Mesopotamian journal of Computer Science
Vol. 2024, **pp**. 214–223
DOI: https://doi.org/10.58496/MJCSC/2024/017; ISSN: 2958-6631
https://mesopotamian.press/journals/index.php/cs

Research Article

An optimized model for Liver disease classification based on BPSO Using Machine learning models.

El-Sayed M. El-Kenawy [1,2,3,*], Nima Khodadadi [4], Abdelhameed Ibrahim [1], Marwa M. Eid [5], Ahmed M. Osman [6], Ahmed M. Elshewey [7]

[1] *School of ICT, Faculty of Engineering, Design and Information & Communications Technology (EDICT), Bahrain Polytechnic, PO Box 33349, Isa Town, Bahrain*

[2] *Jadara University Research Center, Jadara University, Jordan.*

[3] *Applied Science Research Center, Applied Science Private University, Amman, Jordan.*

[4] *Department of Civil and Architectural Engineering, University of Miami, Coral Gables, FL, USA.*

[5] *Faculty of Artificial Intelligence, Delta University for Science and Technology, Mansoura 11152, Egypt.*

[6] *Department of Information Systems, Faculty of Computers and Information, Suez University, P.O.Box:43221, Suez, Egypt.*

[7] *Department of Computer Science, Faculty of Computers and Information, Suez University, P.O.Box:43221, Suez, Egypt.*

ARTICLE INFO

Article History

Received 18 Aug 2024
Revised: 19 Sep 2024
Accepted 18 Nov 2024
Published 10 Dec 2024

Keywords

Health risk prediction

Machine Learning

Liver Disease

BPSO

ABSTRACT

Liver disease (LD) is a world health concern that requires accurate diagnostic methods. This study proposes an optimized machine learning model (ML) based on BPSO for LD classification using a shared public dataset from kaggle Indicates to liver patients from India. The paper used six ML models such as Random Forest (RF), Support Vector Machine (SVM), Dummy Classifier (DC), Extra Trees Classifier (ET), K-Nearest Neighbors (KNN), and Logistic Regression (LT) to evaluate the performance. Through observations we detected that ET achieved an accuracy of 79.82%. The BPSO hyperparameter optimization optimized ET to enhance accuracy to reach 85%. The paper used metrics such as accuracy, precision, recall, F1 score, and AUC. The results indicate that ML techniques with optimization have the potential to develop reliable diagnostic tools for LD.

1. INTRODUCTION

There are millions of fatalities that occur every year due to LD, which is a worldwide health problem. An infection caused by the chronic hepatitis B virus (HBV) is a widespread illness that affects 257 million people globally and more than forty percent of cases have the potential to develop into more serious illnesses [1]. Early identification and antiviral therapy have the potential to stop the course of the disease; however, not all individuals require their treatment to be effective. It is possible that patients with liver fibrosis in the early stages merely require follow-up care, but individuals with substantial fibrosis require active intervention. Consequently, there is a want for biomarkers that are easily available, cost-effective, and trustworthy in order to evaluate the phases of liver fibrosis through the use of non-invasive procedures.

The work uses Binary Particle Swarm Optimization to enhance liver disease categorization through machine learning. The shared Indian Kaggle liver patient dataset is used though the study. RF, SVM, DC, ET, KNN, and LR are utilized for evaluating the performance of classification. we noticed that the accuracy of classification improved by using BPSO.

The findings offer some insights into the potential of ML techniques, when combined with optimization strategies, in the accurate classification of liver failure. The findings make a contribution to the ongoing efforts to develop diagnostic tools that are reliable and non-invasive, and that can assist medical professionals in the process of finding liver disease early on and making decisions regarding its management.

*Corresponding author. Email: skenawy@ieee.org

This paper examines the remaining sections as follows: Section 2 introduces relevant work for several studies that have been conducted on machine learning-based liver disease classification. Section 3 elucidates the suggested methodology, providing details on the application of the optimized approach to liver disease classification. Section 4 examines both the simulations and analysis that have been employed for the suggested model's performance testing. Eventually, the conclusion is discussed in Section 5, covering some potential concerns for future research.

2. RELATED WORK

Dritsas and Trigka [2] employed various ML models and examined a combination of approaches, drawing a comparison with respect to accuracy, precision, recall, F-measure, and area under the curve (AUC) for estimating the probability of liver disease. The experiments revealed the effectiveness of the Voting classifier, which could surpass other models and achieve an accuracy, recall, and F-measure of 80.1%, a precision of 80.4%, and an AUC of 88.4% assuming SMOTE with 10-fold cross-validation.

Shaban [3] highlighted the vital role of extracting the most significant features in the feature selection phase. This approach officially goes by IB2OA, which is short for Improved Binary Butterfly Optimization Algorithm. IB2OA is composed of two distinct phases: Primary Selection (PS) step and Final Selection (FS) step. Furthermore, they introduced two improvements to their model; the first is applied to the initial reduction of features, using the Information Gain (IG) technique, while the second is applied to BOA's initialization, employing Optimization Based on Opposition (OBO). They eventually applied five distinct classifiers— SVM, KNN, NB, DT, and RF—to LD patient recognition in the detection phase. A wide range of various experiments reveals the effectiveness of the proposed IB2OA, which could achieve higher precision, accuracy, recall, and F-score compared to other current top models. Furthermore, it has an average chosen feature score of 4.425, surpassing other models. They found that the KNN classifier achieved the highest accuracy in classification on those tested datasets, compared to all other classifiers examined.

A combination of new ML approaches has been applied by Badvath et al. [4] to the classification phase. Furthermore, they conducted label encoding and data normalization, employing the min-max approach in the preprocessing phase. In the meantime, they applied the ConvNeXt approach to returning features, including age, gender, and Liver function tests (AST, ALT, alkaline phosphatase, bilirubin) in addition to Medical history and comorbidities. They also identified the boasted Grasshopper optimization algorithm (IGOA) and those critical features meant for the accuracy enhancement. They eventually employed the combination of the optimized ML approaches: naïve Bayes and logistic regression (ONBLR) for the liver cirrhosis disease classification in addition to the Harris Hawks optimization algorithm for enhancing the hyperparameters. The introduced ML approach demonstrated effective performance compared to other existing top approaches, achieving an accuracy of 99.18%, a Sensitivity of 99.12%, and a specificity of 98.92%.

Suárez et al. [5] have introduced an approach designed for 191 patients in total in their study. Out of the total, 29 cases represented NAFLD-related HCC. Furthermore, they applied the extreme gradient boosting (XGB) approach for designing the reference predictive model. NAFLD-related HCC patients had a poorer prognosis than those with other potential etiologies related to HCC. They also proved alcohol consumption among NAFLD patients to be the most significant variable for the introduced predictive model, which demonstrates the poorest prognosis. Their study revealed the high effectiveness of XGB compared to other studied approaches, which managed to achieve the greatest values of the analyzed performance metrics. In summary, the study indicated that alcohol consumption, obesity, cirrhosis, and clinically significant portal hypertension (CSPH) among NAFLD-related HCC patients reflected a poorer prognosis than other etiologies and features. That predictive model designed by XGB for patient assessment could also demonstrate high effectiveness and performance.

Theerthagiri and Siddalingaiah [6] have introduced RG-SVM, drawing a comparison with a number of existing algorithms, including LR, DT, KNN, and NB, as well as the introduced RG-SVM algorithms. They found that the algorithms LR, DT, KNN, and NB, and the introduced RG-SVM, could achieve an accuracy of 73, 80, 81, 54, and 93%, respectively. The introduced RG-SVM, which involves a recursive feature selection algorithm, demonstrated highly effective performance compared to other existing algorithms, featuring accuracy enhancement of 14–39% and MSE error reduction of 12–20%, while the sensitivity and specificity enhancements reached 5–26% and 34–72%, respectively, surpassing the results of the compared existing algorithms. Those conclusions can greatly support optimized decision-making and clinical tests made for liver disease patients.

Zhang et al. [7] developed a DT model employing a combination of five serological biomarkers, which comprise HBV-DNA, platelet, thrombin time, international normalized ratio, and albumin, along with the area under curve (AUC) values meant for the assessment of liver fibrosis phases (F0-1, F2, F3, and F4) of 0.898, 0.891, 0.907, and 0.944, respectively, within the training cohort, while the external validation cohort features AUC values of 0.906, 0.876, 0.931, and 0.933 for the liver fibrosis phases (F0-1, F2, F3, and F4), respectively. Depending on the cutoff value, the simulated risk classification highlighted how the DT model performance can accurately align with the pathological diagnosis outcomes when

distinguishing between hepatic fibrosis phases. The ML model, which applies five serum markers, can support accurate diagnosis of hepatic fibrosis phases and effective clinical tracking and treatment of CHB patients.

Hendi et al. [8] developed an innovative approach named CNN+LSTM, which integrates the CNN along with LSTM networks. Their study concluded that the application of ML and DL can enhance the liver disease diagnosis and prognosis. The CNN+LSTM model could attain a high accuracy of 98.73% compared to other models, including CNN, RNN, and LSTM, demonstrating highly effective outcomes. In addition to its high accuracy, it could also achieve a precision of 99%, a recall of 98%, an F1 score of 98%, and an AUC of 99%, featuring high effectiveness in predicting liver diseases and supporting accurate disease diagnosis as well as prognosis.

Agbozo and Balungu [9] stressed the significance of XAI in providing detailed demonstrations of those decisions made by AI to professionals in the biomedical application field, including physicians engaged in AI-powered clinical decisions on disease diagnosis, treatment, or prognosis. Their study has analyzed the Indian Patient Liver Dataset (IPLD), extracted from the Andhra Pradesh region. The deep learning model has achieved an accuracy score of 0.81 and been built on Keras-Tensorflow, while the hyperparameter-tuned model had a 0.82 accuracy score. They applied GANs to the dataset oversampling to manage the imbalance existing in the target values. Furthermore, they incorporated the XAI of Shapley Values to elucidate the liver disease detection model as well as its predictive outcomes.

Alizargar et al. [10] indicated the efficacy of the composite model in NAFLD diagnosis, involving random forest, XGBoost, and SVM approaches. That composite model has achieved exceptional performance metrics, featuring an accuracy of 0.99 and an AUC of 1.00 while also demonstrating high precision and reliability. Their analysis revealed the significance of gender as a variable for NAFLD prediction and how highly it is associated with this disease. The introduced composite model featured improved diagnostic qualities in addition to highly optimized outcomes compared to other indicators, particularly HIS. These outcomes revealed the high effectiveness of the introduced composite model in early NAFLD detection and diagnosis, as well as screening tasks.

Pumhirunroj et al. [11] have applied the ordinary least (OLS) model approach to the variable screening. They have also chosen the most effective combination of variables in order to create an innovative variable set using the principal of component analysis (PCA) technique. Their study demonstrated the effectiveness and accuracy of the forest classification and regression (FCR) model in predicting infection rates, which featured a 0.915 reliability score, followed by 0.794, 0.741, and 0.632 scores, respectively, based on the PCA factor. Furthermore, it provided comprehensive information regarding the factors associated with water body infection, including the length and density of hexagonally formed water flow lines, and it could also track the depth of each process.

Table 1 shows the summary of some related LD classfication studies.

TABLE I. SUMMARY OF THE MOST RECENT STUDIES ON LD CLASSIFICATIONS.

Ref.	**Methodology/Algorithm**	**Feature Selection/Optimization**	**Models**
Dritsas & Trigka [2]	Voting Classifier with SMOTE & 10-fold CV	-	Voting Classifier
Shaban [3]	Improved Binary Butterfly Optimization Algorithm (IB2OA)	Information Gain (IG) & Optimization Based on Opposition (OBO)	SVM, KNN, NB, DT, RF
Badvath et al. [4]	ConvNeXt, ONBLR, Harris Hawks Optimization	Label encoding, Min-max normalization, Grasshopper optimization	Naive Bayes, Logistic Regression
Suárez et al. [5]	Extreme Gradient Boosting (XGB)	-	XGB
Theerthagiri & Siddalingaiah [6]	RG-SVM with Recursive Feature Selection	Recursive Feature Selection	RG-SVM, LR, DT, KNN, NB
Zhang et al. [7]	Decision Tree (DT) with Serological Biomarkers	Selection of biomarkers (e.g., HBV-DNA, platelets)	Decision Tree
Hendi et al. [8]	CNN + LSTM	-	CNN, RNN, LSTM
Agbozo & Balungu [9]	Deep Learning Model + GANs + XAI (Shapley Values)	GANs for oversampling, Shapley Values for explainability	Deep Learning
Alizargar et al. [10]	Composite Model (RF, XGBoost, SVM)	Feature combination, hyperparameter tuning	Composite Model
Pumhirunroj et al. [11]	Forest Classification and Regression (FCR) with OLS & PCA	Principal Component Analysis (PCA)	FCR

3. MATERIALS AND METHODS

3.1. Dataset Description

Kaggle's LD shared dataset of Indian liver patients is available at following https://www.kaggle.com/datasets/fatemehmehrparvar/liver-disorders . Age, gender, total bilirubin, alkaline phosphatase, albumin, and other various variables can be employed as clinical variables for the dataset classification. Liver disease patients are characterized by some target features. In dataset analysis, it is crucial to check for missing values, data distribution, and outliers that affect data quality.

TABLE II. SOME TARGET FEATURES

	age	gender	tb	db	alkphos	sgpt	sgot	tp	alb	a/g_ratio	selector
count	570	570	570	570	570	570	570	570	570	570	570
mean	1.25E-17	0.754386	4.67E-18	3.12E-17	7.48E-17	-1.87E-17	1.25E-17	-4.49E-16	-3.99E-16	1.54E-16	1.287719
std	1.000878	0.430829	1.000878	1.000878	1.000878	1.000878	1.000878	1.000878	1.000878	1.000878	0.453097
min	-2.51721	0	-0.46655	-0.4937	-0.93339	-0.38457	-0.34195	-3.49136	-2.82491	-2.03585	1
25%	-0.73017	1	-0.40268	-0.45838	-0.4723	-0.31287	-0.29034	-0.64038	-0.68953	-0.77891	1
50%	0.009297	1	-0.37074	-0.42305	-0.34173	-0.24669	-0.23529	0.095355	-0.06148	-0.05618	1
75%	0.810385	1	-0.11525	-0.06979	0.025499	-0.10881	-0.07787	0.647158	0.817788	0.478022	2
max	2.782295	1	11.44574	6.430271	7.419106	10.59096	16.58361	2.854369	2.953162	5.820003	2

Figure 1 shows the correlation matrix of the utilized LD dataset extracted from Kaggle.

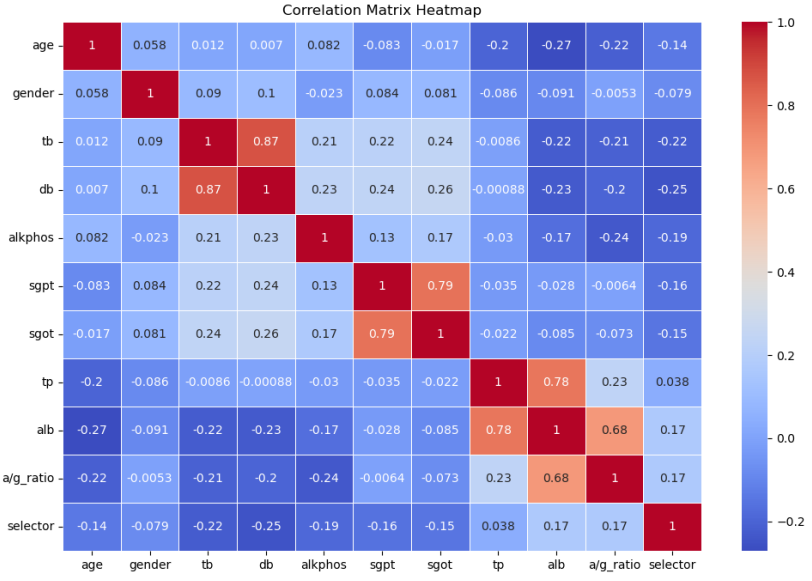

Fig 1. heatmap of Kaggle kidney dataset.

Figure 2 displays the distribution of numerical features. Figure 3 shows the boxplot of the shared LD dataset obtained from Kaggle.

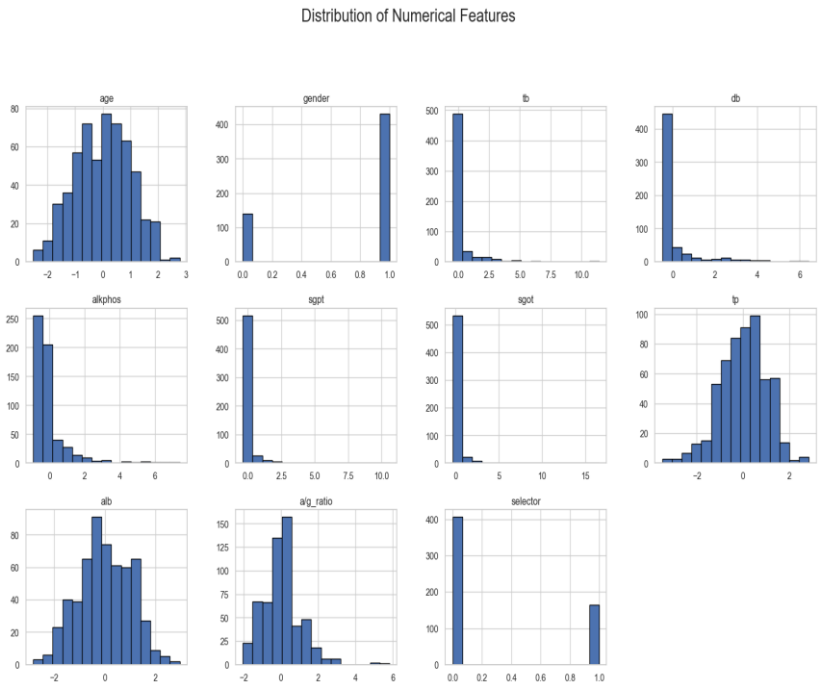

Fig 2. distribution of numerical features.

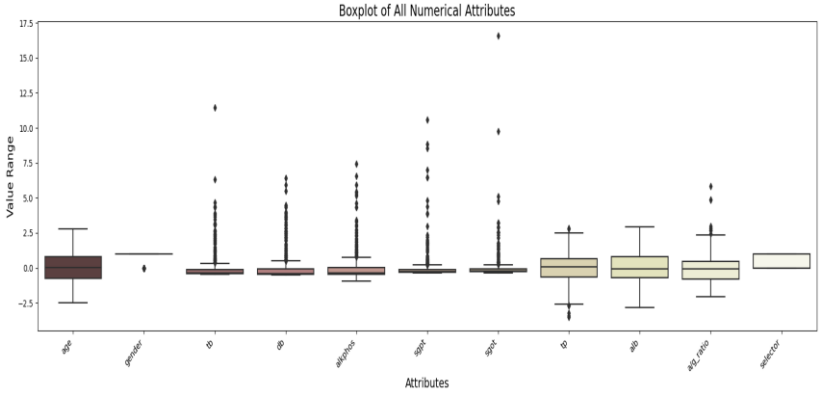

Fig 3. boxplot of LD dataset attributes.

Figure 4 displays the density of dataset attributes.

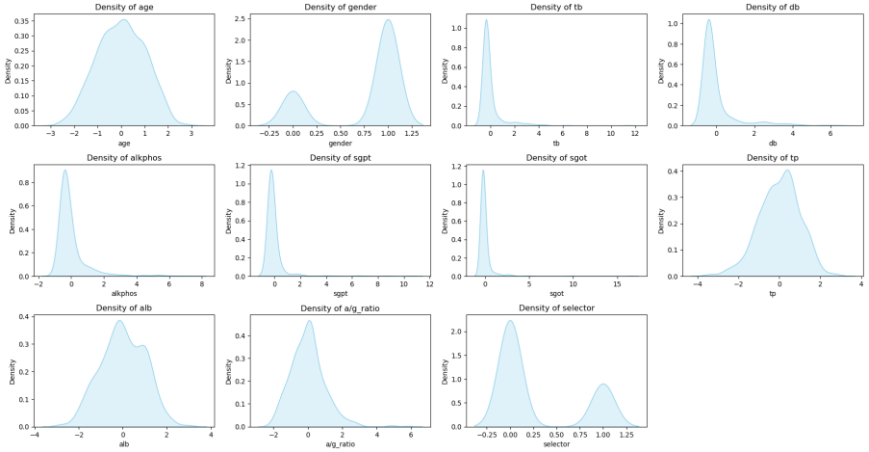

Fig 4. density of numerical features.

Figure 5 displays the pair plot of the LD dataset features.

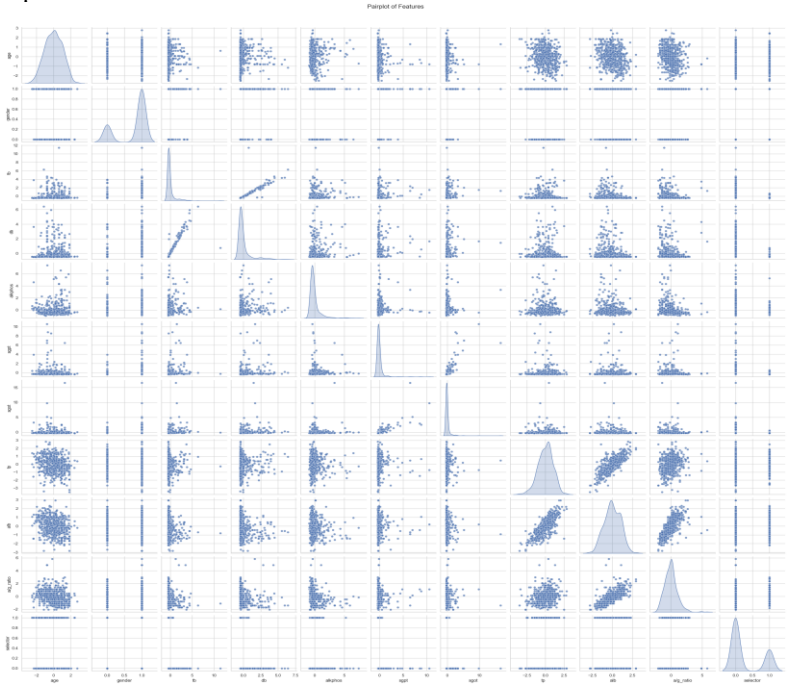

Fig. 5. pair plot of the features.

3.2. Data Preprocessing

The various preprocessing steps, including eliminating null values through imputation, normalizing data by Min-Max Scaling, and encoding categorical variables, particularly gender, with either one-hot encoding or label encoding, optimize the LD dataset for model training. Those steps are intended for keeping dataset quality and upscaling model performance, as well as maintaining appropriate feature scaling.

3.3. Machine Learning Models

1. Random Forest (RF)
A composite model gathering predictions from several decision trees for classification performance enhancement.
2. Support Vector Machine (SVM)
A robust classifier determining the optimal hyperplane for class separation.
3. Dummy Classifier (DC)
Serves as a baseline to compare the performance of other models.
4. Extra Trees Classifier (ET)
A composite method that generates several decision trees, similar to Random Forest, but with improved performance due to randomness in feature selection.
5. K-Nearest Neighbors (KNN)
A distance-based algorithm that classifies data points based on the majority class of their neighbors.
6. Logistic Regression (LR)
A linear model used for binary classification, providing probability-based predictions.

3.4. Model Training and Evaluation

A preprocessed dataset and 80-20 split as well as the BPSO-optimized model, which can be fine-tuned by optimizing hyperparameters, are required for the model training phase. The accuracy, precision, recall, F1 score, and AUC serve as performance metrics, designed for measuring the right prediction and true positive percentages as well as evaluating the ability to differentiate between various thresholds. Furthermore, specially designed formulas are used to return those metrics [12-15].

$$Accuracy = \frac{TP + TN}{TP + FP + FN + TN} \quad (1)$$

$$Sensetivity = \frac{TP}{TP + FN} \quad (2)$$

$$Specificity = \frac{TN}{TN + FP} \quad (3)$$

$$F - score = \frac{2 \times Recall \times Precision}{Recall + Precision} \quad (4)$$

$$AUC = 1/2\left(\frac{TP}{TP + FN} + \frac{TN}{TN + FP}\right) \quad (5)$$

3.5. Simulation Environment

Python is used to do experiments, and libraries, particularly Scikit-learn, NumPy, and Pandas, are also applied to implementing models and manipulating data, as well as evaluating the outcomes. The outcomes are examined to reveal the most effective model, featuring the best liver disease classification.

3.6 The proposed framework

Several ML approaches are integrated with Binary Particle Swarm Optimization (BPSO) in the introduced liver disease classification model for selecting features and fine-tuning hyperparameters. When selecting the most significant features and the most consistent set of parameters, this integrated approach tries to upscale the classification accuracy and also reduce the model complications. The model is structured, especially for prediction performance enhancement and liver disease detection robustness. Figure 6 shows the proposed approach.

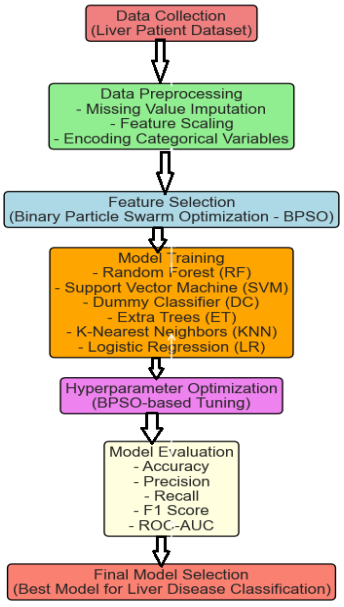

Fig. 6. The proposed framework.

4. RESULTS AND DISCUSSION

This section examined the performance of six machine learning models developed for liver disease classification prior to and following the implementation phase of the Binary Particle Swarm Optimization (BPSO) proposed for the purpose of feature selection and hyperparameter fine-tuning. The examined models were particularly Random Forest (RF), Support Vector Machine (SVM), Dummy Classifier (DC), Extra Trees Classifier (ET), K-Nearest Neighbors (KNN), and Logistic Regression (LR). The performance metrics essentially include accuracy, precision, recall, F1 score, and ROC-AUC, being measured and analyzed for highlighting the most effective liver disease classification model.

Table 3 displays the performance of the ML models before BPSO hyperparameters tunning.

TABLE III. THE PERFORMANCE OF THE SUGGESTED ML MODELS BEFORE BPSO.

Model	Accuracy	Precision	Recall	F1 Score	AUC
Extra Trees Classifier	**0.7982**	**0.6364**	**0.4828**	**0.549**	**0.758**
Random Forest	0.7456	0.5	0.3793	0.4314	0.7365
Support Vector Machine (SVM)	0.7456	1	0	0	0.7014
Dummy Classifier	0.7456	1	0	0	0.5
K-Nearest Neighbors (KNN)	0.6579	0.3611	0.4483	0.4	0.6398
Logistic Regression	0.7105	0.3889	0.2414	0.2979	0.7399

The outcomes revealed the significant role of BPSO in supporting feature selection and hyperparameter optimization as well as improving the performance of most models examined, such as the Extra Trees Classifier. The accuracy of ET upscaled from 79.82% to 85%, indicating the high effectiveness of the introduced liver disease classification model. Furthermore, the optimized performance metrics indicated the vital role of BPSO in mitigating feature redundancy and then supporting more accurate and effective classification.

The analysis also highlighted that the Dummy Classifier lacks powerful predictive features, though it is considered an appropriate baseline. On the contrary, models, particularly RF and ET, are significantly affected by optimization, indicating how pivotally the model tuning impacts the medical diagnostics.

Figure 7 displays the confusion matrices of the suggested ML models.

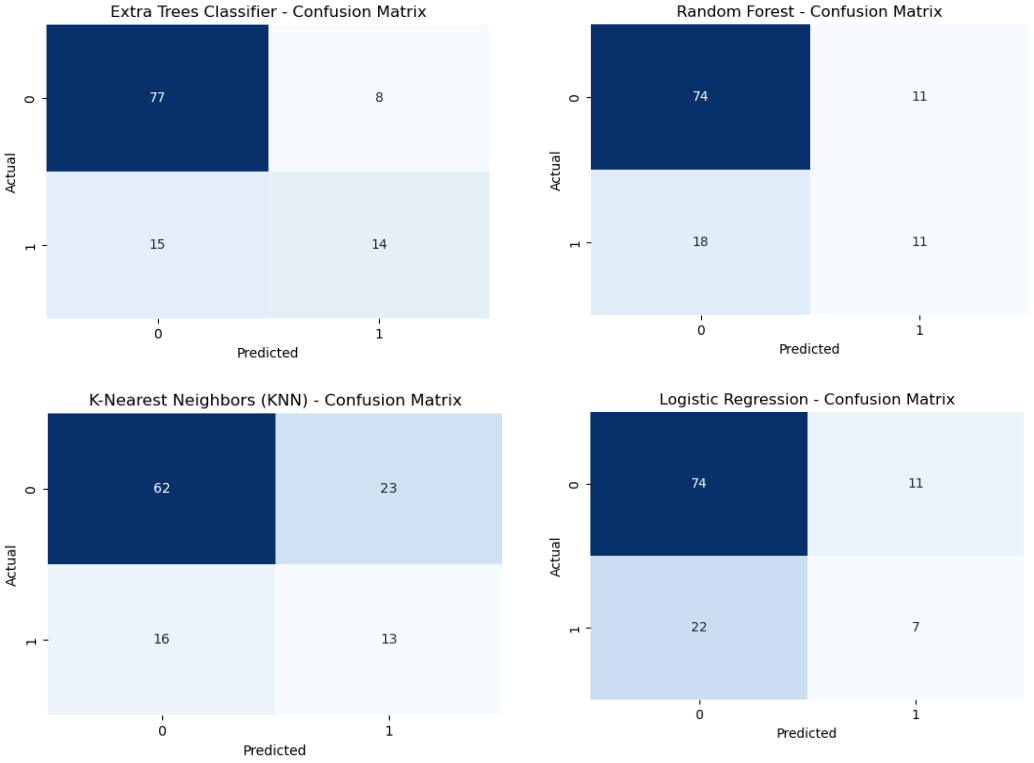

Fig. 7. confusion matrices of the suggested ML models.

Figure 8 displays the AUC of the suggested models.

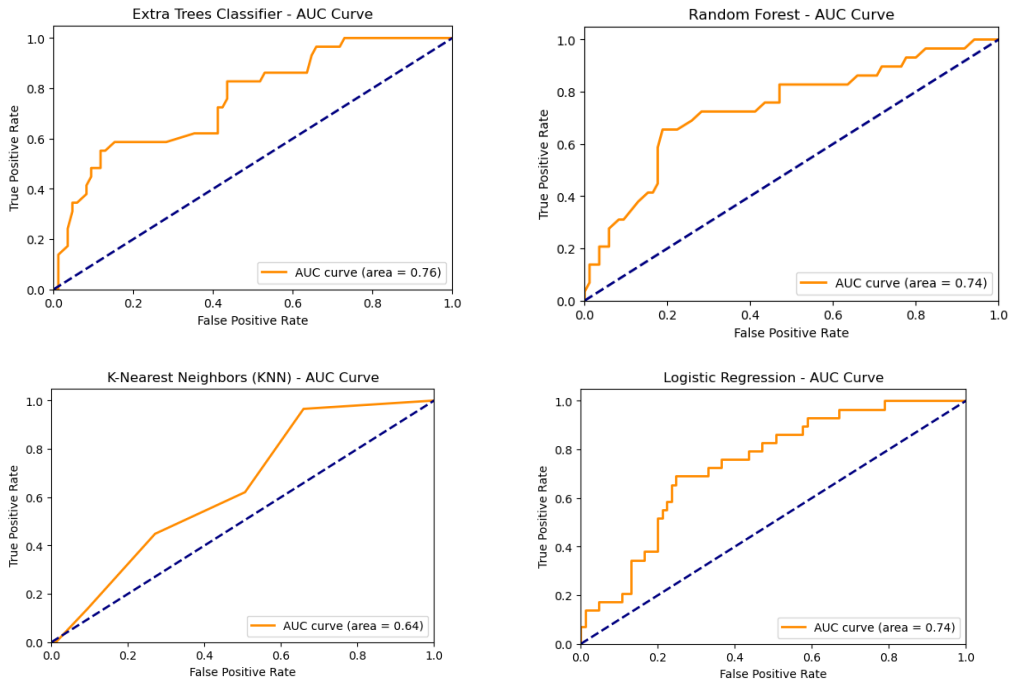

Fig. 8. The AUC of the suggested ML models.

After ET hyperparameter optimization, the accuracy enhanced with 85%. Table 4 displays BPSO hyperparameter configuration.

TABLE IV. PARAMETERS FOR ET OPTIMIZATION WITH BPSO.

Algorithm	Parameter	Values
BPSO	W	[0.5, 0.7]
	Particles	45
	Iterations	100

5. CONCLUSIONS AND FUTURE WORK

This study introduced an optimized liver disease classification ML model and proposed Binary Particle Swarm Optimization (BPSO) for selecting features and fine-tuning hyperparameters. Furthermore, the study highlighted the significance of BPSO, improving the model performance. Along with the Extra Trees Classifier, it could also achieve a high accuracy of 85% and surpass the precision, recall, and F1 score outcomes of other models. Enhanced classification accuracy and reduced features indicated the effectiveness of the introduced approach, managing those complicated medical datasets, including liver patient records.

The study indicated how effectively ML models, along with optimization approaches, can develop reliable and non-invasive diagnostic techniques for liver disease. The introduced model could provide an effective classification system meant for informed and timely decision-making, eventually demonstrating improved outcomes. This paper demonstrated the effectiveness of optimized ML techniques in supporting early disease detection and maintaining efforts in medical diagnosis. Future studies will widely assess the model's generalizability on larger, more diverse datasets extracted from various geographical regions and provide solutions to the class imbalance through applying techniques either oversampling or synthetic data generation. The future research will also study composite learning methods for the purpose of further classification performance enhancement. Furthermore, the real-time applications will be performed to evaluate the practicability in clinical setups. Additionally, other optimization approaches, particularly Genetic Algorithms or Differential Evolution, will be studied for proposing improvements in feature selection and hyperparameter fine-tuning. It is expected that future research will significantly focus on incorporating the optimized model along with deep learning methods for managing dataset complications.

References

[1] M. Jain, et al., "A K-Means clustering and SVM based hybrid concept drift detection technique for network anomaly detection," Expert Systems With Applications, vol. 193, p. 116510, Jan. 2022. DOI: 10.1016/j.eswa.2022.116510.

[2] H. abud-allah, “Multi-view Deep Learning-Based COVID-19 Diagnosis with Chest X-Ray Images: A Comparative Study of SVM and KNN Classifiers”, EDRAAK, vol. 2024, pp. 59–77, Jul. 2024, doi: 10.70470/EDRAAK/2024/009.

[3] X. Duan, et al., "Network traffic anomaly detection method based on multi-scale residual classifier," Computer Communications, vol. 198, pp. 206–216, Nov. 2022. DOI: 10.1016/j.comcom.2022.10.024.

[4] I. I. Al Barazanchi, W. Hashim, R. Thabit, and N. A.-H. K. Hussein, “Advanced Hybrid Mask Convolutional Neural Network with Backpropagation Optimization for Precise Sensor Node Classification in Wireless Body Area Networks”, KHWARIZMIA, vol. 2024, pp. 17–31, Mar. 2024, doi: 10.70470/KHWARIZMIA/2024/004.

[5] B. Yu, et al., "A Network Traffic Anomaly Detection Method Based on Gaussian Mixture Model," Electronics, vol. 12, no. 6, p. 1397, Mar. 2023. DOI: 10.3390/electronics12061397.

[6] A. A. Q. I. . Alnajem, Y. . Niu, and A. . Ibrahim, “Evaluating Educational Strategies Focused on Achieving Sustainable Development Goals in Diverse Global Contexts”, ESTIDAMAA, vol. 2024, pp. 60–71, May 2024, doi: 10.70470/ESTIDAMAA/2024/009.

[7] R. Patil, et al., "Network traffic anomaly detection using PCA and BiGAN," Internet Technology Letters, vol. 5, no. 1, Sept. 2020. DOI: 10.1002/itl2.235.

[8] W. Song, et al., "A Software Deep Packet Inspection System for Network Traffic Analysis and Anomaly Detection," Sensors, vol. 20, no. 6, p. 1637, Mar. 2020. DOI: 10.3390/s20061637.

[9] Y. Feng, et al., "An improved X-means and isolation forest based methodology for network traffic anomaly detection," PLoS ONE, vol. 17, no. 1, p. e0263423, Jan. 2022. DOI: 10.1371/journal.pone.0263423.

[10] L. Vigoya, et al., "IoT Dataset Validation Using Machine Learning Techniques for Traffic Anomaly Detection," Electronics, vol. 10, no. 22, p. 2857, Nov. 2021. DOI: 10.3390/electronics10222857.

[11] M. Jishkariani , Tran., “Assessment of Cyber Security Risks in Electrical Engineering System of Georgia”, Mesopotamian Journal of Civil Engineering, vol. 2024, pp. 54–59, Aug. 2024, doi: 10.58496/MJCE/2024/008.

[12] A. I. Gide and A. A. Mu’azu , Trans., “A Real-Time Intrusion Detection System for DoS/DDoS Attack Classification in IoT Networks Using KNN-Neural Network Hybrid Technique ”, BJIoT, vol. 2024, pp. 60–69, Jul. 2024, doi: 10.58496/BJIoT/2024/008.

[13] Z. Ahmad, et al., "Anomaly Detection Using Deep Neural Network for IoT Architecture," Applied Sciences, vol. 11, no. 15, p. 7050, July 2021. DOI: 10.3390/app11157050.

[14] T. K. Dang, et al., "Machine Learning Based on Resampling Approaches and Deep Reinforcement Learning for Credit Card Fraud Detection Systems," Applied Sciences, vol. 11, no. 21, p. 10004, Oct. 2021. DOI: 10.3390/app112110004.

[15] I. D. Mienye and T. G. Swart, "A Hybrid Deep Learning Approach with Generative Adversarial Network for Credit Card Fraud Detection," Technologies, vol. 12, no. 10, p. 186, Oct. 2024. DOI: 10.3390/technologies12100186.

[16] Y. M. Mohialden, N. M. Hussien, S. A. Salman, Ahmed Bahaaulddin A. Alwahhab, and Mumtaz Ali , Trans., “Enhancing Agriculture Crop Classification with Deep Learning”, Babylonian Journal of Artificial Intelligence, vol. 2024, pp. 20–26, Mar. 2024, doi: 10.58496/BJAI/2024/004.

[17] Y. ALAIWI and T. Ahmed , Trans., “Solar Air Heaters Classifications and Enhancement: A Review ”, Babylonian Journal of Mechanical Engineering, vol. 2024, pp. 71–80, Sep. 2024, doi: 10.58496/BJME/2024/009.

[18] G. Zioviris, et al., "Credit card fraud detection using a deep learning multistage model," The Journal of Supercomputing, vol. 78, no. 12, pp. 14571–14596, Apr. 2022. DOI: 10.1007/s11227-022-04465-9.

[19] A. Alharbi, et al., "A Novel text2IMG Mechanism of Credit Card Fraud Detection: A Deep Learning Approach," Electronics, vol. 11, no. 5, p. 756, Mar. 2022. DOI: 10.3390/electronics11050756.

[20] E. Ileberi, et al., "A machine learning based credit card fraud detection using the GA algorithm for feature selection," Journal of Big Data, vol. 9, no. 1, Feb. 2022. DOI: 10.1186/s40537-022-00573-8.

[21] Tuljaram Meghwar, S. khan Awan, Muhammad Tariq, M. Suleman, and A. A. Shaikh , Trans., “Substitutional Based Gauss-Seidel Method For Solving System of Linear Algebraic Equations”, Babylonian Journal of Mathematics, vol. 2024, pp. 1–12, Jan. 2024, doi: 10.58496/BJM/2024/001.

[22] A. R. Khalid, et al., "Enhancing Credit Card Fraud Detection: An Ensemble Machine Learning Approach," Big Data and Cognitive Computing, vol. 8, no. 1, p. 6, Jan. 2024. DOI: 10.3390/bdcc8010006.