Healthcare Disease Diagnosis by Rahul Wagh

Introduction

The Healthcare Disease Diagnosis initiative aims to accurately diagnose diseases by analyzing a variety of patient data and symptoms. The dataset used in this project contains the following features:

- **Disease:** The name of the disease or medical condition under consideration.
- **Fever:** Indicates whether the patient has a fever (Yes/No).
- **Cough:** Indicates whether the patient has a cough (Yes/No).
- **Fatigue:** Indicates whether the patient experiences fatigue (Yes/No).
- **Breathing Difficulty:** Indicates whether the patient has difficulty breathing (Yes/No).
- **Age:** The age of the patient in years.
- **Gender:** The gender of the patient (Male/Female).
- **BP:** The blood pressure level of the patient (Normal/High).
- **Cholesterol:** The cholesterol level of the patient (Normal/High).
- **Result:** The outcome variable indicating the result of the diagnosis or assessment for the specific disease (Positive/Negative).

Methodology

The project uses machine learning algorithms to diagnose diseases based on the given symptoms and patient information. The dataset is pre-processed by removing duplicate rows and encoding categorical variables using label encoding. The Random Forest Classifier algorithm is used to train the model, and the accuracy of the model is evaluated using various metrics such as accuracy score, classification report, and confusion matrix.

Results

The project provides various visualizations such as count plots, scatter plots, and histograms to analyze the dataset and understand the relationship between different features. The trained model is saved to a file using joblib, and the user can input patient information to get a diagnosis for a specific disease.

Conclusion

The Healthcare Disease Diagnosis project accurately diagnoses diseases based on symptoms and patient information, improving disease diagnosis in the healthcare industry.