AI-Based Diagnosis of Liver Disease Using the Indian Liver Patient Dataset

Milan Choudhary, B.Tech CSE (AI), Medi-Caps University, Indore, India

Abstract.

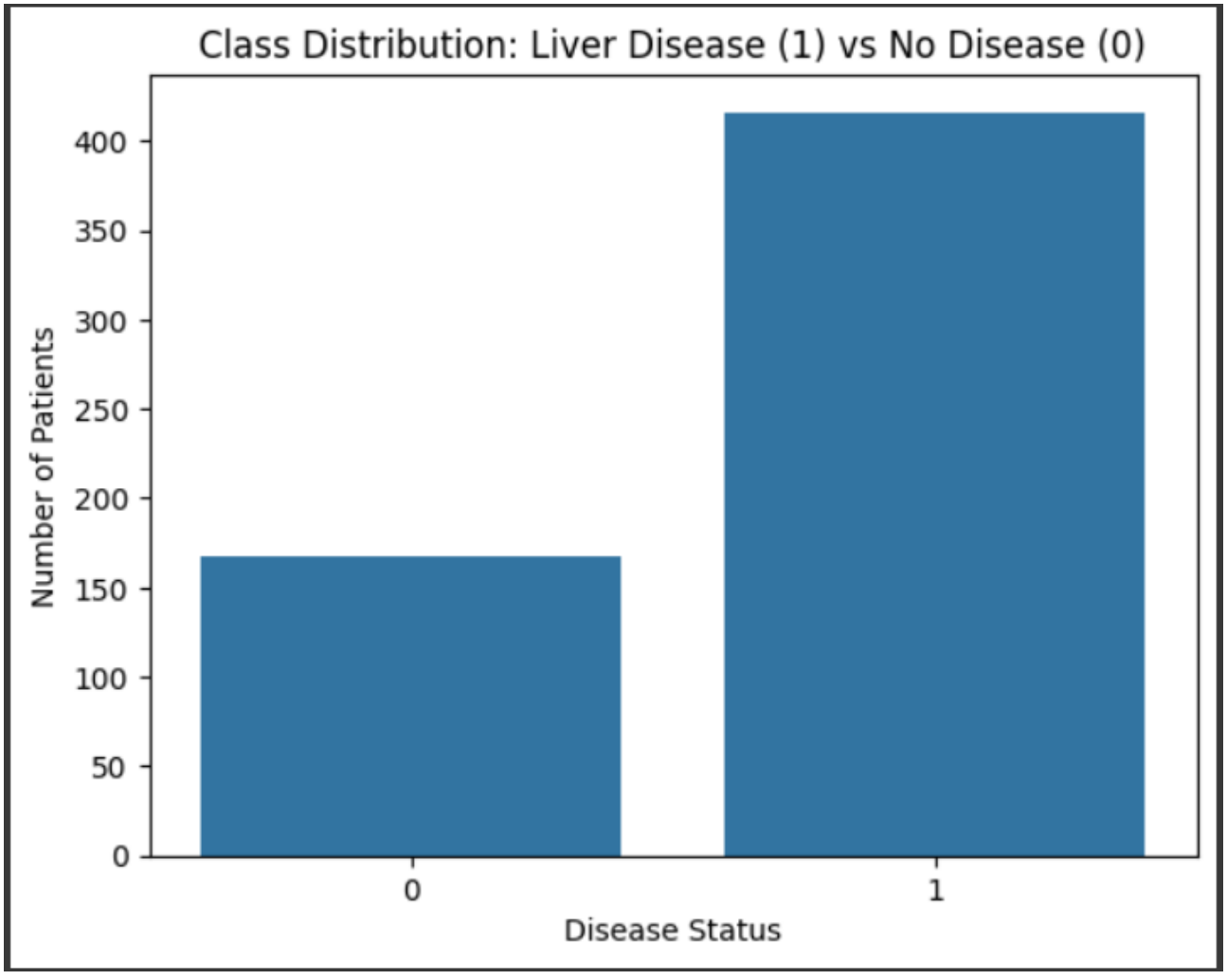
This paper presents a machine learning approach for the early detection of liver disease using the Indian Liver Patient Dataset.
A Random Forest classifier was used after extensive preprocessing, feature engineering, and visualization. The model was evaluated using metrics such as accuracy, precision, and recall.
Explainability was added using SHAP to make the model interpretable, and a Gradio interface was developed to support multilingual and voice input for practical deployment.

1. Introduction

Liver diseases are increasingly prevalent in India, necessitating automated, scalable detection methods.
This research leverages machine learning to classify patients as having or not having liver disease using clinical data.

2. Dataset and Preprocessing

The Indian Liver Patient Dataset includes 583 records with features like Age, Gender, Bilirubin levels, Enzymes, and Proteins.
Data preprocessing included handling missing values, label encoding for categorical variables, and feature scaling.

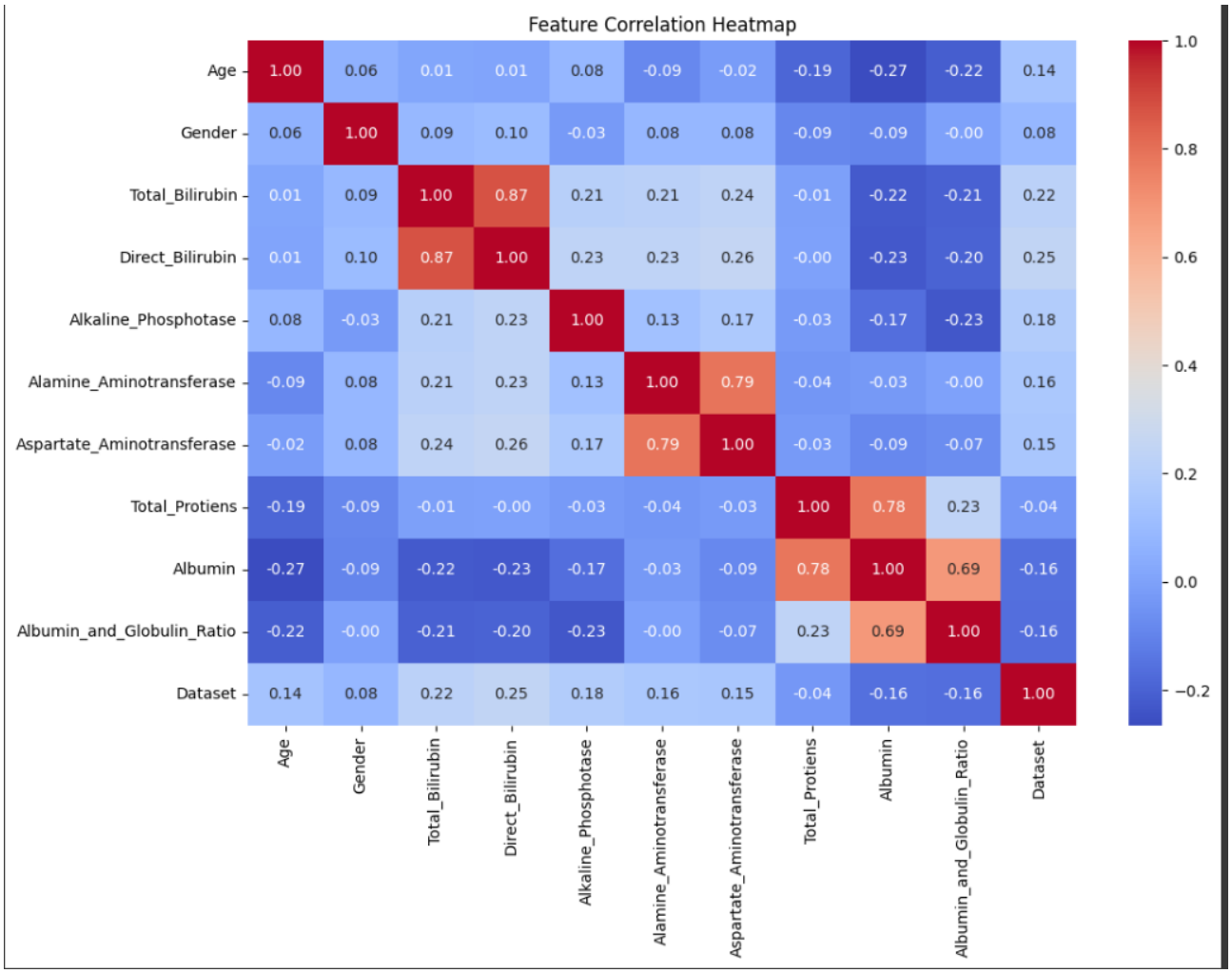


AI-Based Diagnosis of Liver Disease Using the Indian Liver Patient Dataset

Milan Choudhary, B.Tech CSE (AI), Medi-Caps University, Indore, India

3. Exploratory Data Analysis

We examined correlation between features and plotted their statistical relationships to identify informative variables.

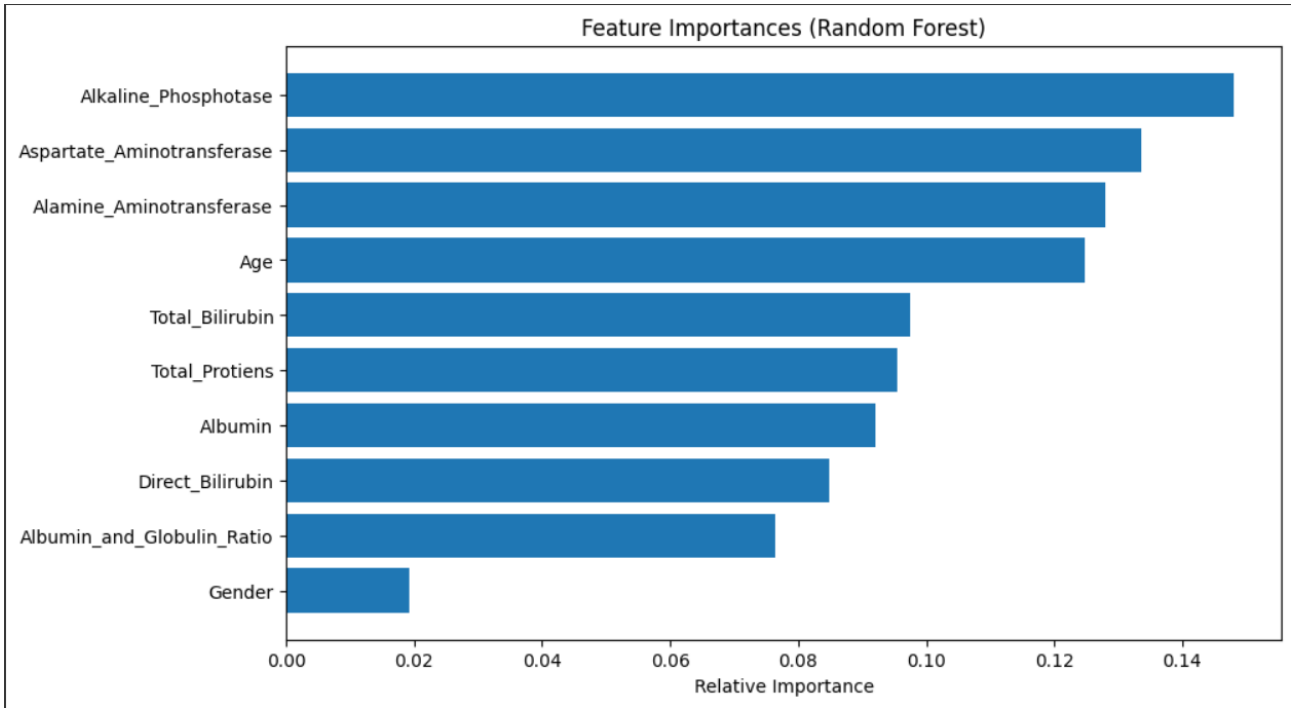


4. Model Training and Evaluation

A Random Forest Classifier was trained using an 80-20 train-test split. Feature importance was extracted to identify impactful predictors.
Evaluation metrics confirmed reliable performance with high accuracy on the test set.

AI-Based Diagnosis of Liver Disease Using the Indian Liver Patient Dataset

Milan Choudhary, B.Tech CSE (AI), Medi-Caps University, Indore, India

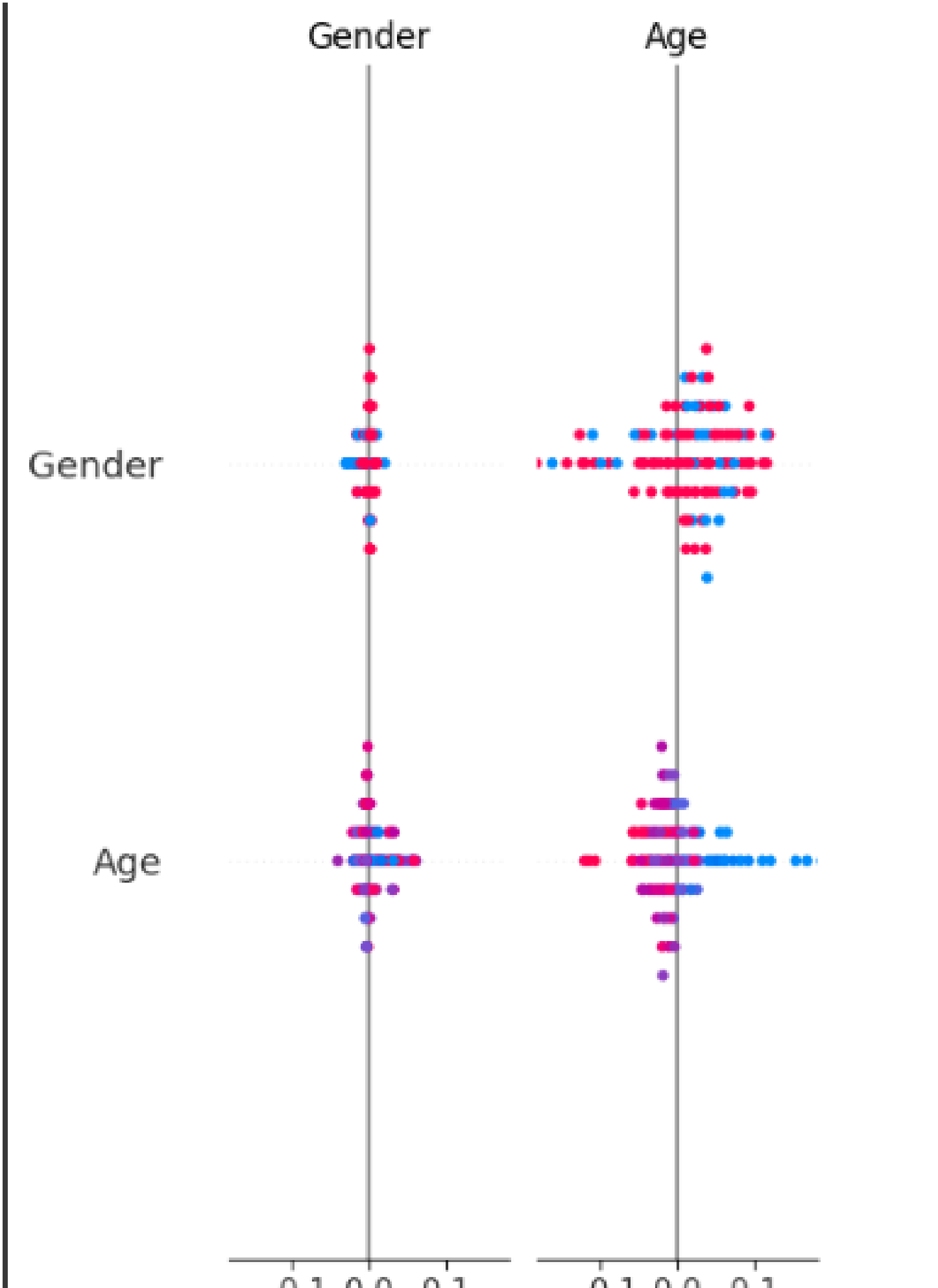


5. Explainability with SHAP

SHAP was used to explain feature-level interactions and contributions, enhancing trust in the model's decisions for healthcare use.

AI-Based Diagnosis of Liver Disease Using the Indian Liver Patient Dataset

Milan Choudhary, B.Tech CSE (AI), Medi-Caps University, Indore, India



AI-Based Diagnosis of Liver Disease Using the Indian Liver Patient Dataset

Milan Choudhary, B.Tech CSE (AI), Medi-Caps University, Indore, India

6. Gradio Interface

A user-friendly Gradio interface was developed, enabling multilingual and voice-based patient interaction to improve accessibility.

7. Conclusion and Future Scope

The proposed system demonstrated reliable detection of liver disease using clinical parameters.
Future work involves integrating this with EHR systems and expanding it to more diseases with real-time input validation.

References.

[1] UCI Machine Learning Repository - Indian Liver Patient Dataset
[2] Lundberg, S. M., & Lee, S.-I. (2017). A Unified Approach to Interpreting Model Predictions (SHAP)
[3] Breiman, L. (2001). Random Forests