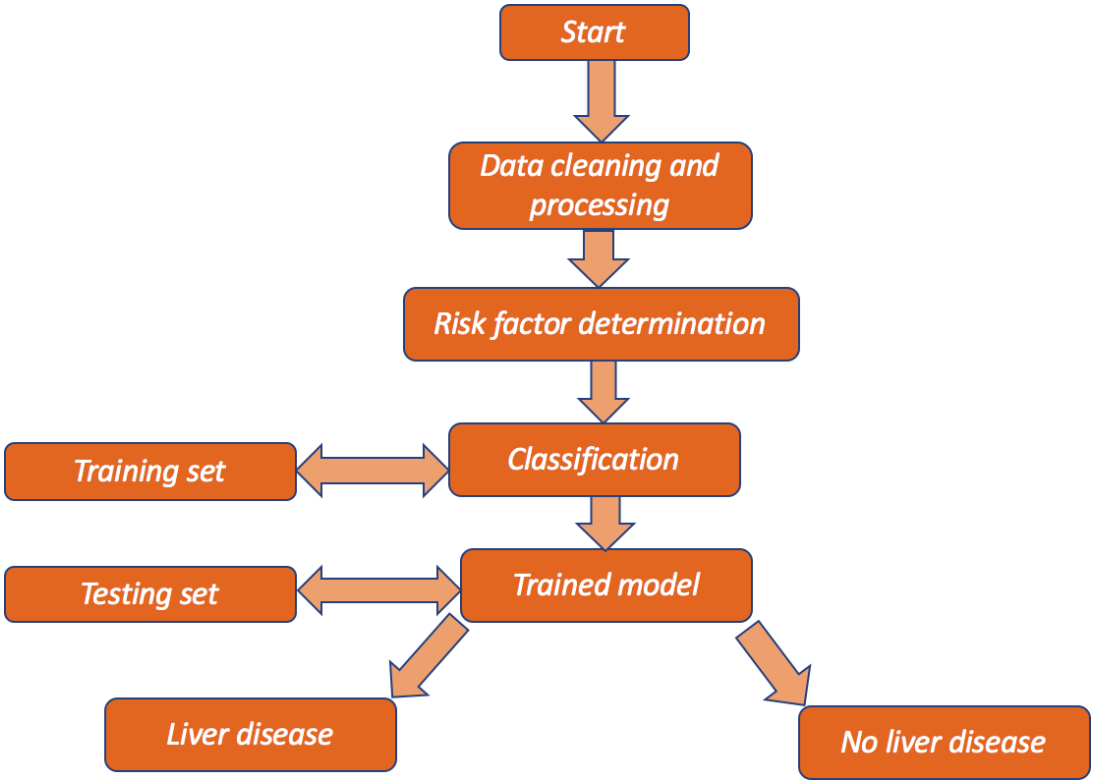
Project Design Phase-I
Solution Architecture

Date	19 September 2022
Team ID	PNT2022TMID15966
Project Name	Project - Statistical Machine Learning Approaches to Liver Disease Prediction
Maximum Marks	4 Marks

Solution Architecture:



The liver has many functions such as glucose synthesis and storage, detoxification, production of digestive enzymes, erythrocyte regulation, protein synthesis, and various other features of metabolism. Diagnosis of a certain liver disease and discovery of its origin are made by interpreting the patterns and ratios of circulating liver-associated molecules measured with the CMP test and compared to values normalized with a patient's age, sex, and BMI.

IMPORT DATASET

After downloading the dataset from Kaggle, we saved it to my working directory with the name dataset.csv. Next, we used read_csv () to read the dataset and save it to the dataset variable. Before any analysis, I just wanted to take a look at the data. So, we used the info () method. There are a total of 13 features and 1 target variable. Also, there are no missing values so we don't need to take care of any null values. Next, we used to describe () method

Parameters:

1.Age

2. Sex

3. Stomach pain

4.BP

5.Cholesterol

6. Swell Legs

7. Itchy skin

8. Infection

9.Genetics

10.ST depression

11.LiverDisease

Age: Age is the most important risk factor in developing cardiovascular or heart diseases, with approximately a tripling of risk with each decade of life. Coronary fatty streaks can begin to form in adolescence. It is estimated that 82 percent of people who die of coronary heart disease are 65 and older. Simultaneously, the risk of stroke doubles every decade after age 55.

Sex: Men are at greater risk of heart disease than pre-menopausal women. Once past menopause, it has been argued that a woman's risk is like a man's although more recentdata from the WHO and UN disputes this. If a female has diabetes, she is more likely to develop heart disease than a male with diabetes.

Stomach Pain: Functional reasons for a distended abdomen tend to involve digestive problems that cause gas and/or digestive contents to accumulate.

Blood Pressure: Over time, high blood pressure can damage arteries that feed your heart. High blood pressure that occurs with other conditions, such as obesity, high cholesterol or diabetes, increases your risk even more.

Cholesterol: A high level of low-density lipoprotein (LDL) cholesterol (the "bad" cholesterol) is most likely to narrow arteries. A high level of triglycerides, a type of blood fat related to your diet, also ups your risk of a heart attack. However, a high level of high-density lipoprotein (HDL) cholesterol (the "good" cholesterol) lowers your risk of a heart attack.

Swell Legs: Leg swelling caused by the retention of fluid in leg tissues is known as peripheral edema. It can be caused by a problem with the venous circulation system, the lymphatic system or the kidneys

Itchy skin: Pruritus can develop at any stage of cholestatic liver disease, and it should be particularly noted that severity of cholestatic itch is independent of the duration, biochemical severity, and histological stage of the underlying liver disease.

Hepatitis Virus: Parasites and viruses can infect the liver, causing inflammation that reduces liver function. The viruses that cause liver damage can be spread through blood or semen, contaminated food or water, or close contact with a person who is infected. The most common types of liver infection are hepatitis viruses, including:

- Hepatitis A
- Hepatitis B
- Hepatitis C

Genetics Problem: An abnormal gene inherited from one or both of your parents can cause various substances to build up in your liver, resulting in liver damage. Genetic liver diseases include:

- Hemochromatosis
- Wilson's disease
- Alpha-1 antitrypsin deficiency

Liver Disease: Chronic liver disease is detected by clinicians who are well trained in identifying significant observations and classifying them as normal or abnormal using background information and other context clues. ML algorithms can be trained to detect the possibility of liver disease in a similar way to assist healthcare workers. Using the correlation of each variable with the risk of liver disease to train the model, ML methods were able to identify which blood donors were healthy and which had liver disease with high accuracy. In the future, the local interpretable model-agnostic explanation (LIME) method will be used to understand the model's interpretability. Instead of binary classification, one may use multinomial classification by separating the types of liver disease. In this way, each model's performance can be compared. The described ML methods can assist health sectors to achieve a better diagnosis providing effective results in identifying groups or levels within medical data to facilitate healthcare workers. Moreover, ML methods are data driven, and they directly use diagnostic variables from patients' medical tests. Thus, it is a more reliable process. The applied ML methods in this article can save time, costs, and potentially lives for the betterment of disease diagnosis. The machine learning algorithms presented in this study can support medical experts but are not the alternative when making decisions from ML classifiers for diagnostic pathways. These methods can reduce many of the limitations that occur in healthcare associated with inaccuracy in diagnoses, missing data, cost, and time. Application of the ML methods can help reduce the total burden of liver disease on public health worldwide by improving recognition of risk factors and diagnostic variables. More importantly, for chronic liver disease, detecting liver disease at earlier stages or in hidden cases by ML could decrease liver-related mortality, transplants, and/or hospitalizations. Early detection improves prognosis, since treatment can be given before progression of the disease to later stages. Invasive tests, such as biopsy, would occur less in this case as well. Although this study focused on hepatitis and chronic liver disease variables for ML training, it can be hypothesized that the methods can be used to distinguish other types of liver disease from

healthy individuals. Applying all of the mentioned methods to other areas of medicine could open the doors for AI/ML-facilitated diagnosis.