Optimization Algorithms for Deep Learning Prediction of Liver cirrhosis: A Survey

Aya Ebrahim[1], Asmaa H. Rabie [2], El-Sayed M. El-Kenawy [3], Hossam El-Din Moustafa[4]

[1] Department of Applied Health Sciences, Higher Technological Institute of Applied Health Sciences, Mansoura, Egypt
[2] Department of Computers Engineering and Control Systems, Faculty of Engineering, Mansoura University, Mansoura 35516, Egypt
[3] Department of Communications and Electronics, Delta Higher Institute of Engineering and Technology, Mansoura 35111, Egypt
[4] Professor at the Department of Electronics and Communications Engineering, Faculty of Engineering, Mansoura University, Egypt
Emails: eng2008aya@gmail.com; asmaahamdy@mans.edu.eg; skenawy@ieee.org; hossam_moustafa@mans.edu.eg

Abstract

Today, new Artificial Intelligence (AI) techniques are utilized to help doctors forecast the occurrence of diseases because of the necessity of sustaining public health and early disease diagnosis. One significant kind of liver damage is liver cirrhosis, which typically results from long-term liver damage brought on by a variety of liver conditions and diseases, including hepatitis, persistent alcoholism, or heredity. We created this review to provide an overview of liver cirrhosis since it is essential to identify it early and prevent the damage from spreading throughout the liver tissues. In order to identify liver cirrhosis from biomedical markers rather than images, this study has recently conducted nine studies overlaying it with various artificial intelligence deep learning techniques. Our suggested approach used various Machine Learning (ML) models to predict the signs of cirrhosis in conjunction with other illnesses. Because this condition is so important, it is important to summarize these studies based on the methodology and findings of detection accuracy and precision.

Keywords: Liver Cirrhosis; artificial intelligence; deep learning; Optimization algorithms.

1. Introduction

With over a million fatalities annually, liver cirrhosis ranks as the fourteenth most prevalent cause of death globally. Cirrhosis is a severe form of liver disease. Cirrhosis typically results in irreversible liver damage. However, the extent of damage might be minimized if liver cirrhosis is identified early. In the modern world, a number of connected technologies, including artificial intelligence, deep learning, and machine learning, are crucial for identifying disease and offering medical support. Because of the participation of both healthy and infected individuals worldwide, researchers are able to gather a massive dataset of diseases. For this reason, prediction plays a very important role in correctly predicting diseases. Lately, early disease detection has

DOI: https://doi.org/10.54216/MOR.030101
Received: October 16, 2024 Revised: December 01, 2024 Accepted: December 13, 2024

decreased the disease's complications for patients. While some people with cirrhosis have severe signs of end-stage liver disease and a low likelihood of survival, others are completely asymptomatic and have a very typical life expectancy [1].

Improving patient outcomes and lowering the morbidity and mortality rates linked to liver cirrhosis depend heavily on the ability to predict and detect the disease in its early stages. Symptoms of liver cirrhosis show up much later in the progression of the disease. In the early stages of infection, over 80% of infected patients show no symptoms, which leads to increased liver damage and high mortality rates. The most effective method for detecting liver cirrhosis is liver biopsy; nevertheless, it is an expensive and intrusive technique that is impractical for recurrent sampling. Cirrhosis is detected by biopsy techniques, imaging studies, and blood testing. The patients' treatment cannot begin right away due to the cost of these tests and the time it takes to get the test results [2, 3].

Utilizing test data and medical history to develop machine learning-based early detection methods that are non-invasive and more accurate. Improvements in machine learning help physicians diagnose and treat patients more quickly. The best chance of healing is only attainable when the sickness is identified early on, as early identification is difficult without machine learning because symptoms are absent. Initially, the quantity of patient datasets is acquired. Second, the data preprocessing step combines two distinct jobs. Data cleaning is one of them; data transformation and integration are the other two. Noise data and data that is not important to the study are eliminated from the obtained data during the data cleaning (data cleansing) phase. Several data sources are merged into a single source during the data integration process. Thirdly, utilize the pre-trained model and display the classification performance using ROC curves and confusion metrics or performance metrics (accuracy, precision, F1-Score, AUC, etc.). By comparing the outcomes with those of other studies, we can determine whether our methods are accurate and able to generalize our model for disease diagnosis [4].

Serum biomarkers have received little attention in previous research, which has mostly focused on comparative studies and the creation of machine learning models incorporating imaging modalities including ultrasound, magnetic resonance imaging (MRI), and elastography. By using the Explainable AI algorithm to close the gap between AI models and human understanding by offering insights into the complex decision-making process of the suggested machine learning model, researchers aim to address the overlooked aspects of liver cirrhosis assessment and improve transparency and reliability. The novel approach aims to overcome the limitations of earlier research and advance the detection of liver cirrhosis. Establishing a link between AI models and human understanding in the healthcare industry fosters confidence and openness in model results. Using biomarkers' ability to accurately detect liver cirrhosis in its early stages and assist medical professionals in determining the underlying cause. analyzing biomarkers in-depth and using a range of machine-learning techniques for a thorough analysis [5].

2. Related Work

In [6] the work suggested a novel model known as LivMarX.LivMarX's non-invasiveness, excellent accuracy, and efficiency make it a promising prognostic model for liver cirrhosis in clinical settings. LivMarX can offer useful insights into the stages of liver cirrhosis by utilizing easily accessible clinical and laboratory data, ensuring that healthcare providers make well-informed decisions about patient management and treatment regimens. By delaying the advancement of the disease and lowering the risk of complications, early detection and treatment can improve patient outcomes. This is encouraged by the model's ability to properly forecast the stages of liver cirrhosis. LivMarX is a cost-effective technology because it just uses blood samples, eliminating the need for costly and expensive imaging methods.

Patient ID, the number of hospital days (N_Days), patient status, prescribed medications, age, gender, presence of ascites, hepatomegaly, spider angiomas, edema, bilirubin, cholesterol, albumin, copper, serum glutamic

DOI: https://doi.org/10.54216/MOR.030101
Received: October 16, 2024 Revised: December 01, 2024 Accepted: December 13, 2024

oxaloacetic transaminase (SGOT), triglycerides, platelet counts, alkaline phosphatase, prothrombin time, and disease stage are all included in the dataset. The dataset had a number of attributes that needed to be transformed and cleaned up prior to data preprocessing. Some variables were found to have missing values prior to the use of any preprocessing techniques.

The study used a range of machine learning methods. LivMarX's performance was improved by the use of hyperparameter optimization. The accuracy was further increased to 86% by this optimization phase. LivMarX's accuracy exceeded that of earlier research using the same dataset by more than 8%, demonstrating the method's effectiveness. After feature engineering and data preprocessing, the dataset was split into 70% training, 20% test, and 10% validation sets. In order to ensure that the models are not overfit to the training data and that they generalize well to new data, the validation set provides access to hyperparameter adjustment and model selection. A number of machine learning methods were then used to forecast the stage of liver cirrhosis. Performance matrices are then computed by various deep learning models.

:

- Logistic Regression (LR) l is a linear classification model. This model achieved 39% accuracy.
- The Decision Tree (DT) Classifier is a non-linear model that divides the data into subgroups according to the values of the attributes in order to provide predictions. This model achieved 73.49% accuracy.
- Random Forest (RF)Classifier where several decision trees are built using an ensemble learning technique, which are then aggregated to predict. This model achieved 84.33% accuracy.
- A Boosting approach called the Gradient Boosting Classifier which combines the predictions of several weak learners to produce a strong ensemble model was applied . This model was chosen for its ability to fine-tune the model by focusing on mistakes made in previous rounds of training. Gradient Boosting is known for its high predictive. This model achieved 79.51% accuracy.
- The CatBoost Classifier, a gradient-boosting variant specifically designed for categorical data was applied. This model was chosen for its superior ability to handle categorical variables efficiently without needing extensive pre-processing. This model achieved 78.31% accuracy.

In [7] the study suggested an approach classification method based on ensemble machine learning. Label encoding and min-max data normalization are done during preprocessing. Next, characteristics like age, gender, medical history, multiple medical conditions, and liver function tests (AST, ALT, bilirubin, and alkaline phosphatase) are obtained using the ConvNeXt technique. Using the Improved Grasshopper optimization algorithm (IGOA), the key elements that further improve accuracy were chosen. Lastly, the liver cirrhosis disease is classified using the optimized ensemble machine learning method known as naïve Bayes and logistic regression (ONBLR). The hyperparameters are optimized using the Harris Hawks optimization technique.

Pre-processing, feature extraction, feature selection, and classification are the four primary phases in this paper. The main objective of this effort is to create a classification model and increase classification efficiency. Following the preprocessing stage, we use the ConvNeXt technique to extract the key features. Next, using the enhanced Grass Hopper that has been suggested.In order to successfully classify liver disease, this study used a

DOI: https://doi.org/10.54216/MOR.030101
Received: October 16, 2024 Revised: December 01, 2024 Accepted: December 13, 2024

novel feature selection and classification technique. The following are the main contributions of this study:

- The preparation stage involves label encoding and data normalization.
- Features that reduce dimensionality are extracted with the help of the ConvNeXt technique.
- The IGOA model is used to choose the important features. The classifier's performance is further enhanced by the chosen features.
- Liver cirrhosis is classified using an ensemble method that combines logistic regression and naïve Bayes (ONBLR). wherein the Harris Hawks optimization technique is used to optimize the parameters.
- The proposed method enhances classification accuracy by accurately classifying patients with liver cirrhosis and healthy individuals in comparison to current methods.
- When compared to current methods, the suggested model produces better results, with 99.18% accuracy, 99.12% sensitivity, and 98.92% specificity.

In [8] this work intended to use Machine Learning (ML) algorithms, such as Random Forest (RF), Extra Trees (ET), and Support Vector Machine (SVM), to create an intelligent automated system that can forecast the stages of cirrhosis. The first usage of the publicly available data, which includes 10,000 records and 70 attributes. Before making the prediction, the data was also analyzed using data mining techniques. Because of the notable imbalance in the dataset's classifications, data balancing was required. In order to reduce a bias issue in a machine learning model, the researcher used the Synthetic Minority Oversampling Technique (SMOTE).

In order to choose important features, feature selection methods were used, including Chi-Square, Mutual Information (MI), and Recursive Feature Elimination and Cross-Validation (RFECV) based on classifiers RF and SVM (RF-RFECV, SVM-RFECV). The Extra-Trees model using the Chi-square feature selection approach (ET-Chi-Square) reached the highest degree of accuracy of 93.87%, according to the experimental results, which concluded the study.

- Extra-Trees (ET) is an ensemble learning method that is based on DTs. Multitudes of trees are constructed, and the nodes are divided into groups according to randomly selected subsets of features.ET achieves accuracy 93.87%
- Support vector machine (SVM) is a method for identifying a hyperplane in a space with N dimensions. Hyperplanes serve as decision boundaries, enabling the classification of data points. SVM achieves accuracy 91.50%.
- Random forest achieves the balance between robustness and accuracy. Rf achieves accuracy 93.08%.

DOI: https://doi.org/10.54216/MOR.030101
Received: October 16, 2024 Revised: December 01, 2024 Accepted: December 13, 2024

In Paper[9] This work was applying machine learning methods to liver cirrhosis prediction, with an emphasis on improving accuracy rates, is presented in this publication. The study shows how well the suggested methodology fills the research gap and produces remarkably accurate predictions through extensive testing and thorough review utilizing liver cirrhosis datasets. The innovation is found in the thorough testing and performance analyses that show notable gains in prediction accuracy rates over current techniques. The study is using machine learning algorithms to predict liver cirrhosis. The dataset variables include a number of important indicators, including age, gender, different bilirubin levels, such as total and direct bilirubin, and enzymes, such as alkaline phosphatase, alanine aminotransferase, and aspartate aminotransferase. Additionally, the dataset comprises

Additionally, the dataset contains binary variables that indicate the existence or absence of liver disease and liver cirrhosis, as well as indicators such as total proteins, albumin, and the albumin/globulin ratio. The study's primary objective is to use machine learning algorithms to predict liver cirrhosis specifically, Support Vector Machine, Decision Tree Classification, and Random Forest Classification—based on the factors in the dataset.

- During training, the Random Forest (RF) algorithm creates many decision trees and returns the class that represents the average prediction (regression) or the class mode (classification) of each tree separately.

- Gradient boosting incorporates decision trees into the model in a progressive manner. The residuals or mistakes of the aggregate ensemble of earlier trees are predicted by each new tree during training. In order to move the model closer to the right predictions, subsequent trees are built with an emphasis on minimizing the residuals from the earlier model iterations.

- Data moves across the network from input to output layers using the feed forward process used by the MLP. Each neuron in the MLP applies a weighted sum of its inputs, followed by an activation function and a bias term to produce non-linearity. Activation functions include the sigmoid, rectified linear unit (ReLU), and hyperbolic tangent (tanh).

- Extra Trees uses two degrees of randomness in the tree generation process to produce an ensemble of decision trees. First, it makes many subsets of the dataset for every tree using bootstrapping. Secondly, at each node of the decision tree, instead of choosing the optimal split among the randomly picked attributes.

- By employing various kernel functions (such as linear, polynomial, or radial basis function) to convert data into higher dimensions, the Support Vector Machine (SVM) algorithm seeks to solve the optimization problem by determining the hyperplane that maximizes the margin while minimizing classification errors. This makes it simpler to identify a separating hyperplane.

- The KNN algorithm is a non-parametric, instance-based learning technique that looks at a data point's neighbors in the feature space to predict its value or class membership. In theory, when a new data point is introduced, KNN finds the 'K' closest data points from the training dataset using a selected distance metric. Then, it classifies or predicts the new data point's target by assigning the majority class label (for classification) or calculating the average value (for regression) of those neighbors.

- The Decision Tree technique is simple to understand, can handle both category and numerical data, and can identify intricate links in the dataset. However, if they are not

DOI: https://doi.org/10.54216/MOR.030101
Received: October 16, 2024 Revised: December 01, 2024 Accepted: December 13, 2024

properly pruned or limited, they may not generalize well to unknown data and are prone to overfitting, particularly with deep trees.

- Neural Network Feedforward FNNs provide non-linear transformations that are essential for identifying intricate relationships in the data. Each neuron in a FNN takes inputs, applies weights and biases, and then sends the altered information through an activation function.

The Precision-Recall Curve (PRC) is a statistic used in this study to assess machine learning algorithms. In situations where class disparities are common, this statistic is very important. The Receiver Operating Characteristic (ROC) curve, is another important parameter used in this work to evaluate the effectiveness of the machine learning model. The ROC curve, which illustrates the trade-off between sensitivity and specificity across various decision thresholds, is a frequently used tool in binary classification problems.

In [10] Liver failure is usually the result of cirrhosis getting worse. In addition to gastrointestinal bleeding, patients with cirrhosis may also suffer damage to their brains and nerve systems. The goal of cirrhosis treatment is to stop the condition from getting worse. Therefore, early detection of cirrhosis is essential to preventing consequences. Machine learning (ML) has been shown to be helpful in producing precise and accurate data that can be used to diagnose a number of illnesses. However, to date, no research has employed machine learning to identify cirrhosis in hepatitis C patients. The dataset was used to train Four machine learning algorithms: the Extra Trees model, the Random Forest, the Gradient Boosting Machine, and the Extreme Gradient Boosting to diagnose cirrhosis in hepatitis C patients. With just 16 of the 28 features, the Extra Trees model achieved 96.92% accuracy, 94.00% recall, 99.81% precision, and 96% area under the receiver operating characteristic curve, outperforming the other models.

Prior to creating the models, pre-processing methods were used. Initially, binarization was employed to predict just cirrhosis-affected HCV patients. As a result, all other classes were changed to 0 and the cirrhosis class to 1. The min-max scaler was then used to standardize the data. The problem of data imbalance has emerged as a result of the target class's conversion. Random oversampling was used as a result. The HCV dataset, which includes statistical, numerical, and categorical features, serves as the foundation for this investigation. Strategy for Optimization to create models that can solve issues as efficiently as possible, an algorithm's hyperparameters must be changed. For this, the study used grid search with stratified 10-fold cross-validation. A search space can be defined using grid search by defining the hyperparameters and the range of possible values for them. The GridsearchCV technique finds the ideal set of hyperparameters by generating every conceivable combination of values after the hyperparameter grid has been defined. The 10-fold cross-validation is used to assess the model's performance to determine the effectiveness of each combination of hyperparameters.

- . For the original data, the RF's accuracy is 74.22%; however, for the over-sampled data, it is 96.48%.
- For the original data, the GBM's accuracy is 73.93%; however, for the over-sampled data, it is 95.70%.
- using the original data, the XGBoost obtains an accuracy of 73.93%; however, using the over-sampled data, it achieves an accuracy of 90.99%.
- For the original data, the ET's accuracy is 74.22%; however, for the over-sampled data, it is 96.82%.

In [11] A deep learning model based on MLP has been created in this study to predict cirrhosis. DT, kNN, LR, NB, RF, and SVM have all been compared to the constructed model. The suggested model exceeded the

DOI: https://doi.org/10.54216/MOR.030101
Received: October 16, 2024 Revised: December 01, 2024 Accepted: December 13, 2024

models that were compared, according to experimental investigations that used accuracy, precision, recall, and F1-score. The purpose of this study was to ascertain if participants had cirrhosis based on their blood test results and demographic traits. A deep learning model based on MLP was created for this reason.

In the framework of MLP, several neurons with non-linear activation functions are hierarchically connected. MLP consists of an input layer, an output layer, and one or more intermediate layers. The input strings are sent to the input layer for processing. The created MPL-based model predicts if a patient has cirrhosis based on the input of their blood test results. Data from blood tests and patient demographics are displayed as input in the input layer of the generated model. There are two hidden layers in the model. GridSearchCV was used to conduct hyperparameter analysis investigations in order to ascertain the number of neurons and epochs in the hidden layers. In the input layer, the ReLU activation function is employed. The ReLu activation function is utilized in hidden layers to sort the layers and perform nonlinear computations. Since the binary classification is finished, the output layer employs the sigmoid activation function.

A deep learning model based on MLP was created in this work to identify patients with cirrhosis. The created model's experimental results were thoroughly compared to those of DT, kNN, LR, NB, RF, and SVM. Each model's F1-score, recall, accuracy, and precision were compared. Model parameters were determined by parameter analysis studies using GridSearchCV.

- DT's accuracy, precision, recall, and F1-score are 0.6585, 0.7692, and 0.7142, respectively.

- kNN's accuracy, precision, recall, and F1-score are 0.5365, 0.6666, 0.6428, and 0.6544, respectively.
- LR's accuracy, precision, recall, and F1-score are 0.6341, 0.7407, and 0.7142, respectively.

- NB's accuracy, precision, recall, and F1-score are 0.4146, 0.5769, and 0.5357, respectively.

- The F1-score is 0.7999, the recall is 0.7857, the precision is 0.8148, and the accuracy of RF is 0.7317.

- SVM's accuracy, precision, recall, and F1-score are 0.7317, 0.8148, and 0.7857, respectively.

- The suggested model's accuracy, precision, recall, and F1-score are all 0.8048, 0.8571, and 0.8571, respectively.

In [12] the research provided a thorough and empirical review of a machine learning approach for the early diagnosis of liver cirrhosis. The goal of the study is to create a novel architecture for noninvasively early detection of liver cirrhosis. There are five steps in the experiment. Following data acquisition, the following data processing steps are carried out: feature selection, class balancing, data preprocessing, categorical data transformation, and cross validation. LR and XGBoost classifiers are used in this investigation.In order to aid medical professionals diagnose patients, this study also suggested a way to gather information and thorough descriptions of the factors that go into prediction.

- When voting for cirrhotic and non-cirrhotic liver, for example, LR determines the probability of a particular result. The dependant variable has an output between '0' and '1' since the result is a probability.It has a 68% accuracy rate.

- A scalable tree boosting machine learning tool that optimizes gradient boosting training is called XGBoost. The tree-boosting model for a sample space S with b samples and d features is constructed

DOI: https://doi.org/10.54216/MOR.030101
Received: October 16, 2024 Revised: December 01, 2024 Accepted: December 13, 2024

by adding the scores from each decision tree. It has a 90.5% accuracy rate.

In [13] seven classification algorithms are used in this study: ANNs-MLP, RF, DT, NB, K-NN, LR, and SVM. The study's original data, which included 25 factors, came from 2000 actual patients . As is well known, the dataset that is employed affects the accuracy rates of the machine learning algorithms. Thus, the initial step should be to decide which algorithm would be utilized to categorize the study's data. First, the liver dataset was used to compare these seven classification methods. The findings indicate that the DT algorithm performs better than all other approaches in the majority of cases. The Boruta approach, another data mining technique, was used to select features.

Both with and without liver cirrhosis, were included in the data of this study. First, missing-value data has been identified. AFP, sodium, LDH, and ferritin are not included in the dataset since their missing data rates exceed 40%. Twenty variables are hence input variables. The density-based spatial clustering of applications with noise technique was then applied to find anomalies in the liver cirrhosis dataset, and the extreme values that were found were eliminated.

- The accuracy of MLP-ANN ,KNN,LR,NB,RF,SVM and DT is 84.90%,85.40%,83.30%,67.80%,84.35%,85.20% and 87.75% respectively.
- The accuracy of DT is the best result.

In[14] the suggested model used Nave Bayes, Support Vector Machine (SVM) with 10-fold cross-validation, and Classification and Regression Tree (CART) to predict liver cirrhosis sickness. Model performance was assessed using F1 Score, recall, accuracy, and precision. Predicting how liver cirrhosis illness may manifest is necessary for effective treatment and averting disastrous health consequences. Consequently, the researchers utilize machine learning approaches to develop a prediction model.

The dataset comes after data preparation, which is a processing phase that combines two distinct jobs. Data cleaning is one of them; data transformation and integration are the other two. Noise data and data that is not important to the study are eliminated from the obtained data during the data cleaning (data cleansing) phase. Several data sources are merged into a single source during the data integration process. The training dataset is one that is used to fit the parameters (e.g., weights) in order to learn. The majority of methods that look for empirical links in training data have a tendency to overfit the data, which allows them to find and take advantage of associations that appear to be there but are not generally true.

The training and the test dataset have the same probability distribution whereas the test dataset is created on its own. If a model that fits the training dataset also fits the test dataset well, then there has been little overfitting. A larger fit between the training and test datasets is usually a sign of overfitting. Thus, a test set is a collection of instances that are solely used to evaluate a fully described classifier's performance (i.e., generalization).

- Naïve Bayes achieved accuracy 60%,F1score 59% ,recall 45% and 98% precision.
- CART achieved 73%accuracy ,F1score 80% ,recall 81% and 79% precision.
- With an accuracy of 75%, precision of 73%, recall of 100%, and F1 Score of 84%, the Support Vector Machine (SVM) technique yields the best results out of all the strategies employed in this study.

DOI: https://doi.org/10.54216/MOR.030101
Received: October 16, 2024 Revised: December 01, 2024 Accepted: December 13, 2024

3. Conclusions

As stated by the summary of these studies that data classification is crucial for achieving good illness detection performance. The data, which is divided into two classes (cirrhosis and normal) persons, is the most crucial component in predicting the disease [15,16]. In order to prevent overfitting and enable classifiers to detect diseases accurately, data optimization is also crucial. Many machine learning models were available, including RF, DT, SVM, ET, GMB, XGBoost, MLP, and KNN. It assists us in identifying the illness at its earliest stage and preventing its spread. With the aid of optimization techniques, the results for disease identification were excellent, reaching over 100%.

Table 1: Different works in Liver cirrhosis prediction summary .

Author	**year**	**AI Techniques**	**Optimization/ Feature Selection**	**Metrics**	**Data Set**	**Approach**	**No. cross validation**
Kamath, S. K. [6]	**2024**	**LR, RF, , DT, XGBoost,CatBoost**	**Genetic Algorithm, GridSearchCV**	**Accuracy: 86%**	**424 patient** **11 variable**	**2**	**5-fold cross validation**
Badvath, D [7]	**2024**	**NB,LR**	**Grasshopper optimization algorithm (IGOA)**	**Accuracy: 99.18%**	**-**	**2**	**-**
Ali, D. S., [8]	**2024**	**RF,ET,SVM**	**Chi square** **Feature selection**	**Accuracy** **93.87%.**	**10000 record** **70 feature**	**2**	**5- fold cross validation**
Choi, Y.-S [9]	**2024**	**RF,MLP,DT,KNN,ET, SVM,FNN**	**stochastic gradient descent (SGD)**	**Precision:0.68** **Recall:0.5** **F1score:0.53**	**616 patient** **13 variable**	**2**	**-**
Alotaibi, A.[10]	**2023**	**RF,GBM,XGBOOST,ET**	**GridsearchCV**	**Accuracy: 96.82%**	**2038 patient** **28 variable**	**4**	**10-fold cross validation**
Utku, A. [11]	**2023**	**DT,KNN,LR,RF,SVM**	**GridSearchCV**	**Accuracy: 80.48%**	**418 patient** **19 variable**	**2**	**10-fold cross validation**

DOI: https://doi.org/10.54216/MOR.030101
Received: October 16, 2024 Revised: December 01, 2024 Accepted: December 15, 2024

Arya, G., [12]	**2023**	**LR,XGBOOST**	**Genetic Algorithm**	**Accuracy: 90.5%**	**424 patient** **18 variable**	**4**	**20-fold cross validation**
Güneş, O. M., [13]	**2023**	**MLP-ANNs, DT, (RF), (NB), SVM, K-NN, and (LR).**	**Boruta method,** **Feature selsction**	**87.75%)**	**2000 patient** **25 variable**	**2**	**10-fold cross validation**
Ganty, J.[14]	**2022**	**NB,SVM,CART**	**Swarm Optimization algorithm**	**Accuracy: 75%**	**583 patient** **11 variable**	**2**	**10-fold cross validation**

Funding: "This research received no external funding"

Conflicts of Interest: "The authors declare no conflict of interest."

References

[1] Dickson, E., Grambsch, P., Fleming, T., Fisher, L., & Langworthy, A. (2023). Cirrhosis patient survival prediction. UCI Machine Learning Repository. https://doi.org/10.24432/C5R02C5R02G.

[2] Chen, Y.-Y., Lin, C.-Y., Yen, H.-H., Su, P.-Y., Zeng, Y.-H., Huang, S.-P., & Liu, I.-L. (2022). Machine-learning algorithm for predicting fatty liver disease in a Taiwanese population. Journal of Personalized Medicine, 12(7), Article 1026. https://doi.org/10.3390/jpm12071026.

[3] de Mello, M. L., de Oliveira, J. C. L., de Souza, M. C., & de Almeida, H. M. (2018). Predictive modeling of fatty liver disease using machine learning techniques. Journal of Biomedical Informatics, 83, 188-196. https://doi.org/10.1016/j.jbi.2018.05.017

[4] Zemouri, R., Zerhouni, N., & Racoceanu, D. (2019). Deep learning in the biomedical applications: Recent and future status. Applied Sciences, 9(15), 1526. https://doi.org/10.3390/app9081526.

[5]Sun, S., Cao, Z., Zhu, H., & Zhao, J. (2020). A survey of optimization methods from a machine learning perspective. IEEE Transactions on Cybernetics, 50(8), 3668-3681. https://doi.org/10.1109/TCYB.2019.2950779

[6] Kamath, S. K., Pendekanti, S. K., & Rao, D. (2024). LivMarX: An optimized low-cost predictive model using biomarkers for interpretable liver cirrhosis stage

classification. IEEE Access, 12, 92506–92522. https://doi.org/10.1109/ACCESS.2024.3422451

[7] Badvath, D., Miriyala, A. S., Gunupudi, S. C. K., & Kuricheti, P. V. K. (2024). ONBLR: An effective optimized ensemble ML approach for classifying liver cirrhosis disease. Biomedical Signal Processing and Control, 89, 105882. https://doi.org/10.1016/j.bspc.2023.105882

[8] Ali, D. S., & Aljabery, M. A. (2024). Predicting liver cirrhosis stages using Extra Trees, Random Forest, and SVM with data mining techniques. Informatica, 48(21), 15–26. https://doi.org/10.31449/inf.v48i21.6752

[9]Choi, Y.-S., & Oh, E. (2024). Investigating machine learning-based algorithms for liver cirrhosis prediction. Advances in Engineering and Intelligence Systems, 3(1), 117-126. https://aeis.bilijipub.com

[10]Alotaibi, A., Alnajrani, L., Alsheikh, N., Alanazy, A., Alshammasi, S., Almusairii, M., Alrassan, S., & Alansari, A. (2023). Explainable ensemble-based machine learning models for detecting the presence of cirrhosis in hepatitis C patients. Computation, 11(6), 104. https://doi.org/10.3390/computation11060104

[11] Utku, A. (2023). Deep learning based cirrhosis detection. Operational Research in Engineering Sciences: Theory and Applications, 6(1), 95-114. https://doi.org/10.31181/oresta/060105Utku/Oper.
Res. Eng. Sci. Theor. Appl. 6(1)2023 95-114

[12] Arya, G., Bagwari, A., Saini, H., Thakur, P., Rodriguez, C., & Lezama, P. (2023). Explainable AI for enhanced interpretation of liver cirrhosis biomarkers. IEEE Access, 11, 123729-123745. https://doi.org/10.1109/ACCESS.2023.3329759

[13] Güneş, O. M., Kasap, P., & Çorba Zorlu, B. S. (2023). The comparison of machine learning classification algorithms used to diagnose liver cirrhosis disease and a brief review. Concurrency and Computation: Practice and Experience, 35, e7628. https://doi.org/10.1002/cpe.7628

[14] Ganty, J., Wajiga, G. M., Malgwi, Y. M., & Maidabara, A. H. (2022). A diagnostic model for the prediction of liver cirrhosis using machine learning techniques. Computer Science & IT Research Journal, 3(1), 36-51. https://doi.org/10.51594/csitrj.v3i1.296

[15] https://www.kaggle.com/code/varshaakshir/liver-cirrhosis-prediction/input

[16] Fedesoriano. (Aug. 2021). Cirrhosis Prediction Dataset. [Online]. Available: https://www.kaggle.com/fedesoriano/cirrhosis-prediction-dataset

DOI: https://doi.org/10.54216/MOR.030101
Received: October 16, 2024 Revised: December 01, 2024 Accepted: December 13, 2024