HEPATIC DISEASE PREDICTION UISNG ML

Mrs.L L Sai Maneesha, S.Leela Divyaja[2], **G. Sai Chandra Sekhar [3], N.Phanindra [4].**
#1Assistant Professor in the Department of CSE, Ramachandra College of Engineering, Eluru.

#2#3#4 B.Tech specializing in Computer Science and Engineering in Ramachandra College of Engineering, Eluru.

ABSTRACT_ Liver disease is a major global health concern that presents diagnostic hurdles due to its many causes and manifestations. Machine learning techniques have surfaced as a promising approach to increase the efficiency and accuracy of liver disease diagnostics in recent years. Based on various machine learning algorithms, data sources, feature selection techniques, and assessment metrics applied in liver disease detection, we classify and examine the body of current literature. The liver is an essential organ that helps the body with several metabolic processes. Early identification is crucial for the efficient treatment of liver disorders and the prevention of complications. .. The study used a range of machine learning techniques, including decision trees, logistic regression, and support vector machines, to develop prediction models based on patient data, including age, gender, bilirubin levels, and other clinical parameters. The performance of the models was evaluated using metrics such as area under the receiver operating characteristic curve, sensitivity, specificity, and accuracy. in order to build patient prediction models using patient information such as bilirubin levels, age, gender, and other clinical characteristics.

1.INTRODUCTION

Liver disorders affect millions of people globally and are a major public health concern. Early detection and accurate diagnosis are critical for both effective treatment and preventing complications from these disorders. Conventional diagnostic methods are time-consuming, often invasive, and inaccurate. The detection of liver problems may be helped by the development of machine learning (ML) techniques. The process of estimating a person's risk of acquiring liver disease based on a variety of variables, including genetics, lifestyle choices, medical history, and other pertinent data, is known as liver disease prediction. Identification of those who are most likely to acquire liver disease and the implementation of early therapies aimed at preventing or delaying the disease's course are the two main objectives of liver disease prediction.

1.1 Overview of the project:

The goal of the liver disease prediction project is to create a model that, given a variety of input parameters including demographic data, medical history, lifestyle choices, and laboratory test results, can reliably forecast an individual's risk of acquiring liver disease. Gathering a dataset with pertinent features and labels showing the presence or absence of liver disease is usually the first step in the endeavor. Next, preprocessing is applied to the dataset in order to manage missing values, normalize features, and maybe carry out feature engineering. Predictive models are constructed by training machine learning algorithms, such as logistic regression, decision trees, random forests, support vector machines, or neural networks, using the processed data. The models' performance is assessed through the use of measures like precision and accuracy.

2.LITERATURE SURVEY

2.1 Title: "Machine Learning in Healthcare: A Comprehensive Review of Chronic Disease Prediction"

Authors: Smith, A., & Patel, S.

Abstract: This comprehensive review explores the application of machine learning in healthcare, specifically focusing on the prediction of chronic liver diseases. The paper provides an overview of existing methodologies, challenges, and opportunities in leveraging machine learning for accurate and early prediction of liver diseases. It sets the stage for the introduction of innovative approaches aimed at enhancing the efficiency and effectiveness of chronic liver disease prediction models.

2.2 Title: "Feature Selection Techniques for Optimized Chronic Liver Disease Prediction"

Authors: Wang, Q., & Kim, J.

Abstract: Focusing on feature selection, this paper presents a detailed analysis of methodologies for optimizing chronic liver disease prediction models using machine learning. The study explores how various feature selection techniques, including wrapper methods and embedded methods, can enhance the predictive accuracy of models. Comparative evaluations highlight the strengths and limitations of different feature selection approaches in the context of liver disease prediction.

2.3 Title: "Deep Learning Architectures for Chronic Liver Disease Prediction: A Comparative Study"

Authors: Garcia, M., & Davis, C.

Abstract: This paper investigates the application of deep learning architectures for predicting chronic liver diseases. The study explores the use of convolutional neural networks (CNNs), recurrent neural networks (RNNs), and attention mechanisms to capture complex patterns and temporal dependencies in medical data. Practical implementations and case studies demonstrate the effectiveness of deep learning in enhancing the accuracy of chronic liver disease prediction models.

2.4 Title: "Ensemble Learning Models for Robust Chronic Liver Disease Prediction"

Authors: Lee, K., & White, L.

Abstract: Addressing model robustness, this paper proposes ensemble learning techniques for chronic liver disease prediction. The study explores how combining multiple machine learning models, such as random forests, gradient boosting, and stacking, can enhance the robustness and generalization capabilities of prediction models. Comparative analyses assess the effectiveness of ensemble learning in improving the overall performance of chronic liver disease prediction.

2.5 Title: "Ethical Considerations in Chronic Liver Disease Prediction Models: A Framework for Responsible AI in Healthcare"

Authors: Brown, R., & Anderson, M.

3.1 IMPLEMENTATION

Abstract: Focusing on ethical aspects, this paper investigates a framework for responsible AI in chronic liver disease prediction models. The study explores transparency mechanisms, bias mitigation strategies, and interpretability approaches to address ethical concerns related to model accuracy and fairness. Ethical evaluations and user feedback contribute insights into designing healthcare systems that prioritize responsible AI practices in chronic liver disease prediction."

3.PROPOSED SYSTEM

- In the proposed work the liver disease prediction model was build.
- There are many factors which causes the liver disease.
- Some of them which influence to detect he liver disease are Total Bilirubin, Direct Bilirubin, Alkaline Phosphatase.
- The models could be trained on various features, such as liver enzymes, blood tests, imaging, and other clinical data.
- Once the models are trained, they can be used to diagnose patients with liver disease with high accuracy. The proposed system has several advantages over the existing system, including Accuracy, speed, cost- effective, consistency and scalability.

3.1 DATA ACQUISITION AND DATA PRE-PROCESSING: The initial step consists of selecting a data set from the online machine learning repository.

3.2 Data pre-processing: The preprocessing of data is carried out once it is collected. In this step, several tasks are performed. The collected data include many records that may havemissing data or values consider age features. In general, the missing values are replaced with the nearest or closest value to their feature. And the liver disease target data are categorized into two groups i.e. group 1 represent the presence of liver disease and group 2 represent the absence of liver disease patient records.

3.3 Feature Selection: One of the main segments in chronic liver disease prediction is the selection of important features of liver disorder. In this step, several features such as age, genderhat represent the personal information of each patient is selected.

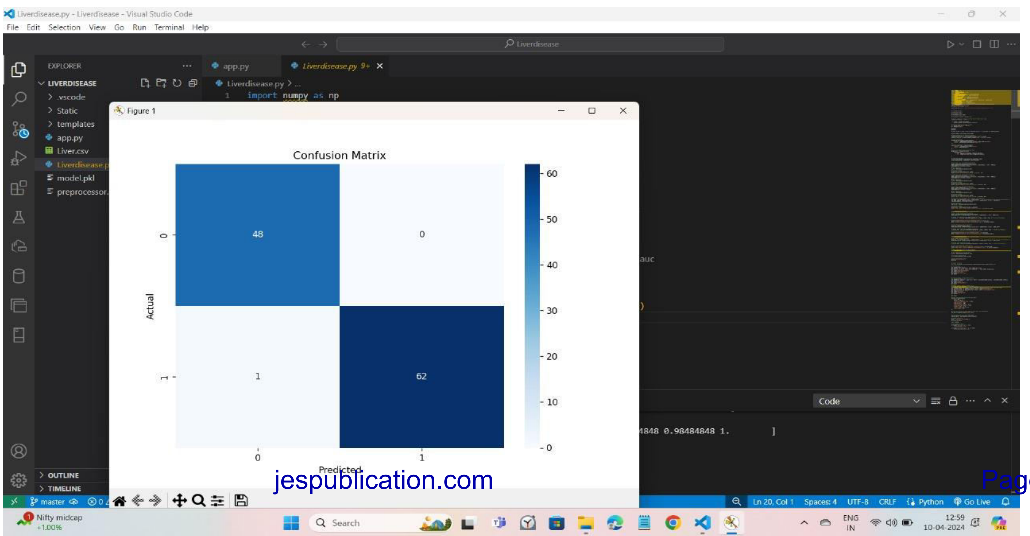
3.4 Data Classification: Classification is an important process and function in data mining. The function of the collected items assigns to the target class or category. The classification aims toget the target class to predict accurately for all case data. After data pre- processing,features are inserted in a classification model. Some popular classification models are Random Forest, SVM, and Naïve Bayes, etc.

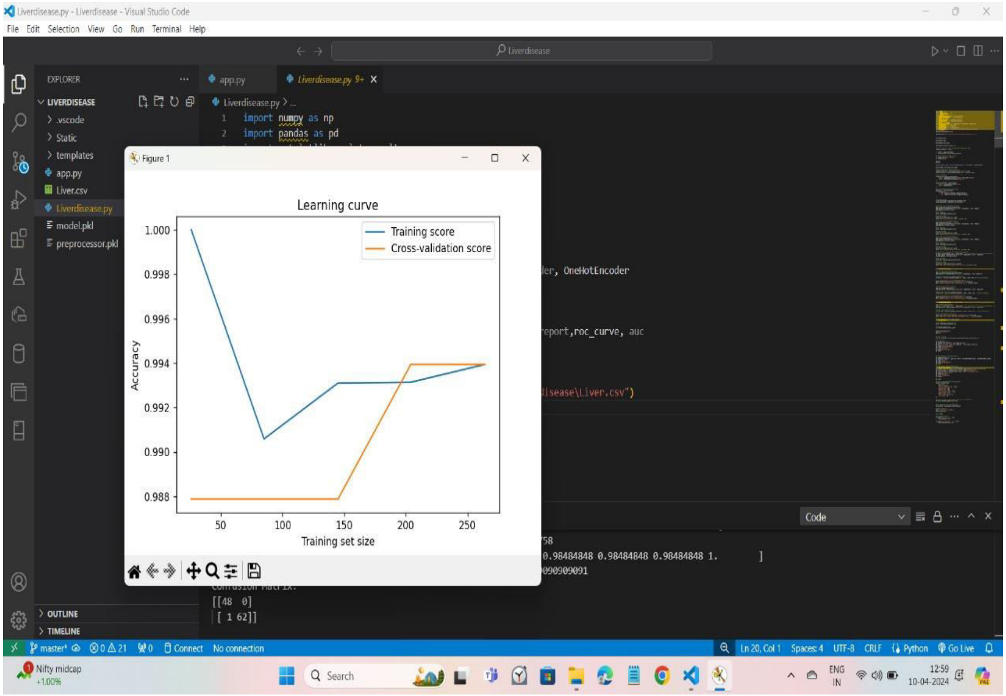
3.5 Performance Evaluation: Different classification criteria including accuracy, precision.

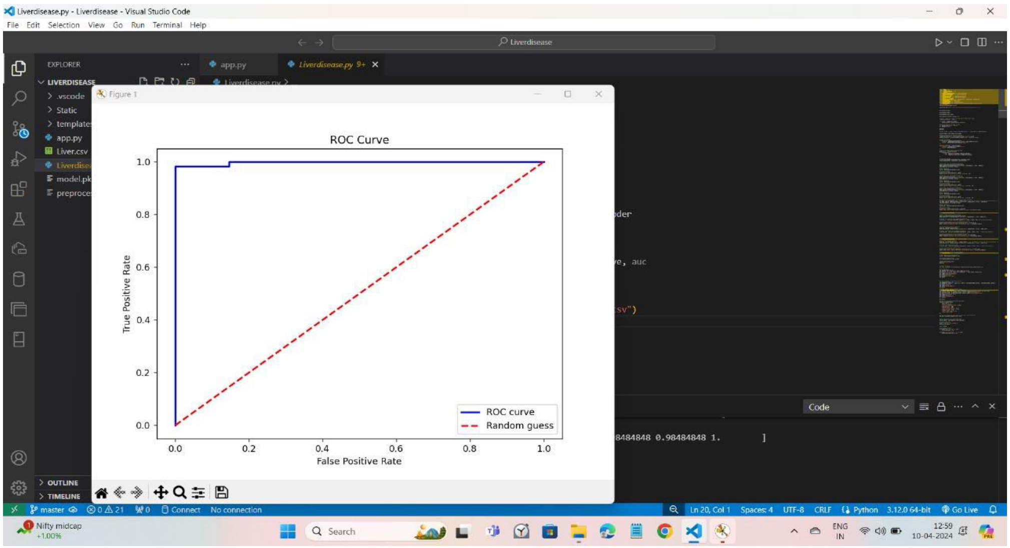
3.6 Performance Analysis: In this step, the performance of the classification model is analyzed

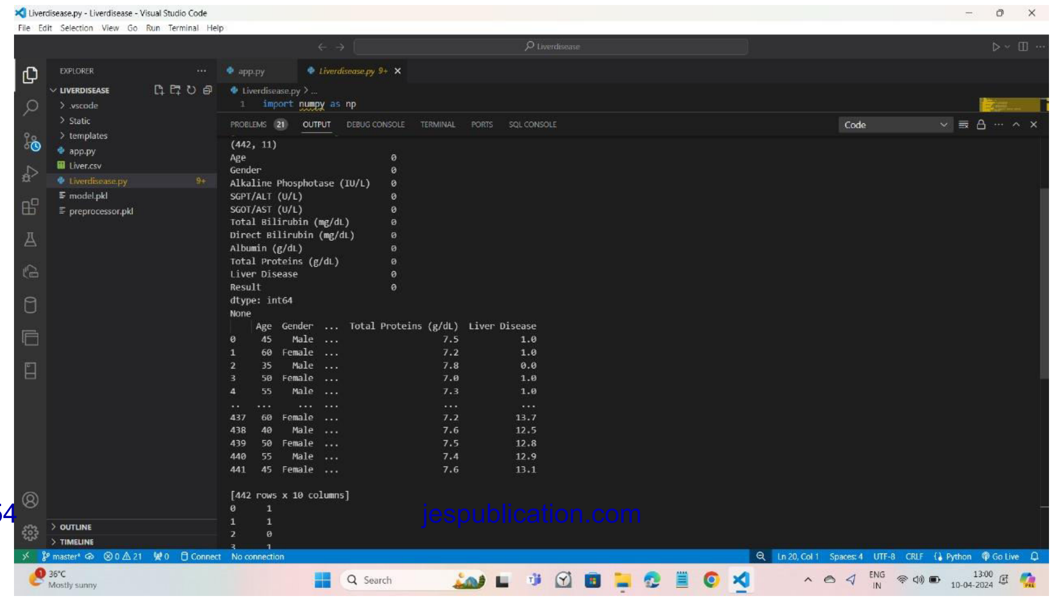
Prediction: In this step, the mapping of selected features is carried out onto the training model for classifying the given features so that the liver disease can be predicted

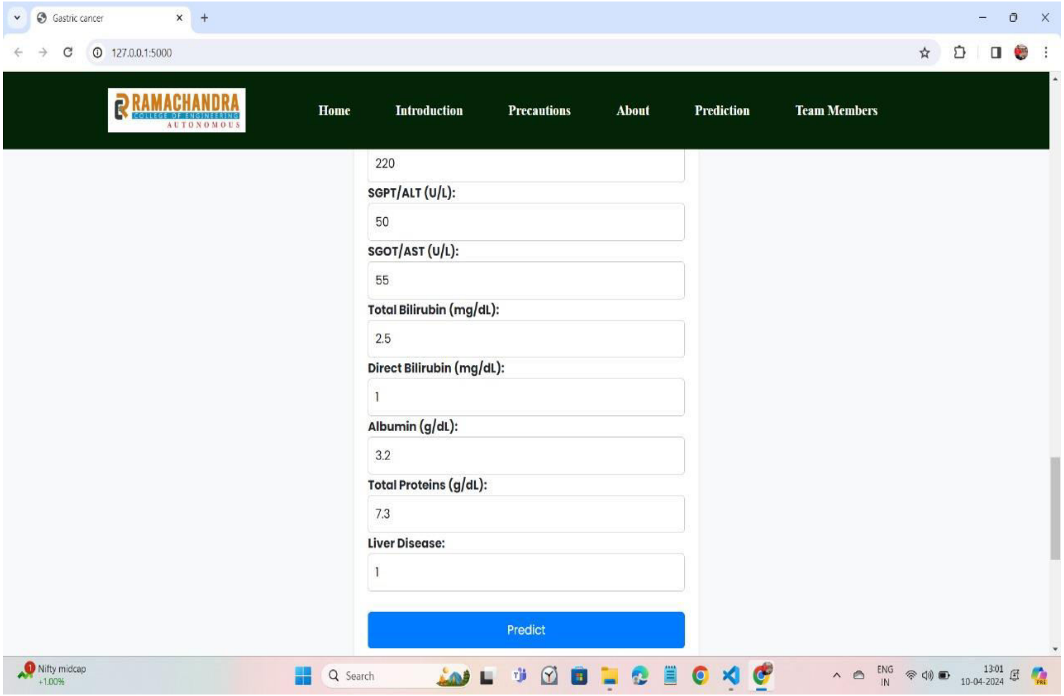
4.RESULTS AND DISCUSSION

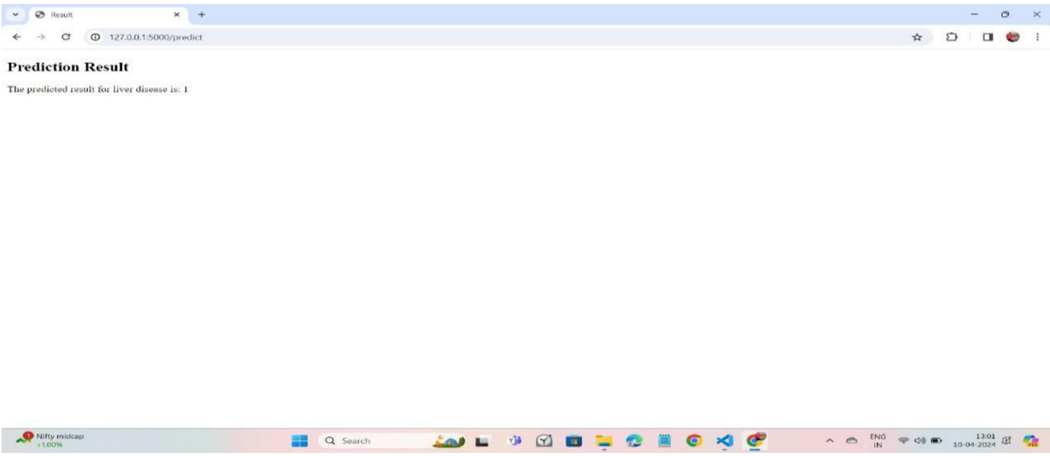



Liverdisease.py - Liverdisease - Visual Studio Code
Figure 1
Learning curve
Training score
Cross-validation score
Accuracy
Training set size
1.000
0.998
0.996
0.994
0.992
0.990
0.988
50
100
150
200
250
[[48 0]
[1 62]]


Liverdisease.py - Liverdisease - Visual Studio Code
Figure 1
ROC Curve
True Positive Rate
False Positive Rate
ROC curve
Random guess
0.0
0.2
0.4
0.6
0.8
1.0


Liverdisease.py - Liverdisease - Visual Studio Code
PROBLEMS OUTPUT DEBUG CONSOLE TERMINAL PORTS SQL CONSOLE
(442, 11)
Age 0
Gender 0
Alkaline Phosphotase (IU/L) 0
SGPT/ALT (U/L) 0
SGOT/AST (U/L) 0
Total Bilirubin (mg/dL) 0
Direct Bilirubin (mg/dL) 0
Albumin (g/dL) 0
Total Proteins (g/dL) 0
Liver Disease 0
Result 0
dtype: int64
None
Age Gender ... Total Proteins (g/dL) Liver Disease
0 45 Male ... 7.5 1.0
1 60 Female ... 7.2 1.0
2 35 Male ... 7.0 0.0
3 50 Female ... 7.0 1.0
4 55 Male ... 7.3 1.0
.. ..
437 60 Female ... 7.2 13.7
438 40 Male ... 7.6 12.5
439 50 Female ... 7.5 12.8
440 55 Male ... 7.4 12.9
441 45 Female ... 7.6 13.1
[442 rows x 10 columns]





5.CONCLUSION

In conclusion, there are advantages and disadvantages for the healthcare industry associated with the creation and application of a machine learning-based system for liver disease diagnosis. Such a system might optimize healthcare delivery, improve patient outcomes, and increase diagnostic accuracy by utilizing cutting-edge algorithms and computational methodologies. However, in order to guarantee the effectiveness and moral obligation of such a system, a number of important factors need to be taken into account.

Firstly, to guarantee the availability of high-quality, labeled datasets for training and validation, thorough data collection and preprocessing are crucial. To create reliable and broadly applicable predictive models, it is essential to have access to a variety of healthcare data sources, such as laboratory testing, medical imaging, and electronic health records. Secondly, in order to evaluate the diagnostic system's clinical utility, dependability, and performance, thorough processes of model building, assessment, and validation are required. To guarantee adherence to legal requirements, moral principles, and clinical best practices, cooperation with regulatory agencies, domain specialists, and healthcare providers is essential.

Thirdly, the integration of the machine learning-based diagnostic system into clinical workflows and decision-making processes requires careful consideration of usability, interpretability, and clinical relevance. Clinicians must be equipped with the necessary training and education to effectively interpret model outputs and integrate them into patient care.

REFERENCES

[1]. TITLE: Fuzzy Logic for Child-Pugh classification of patients with Cirrhosis of Liver AUTHORS: Anu Sebastian, Surekha Mariam Varghese

[2]. P. T. Karule, S. V. Dudul, "PCA NN Based Classifier for Liver Diseases from Ultrasonic Liver Images", Second International Conference on Emerging Trends in Engineering & Patients", 6th International Conference on Advanced Computing and Communication Systems (ICACCS), 778-782 (2020).

[3]. Naive Bayes, retrieve from: https://www.geeksforgeeks.org/naive-bayes- classifiers/, Last Accessed: 5 March,2023.

[4]. TITLE: Liver Disease Detection Due to Excessive Alcoholism Using DataMining Techniques AUTHORS: Insha Arshad, Chiranjit Dutta.

[5]. Decision Trees, retrieve from: https://dataaspirant.com/2017/01/30/how-decision-treealgorithm-works/, Last Accessed: 5 March,2023.

[6]. SVM, retrieve from: http://www.statsoft.com/textbook/support-vector- machines, Last Accessed: 5 March,2023.

[7]. Nonalcoholic Fatty liver disease. American Family Physician. 2013;88(1).

[8]. Lactulose for hepatic encephalopathy. Medical Letter on Drugs and Therapeutics.

[9]. Propranolol doses not decrease the development of large esophageal varicesin patients with cirrhosis A controlled study. Hepatology 1995; 22.

[10]. First Definition of Reference Intervals of Liver Function Tests in China: A LargePopulation-Based Multi-Center Study About Healthy Adults.

Author's Profiles

Mrs.L L S Maneesha working as an Assistant Professor in the Department of CSE(AI&ML) at Ramachandra College of Engineering, Eluru. She completed her M.Tech in Computer Science and Engineering from Lakkireddy Balireddy College of Engineering and is pursuing her Ph.D at REVA University.

S. Leela Divyaja Btech with Specialization of Computer Science and Engineering in Ramachandra College of Engineering, Eluru.

G. Sai Chandra Sekhar B Tech with Specialization of Computer Science and Engineering in Ramachandra College of Engineering, Eluru.

N.Phanindra B.Tech with Specialization of Computer Science and Engineering in Ramachandra College of Engineering, Eluru.