Evaluation Based Approaches Liver Diseases Prediction Using Machine Learning Algorithm

[1]Dr.Sneha Jondhle, [2]Mr.Harsh Jondhle, [3]Mr. Devesh Pandurang Patil, [4]Mr. Ajay Shivaji Bangar, [5]Mr. Subodh Manik Gawali.

[1]Associate Prof, [2]Asst.Prof, [3,4,5]UG Student, [1,2]EXTC Engg, [3,4,5]Computer Engg.

[1]Shivajirao S. Jondhale College of Engineering, Dombivili, [2,3,4,5]Shivajirao S. Jondhle College of Engineering & Technology, Asangaon, Maharashtra, India.

[1]snehajondhle@gmail.com, [2]harshjondhle@gmail.com, [3]patildevesh943@gmail.com, [4]ajaybangar673@gmail.com, [5]subodhgawali2003@gmail.com

Abstract - One of the most important aspects of human existence is the well-being of those free of liver tumors. Consequently, early diagnosis of liver illness is essential for improved therapy. It is extremely challenging for medical professionals to diagnose an ailment in its early stages based on minor symptoms. For many, the indicators become apparent only after it is too late. In order to combat this epidemic, the current effort attempts to improve the understanding of liver disease through the use of machine learning techniques. The main goal of the current work was to develop algorithms for liver dataset classification of healthy individuals. This study compares the categorization algorithms and provides prediction accuracy results based on their success variables.[1]

Key Word- Support vector machines, Machine learning algorithms, Liver disease, Prediction algorithms, Classification, Tumors.

I. INTRODUCTION

The scale of patient medical records increases day by day in the health care sector. Data mining is the method of using a computer-based information system (CBIS), using modern tactics, to uncover insights from data. Ensemble learning is a well-known method used for prediction by integrating multiple ensemble models of machine learning[5]. Aggregations of various classifiers are J48, C4.5 and Naive Bayes, etc. . Ensembles search for better outcomes than all of the simple classifiers[4]. If a virus infects the liver, substances that harm the liver are swallowed, or the immune system malfunctions, significant liver damage or dysfunction may develop, potentially leading to death.[7] Here is a high probability of liver failure among Indians. India is expected to become the World Capital of Liver Diseases by 2025. Liver disease patient are rapidly increasing due to the consumption of alcohol, intake of contaminated food and drugs that is widespread global. Because of the deskbound lifestyle, increased alcohol intake and smoking, the pervasive prevalence of infection inside liver in India is contributing around 100 forms of liver infections are present. It would also be of great value in the medical field to build a computer that will increase the diagnosis of the disease. systems that can assist physicians in making correct treatment choices, and the patient queue will also be minimized by liver specialists such as endocrinologists assisted by Automated categorization Methods for Disorders in Liver part. In medical diagnosis and disease prediction, classification techniques are widely common.[1]

II. AIMS AND OBJECTIVE

a) Aim

The aim of evaluation-based Approaches to predict liver disease via machine learning algorithms is to assess and improve the reliability and performance of models of machine learning in predicting liver diseases. These approaches are crucial for developing accurate, clinically relevant, and practical tools for diagnosing liver conditions.

b) Objective

Develop models for machine learning that can

accurately identify the presence of liver disease early on, allowing for timely medical intervention and treatment.

- Create prediction models that can assess the risk of liver disease for individuals based on their medical history, demographics, and other relevant factors, aiding in proactive healthcare management.
- Educate patients about their liver disease factors and potential risks interventions, promoting healthier lifestyles and increased awareness.

 DOI : 10.35291/2454-9150.2024.0270

III. LITERATURE SURVEY

Paper 1: Liver Cancer Prediction in a Viral Hepatitis Cohort: A Deep Learning Approach

Using Deep Learning to Predict Liver Cancer in a Viral Hepatitis Cohort. The main cause of liver illnesses, including liver cancer, which is the main cause of cancer-related mortality, is viral hepatitis. Unfortunately, therapy for this disease is typically challenging or downright impractical because it is discovered in its latter stages. Using a cohort of patients with hepatitis, this study used deep learning (DL) models to identify liver cancer early. In order to examine viral hepatitis patients from 2002 to 2010, it is polled one million randomly selected samples from the National Health Insurance Research Database (NHIRD). Then, based on the hepatitis cohort's medical history, this employed DL models to forecast instances of liver diseased In a cohort of people with hepatitis, the CNN model can accurately predict the development of liver cancer[2].

Paper 2: A hybrid intelligent system for medical data classification.

In order icd decision assistance for the categorization of medical data, this study looks at the effectiveness of An intelligent hybrid system that integrates the model of Random Forest, the categorization and Regression Tree, and the Fuzzy Min-Max neural network. Combining the finest features of each component model with its limitations is the aim of the hybrid intelligent system. It is said that the Fuzzy Min-Max neural network is responsible for its capacity to learn gradually from data samples; its ability to explain expected outputs is enabled by the Classification and Regression Trees; and its good classification performances are facilitated by the Random Forest. The UCI Repository of Machine Learning's Breast Cancer Wisconsin, Pima Indians Diabetes, and Liver Disorders benchmark collections of medical data are used to evaluate The hybrid intelligent system's effectiveness. The calculation of several practical performance measures, including area, sensitivity, specificity, and accuracy The Receiver's Operating Characteristic curve is employed in medical applications.

The outcomes undergo analysis and comparison with other methods' published results from the literature. The results of the experiments show that the hybrid intelligent system performs well when it comes to classifying medical data[3].

Paper 3: Survey on Machine Learning Techniques for the Diagnosis of Liver Disease.

Artificial Intelligence have several branches which includes expert systems, fuzzy systems, machine learning, etc. Machine Learning is of great use in the areas of engineering. Similar to expert systems, fuzzy systems also store experts knowledge and uses it in their system for processing the input and generating outputs. IN, a Fuzzy system is used to control the two state variables using some membership functions which are defined by experts. IN a Fuzzy Cluster Means (FCM) method for diagnosing Liver Disease which is the global health problem, is presented. FCM performs an important role for classification, and evaluation. The Liver is a organ which plays an vital role in all human being and presently there is no way

For restoring the lack of liver function. Liver disease patient are rapidly increasing due to the consumption of alcohol, intake of contaminated food and drugs that is widespread global.

Machine learning uses AI for generating predictive models more effectively and efficiently than conventional methods through detection of hidden patterns within the large data sets. Naive Bayes classifiers predicts the class labels by computing the likelihood of the observed features under each class, returning the class with maximum likelihood. K- nearest determines output based on the value of the classes of K-nearest training samples.[6]

IV. EXISTING SYSTEM

The scale of patient medical records increases day by day in the health care sector. Major issues deliberated on patients with liver disease are not readily detected at starting phase since that can usually operate even though it is partly impaired. An early detection of liver disorders will improve the rate of survival of the patient. There is a high probability of liver failure among Indians. it is very difficult to detect in early stages of the disease with high accuracy recovery of the disease.[1]

DISADVANTAGES OF EXISTING SYSTEM:

- cannot predict proper accuracy results.
- cannot predict liver disease in early stages.

Algorithm: KNN ,Random forest

V. COMPARATIVE STUDY

Table.1: Comparative Analysis

Paper Title	Author	Year	Publication	Description
Evaluation based Approaches for Liver Disease Prediction Machine Learning Algorithms	C.Geetha Dr. AR. Arunachalam	**2021**	ICCCI	Model for predicting liver illness using logical regression and a support vector machine.
Survey on Machine Learning Techniques for The Diagnosis of Liver Disease	Prof. Satish Manje, Miss.Tannaz Inamdar, Miss.Shivani Kamalekar, Miss. Akanksha Shaha	**2022**	IJREAM	Compares machine learning techniques for diagnosing and predicting liver disease in medical fields, provides insights on different algorithms.

 DOI : 10.35291/2454-9150.2024.0270

Liver Cancer Prediction in a Viral Hepatitis Cohort: A Deep Learning Approach	Dinh-Van Phan.Chien-Lung Chan,Ai-Hsien Adams Li , Ting-Ying Chien,Van-Chuc Nguyen	**2020**	IJC	The results suggested that the Convolution Neural Networks model may accurately identify liver cancer.
Hybrid intelligent system for medical data classification	Manjeevan, Seera	**2014**	International Journal	Tasks involving the categorization of medical data are successfully completed by the hybrid system.

VI. PROBLEM STATEMENT

The main goal of this project is to develop a machine learning model that can predict the presence or risk liver illness in individuals based on relevant clinical and demographic factors using SVM and Logical regression Algorithm. High levels of precision, recall, accuracy, and generalizability across a range of patient groups should all be attained by the model.

Liver disease is a serious worldwide health concern that impacts millions of individuals worldwide. Effective treatment and better patient outcomes depend on early identification and precise diagnosis of liver disorders. Machine learning offers the potential to assist healthcare professionals in identifying individuals at risk of liver disease and providing timely interventions.[1]

VII. PROPOSED SYSTEM

The proposed work aims to enhance the predictive and classification quality of healthcare data by developing a hybrid predictive classifier model using the classifier ensemble. This project can help doctors make correct treatment choices, and the patient queue will also be minimized by liver specialists such as endocrinologists assisted

by Automated categorization Methods for Disorders in Liver part. In medical diagnosis and disease prediction, classification techniques are widely common. Michael J Sorich described on chemical datasets, the classification (SVM) Support Vector Machine and logistic regression provides better prediction results.

VIII. ALGORITHM

Step 1-:start

Step2-: For SVM

```
def start_svm():
 from sklearn import svm
 def svc_param_selection(X, y, nfolds):
 Cs = [0.001, 0.01, 0.1, 1, 10]
 gammas = [0.001, 0.01, 0.1, 1]
 param_grid = {'C': Cs, 'gamma': gammas}
 grid_search =
GridSearchCV(svm.SVC(kernel='rbf'), param_grid,
cv=nfolds)
grid_search.fit(X_train, y_train)
grid_search.best_params_
return grid_search.best_params_
 svClassifier = SVC(kernel='rbf')
 svClassifier.fit(X_train, y_train)
 svc_param_selection(X_train, y_train, 5)
 model = SVC(C=1, gamma=1)
 model.fit(X_train, y_train)
```

Step3-:logistic regression

```
def start_logistic_regression():
 tuned_params = {'C': [0.0001, 0.001, 0.01, 0.1, 1, 10,
100, 1000, 10000], 'penalty': ['l1', 'l2']}
model = GridSearchCV(LogisticRegression(),
tuned_params, scoring='roc_auc', n_jobs=-1)
model.fit(X_train, y_train)
model.best_estimator_
## Predict Train set results
y_train_pred = model.predict(X_train)
## Predict Test set results
y_pred = model.predict(X_test)
# Get just the prediction for the positive class
y_pred_proba =
model.predict_proba(X_test)[ 1]
```

IX. MATHEMATICAL MODEL

1) SVM

The Support Vector Machine (SVM) technique is a simple yet powerful supervised machine learning approach that may be used to develop regression and classification models. The SVM algorithm works effectively with both continuously and inconsistently separable datasets. Even with a small amount of data, the svm algorithm continues to demonstrate its brilliance.[5]

Length of Vectors

Length of vectors are also called as norms. It tells how far vectors are from the origin. Length of vector x(x1 , x2 , x3) is calculated as : ||x||+√(x 2+x2 2+x32) **Hyperplane** is plane that linearly divide the n-dimensional data points in two component. In case of 2D, hyperplane is line, in case of 3D it is plane.It is also called as n-dimensional line.

2) Logistic Regression

Length of vectors are also called as norms. It tells how far vectors are from the origin. Length of vector x(x1 , x2 , x3) is calculated as : ||x||+√(x 2+x2 2+x32) **Hyperplane** is plane that linearly divide the n-dimensional data points in two component. In case of 2D, hyperplane is line, in case of 3D it is plane.It is also called as n-dimensional line. $0 \leq hg(x) \leq$.[1]

 DOI : 10.35291/2454-9150.2024.0270

X. SYSTEM ARCHETECTURE

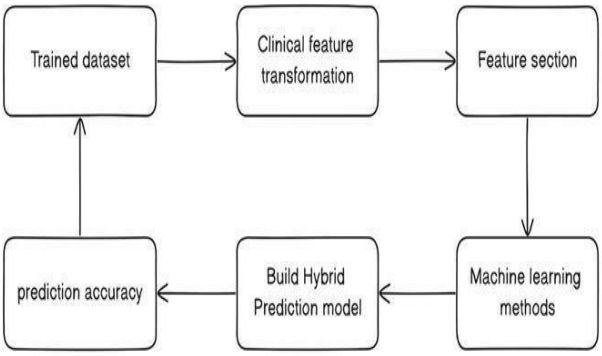

Fig.1: System Architecture

Explanation: The entire set of patient features and their corresponding values will be included in the dataset that is taken from the repository; correlated features that are essential for disease detection will be chosen from the dataset. Every classification model will receive the chosen dataset, and each model's performance measure

will be determined. Calculations for accuracy, precision, F1 score, recall, and ROC will be made and used in the analysis.

XI. ADVANTAGES

- The primary benefit of Algorithms for Machine Learning place of traditional predictive models is their ability to identify new patterns between variables and produce predictions by learning from preexisting data.
- It has been proved that MLAs raise the precision of identifying people who are at risk of illness.
- Learning methods with the assistance of a supervisor, teacher, or instructor are called as supervised learning. It consists of a training set of patterns linked to label data, which facilitates the learning and prediction of the algorithm from input to output.

XII. DESIGN DETAILS

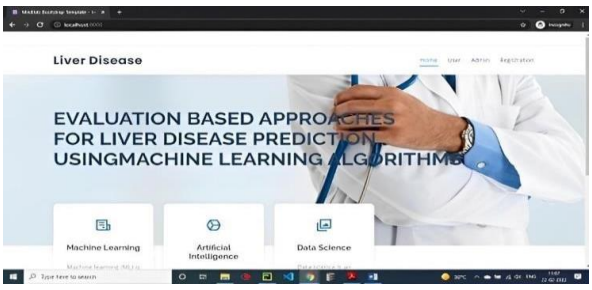

Fig 2 : home screen

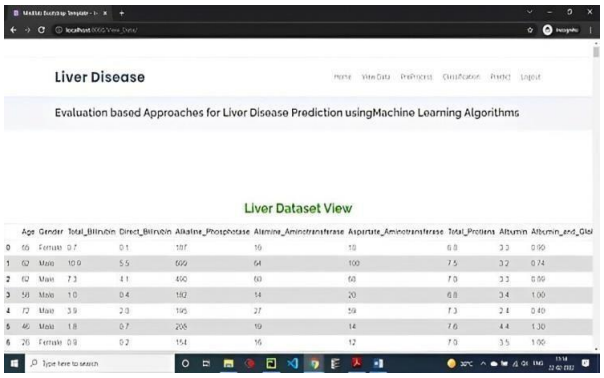

Fig 3: Liver Dataset view

The initiative intends to find the classification algorithm to predict hepatitis disease with the aid of algorithms for machine learning. Different machine learning models are used in this project for categorization and forecasting. The result analysis is done on the basis on the scores obtained by classification models of performance measures. Each models have different score. The prototype featuring theoverall highest score will be the best classifier and will be used for prediction.

XIII. CONCLUSION

Thus, we have tried to implement the paper "C. Geetha, Dr.AR.Arunachalam" "Evaluation based Approaches for Liver Disease Prediction Machine Learning Algorithms" 2021 IEEE Access and conclusion is as follows methods for identifying and evaluating liver disease in patients have been proposed and assessed utilizing machine learning methods. SVM, or logitistical regression, is the combination of two primary machine learning methods. The prediction analysis has been carried out using these models.

REFERENCES

[1] C. Geetha and A. Arunachalam, "Evaluation based Approaches for Liver Disease Prediction using Machine Learning Algorithms," 2021 International Conference on Computer Communication and Informatics (ICCCI), Coimbatore, India, 2021, pp.1-4.

[2] Dinh-Van Phan.Chien-Lung Chan,Ai-Hsien A Adams Li , Ting-Ying Chien,Van-Chuc Nguyen," Liver Cancer Prediction in a Viral Hepatitis Cohort: A Deep Learning Approach"2020 International journal of cancer(IJC) 10.1002/ijc.33245

[3] Manjeevan Seera, Chee Peng Lim, A hybrid intelligent system for medical data classification, Expert Systems with Applications 41 (2014) 2239– 2249.

[4] Zhi-Hua Zhou, Ensemble Methods: Foundations and Algorithms, Machine Learning & Pattern Recognition Series, 2012.

[5] Rakshith D B , Mrigank Srivastava , Ashwani Kumar, Gururaj S P, 2021, Liver Disease Prediction System using Machine Learning Techniques, (IJERT) Volume 10, Issue 06 (June 2021),

[6] Prof. Satish Manje, Miss.Tannaz Inamdar, Miss.Shivani Kamalekar, Miss. Akanksha Shaha, Survey on Machine Learning Techniques for the Diagnosis of Liver Disease, IRJEAM, ISSN : 2454-9150 Vol-08, Issue-01, APR 2022.

[7] Bhupathi, Deepika & Tan, Christine Nya-Ling & Sremath Tirumala, Sreenivas & Ray, Sayan. Liver disease detection using machine learning techniques, CITRENZ (2022).

 DOI : 10.35291/2454-9150.2024.0270