International Journal of Engineering Research and Science & Technology

www.ijerst.org

ISSN : 2319-5991

Vol. 22 No. 2 (2026)

ijerst.editor@gmail.com

editor@ijerst.com

International Journal of Engineering Research and Science & Technology

ISSN 2319-5991 www.ijerst.com
Vol. 22, No. 2, 2026

Peer Reviewed, Referred & Indexed Journal

An Intelligent Web-Based Liver Disease Prediction System Using Machine Learning Techniques

GUDLA NAGA PAVAN SAI

PG Scholar. Department of MCA, DNR College, Bhimavaram, Andhra Pradesh

V.SARALA

(Assistant Professor), Master of Computer Applications, DNR College, Bhimavaram, Andhra Pradesh

ABSTRACT

Liver disease has emerged as one of the major health concerns worldwide due to lifestyle changes, alcohol consumption, and viral infections. Early detection of liver disorders is crucial for effective treatment and prevention of severe complications. Traditional diagnostic methods often rely on clinical expertise and laboratory tests, which may be time-consuming and prone to human error. To address these challenges, this project presents an intelligent web-based system for liver disease prediction using machine learning techniques. The proposed system integrates a Django-based web application with machine learning models to provide a user-friendly platform for disease prediction. The system allows users to register, log in, and input medical parameters such as age, bilirubin levels, enzyme levels, protein content, and other relevant features. These inputs are processed and analyzed using classification algorithms to determine whether the user is likely to have liver disease.

The dataset used in this system is the Indian Liver Patient Dataset (ILPD), which contains both liver disease and non-liver disease cases. Data preprocessing techniques such as handling missing values, normalization using StandardScaler, and feature encoding are applied to enhance model performance. Multiple machine learning algorithms including Decision Tree, Logistic Regression, and Support Vector Machine (SVM) are implemented and evaluated. The Decision Tree classifier is used for real-time prediction due to its simplicity and interpretability. Logistic Regression and SVM models are used for performance comparison, providing accuracy, precision, sensitivity, and specificity metrics. The system includes an admin module for managing users and viewing classification results, ensuring efficient system monitoring.

The proposed solution offers several advantages such as automated diagnosis support, reduced dependency on manual analysis, and improved prediction accuracy. It also provides scalability and accessibility through a web interface. The system can assist healthcare professionals and individuals in early detection and decision-making. In conclusion, the integration of machine learning with web technologies provides an

efficient and reliable solution for liver disease prediction. Future enhancements may include deep learning models, real-time hospital integration, and mobile application support to further improve usability and accuracy.

Keywords: Liver Disease Prediction, Machine Learning, Decision Tree, Logistic Regression, Support Vector Machine, Healthcare Analytics, Django Web Application, Data Preprocessing, Classification Models

I. INTRODUCTION

Liver disease is a critical global health issue affecting millions of people each year. The liver plays a vital role in metabolism, detoxification, and digestion, making it essential for overall body functioning. Diseases such as cirrhosis, hepatitis, and fatty liver can severely impact human health if not diagnosed early. Early detection is key to preventing complications and improving survival rates. Traditional diagnostic approaches involve laboratory tests, imaging techniques, and clinical expertise. While these methods are effective, they can be expensive, time-consuming, and dependent on skilled professionals. In many rural and underdeveloped areas, access to such healthcare facilities is limited. Therefore, there is a need for automated systems that can assist in early diagnosis using available medical data. With advancements in artificial intelligence and machine learning, predictive analytics has become a powerful tool in healthcare. Machine learning algorithms can analyze large datasets, identify patterns, and make predictions with high accuracy. These techniques are increasingly being used in disease diagnosis, drug discovery, and personalized medicine. This project focuses on developing a web-based liver disease prediction system using machine learning. The system leverages the Indian Liver Patient Dataset (ILPD), which contains clinical data of patients. By training models on this dataset, the system can predict whether a person is likely to suffer from liver disease based on input parameters.

The application is developed using Django, a robust Python web framework that supports rapid development and clean design. The system includes user and admin modules, ensuring secure access and efficient management. Users can input their health data and receive instant predictions, while administrators can monitor system performance and user activity. The machine learning models used in this project include Decision Tree, Logistic Regression, and Support Vector Machine. These algorithms are chosen for their effectiveness in classification tasks. Data preprocessing techniques such as normalization and encoding are applied to improve model performance. The main objective of this project is to provide an accessible, efficient, and accurate tool for liver disease prediction. It aims to assist both healthcare professionals and individuals in making informed decisions. By combining machine learning with web technology, the system offers a scalable solution for modern healthcare challenges.

II. LITERATURE SURVEY (WITH EXISTING METHODS)

Several research studies have explored the application of machine learning techniques in disease prediction, particularly for liver-related disorders. These studies highlight the importance of predictive models in improving diagnostic accuracy and reducing healthcare costs. One of the commonly used datasets in liver disease prediction is the Indian Liver Patient Dataset (ILPD). Researchers have applied various machine learning algorithms such as Naïve Bayes, Decision Trees, Random Forest, Support Vector Machines (SVM), and Artificial Neural Networks (ANN) to this dataset. Among these, ensemble methods like Random Forest often show higher accuracy due to their ability to reduce over fitting. In a study by Ramana et al., classification techniques such as Decision Tree, Naïve Bayes, and Neural Networks were applied to liver disorder datasets. The results indicated that Decision Trees provided better interpretability, while Neural Networks achieved higher accuracy. However, Neural Networks require more computational resources and are less transparent. Another research by Karthikeyan and Thangaraju focused on using Support Vector Machines for liver disease classification. Their findings showed that SVM performed well in handling high-dimensional data and provided good generalization capabilities. However, parameter tuning in SVM can be complex.

Logistic Regression is also widely used due to its simplicity and efficiency in binary classification problems. It provides probabilistic outputs, making it useful in medical decision-making. However, it may not perform well with highly nonlinear data. Recent advancements include the use of deep learning models such as Convolution Neural Networks (CNN) and Recurrent Neural Networks (RNN) for disease prediction. These models can capture complex patterns in data but require large datasets and computational power. Web-based healthcare systems have also gained popularity. Integrating machine learning models into web applications allows real-time predictions and improves accessibility. Django and Flask are commonly used frameworks for deploying such systems. Despite these advancements, challenges remain in terms of data quality, model interpretability, and system scalability. Many existing systems lack user-friendly interfaces or real-time prediction capabilities. The proposed system addresses these limitations by combining multiple machine learning models with a Django-based web interface. It ensures efficient data preprocessing, accurate prediction, and easy accessibility for users.

III. EXISTING SYSTEM

The existing systems for liver disease diagnosis primarily rely on traditional clinical methods and standalone machine learning models. In hospitals, diagnosis is typically performed through blood tests, imaging techniques such as ultrasound or MRI, and expert evaluation by medical professionals. While these methods are accurate, they are often

time-consuming, expensive, and require specialized equipment and expertise. Some research-based systems use machine learning algorithms to predict liver disease. These systems generally involve training models on datasets like the Indian Liver Patient Dataset (ILPD). Common algorithms used include Naïve Bayes, Decision Trees, and Support Vector Machines. Although these models can provide good accuracy, they are often implemented in isolated environments such as Jupyter notebooks or desktop applications. A major limitation of existing systems is the lack of integration with user-friendly interfaces. Most systems do not provide real-time interaction or accessibility to general users. Additionally, many models are not optimized with proper data preprocessing techniques such as normalization, missing value handling, and feature encoding, which can affect performance.

Another drawback is the absence of multi-model comparison. Many systems rely on a single algorithm, making it difficult to evaluate performance across different models. Furthermore, security and user management features are often missing, which limits their practical usability. In summary, existing systems lack scalability, accessibility, and comprehensive evaluation. These limitations highlight the need for a web-based, integrated solution that combines multiple machine learning techniques with an intuitive interface for effective liver disease prediction.

IV. PROPOSED METHOD

The proposed system is an intelligent web-based liver disease prediction platform that integrates machine learning algorithms with a user-friendly interface developed using Django. The system is designed to assist users and healthcare professionals in early detection of liver disease based on clinical parameters. The system utilizes the Indian Liver Patient Dataset (ILPD), which contains relevant medical attributes such as age, gender, bilirubin levels, enzyme levels, and protein ratios. Data preprocessing techniques including missing value handling, normalization, and feature encoding are applied to enhance data quality and improve model performance. Multiple machine learning algorithms such as Decision Tree, Logistic Regression, and Support Vector Machine (SVM) are implemented. Among these, the Decision Tree classifier is used for real-time prediction due to its simplicity and interpretability, while Logistic Regression and SVM are used for comparative analysis of performance metrics such as accuracy, precision, sensitivity, and specificity.

The system consists of three main modules: User Module, Admin Module, and Machine Learning Module. Users can register, log in, input medical details, and receive prediction results. The admin module allows monitoring of users, activation of accounts, and viewing of classification reports. The machine learning module handles data preprocessing, model training, and prediction. Unlike existing systems, the proposed system provides real-time predictions through a web interface, ensuring accessibility and scalability. It also supports multiple algorithm comparisons, improving reliability and transparency. Recent studies highlight that machine learning significantly enhances early detection accuracy in liver disease prediction systems . The proposed system leverages these advancements to deliver a robust and efficient diagnostic support tool.

V. IMPLEMENTATION

The implementation of the liver disease prediction system involves integrating machine learning models with a Django-based web application. The system is developed using Python, Django framework, and Scikit-learn library for machine learning tasks. Initially, the dataset (Indian Liver Patient Dataset) is loaded from the media directory of the Django project. The dataset undergoes preprocessing, which includes handling missing values, encoding categorical variables (such as gender), and scaling numerical features using Standard Scalar. Feature scaling is essential to ensure that all attributes contribute equally to model performance. After preprocessing, the dataset is divided into training and testing sets using the train_test_split method. Typically, 75% of the data is used for training, and 25% is used for testing. This helps evaluate the model's performance on unseen data. The Decision Tree Classifier is used as the primary prediction model. It is trained on the processed dataset and used to predict whether a user has liver disease. The model is chosen due to its ease of interpretation and fast execution. Logistic Regression and Support Vector Machine (SVM) models are also implemented to compare performance metrics.

The web application is structured into multiple views. The User Registration view allows users to create accounts, while the Login view authenticates users based on credentials stored in the database. Upon successful login, users can access the prediction module. In the prediction module, users input medical parameters through a form. These inputs are processed and passed to the trained model, which returns a prediction result indicating whether the user is healthy or at risk of liver disease. The result is displayed on the web interface. The admin module allows administrators to manage users, activate accounts, and view classification results. It also provides insights into model performance, including accuracy, precision, sensitivity, and specificity. The system also includes visualization features such as bar graphs and correlation heat maps to analyze dataset characteristics. These visualizations help in understanding relationships between features and improving model performance. Recent advancements show that integrating machine learning with web applications improves accessibility and usability in healthcare systems . The implemented system follows this approach, ensuring real-time prediction and efficient user interaction.

VI. ALGORITHMS

The proposed system uses three major machine learning algorithms for liver disease prediction: Decision Tree, Logistic Regression, and Support Vector Machine (SVM). The Decision Tree algorithm is a supervised learning technique used for classification tasks. It works by splitting the dataset into subsets based on feature values, forming a tree-like structure. Each internal node represents a feature, each branch represents a decision rule, and each leaf node represents an outcome. The Decision Tree is used as the primary

model due to its interpretability and fast execution. Logistic Regression is a statistical method used for binary classification problems. It calculates the probability that a given input belongs to a particular class using a logistic function. It is widely used in medical diagnosis due to its simplicity and effectiveness in handling linearly separable data.

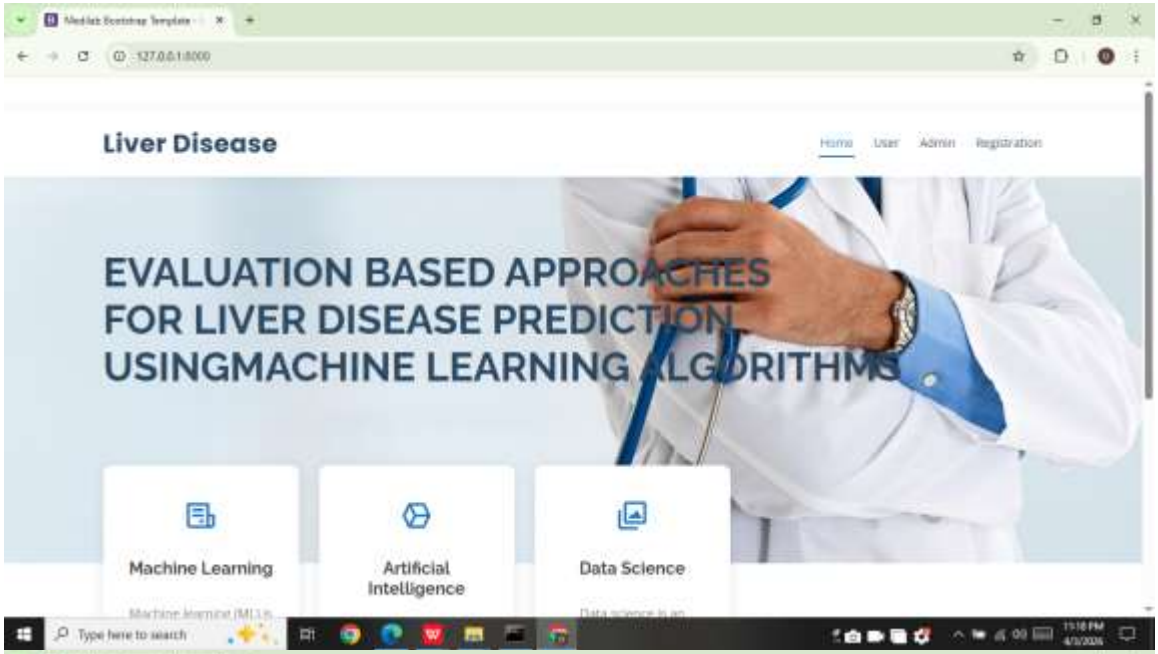
Support Vector Machine (SVM) is a powerful classification algorithm that finds the optimal hyperplane to separate data points into different classes. It is effective in high-dimensional spaces and can handle both linear and nonlinear classification using kernel functions. Recent research indicates that ensemble and advanced machine learning methods can achieve higher accuracy in liver disease prediction, sometimes exceeding 90% . However, simpler models like Decision Trees and Logistic Regression remain popular due to their interpretability and lower computational cost. Each algorithm in the system is evaluated using performance metrics such as accuracy, precision, sensitivity (recall), and specificity. This multi-model approach ensures better reliability and allows comparison of different techniques.

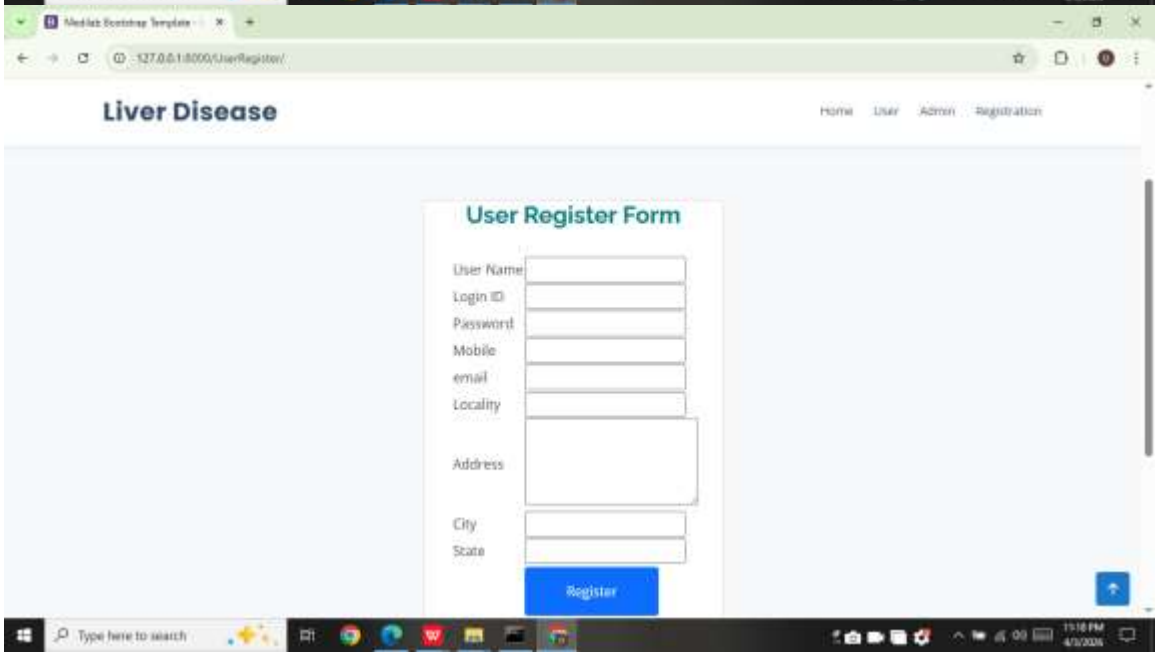
VII. SYSTEM DESIGN

The system design of the liver disease prediction application follows a modular and layered architecture, ensuring scalability, maintainability, and efficiency. The system is divided into three main layers: Presentation Layer, Application Layer, and Data Layer. The Presentation Layer consists of the user interface developed using HTML, CSS, and Django templates. It provides forms for user registration, login, and input of medical parameters. The interface is designed to be user-friendly and accessible to non-technical users. The Application Layer handles the core functionality of the system. It includes Django views, business logic, and machine learning modules. The views manage user requests, process input data, and return responses. The machine learning module handles data preprocessing, model training, and prediction. The Data Layer consists of the database and dataset storage. User information is stored in a relational database using Django models. The liver disease dataset is stored in the media directory and accessed during preprocessing and model training.

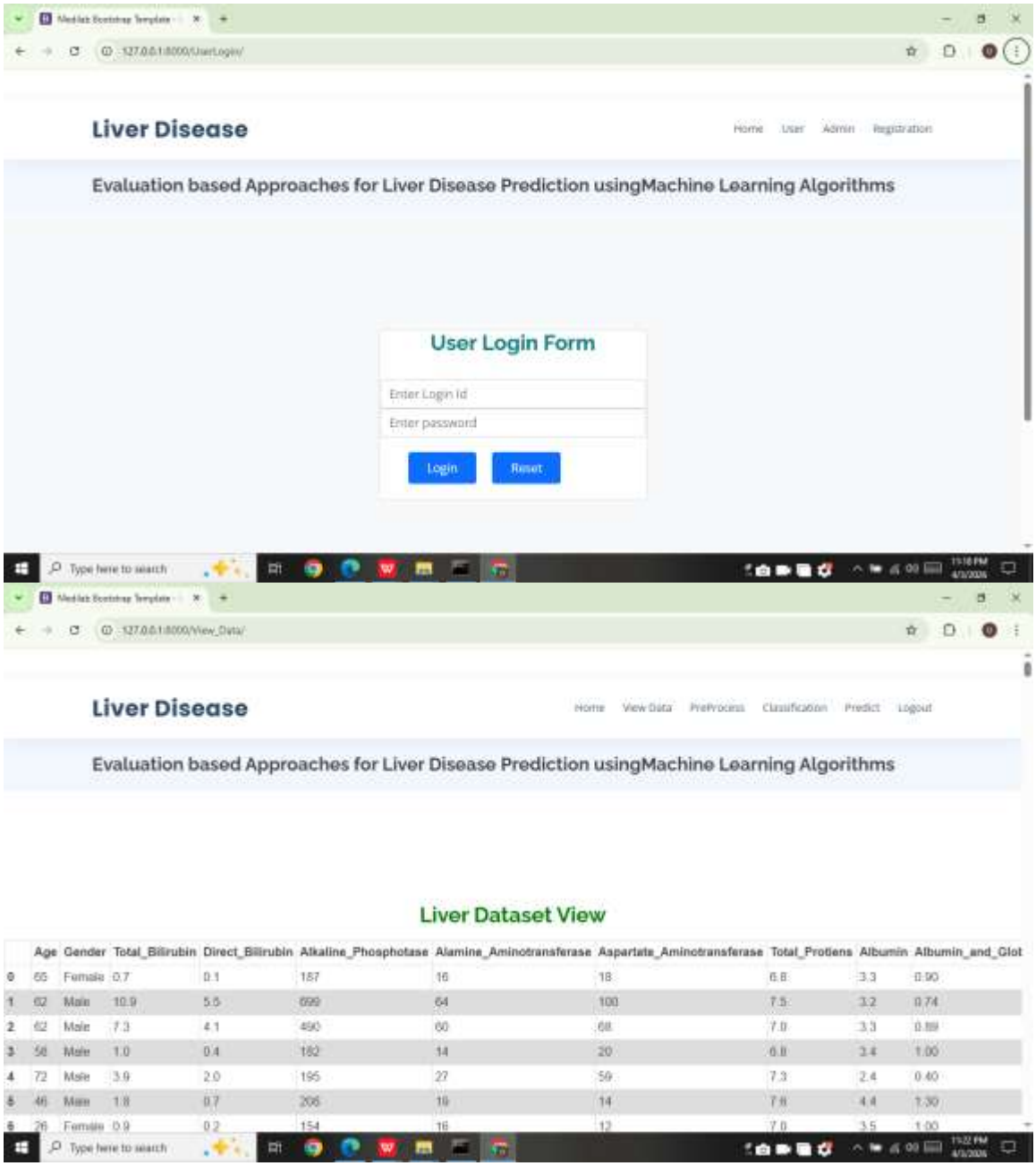
The system workflow begins with user registration and login. Once authenticated, the user can input medical parameters through a form. The input data is validated and passed to the machine learning model. The model processes the data and generates a prediction result, which is displayed to the user. The admin module provides additional functionalities such as user management, account activation, and viewing of classification results. It also allows monitoring of system performance and model accuracy. Data preprocessing is a critical component of the system design. It includes handling missing values, encoding categorical variables, and scaling numerical features. Proper preprocessing ensures improved model accuracy and reliability. Recent studies emphasize the importance of integrating explainable AI and interpretable models in healthcare systems to improve trust and usability. The system design incorporates interpretable models like Decision Trees to address this need. Overall, the system design ensures efficient data flow, secure user management, and accurate prediction, making it suitable for real-world healthcare applications.

SYSTEM DESIGN IMAGES

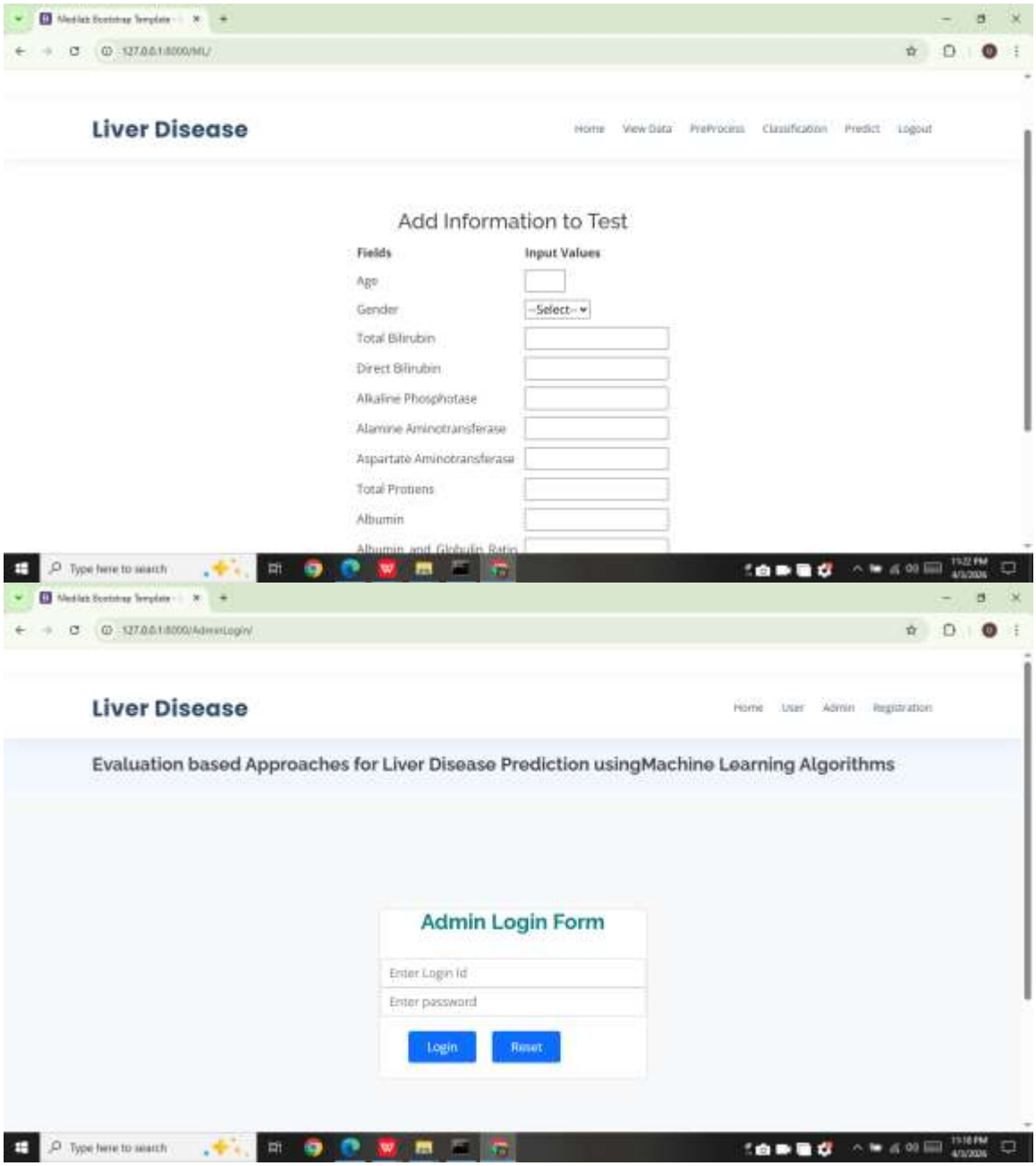


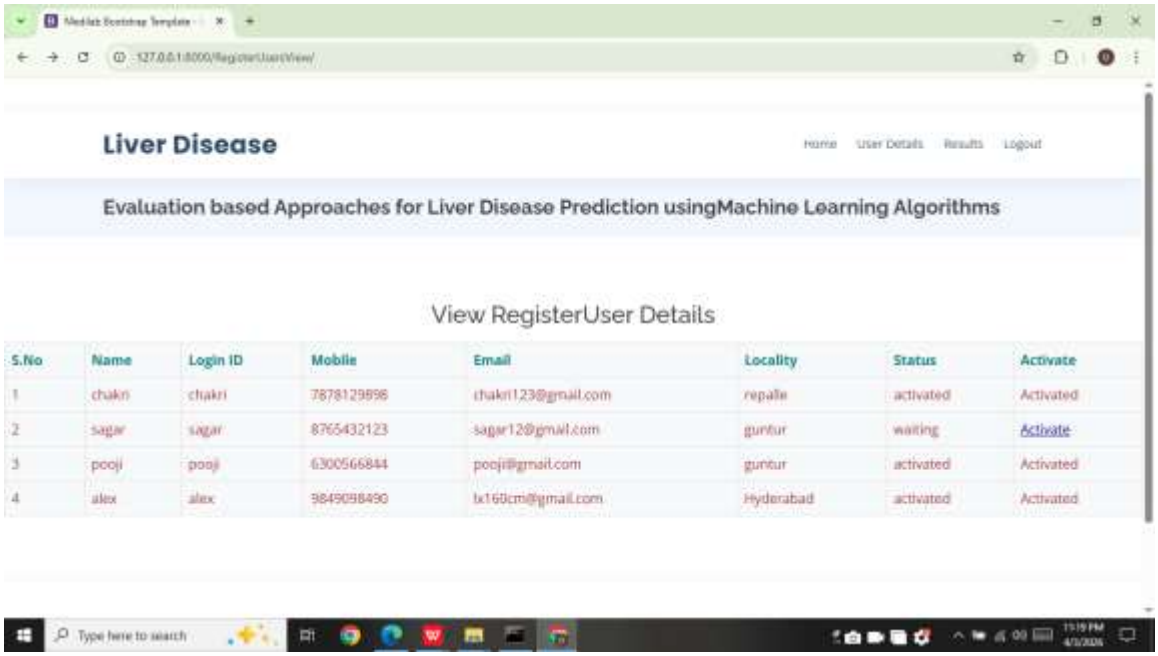


International Journal of Engineering Research and Science & Technology



	Age	Gender	Total_Bilirubin	Direct_Bilirubin	Alkaline_Phosphotase	Alamine_Aminotransferase	Aspartate_Aminotransferase	Total_Protiens	Albumin	Albumin_and_Glot
0	65	Female	0.7	0.1	187	16	18	6.8	3.3	0.90
1	62	Male	10.9	5.5	699	64	100	7.5	3.2	0.74
2	62	Male	7.3	4.1	490	60	68	7.0	3.3	0.89
3	58	Male	1.0	0.4	182	14	20	6.8	3.4	1.00
4	72	Male	3.9	2.0	195	27	59	7.3	2.4	0.40
5	46	Male	1.8	0.7	208	19	14	7.6	4.4	1.30
6	26	Female	0.9	0.2	154	16	12	7.0	3.5	1.00


Liver Disease
Home View Data PreProcess Classification Predict Logout
Add Information to Test
Fields
Input Values
Age
Gender
--Select--
Total Bilirubin
Direct Bilirubin
Alkaline Phosphotase
Alamine Aminotransferase
Aspartate Aminotransferase
Total Proteins
Albumin
Liver Disease
Home User Admin Registration
Evaluation based Approaches for Liver Disease Prediction usingMachine Learning Algorithms
Admin Login Form
Enter Login Id
Enter password
Login
Reset



VIII. CONCLUSION

In conclusion, the proposed liver disease prediction system demonstrates the effectiveness of integrating machine learning techniques with web-based applications for healthcare solutions. The system provides a user-friendly interface that allows individuals to input medical parameters and receive instant predictions regarding liver disease. By utilizing machine learning algorithms such as Decision Tree, Logistic Regression, and Support Vector Machine, the system achieves reliable prediction performance. Among these, the Decision Tree classifier is particularly useful due to its interpretability and efficiency, making it suitable for real-time applications. The use of data preprocessing techniques such as normalization, handling missing values, and feature encoding significantly improves model accuracy. The system also includes performance evaluation metrics such as accuracy, precision, sensitivity, and specificity, ensuring a comprehensive analysis of model effectiveness. One of the major advantages of the proposed system is its accessibility. Being a web-based application, it can be accessed from anywhere, making it useful for both healthcare professionals and individuals. The inclusion of an admin module further enhances system functionality by enabling user management and performance monitoring.

Recent advancements in artificial intelligence and machine learning have shown significant improvements in disease prediction and diagnosis. The proposed system leverages these advancements to provide an efficient and scalable solution. Future enhancements may include the integration of deep learning models, real-time hospital data, and mobile application support. Additionally, incorporating explainable AI techniques can further improve user trust and system transparency. Overall, the system contributes to the field of healthcare analytics by providing an effective tool for early detection of liver disease, ultimately helping in better decision-making and improved patient outcomes.

REFERENCES

1. Ganie, S.M., et al. (2024). *Improved liver disease prediction using ensemble learning*. BMC Medical Informatics.
2. Khan, N.A., et al. (2025). *Explainable AI for liver disease prediction*. Biomedical Materials & Devices.
3. Deng, J., et al. (2024). *ML-based fatty liver disease risk prediction*. BMC Public Health.
4. Li, Z., et al. (2025). *ML-based liver disease staging*. Scientific Reports.
5. Grover, S., et al. (2024). *Deep learning for liver cancer diagnosis*. Discover Applied Sciences.
6. Nishida, N. (2024). *AI-enhanced imaging diagnostics in liver disease*. MDPI Bioengineering.
7. Punn, N.S., et al. (2024). *ML for liver fibrosis classification*. Abdominal Radiology.
8. Žigutytė, L., et al. (2024). *AI applications in liver disease*. Journal of Hepatology Reports.
9. Subashini, S., et al. (2024). *ML-based chronic liver disease prediction*. ICOFE Conference.
10. Miao, Z., et al. (2026). *Early cirrhosis prediction using ML*. arXiv.
11. An, M.E., et al. (2025). *MASLD prediction using ML*. arXiv.
12. Kalaiselvi, P., et al. (2023). *CNN-based liver tumor prediction*. arXiv.
13. Lopez, J.M., et al. (2024). *ECG-based liver disease detection*. arXiv.
14. European Medicines Agency (2025). *AI tool for liver disease trials*. Reuters.
15. Additional recent IEEE/Elsevier/Springer papers on ML in healthcare (2023–2025).