TRAIN THE MODEL ON IBM

INTEGRATE FLASK WITH SCORING END POINT

DATE	20- NOVEMBER- 2022
TEAM MEMBER	Haripriyan S, Gowthaman M , Maigokul P , Guna Bharathi J
TEAM ID	PNT2022TMID15752
PROJECT NAME	STATISTICAL MACHINE LEARNING APPROACHES TO LIVER DISEASE PREDICTION

ABSTRACT

One of the most vital causes of death worldwide is liver disease. We, human come long way in the medical field and scientific advancements to treat diseases and it's evident that when these liver diseases are detected early, they can be treated easily.

In order to be able to accurately predict if there's a chance of the liver disease it is imperative to identify the features/symptoms which play a significant role in causing the Liver Disease. In order to improve the performance of the prediction models, it is important to choose the right combination of significant features.

A new system is proposed that identifies the significant features and then predicts whether or not a person may suffer or is suffering from Liver Disease using the identified features. Our system ought to be used as a supplementary tool in diagnosis. Data is essential and we will be using the dataset available on the UIC repository. We will be using genetic algorithms to identify.

INTRODUCTION

Problem Definition including the significance and objective
Liver is the cleaning and detoxification mechanism of our body. If there is any problem with our livers, our bodies cannot properly dispose of its wastes This can lead to several other problems. Liver diseases are responsible for around 2% of the world's deaths. Early diagnosis of these diseases helps in preventing from deaths. This project tries to help medical professionals detect liver diseases in its early stages and help reduce the rate of liver diseases.

Liver disease is a category of disease that has the negative effect on the usual working of the liver. The detection of a liver disease in its early stages is very important to prevent any adverse effects in the future. Thus, the main aim of the proposed model is to detect the liver disease using binary classification. For this purpose we will be using various machine learning classification algorithms. We will also be using genetic algorithms and deep learning techniques to improve our classification accuracy.

Liver disease is a very broad term. When we talk about liver disease there can be many diseases such as fatty liver disease, non-alcoholic fatty liver disease, hepatitis, etc. Liver diseases are responsible for over 2% of the world's deaths. So, detection of the disease in its early stages is very important. Diagnosing Liver disease involves looking at various statistics about the liver including age, total bilirubin, direct bilirubin, alkaline phosphatase, total protein, albumin and many more . A thorough liver examination is currently required to tell if a person has liver disease or not. This is not accessible to many people around the world. For this reason, we need a cost effective way to tell if a person suffers from liver disease with confidence. We planned to develop a web application where a medical professional will input various liver functioning data to know whether the person suffers from liver disease or not based on the algorithm used for prediction. Based on the results of various algorithms, the medical professional will determine if the person suffers from liver disease. Methodologies.

There are several steps involved in the process of classification of liver disease.

- Data collection
- Data preparation
- Choose a model
- Train the model
- Evaluate the model
- Parameter Tuning
- Formulate predictions

1. Data Collection Data collection refers to the collection of the required datasets for the purpose of training the models. This can be done from reputed, peer reviewed websites like Kaggle, UCI repository, etc.

2. If the required data is not available, it is our responsibility to collect the required data in sufficient quantity to train the model effectively.

Data Preparation

- Data preparation involves in transformation of our raw data into the suitable form that can beused in training our machine learning model.
- This step is also a good time to perform various visualizations of our data to look for any relevant relationships between our variables can lead to several other problems.
- Liver diseases are responsible for around 2% of the world’s deaths.

PROPOSED MODEL:

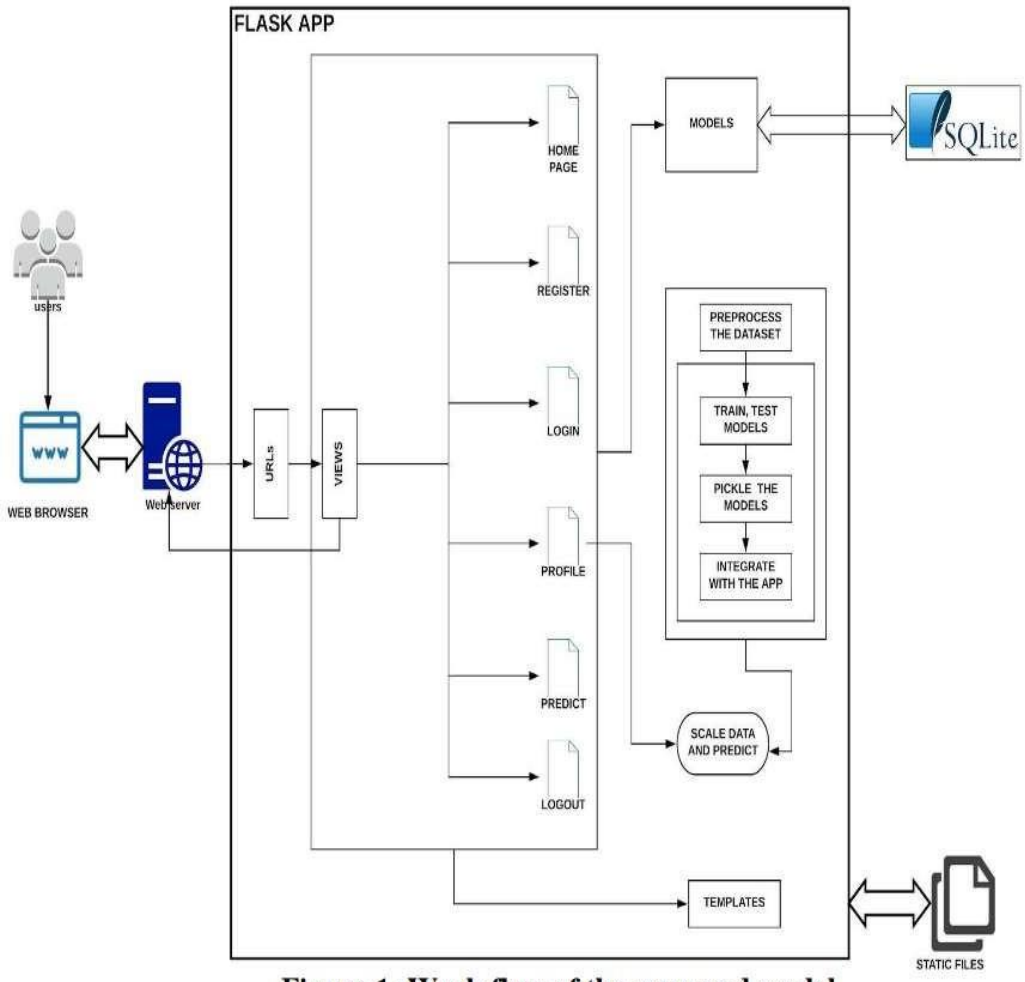

Figure-1: Work flow of the proposed model