Predictive Modeling and Survival Analysis of Primary Biliary Cirrhosis: A Comparative Study of Machine Learning Approaches

El-Sayed M. El-kenawy
Delta Higher Institute of Engineering and Technology
Department for Communications and Electronics
Mansoura 35511, Egypt

Applied Science Research Center
Applied Science Private University
Amman, Jordan
skenawy@ieee.org

Manish Kumar Singla
Department of Biosciences,
Saveetha School of Engineering,
Saveetha Institute of Medical and Technical Sciences,
Chennai, India-602 105
msingla0509@gmail.com

Amel Ali Alhussan
Department of Computer Sciences
College of Computer
& Information Sciences
Princess Nourah bint Abdulrahman University
Riyadh, Saudi Arabia
aaalhussan@pnu.edu.sa

Doaa Sami Khafaga
Department of Computer Sciences
College of Computer
& Information Sciences
Princess Nourah bint Abdulrahman University
Riyadh, Saudi Arabia
dskhafga@pnu.edu.sa

Amal H. Alharbi
Department of Computer Sciences
College of Computer
& Information Sciences
Princess Nourah bint Abdulrahman University
Riyadh, Saudi Arabia
ahalharbi@pnu.edu.sa

Marwa M. Eid
Faculty of Artificial Intelligence
Delta University for Science
& Technology Mansoura 35111, Egypt
Jadara University Research Center
Jadara University
Jordan
mmm@ieee.org

***Abstract*— Primary biliary cirrhosis (PBC)—an immune-mediated chronic liver disease destroying intrahepatic bile ducts, leading to cirrhosis/failure—was analysed using Mayo Clinic trial data (424 patients, 1974-1984): 312 in randomized D-penicillamine/placebo trial, 130 observational. Machine learning models trained on demographic/biochemical features for survival/prognosis prediction identified MLP as optimal (96.07% accuracy, 94.34% sensitivity, 96.99% specificity), outperforming KNN and Decision Tree. Serum bilirubin and liver disease stage emerged as key prognostic indicators. These models support clinical decision-making for PBC identification/management.**

Index Terms—Primary biliary cirrhosis, predictive modelling, machine learning, survival analysis, clinical indicators, disease progression.

I. Introduction

Primary biliary cirrhosis is a chronic autoimmune liver disease that predominantly affects middle-aged women, leading to progressive bile duct destruction, fibrosis, cirrhosis, and potentially liver failure. Traditional prognostic tools like MELD and Child–Pugh scores, though useful, rely on linear assumptions and struggle to integrate complex, multidimensional clinical data.

This study uses ML algorithms (MLP, KNN, Decision Trees, Random Forests) on the 424-patient Mayo Clinic PBC cohort (1974–1984) to predict prognosis and survival, aiming to build accurate models, identify key predictors, compare methods, and assess clinical utility. By modeling nonlinear relationships and handling high-dimensional features, these ML approaches are positioned to improve risk stratification, monitoring, and precision medicine in hepatology.

II. Literature Review

Machine learning outperforms traditional scores (like MELD/MELD-Na) for predicting cirrhosis outcomes such as mortality, portal hypertension, readmission, HCC, fibrosis, microbiome-based severity, and encephalopathy, often reaching AUC/C-index around 0.8–0.9. However, these gains highlight the parallel need for better interpretability so models can be trusted and used in real-world clinical practice.

III. Discussion and Results

The study uses the 424-patient Mayo Clinic PBC cohort and a Kaggle-formatted cirrhosis dataset with rich clinical and demographic features to build survival and prognosis models. Figure 1 indicates that more advanced cirrhosis stages cluster in older patients, suggesting a relationship between age and disease severity.

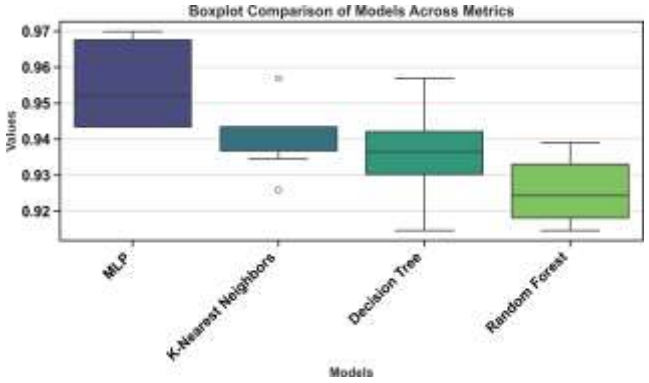

Fig. 1. Stage vs Age Plot

Table I reports the performance of several machine learning models on the cirrhosis dataset using accuracy, sensitivity, specificity, and F-score as evaluation metrics. The **MLP** model achieves the highest accuracy of 0.9607, indicating superior ability to correctly predict patient outcomes and distinguish cirrhosis status in this cohort.

Figure 2 shows that the MLP model has the lowest and most consistent errors, with a narrow spread and maximum error around 0.20, confirming it as the most reliable model for PBC outcome prediction.

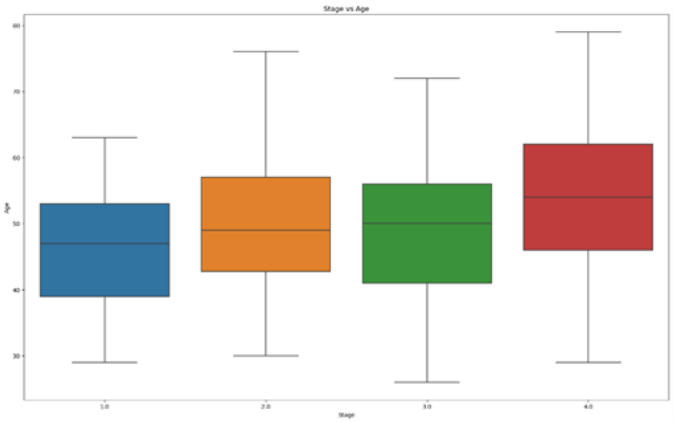

Fig. 2. Comparison of Machine Learning Model Performance in Predicting PBC Outcomes: Boxplot Analysis

Figure 3 shows that the MLP model has the highest Precision (around 0.95) and a similarly strong NPV, meaning it is best at correctly identifying both PBC and non-PBC cases.

Figure 4 KDE plot shows the top model's accuracy tightly concentrated around 0.87–0.89, indicating consistently high and reliable survival prediction performance.

TABLE I
PERFORMANCE METRICS OF MACHINE LEARNING MODELS ON THE CIRRHOSIS DATASET

Model	Accuracy	Sensitivity (TPR)	Specificity (TNR)	Precision (PPV)	Negative Predictive Value (NPV)	F1-Score
MLP	0.9607	0.9434	0.9699	0.9434	0.9699	0.9434
K-Nearest Neighbors	0.9434	0.9259	0.9569	0.9434	0.9434	0.9346
Decision Tree	0.9383	0.9146	0.9569	0.9434	0.9346	0.9288
Random Forest	0.9284	0.9146	0.9390	0.9202	0.9346	0.9174

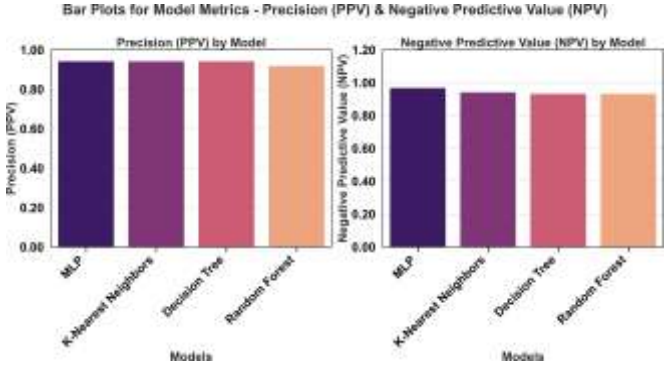

Fig. 3. Precision and Negative Predictive Value Comparison of Machine Learning Models for Primary Biliary Cirrhosis Prediction

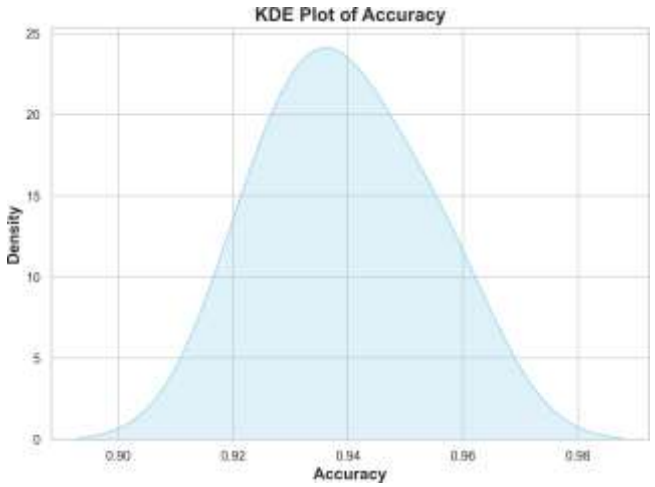

Fig. 4. KDE Plot of Model Accuracy in Predicting PBC Survival

Figure 5 shows that the best PBC model consistently achieves high accuracy, sensitivity, and specificity across runs, with CDFs rising sharply at strong performance levels, indicating stable and reliable predictive behavior.

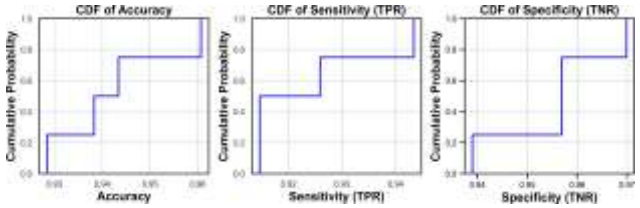

Fig. 5. Cumulative Distribution Functions of Accuracy, Sensitivity, and Specificity for PBC Prediction Model

Figure 6 shows density plots of the best model's metrics for PBC prediction. Accuracy values cluster at the high end, specificity peaks near 0.95, and F1-score centers around 0.9, indicating consistently strong overall performance.

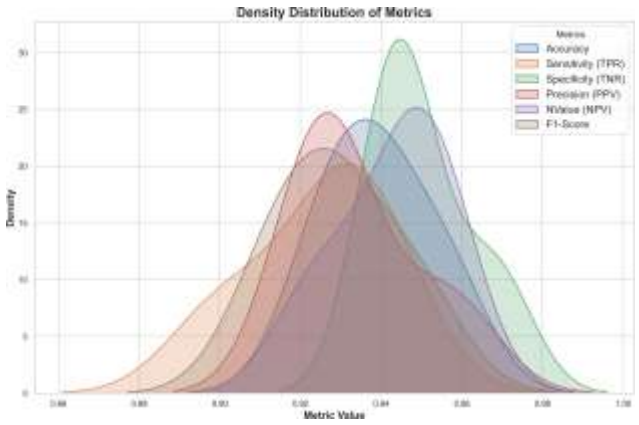

Fig. 6. Density Distributions of Performance Metrics for PBC Prediction

Figure 7 shows that, among all algorithms, the MLP's performance values are most densely concentrated at the high end of the distribution, indicating it most frequently achieves superior predictive accuracy for PBC.

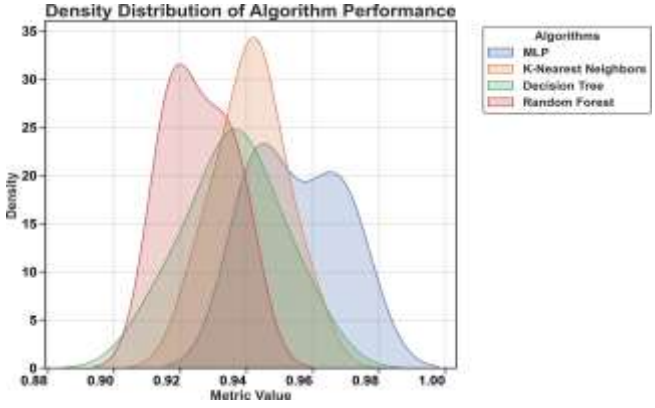

Fig. 7. Density Distribution of Algorithm Performance in Predicting Primary Biliary Cirrhosis

Figure 8 compares Accuracy and Sensitivity (TPR) for MLP, KNN, Decision Tree, and Random Forest in predicting PBC outcomes. MLP attains the highest accuracy (≈0.8), indicating strong correct classification of PBC-related cases.

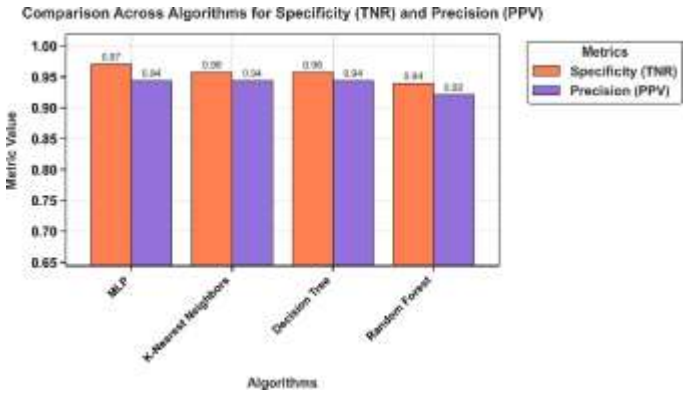

Fig. 8. Accuracy and Sensitivity of Machine Learning Algorithms for PBC Prediction

Figure 9 compares Specificity (TNR) and Precision (PPV) for MLP, KNN, Decision Tree, and Random Forest in PBC prediction. Figure 10 uses a radar plot over Sensitivity, Specificity, Accuracy, F1-score, NPV, and Precision, where the Multilayer Perceptron shows the strongest overall performance profile among the models.

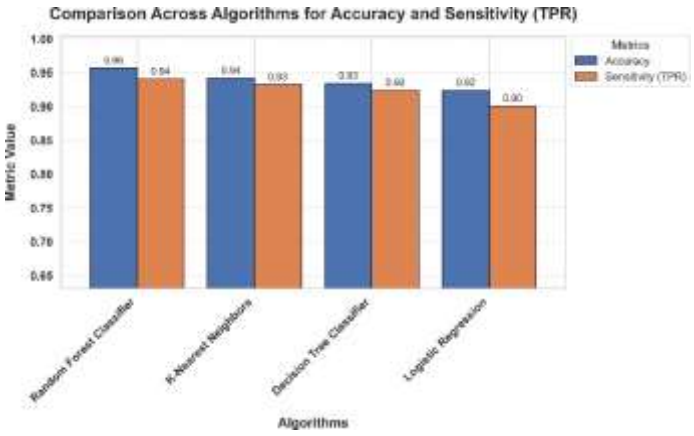

Fig. 9. Specificity and Precision Comparison of Machine Learning Algorithms for PBC Prediction

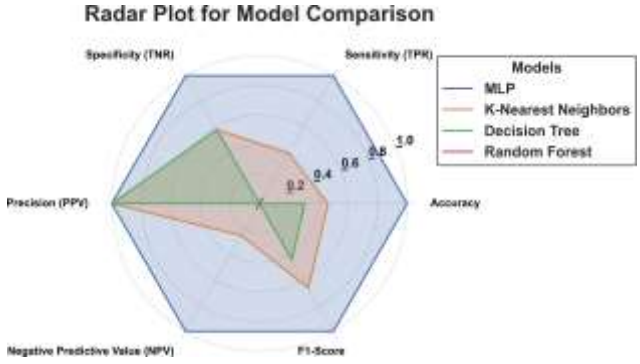

Fig. 10. Radar Plot Comparing Machine Learning Models for PBC Prediction

Figure 11 plots Accuracy (objective function) for MLP, KNN, Decision Tree, and Random Forest on the PBC dataset. MLP attains the highest and most consistent performance, with maximum accuracy of about 0.978.

IV. Conclusion

Machine learning on the Mayo PBC cohort shows that key clinical and biochemical variables can be used to predict prognosis with high accuracy, with an MLP model achieving around 96% accuracy, 94% sensitivity, and 97% specificity. These models both complement existing prognostic scores and highlight the need for external validation and integration with multimodal and EHR data to enable scalable, explainable precision medicine tools in chronic liver disease.

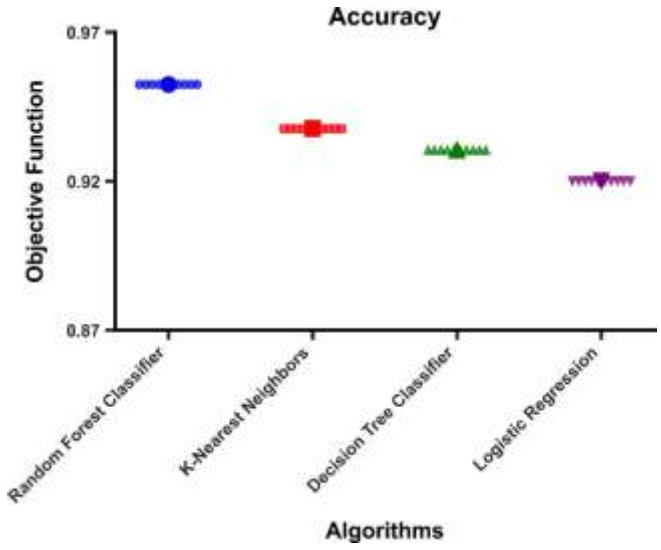

Fig. 11. Accuracy Comparison of Machine Learning Algorithms for Primary Biliary Cirrhosis Prediction

REFERENCES

[1] El-Sayed M El-Kenawy, Abdelhameed Ibrahim, Amel Ali Alhussan, Doaa Sami Khafaga, Ayman EM Ahmed, and Marwa M Eid. Smart city electricity load forecasting using greylag goose optimization-enhanced time series analysis. *Arabian Journal for Science and Engineering*, pages 1–19, 2025.

[2] El-Sayed M Elkenawy, Amel Ali Alhussan, Doaa Sami Khafaga, Zahraa Tarek, and Ahmed M Elshewey. Greylag goose optimization and multilayer perceptron for enhancing lung cancer classification. *Scientific Reports*, 14(1):23784, 2024.

[3] Ahmed El-Sayed Saqr, Mohamed S Saraya, and El-Sayed M El-Kenawy. Enhancing co2 emissions prediction for electric vehicles using greylag goose optimization and machine learning. *Scientific Reports*, 15(1):16612, 2025.

[4] El-Sayed M Elkenawy, Amel Ali Alhussan, Marwa M Eid, and Ab-delhameed Ibrahim. Rainfall classification and forecasting based on a novel voting adaptive dynamic optimization algorithm. *Frontiers in Environmental Science*, 12:1417664, 2024.

[5] Amel Ali Alhussan, El-Sayed M. El-kenawy, Doaa Sami Khafaga, Amal H. Alharbi, and Marwa M. Eid. Groundwater resource prediction and management using comment feedback optimization algorithm for deep learning. *IEEE Access*, pages 1–1, 2025.

[6] F. Kanwal, T. J. Taylor, J. R. Kramer, Y. Cao, D. Smith, et al. Develop-ment, validation, and evaluation of a simple machine learning model to predict cirrhosis mortality. *JAMA Network Open*, 3(11):e2023780, 2020.

[7] J. Reinisˇ, O. Petrenko, B. Simbrunner, B. S. Hofer, F. Schepis, et al. Assessment of portal hypertension severity using machine learning models in patients with compensated cirrhosis. *Journal of Hepatology*, 78(2):390–400, 2023.

[8] C. Hu, V. Anjur, K. Saboo, K. R. Reddy, J. O'Leary, et al. Low predictability of readmissions and death using machine learning in cirrhosis. *Official journal of the American College of Gastroenterology — ACG*, 116(2):336, 2021.

[9] E. Audureau, F. Carrat, R. Layese, C. Cagnot, T. Asselah, et al. Personalized surveillance for hepatocellular carcinoma in cirrhosis – using machine learning adapted to hcv status. *Journal of Hepatology*, 73(6):1434–1445, 2020.

[10] D. Chang, E. Truong, E. A. Mena, F. Pacheco, M. Wong, et al. Machine learning models are superior to noninvasive tests in identifying clinically significant stages of nafld and nafld-related cirrhosis. *Hepatology*, 77(2):546, 2023.

[11] M. Noureddin, Z. Goodman, D. Tai, E. L. K. Chng, Y. Ren, et al. Machine learning liver histology scores correlate with portal hypertension assessments in nonalcoholic steatohepatitis cirrhosis. *Alimentary Pharmacology & Therapeutics*, 57(4):409–417, 2023.

[12] K. Saboo, N. V. Petrakov, A. Shamsaddini, A. Fagan, E. A. Gavis, et al. Stool microbiota are superior to saliva in distinguishing cirrhosis and hepatic encephalopathy using machine learning. *Journal of Hepatology*, 76(3):600–607, 2022.

[13] H. Yang, X. Li, H. Cao, Y. Cui, Y. Luo, et al. Using machine learning methods to predict hepatic encephalopathy in cirrhotic patients with unbalanced data. *Computer Methods and Programs in Biomedicine*, 211:106420, 2021.

[14] A. Guo, N. R. Mazumder, D. P. Ladner, and R. E. Foraker. Predicting mortality among patients with liver cirrhosis in electronic health records with machine learning. *PLOS ONE*, 16(8):e0256428, 2021.

[15] W. M. Shaban. Early diagnosis of liver disease using improved binary butterfly optimization and machine learning algorithms. *Multimedia Tools and Applications*, 83(10):30867–30895, 2024.