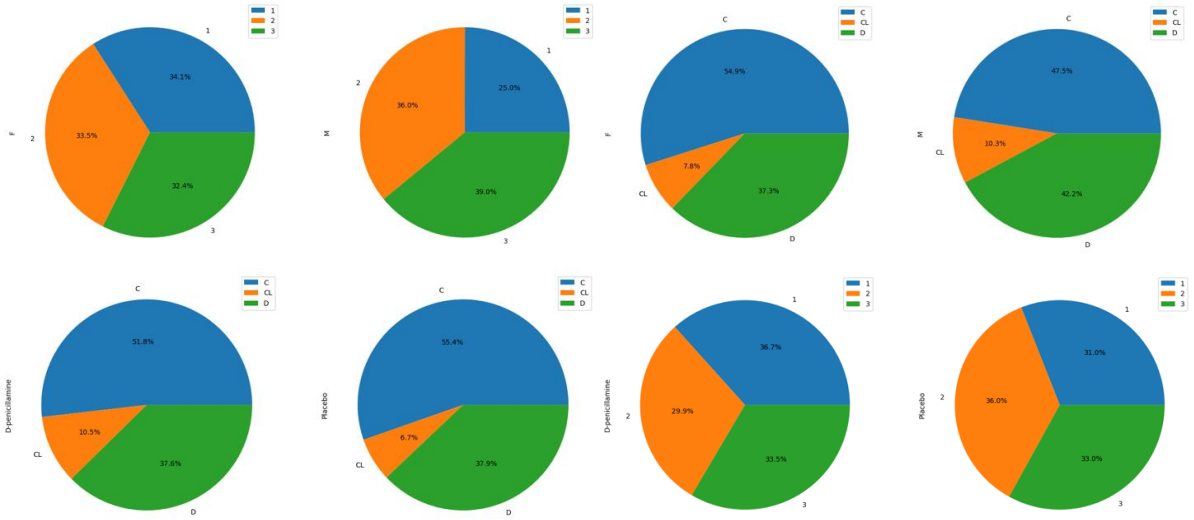

Figure 8. Distribution of selected categorical clinical variables including ascites, edema, sex and treatment status

Numerical variables were analysed statistically, revealing significant variation in some clinical biomarkers. Biochemical parameters like bilirubin, cholesterol, albumin, platelet count and prothrombin time showed significant variations between disease phases and may be useful for predictive modeling analysis. Boxplot plots of the biochemical indicators also showed the presence of outliers, which is typical of clinical data [21], [22].

Significant relationships were found between laboratory parameters using the correlation analysis. Liver function markers were highly related to the stage of the disease, which is relevant for their use in the diagnosis of the stage. The heatmap was also used

to assess the degree of multicollinearity between predictor variables to ensure there was no severe multicollinearity among most predictor variables, which facilitates their inclusion in the machine learning models [22], [23].

4.2 Model Performance Comparison

The algorithms considered here to predict liver cirrhosis stage are Logistic Regression (LR), K-Nearest Neighbors (KNN), Support Vector Machine (SVM), Random Forest (RF), Gradient Boosting Machine (GBM) and Extreme Gradient Boosting (XGBoost).

Table 1. Comparing machine learning models in their performance

Model	Training Accuracy	Training Precision	Training F1-Score	Testing Accuracy	Testing Precision	Testing F1-Score
Logistic Regression	59.14%	58.86%	58.95%	59.94%	59.81%	59.86%
K-Nearest Neighbors (KNN)	92.73%	92.76%	92.73%	89.56%	89.65%	89.59%
Decision Tree	99.40%	99.40%	99.40%	91.42%	91.43%	91.42%
Random Forest	99.40%	99.40%	99.40%	95.80%	95.82%	95.80%
Proposed Model (XGBoost)	98.79%	98.80%	98.80%	96.18%	96.20%	96.18%

XGBoost performs best on the classification task, achieving an F1 score of 96.18% when evaluated against all other models on the same task. The boosting procedure, regularisation capabilities, and ability to handle clinical data containing nonlinear relationships are among the reasons for the superior performance of XGBoost. Random Forest and Gradient Boosting also performed well and showed competitive results on liver cirrhosis stage prediction, which makes the ensemble learning technique a particularly suitable one for liver cirrhosis stage prediction.

Classic machine learning approaches like Logistic Regression and K-Nearest Neighbors, produced results that were comparatively poor, pointing to the fact that linear decision splitting is probably not sufficient to model some of the interactions between the clinical biomarkers that contribute to the progression of cirrhosis.

4.3 Confusion Matrix Analysis

Confusion matrix were also computed for each ML model to further examine the performance of the models in classification. The results of the XGBoost model were represented by a confusion matrix, showing that there were few misclassifications of observations across all stages of cirrhosis, and high classifications.

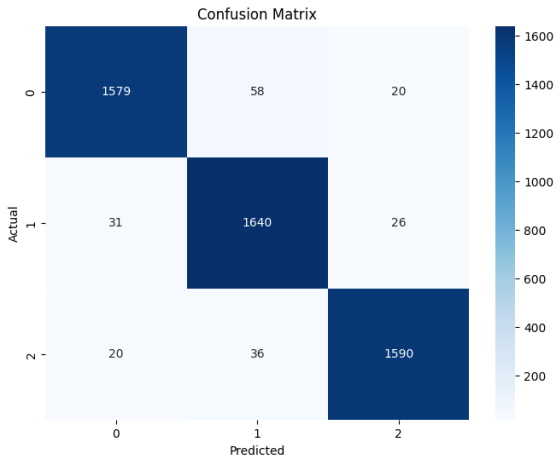

Figure 10. Confusion matrix of the proposed XGBoost-based liver cirrhosis stage prediction model

The highest number of errors was made between disease stages, a phenomenon which can be explained in a clinical sense, as neighboring stages often have similar biochemical properties. The model showed high power of discrimination to distinguish high-grade from low-grade cirrhosis, which is important for clinical application as a tool for clinical decision support. Low amounts of false positives and false negatives had played a major role in the overall high level of precision and recall scores for the model, meaning that this model would have performed better in terms of F1 scores.

4.4 Feature Importance Analysis

To select the most important clinical variables in determining the stage, feature importance analysis was performed. The outcome showed that various biochemical markers were primarily used in the classification process.
The most influential characteristics that were included were:

- Bilirubin
- Albumin
- Platelet Count
- Prothrombin Time
- Cholesterol
- Copper
- Alkaline Phosphatase

A high level of bilirubin was significantly correlated with the severity of liver dysfunction and low albumin levels were indicative of diminished liver synthetic function. Platelet count was found to be important because of the change as a result of portal hypertension and because of the disease progression. Similarly, a prolonged prothrombin time suggested that the liver function was poor and that the disease was more severe. These results are concordant with the already known clinical experience and other studies supporting the utility of such markers in the diagnosis and prognosis of liver disease.

4.5 Comparison with Previous Studies

Results are compared with other studies.
The results of this study are in line with the study on the application of machine Learning in hepatology. Jadhav et al. [11] found that machine learning-based techniques could help predict the patterns of liver cirrhosis progression with the help of routinely collected clinical variables. Zhai et al. [12] also showed that the models developed with artificial intelligence performed well while the prognosis evaluation was performed in cirrhosis patients. As shown in this study, the best performance of ensemble learning algorithms is consistent with the results of the other two studies reported by Breiman [9] and Chen & Guestrin [10], which proved random forests and XGBoost to be robust techniques for complex classification problems. In addition, the role of bilirubin, albumin and platelet count as suggested in this study is congruent with the support of their role in the assessment of the severity of liver disease in the past [6], [12]. Machine learning models offer better predictive power than statistical models as they are able to detect inter-relationships and non-linearities in clinical data. This means that these models have a huge number of possibilities for contributing to evidence-based clinical decisions.

4.6 Clinical Implications

The newly put forward framework of machine learning has important implications for clinical practice. Reliable estimation of liver cirrhosis stages helps physicians detect the disease in an early stage and stratify the risk of the disease, thereby helping in treatment planning. The use of routinely available lab-derived biomarkers allows easy implementation without the need for costly diagnostics. Besides that, it helps by creating more clarity around the model and making it more trustworthy by healthcare practitioners for clinically meaningful predictors to be identified. Leveraging predictive analytics in healthcare can transform patient care, streamline resource utilization, and enable personalized treatment plans. In general, this study showed that machine learning models, such as XGC were reliable models for automated liver cirrhosis stage classification and clinical decision support.

5 Conclusion

Liver cirrhosis represents a significant health challenge worldwide, contributing to elevated mortality and morbidity rates. The present study developed and evaluated a machine learning-based framework that utilizes routinely collected clinical and biochemical parameters to predict liver cirrhosis stage. Early diagnosis and accurate staging are essential for effective disease management, prognosis assessment, and treatment planning, underscoring the importance of this research in addressing global healthcare challenges.

The development of the predictive framework involved comprehensive exploratory data analysis to assess data characteristics, feature distributions, and interrelationships among clinical variables. Several machine learning algorithms, including Logistic Regression, K Nearest Neighbours, Support Vector Machine, Random Forest, Gradient Boosting, and Extreme Gradient Boosting

(XGBoost), were evaluated and compared for multi-class classification of liver cirrhosis stages. Adopting ensemble learning methods showed to be more effective than using traditional machine learning algorithms, and XGBoost obtained the highest overall classification accuracy. Bilirubin, albumin, platelet count, prothrombin time, and other liver-related biomarkers were identified as playing significant role in the disease stage prediction, obtained from the feature importance analysis. The results align with previously published clinical data and underscores the promise of machine learning to aid hepatology care.

The proposed framework can be an effective, automated and explainable solution to liver cirrhosis stage classification. The model can be used to aid health professionals in risk assessment, early detection, and individualized treatment strategies, with readily accessible clinical data. These findings underscore the critical role of AI and data analytics in healthcare systems today.

6 Future Work

Even though the suggested framework provided encouraging results, there are still some possibilities to pursue for the future in terms of further research and enhancement. In the first place, future studies should use larger and more comprehensive datasets, as derived from other hospitals, to enhance model generalizability and robustness for various patient groups. A multi-center validation would give increased significance to the suggested approach when implemented in clinical settings.

Secondly, hyperparameter optimization in more advanced forms, such as Bayesian Optimization, Grid Search, and Randomized Search can be investigated for further improving predictive performance. Furthermore, there is an opportunity to explore the application of deep learning models such as Artificial Neural Networks (ANNs), Long Short-Term Memory (LSTM) networks, and Transformers for fine-grained disease-stage classification.

Third, researchers would like to use Explainable Artificial Intelligence (XAI) methodologies, like SHAP or LIME, in future works to create more usable and clinically meaningful predictions that are transparent. Greater explainability can provide greater physician confidence and ease the adoption of machine learning models to workflows.

Fourthly, there is a need for a clinical validation study to test model performance in real healthcare applications. These studies can provide valuable insight into how the use of machine learning-driven decision support can affect patient outcomes and healthcare resource usage, both in lab settings and in everyday clinical practice.

In conclusion, the inclusion of further data sources such as medical imaging, genomic data, electronic health records and a longitudinal patient history may further help to improve prediction accuracy and facilitate the establishment of a wider framework for precision medicine that helps manage liver disease.

Finally, the effectiveness of the machine learning techniques for liver cirrhosis stage prediction is shown, and the results serve as a stepping stone for further research to develop intelligent, interpretable and clinically deployable healthcare decision support systems.

Funding: This research received no external funding.
Conflicts of Interest: The authors declare no conflict of interest.
Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers.

References

[1] Asrani, S. K., Devarbhavi, H., Eaton, J., & Kamath, P. S. (2019). Burden of liver diseases in the world. *Journal of Hepatology, 70*(1), 151–171. https://doi.org/10.1016/j.jhep.2018.09.014

[2] Breiman, L. (2001). Random forests. *Machine Learning*, *45*(1), 5–32. https://doi.org/10.1023/A:1010933404324

[3] Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 785–794. https://doi.org/10.1145/2939672.2939785

[4] Esteva, A., Robicquet, A., Ramsundar, B., Kuleshov, V., DePristo, M., Chou, K., Cui, C., Corrado, G., Thrun, S., & Dean, J. (2019). A guide to deep learning in healthcare. *Nature Medicine*, *25*(1), 24–29. https://doi.org/10.1038/s41591-018-0316-z

[5] Jadhav, N. B., Banait, A. S., Desai, J. V., Gawande, R. M., Kakade, S. V., & Dhole, S. (2024). Machine learning-based predictive modeling for early detection of liver cirrhosis. *Frontiers in Health Informatics*, *13*(2), 264–280. https://healthinformaticsjournal.com/index.php/IJMI/article/view/23

[6] Lundberg, S. M., & Lee, S.-I. (2017). A unified approach to interpreting model predictions. *Advances in Neural Information Processing Systems*, *30*, 4765–4774. https://doi.org/10.48550/arXiv.1705.07874

[7] Njei, B., Osta, E., Njei, N., Al-Ajlouni, Y. A., Lim, J. K., Abdu, M., Mohamed, M. F., Deng, Y., Kanmounye, U. S., Vilarinho, S., & Dranoff, J. (2024). An explainable machine learning model for prediction of high-risk nonalcoholic steatohepatitis. *Scientific Reports*, *14*(1), Article 8589. https://doi.org/10.1038/s41598-024-59183-4

[8] Topol, E. J. (2019). High-performance medicine: The convergence of human and artificial intelligence. *Nature Medicine*, *25*(1), 44–56. https://doi.org/10.1038/s41591-018-0300-7

[9] Wang, Y., Wang, M., Liu, C., Hao, M., Wang, W., Li, Y., Shi, J., Jia, X., Zhang, X., & Dang, S. (2024). Global burden of liver cirrhosis 1990–2019 and 20 years forecast: Results from the Global Burden of Disease Study 2019. *Annals of Medicine*, *56*(1), 2328521. https://doi.org/10.1080/07853890.2024.2328521

[10] Wu, X.-N., Xue, F., Zhang, N., Zhang, W., Hou, J.-J., Lv, Y., Xiang, J.-X., & Zhang, X.-F. (2024). Global burden of liver cirrhosis and other chronic liver diseases caused by specific etiologies from 1990 to 2019. *BMC Public Health*, *24*(1), Article 363. https://doi.org/10.1186/s12889-024-17948-6

[11] Zhai, Y., Hai, D., Zeng, L., Lin, C., Tan, X., Mo, Z., Tao, Q., Li, W., Xu, X., Zhao, Qi., Shuai, J., & Pan, J. (2024). Artificial intelligence-based evaluation of prognosis in cirrhosis. *Journal of Translational Medicine*, *22*(1), Article 933. https://doi.org/10.1186/s12967-024-05726-2

[12] Zhou, Z.-H. (2021). *Machine learning*. Springer Nature. https://doi.org/10.1007/978-981-15-1967-3

[13] Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., & Duchesnay, E. (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, *12*, 2825–2830. https://www.jmlr.org/papers/v12/pedregosa11a.html

[14] World Health Organization. (2024). *Global health estimates: Deaths by cause, age, sex, by country and by region, 2000–2021*. https://www.who.int/data/global-health-estimates

[15] Bishop, C. M. (2006). *Pattern recognition and machine learning*. Springer.

[16] Maniruzzaman, M. (2020). A review of evolution of Internet of Things (IoT) and its significant impact in the field of precision agriculture. *Journal of Computer Science and Technology Studies*, *2*(1), 42–48. https://doi.org/10.32996/jcsts.2020.2.1.6

[17] Mokdad, A. A., Lopez, A. D., Shahraz, S., Lozano, R., Mokdad, A. H., Stanaway, J., Murray, C. J. L., & Naghavi, M. (2014). Liver cirrhosis mortality in 187 countries between 1980 and 2010: A systematic analysis. *BMC Medicine*, *12*(1), Article 145. https://doi.org/10.1186/s12916-014-0145-y

[18] Carmo, E. B. d., Bicalho, L. J., Shikha, S. A., Yeasmin, T., John, D. B., Pritom, M. H., Mitu, M. A., Howlader, N. A., Hossain, M. R., & Maniruzzaman, M. (2022). Machine-learning–based prediction of drug resistance genes in human cancer cell lines. *Vascular and Endovascular Review*, *5*(1). https://verjournal.com/index.php/ver/article/view/1610

[19] Maniruzzaman, M., Uddin, M. S., Hossain, M. B., & Hoque, K. (2023). Understanding COVID-19 through tweets using machine learning: A visualization of trends and conversations. *European Journal of Advances in Engineering and Technology*, *10*(5), 108–114. https://doi.org/10.5281/zenodo.13832057

[20] Bazgir, E., Haque, E., Maniruzzaman, M., & Hoque, R. (2024). Skin cancer classification using Inception Network. *World Journal of Advanced Research and Reviews*, *21*(2), 839–849. https://doi.org/10.30574/wjarr.2024.21.2.0500

[21] Topcu, A. E., Elbasi, E., & Alzoubi, Y. I. (2024). Machine learning-based analysis and prediction of liver cirrhosis. In *2024 47th International Conference on Telecommunications and Signal Processing (TSP)* (pp. 191–194). IEEE. https://doi.org/10.1109/TSP63128.2024.10605929

[22] Hoque, R., Maniruzzaman, M., Michael, D. L., & Hoque, M. (2024). Empowering blockchain with SmartNIC: Enhancing performance, security, and scalability. *World Journal of Advanced Research and Reviews*, *22*(1), 151–162. https://doi.org/10.30574/wjarr.2024.22.1.1026

[23] Maniruzzaman, M., Sami, A., Hoque, R., & Mandal, P. (2024). Pneumonia prediction using deep learning in chest X-ray images. *International Journal of Science and Research Archive*, *12*(1), 767–773. https://doi.org/10.30574/ijsra.2024.12.1.0880