Liver Disease Detection Using Machine Learning: A Comparative Study of XGBoost, LightGBM, and CatBoost

Akhilesh Thangellapally
Master of science in Data Analytics
National College of Ireland
Dublin, Ireland

***Abstract*—This work aims at identifying a liver disease detection system applying the Method of Machine Learning on clinical data and blood values. The goal for the system is to group the patients into two groups: "Liver Disease," which is coded by (1), and "No Liver Disease," coded by (2). The evidence used for training includes the medical essential parameters like bilirubin level, Alkaline Phosphatase, AST, ALT, proteins, albumin, and the like, which are definitive to the health of the liver. Three ML algorithms, XGBoost, LightGBM, and CatBoost, are utilized to estimate the probability of liver disease. Thus, the goal of the project is to show the advantage over the simple, time-tested conventional approaches by using several models created on the Ensemble learning approach, which is known for its high performance. Evaluation is based on common Indicators like Accuracy, Precision, Recall, F1-Score, ROC-AUC, and RMSE. Thus, the project's goal is to design an efficient and effective system that would help diagnose liver diseases in their initial stages and prevent further evolution of the disease.**

***Index Terms*—Liver Disease Detection, Machine Learning (ML), XGBoost, LightGBM, CatBoost, Ensemble Learning, Classification Models, Predictive Modeling, Clinical Data, Binary Classification, Accuracy, ROC-AUC, Precision, Recall, F1-Score, RMSE, Data Preprocessing, Feature Engineering, Cross-Validation, Health Informatics, Medical Diagnosis, Liver Function Tests (LFTs), Early Disease Detection, Dataset Imbalance, Predictive Analytics.**

I. Introduction

The liver is the largest internal organ of the human body and it is involved in the metabolism of blood chemicals and waste. It is an essential organ since infections with viruses and parasites contribute to inflammation and liver disease (LD). LD is one of the major clinical problems and a leading cause of morbidity and mortality, especially thanks to such comorbidities as diabetes, obesity, and metabolic syndrome. Examining illness causes and as a dual explanation, it can be suggested that early detection is crucial for efficient diagnosis, prevention, and treatment, although the sole reliance on a particular diagnostic method is insufficient. Liver Function Tests (LFTs) are useful in diagnosing the extent of liver damage, even though they cannot by themselves establish a diagnosis. In this regard, diagnosis methods/tools based on the computer, including ML, can contribute to the accurate diagnosis of liver diseases. The blinders of knowledge discovery of ML have enabled dealing with important information to improve the decision, both in the diagnosis and prognosis of diseases. The rationale for undertaking this research is due to the inadequate liver disease diagnosis in the current healthcare delivery systems. This study shall develop a predictive model to categorize the patients into two groups of "Liver Disease" and "No Liver Disease" using machine learning algorithms. Hence, the three chosen models for this study are XGBoost, LightGBM, and CatBoost due to their efficiency and effectiveness in classification models. These algorithms will be tested on the provided dataset with accuracy, precision, recall, F1-score, ROC-AUC, and RMSE. The aim is to recognize the most effective model to distinguish liver diseases from the rest, perhaps leading to improved diagnosis and results of the illness.

II. Related Work

As per the statistics, liver disease continues to be a serious global health issue that results in millions of deaths every year. The early detection is a key factor permitting an effective treatment and a better prognosis. In the past, clinical examination, physical examination, and laboratory tests like the liver function tests (LFTs) such as bilirubin, alkaline phosphatase and alanine aminotransferase are used to diagnose the disease of the liver. Nevertheless, these approaches are limited by overlapping symptoms with other disorders, human error in analysis and lack of forecast of the disease severity [3]. Machine learning (ML) has led to new disease detection and prognosis possibilities in the recent years of artificial intelligence. Historical clinical data of patients can be used to learn by the ML algorithms to make accurate predictions, which turns out to be a beneficial tool in the healthcare domain. This domain has been explored by various studies, and most of them, particularly using Indian Liver Patient Dataset (ILPD) containing clinical data of liver patients and non patients. Authors have also used "decision trees", "support vector machine (SVM)", "logistic regression", "neural networks" to detect liver disease in similar works [1]. For

example, SVM is applied and it gets (about) 75% accuracy. However, Random Forests produced better results using the ensemble techniques. Unfortunately, these models often required thorough hyperparameter tuning and were quite dependent on class imbalance. To tackle the issues posed above, recent approaches have started to use the gradient boosting techniques like "XGBoost", "LightGBM" or "CatBoost". They are known to have high performance, overfitting robustness, and they can deal with missing values and categories as features. Moreover, CatBoost turns out useful in medical datasets, among others, when the features are not numerical. Taking the above advancements in the context, the present project takes the step forward by performing an explicit comparison of many ensemble-based classifiers including XGBoost, LightGBM, CatBoost, and Gradient Boosting on the ILPD dataset. It is to classify the individuals into two categories (with or without liver disease), based on important medical parameters. Specific features in the dataset are "total bilirubin", "direct bilirubin", "alkaline phosphatase", "serum glutamic pyruvic transaminase (SGPT)", "serum glutamic-oxaloacetic transaminase" "(SGOT)"", "total protein, albumin", and "albumin globulin ratio" [4]. The models are trained and evaluated by metrics such as accuracy, precision, recall, F1 score, areas under curve of Receiver Operating Characteristics "(ROC-AUC)" after some preprocessing steps like label encoding and standardization. In contrast to past studies, this work groups together several very powerful ML algorithms and touts them through a comprehensive evaluation process. In addition, it also performs cross validation to avoid overfitting and has a generalizability. It consequently attempts to furnish an improved and accurate liver disease prediction system to support physicians in making decisions.

III. METHODOLOGY

A. *Selection of Methodology*

In using clinical features to diagnose liver diseases, given that this is a clinical problem, an idiographic approach was followed, as a result of which, what was implemented was not strictly categorized as CRISP-DM. According to the principles of the most used and respected process model for data mining – the CRISP-DM, this project is being executed. It is therefore made up of six cyclic steps, namely "Business Understanding", "Data Understanding", "Data Preparation", "Modeling", "Evaluation", and "Deployment". The reason for choosing CRISP-DM was that it is a rigorous framework suitable for healthcare analytics and allows for multiple experimental iterations and improves the model. In the Business Understanding phase, the objective was specified: Predictive model for liver disease diagnosis and patient data. At the Data Understanding step, the "Liver Patient Dataset (LPD)" was reviewed to recognize data types, distribution, and include missing values. The three algorithms, namely, XGBoost, LightGBM, and CatBoost, were used for comparison because the algorithms perform well with tabular clinical data. Optimization of the hyperparameter was done using GridSearchCV and RandomizedSearchCV to improve the model performance with no overfitting [6].

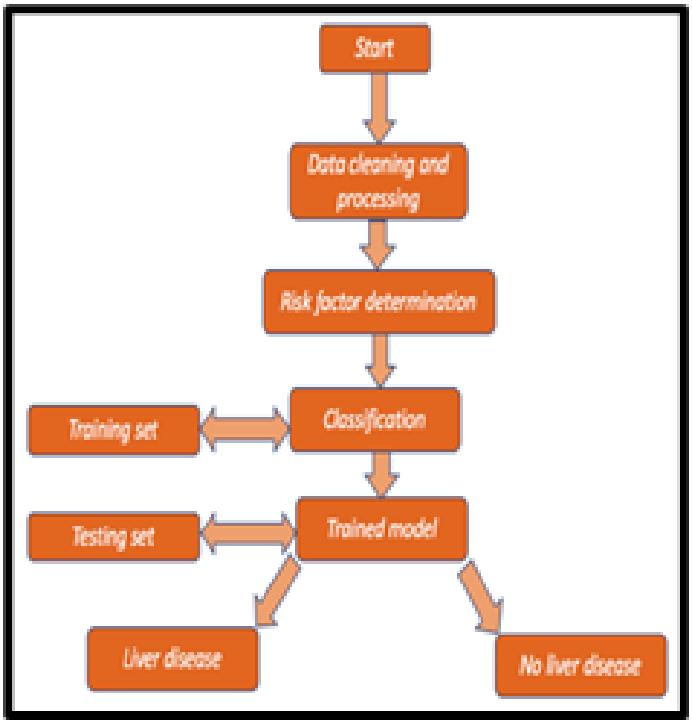


Fig. 1. CRISP-DM Process Flow Applied to Liver Disease Prediction

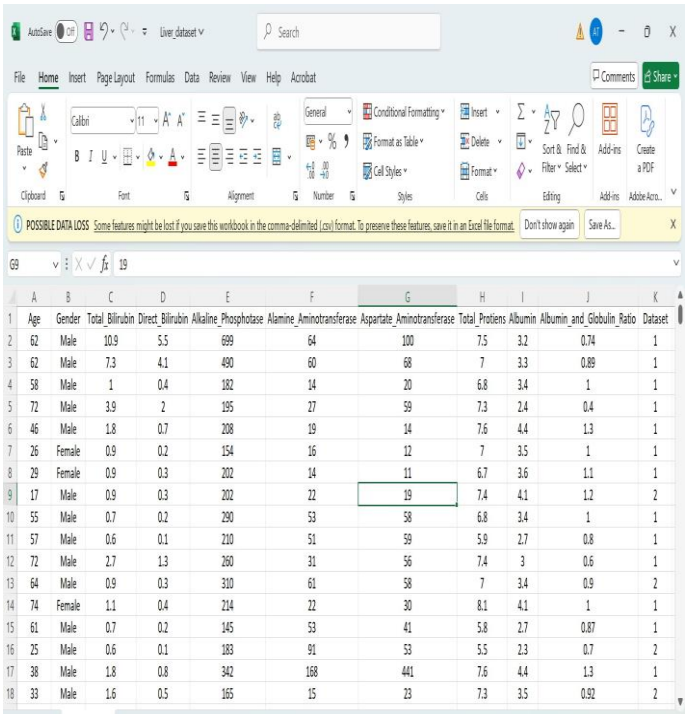
The algorithms' diagnostic performance was assessed by the conventional measures such as Accuracy, Precision, Recall, F1-score, ROC-AUC, and RMSE. This ad hoc methodology fits perfectly with the addressed problem and the present practice in the data science field.

B. *Dataset Description*

Liver disease prediction involves the use of clinical and laboratory data of patients and is included in the dataset. The matrix includes diverse evaluations of the liver functions, which include bilirubin, liver enzymes, proteins, and albumin, among others. These parameters are very essential in the diagnosis of liver diseases. There is information on quantitative and qualitative variables such as Total Bilirubin, Alkaline Phosphatase, and Gender, and the quantitative variable is whether a patient has liver disease. Key features include:

- **Attributes:** The dataset attributes I used to build the model include age, gender, total bilirubin, direct bilirubin, alkaline phosphatase, alanine aminotransferase, aspartate aminotransferase, total proteins, albumin, and albumin/globulin ratio.
- **Target variable:** The Target column of the Dataset shows 1 for disease and 2 for the absence of the disease of the liver. It is a combination of numeric and non-numeric variables, hence making a few modifications in the data, such as handling the missing values, constructing the categories of independent variables, and standardizing the variables.

In the above Figure, basic characteristics of the used dataset are depicted, which include age, sex, bilirubin, liver enzymes,



	A	B	C	D	E	F	G	H	I	J	K
1	Age	Gender	Total_Bilirubin	Direct_Bilirubin	Alkaline_Phosphotase	Alamine_Aminotransferase	Aspartate_Aminotransferase	Total_Protiens	Albumin	Albumin_and_Globulin_Ratio	Dataset
2	62	Male	10.9	5.5	699	64	100	7.5	3.2	0.74	1
3	62	Male	7.3	4.1	490	60	68	7	3.3	0.89	1
4	58	Male	1	0.4	182	14	20	6.8	3.4	1	1
5	72	Male	3.9	2	195	27	59	7.3	2.4	0.4	1
6	46	Male	1.8	0.7	208	19	14	7.6	4.4	1.3	1
7	26	Female	0.9	0.2	154	16	12	7	3.5	1	1
8	29	Female	0.9	0.3	202	14	11	6.7	3.6	1.1	1
9	17	Male	0.9	0.3	202	22	19	7.4	4.1	1.2	2
10	55	Male	0.7	0.2	290	53	58	6.8	3.4	1	1
11	57	Male	0.6	0.1	210	51	59	5.9	2.7	0.8	1
12	72	Male	2.7	1.3	260	31	56	7.4	3	0.6	1
13	64	Male	0.9	0.3	310	61	58	7	3.4	0.9	2
14	74	Female	1.1	0.4	214	22	30	8.1	4.1	1	1
15	61	Male	0.7	0.2	145	53	41	5.8	2.7	0.87	1
16	25	Male	0.6	0.1	183	91	53	5.5	2.3	0.7	2
17	38	Male	1.8	0.8	342	168	441	7.6	4.4	1.3	1
18	33	Male	1.6	0.5	165	15	23	7.3	3.5	0.92	2

Fig. 2. A portion of the Liver_dataset showing all the clinical parameters

proteins, and the target label representing whether there is liver disease or not. All these features are very important in diagnosing conditions affecting the liver and are the foundation in the training and prediction of the model in this study. The ultimate learning objective is to use the aforementioned features to perform a binary classification on patients to indicate whether they have liver diseases.

C. *Data Preprocessing*

Data preprocessing is an important phase of machine learning to make an efficient and accurate model and create more generalizable models. In the present study, utilizing the "Liver Patient Dataset ILPD), there are some preprocessing practices used for predicting liver disease further to the following.

- Firstly, categorical variables were handled. In the first step, the "Gender" variable, containing textual data, Male and Female, was converted into numbers as 0 and 1 by applying LabelEncoder on the "Gender" column using the map method. This conversion helps in the interpretation of the data, which is in the form of numbers that are suitable for the machine learning algorithms.
- Before analyzing the data, the target variable, which is originally given under the name of "Dataset," was renamed to "Target". It was then binarized, where 1 represented patients who have liver diseases while the remaining were represented as 0.
- The numerous missing values were handled by imputing the mean of each column to retain the principal format of the structure and to avoid excluding records.

Features were then defined as the input variables (X), while dependent variables or the prediction targets were considered labels (y). The data set was then divided into the training set and the testing set based on the 80:20 split. As a last preparation step, all of them were normalized through StandardScaler to minimize feature value ranges and enhance the model learning.

D. *Machine Learning Models Applied*

In this case, three selected gradient boosting algorithms were taken into account: XGBoost, LightGBM, and CatBoost. These models are particularly suitable for the analysis of structured medical datasets because they are less sensitive to input noise, fast, and accurate in classification problems [7]. LightGBM and XGBoost are suitable for large datasets with parallel computation options, while, compared to others, CatBoost is better at handling categorical features such as gender, without requiring categorical encoding. In order to reduce overfitting and improve the model's performance, cross-validation for model selection was performed using "GridSearchCV" and "RandomizedSearchCV". The learning rate, maximum tree depth, and the number of estimators, as well as the regularization terms, were adjusted. In the training process, early stopping was adopted to prevent evaluation progress when its outcomes reached their highest level. This regular tuning was useful for generalizing the models and achieving higher efficiency on unseen scenarios.

E. *Evaluation Metrics*

In evaluating the performance of the ML models, various measures of classification are used to guide the assessment. Accuracy gives a general percentage of correctly classified samples and can be thought of as multidimensional cross-validation. Specificity concentrates on non-positives not being classified as positive, or in other words, true negatives are maintained so that only genuine positives are predicted. Recall, also known as sensitivity, measures the model's effectiveness at identifying all actual positives within the data set [8]. The F1 score is better than precision and recall for cases where the number of records in one class is much smaller than the other class. ROC-AUC indicates the separation between the rates of the two classes, and RMSE indicates the prediction error. All these can be said to provide a holistic picture of the effectiveness of models.

IV. Results & Discussion

A. *Results*

The main goal of this project is to develop an ML model for liver disease detection through clinical and blood test features. This means that the variable "Liver disease" can be coded as "Yes" if the patient has liver disease or "No" if the patient has no record of liver disease. In this case, different validation parameters such as accuracy, precision, recall, F1, ROC-AUC, and RMSE were used.

According to the findings, LightGBM has the highest accuracy of 96.70% & 96.42% , a precision of 97.31% &

Model	Accuracy	Precision	Recall	F1-Score	ROC-AUC	RMSE
Gradient Boosting	0.849057	0.841176	0.966216	0.899371	0.940456	0.388514
LightGBM	0.966981	0.973154	0.97973	0.976431	0.983953	0.181711
CatBoost	0.915094	0.911392	0.972973	0.941176	0.969911	0.291386

Fig. 3. Dataset 1:Model Performance Comparison

Model	Accuracy	Precision	Recall	F1-Score	ROC-AUC	RMSE
Gradient Boosting	0.835821	0.85283	0.933884	0.891519	0.89585	0.40519
LightGBM	0.964179	0.96748	0.983471	0.97541	0.976851	0.189264
CatBoost	0.928358	0.929134	0.975207	0.951613	0.961299	0.26766

Fig. 4. Dataset 2:Model Performance Comparison

96.74%and a recall of 97.97% & 98.34%, with an F1-score of 97.64% & 97.54%. It also yields the optimum value of ROC-AUC that is 98.40 & 97.68 percent, which shows that the concept is very efficient in discriminating between patients with liver disease and patients without liver disease. As Fig 3 and 4 shows, LightGBM has the lowest RMSE (0.1817 & 0.1892), which proves the higher accuracy of the model. The model uses the Gradient Boosting method with an accuracy of 84.91% & 83.58%; however, precision rate are lower and equals 84.12% & 85.28% , and the F1-score equals 89.94 & 89.15 percent as a minimum, in contrast to LightGBM. However, its recall (96.62% & 93.38 %) shows that it is sensitive enough to identify the patients with liver disease, though with higher false positive rates. CatBoost's performance is also good, as it achieves 91.51% & 92.83 % of accuracy, though overall, it shows less performance than LightGBM. For the liver disease cases, the model had a recall of 97.30% & 97.52 % and an F1-score of 94.12% & 95.16% ; thus, even though it did not provide the LightGBM level of precision, it was deemed accurate.

The three confusion matrices give information regarding the performance of the models in classification by Gradient Boosting, LightGBM, and CatBoost.

- Gradient Boosting depicted in Figure 1 shows mild performance with TP 143 while TN is 37, and has a problem in labeling negatives as positives, resulting in high FP.
- LightGBM has the best performance with 145 TP and 60 TN, 4 FP and 3 F,N which stands for good model precision and recall.
- CatBoost also gives quite good results with 144 of TP and 50 of TN; however, it is behind LightGBM with 14

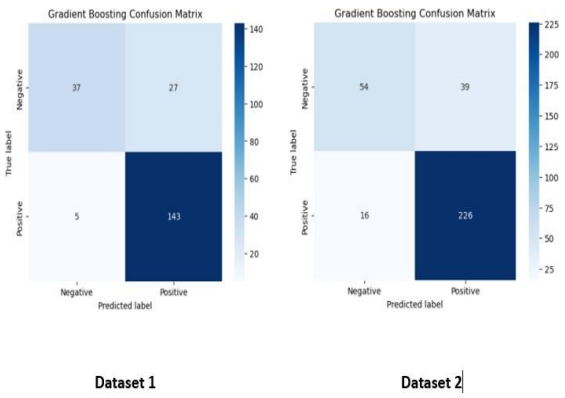


Fig. 5. Gradient Boosting Confusion Matrix

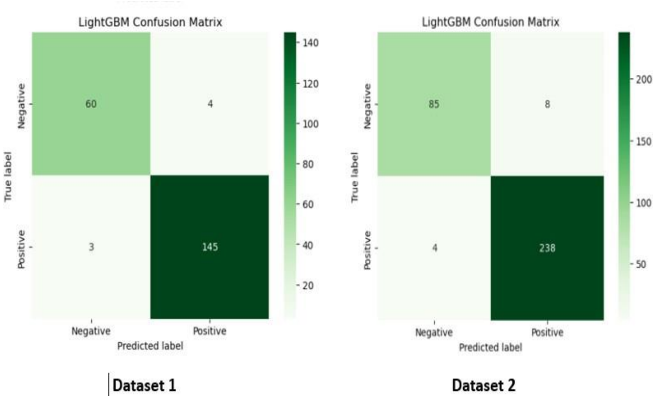


Fig. 6. LightGBM Confusion Matrix

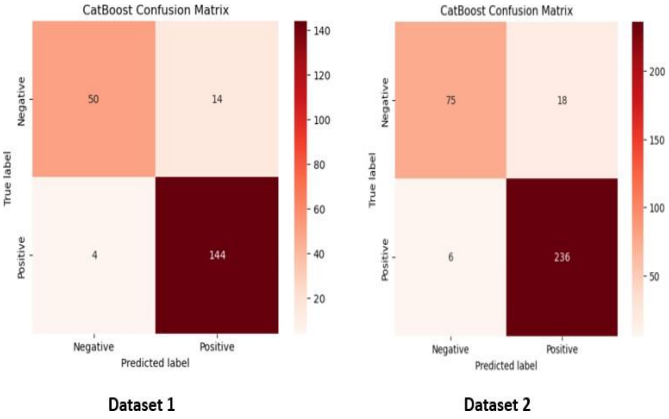


Fig. 7. Catboost Confusion Matrix

FP and 4 FN.

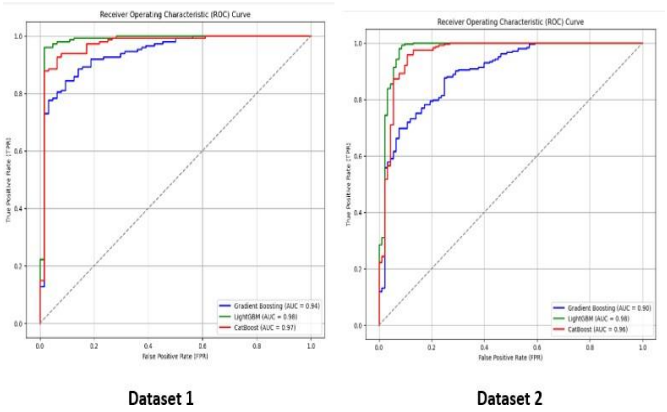


Fig. 8. Comparative ROC Curve for the chosen three ML Models

The above figure shows the classification effectiveness of each model in terms of one class against the other. From the figure, LightGBM has the highest area under the curve (AUC = 0.98) than the other two algorithms, the CatBoost having AUC = 0.97, while "Gradient Boosting" has a slightly lower AUC of 0.94. Again, it is understood that the higher value of AUC indicates better classification of positive and negative situations by the model. The diagonal grey line is a reference line of a random classifier with an AUC value of 0.5, which proves that the models have a higher level of classification. The results at large indicate that blood parameters and clinical data predict the degree of liver disease with machine learning algorithms such as "LightGBM", "CatBoost", and "Gradient Boosting", "LightGBM" demonstrated the highest accuracy and AUC among these models, proving it to be the most efficient model in distinguishing between liver disease and other non-liver disease cases. Two additional parameters, bilirubin, enzymes, and proteins, help the system to make predictions that are more precise and accurate as compared to the traditional models. This study proves that the use of ML models in liver disease diagnosis helps in the early diagnosis of the disease, hence enhancing the patient's quality of life.

*B. **Discussion***

The liver disease detection system developed in this project reflects a practical application of machine learning in clinical diagnosis. We preprocessed the dataset to encode the categorical variables as well as standardize the numerical features to avoid inconsistencies to affect model's performance. Since the classification objective, the target variable was binary encoded to signify the presence or absence of liver disease [9]. XGBoost, LightGBM, CatBoost and Gradient Boosting Classifiers were used as 4 advanced machine learning models. However, ensemble models are particularly effective when the data is structured and of high dimensions with intricate features relationships. CatBoost was among them which stood out since it has native support for categorical and has less sensitivity to overfitting [10]. These models were evaluated using accuracy, precision, recall, F1 score and ROC AUC for the evaluation. The ROC-AUC score and accuracy showed that the CatBoost model is the best out of all models provided, as it is better at differentiating positive and negative classes. Further, the use of StratifiedKFold cross validation helped in maintaining a class balance in training and testing phases itself, thus making these findings safe. The comparative analysis of multi algorithms considered in this project is one of the main strengths of this project because it gives a complete view of the performances of those algorithms. Although the dataset has limited size and could be class imbalanced, it does mean generalizability is possibly influenced

V. CONCLUSION AND FUTURE WORK

The development of a "machine learning" based system for the detection of liver disease using clinical as well as blood parameter data was successful. The integrity of the data was ensured by the preprocessing steps while ensemble classifiers improved the model robustness. LightGBM was found to work best of the evaluated models, since its categorical feature handling and boosting mechanism gave it superior accuracy, generalization ability over the other evaluated models. It classifies patients to the two groups of having liver disease and not having liver disease, which facilitates quick and robust diagnosis aid. It minimizes people's dependence on manual interpretation and automates the diagnostic process, bringing about improvement in medical decision making [11]. However, some limitations such as dataset size and class imbalance are acknowledged by the results. Further strengthening the model might include expanding the dataset with other patient demographics or longitudinal data. On the whole, machine learning is presented in this liver disease detection system as a means by which it can transform clinical practice, helping lay the path for future healthcare AI applications pertaining to early diagnoses and best practice treatment methods.

Using deep learning techniques or hybrid models that com- bine a ML component with a rule-based model that could improve prediction accuracy.

Incorporating the model into a clinical decision support system (CDSS) to allow predictions in real time in an clinical context.

It would also be enhanced by including additional attributes in the system such as patient history, medications etc., through improved predictions of disease progression and/or personal- ized treatment recommendations.

REFERENCES

[1] K. Gupta, N. Jiwani, N. Afreen and D. D, "Liver Disease Prediction using Machine learning Classification Techniques," 2022 IEEE 11th International Conference on Communication Systems and Network Technologies (CSNT), Indore, India, 2022, pp. 221-226, doi: 10.1109/CSNT54456.2022.9787574. keywords: Machine learning algorithms;Liver diseases;Hospitals;Prediction algorithms;Boosting;Feature extraction;Classification algorithms;LR;RF;DT;GB;XGB;LGB;SMOTE,

[2] N. D. Yadav and V. K. Singh, "Proposing Machine Learning as the technology to Deal with the Liver Disease," 2022 OPJU International Technology Conference on Emerging Technologies for Sustainable Development (OTCON), Raigarh, Chhattisgarh, India, 2023, pp. 1-6, doi: 10.1109/OTCON56053.2023.10114018. keywords: Support vector machines;Liver diseases;Machine learning;Programming;Libraries;Encoding;Kernel;SVM;ML;AI;Python;Kaggle,

[3] R. Kalaiselvi, K. Meena and V. Vanitha, "Liver Disease Prediction Using Machine Learning Algorithms," 2021 International Conference on Advancements in Electrical, Electronics, Communication, Computing and Automation (ICAECA), Coimbatore, India, 2021, pp. 1-6, doi: 10.1109/ICAECA52838.2021.9675756. keywords: Proteins;Cloud computing;Machine learning algorithms;Sensitivity;Adaptive systems;Liver diseases;Medical services;Data Mining;Adaptive Neuro-Fuzzy Inference System;KNN;Decision Tree,

[4]] G. S. Mate, P. Sutar, V. Mhaske, O. Jaybhaye and Y. Wagh, "Demystifying Liver Disease Prediction: The Role of PSO Algorithm," 2024 OPJU International Technology Conference (OTCON) on Smart Computing for Innovation and Advancement in Industry 4.0, Raigarh, India, 2024, pp. 1-12, doi: 10.1109/OTCON60325.2024.10685852. keywords: Surveys;Support vector machines;Technological innovation;Machine learning algorithms;Accuracy;Liver diseases;Medical services;Machine Learning;Particle Swarm Optimization (PSO);Disease Prediction;Liver Disease;Healthcare,

[5] Mostafa F, Hasan E, Williamson M, Khan H. Statistical Machine Learning Approaches to Liver Disease Prediction. Livers. 2021; 1(4):294-312. https://doi.org/10.3390/livers1040023

[6] Akter, S., Shekhar, H.U. and Akhteruzzaman, S., 2021. Application of biochemical tests and machine learning techniques to diagnose and evaluate liver disease. Advances in Bioscience and Biotechnology, 12(6), pp.154172. doi: https://www.scirp.org/pdf/abb_2021062409473676.pdf

[7] Fathi, M., Nemati, M., Mohammadi, S.M. and Abbasi-Kesbi, R., 2020. A machine learning approach based on SVM for classification of liver diseases. Biomedical Engineering: Applications, Basis and Communications, 32(03), p.2050018. doi: https://www.worldscientific.com/doi/abs/10.4015/S1016237220500180

[8] Gupta, Ketan & Jiwani, Nasmin & Afreen, Neda & D, Divyarani. (2022). Liver Disease Prediction using Machine learning Classification Techniques. 221-226. 10.1109/CSNT54456.2022.9787574.

[9] M. Hassan Mohamed et al., "Toward an Accurate Liver Disease Prediction Based on Two-Level Ensemble Stacking Model," in IEEE Access, vol. 12, pp. 180210-180237, 2024, doi: 10.1109/ACCESS.2024.3459429. keywords: Liver diseases;Accuracy;Feature extraction;Stacking;Predictive models;Diseases;Classification tree analysis;Machine learning;Ensemble stacking;feature selection;ILPD dataset;liver disease prediction;machine learning,

[10] S. Mathur, P. Karodi and R. Dhanare, "Fatty Liver Disease Prediction Through Machine Learning," 2024 3rd International Conference on Sentiment Analysis and Deep Learning (ICSADL), Bhimdatta, Nepal, 2024, pp. 366-371, doi: 10.1109/ICSADL61749.2024.00066. keywords: Machine learning algorithms;Accuracy;Liver diseases;Artificial neural networks;Predictive models;Prediction algorithms;Data models;Fatty Liver Disease (FLD);XGBoost (XGB);Naive Bayes (NB);Random Forest (RF);Artificial Neural Network (ANN),

[11] V. Srihitha, M. A. Goud, S. Cherishma, M. Yesubabu and G. Suryanarayana, "Enhanced Liver Disease Prediction Using Ensemble Learning," 2024 IEEE 6th International Conference on Cybernetics, Cognition and Machine Learning Applications (ICCCMLA), Hamburg, Germany, 2024, pp. 123-128, doi: 10.1109/ICCCMLA63077.2024.10871784. keywords: Machine learning algorithms;Liver diseases;Data integration;Medical treatment;Predictive models;Prediction algorithms;Data models;Real-time systems;Planning;Reliability;AdaBoost;CatBoost;LightGBM;XGboost,

[12] G. S. Mate, P. Sutar, V. Mhaske, O. Jaybhaye and Y. Wagh, "Demystifying Liver Disease Prediction: The Role of PSO Algorithm," 2024 OPJU International Technology Conference (OTCON) on Smart Computing for Innovation and Advancement in Industry 4.0, Raigarh, India, 2024, pp. 1-12, doi: 10.1109/OTCON60325.2024.10685852. keywords: Surveys;Support vector machines;Technological innovation;Machine learning algorithms;Accuracy;Liver diseases;Medical services;Machine Learning;Particle Swarm Optimization (PSO);Disease Prediction;Liver Disease;Healthcare.

[13] Sharma, V., Samant, S. S., Singh, T., Fekete, G. (2024). An Integrative Framework for Healthcare Recommendation Systems: Leveraging theLinear Discriminant Wolf–Convolutional Neural Network (LDW-CNN) Model. Diagnostics, 14(22), 2511.

[14] Gururaj, H. L., Flammini, F., Srividhya, S., Chayadevi, M. L., Selvam, S. (Eds.). (2024). Computer Science Engineering: Proceedings of the 1st International Conference on Computing and Intelligent Information Systems (ICCIIS 2024), Bangalore, India, 19-20th April, 2024 Volume 1. CRC Press.