Project Design Phase-I
Proposed Solution

Date	24 September 2022
Team ID	PNT2022TMID00067
Project Name	Statistical Machine Learning Approaches to Liver Disease Prediction
Maximum Marks	2 Marks

Proposed Solution:

S.No.	Parameter	Description
1.	Problem Statement (Problem to be solved)	Discovering the existence of liver disease at an early stage is a complex task for the doctors. The main objective of the project is to examine data from liver patients concentrating on relationships between a key list of liver enzymes, proteins, age and gender using them to try and predict the likeliness of liver disease.
2.	Idea / Solution description	Our solution is to build a model by applying various machine learning algorithms and find the best accurate model to predict whether as a liver disease or not. We plan to perform data-pre-processing and data visualization methods to increase the accuracy of the model. And integrate the chosen model into Flask based web application where the User can predict the disease by entering parameters in the web application.
3.	Novelty / Uniqueness	• Data pre-processing which includes Data Cleaning, Data transformation, and Data Reduction is performed to increase the accuracy of the model. • Various Machine model is implemented and the highest accurate model is chosen. • Model output is evaluated using MSE, confusion matrix and other various metrics. • ROC-AUC considers the rank of the output probabilities and intuitively measures the likelihood that model can distinguish between a positive point and a negative point. We will use AUC to select the best model among the various machine learning models.
4.	Social Impact / Customer Satisfaction	Since the likeliness of the liver disease is predicted with high accuracy, user can able to take remedial measures

5.	Business Model (Revenue Model)	Revenue can be made by collaborating with Hospitals and other health related companies and Integrating subscription services to the application
6.	Scalability of the Solution	Accuracy of the model can be increased by training with large data. The model can be made to learn from the user input. Model is deployed in the web where the public from across the world can use to predict the likeliness of liver disease.