**Predicting Liver Cirrhosis Stage Based on Clinical and Biochemical Features**

**Overview of Liver Cirrhosis**

* **Liver cirrhosis** is a **chronic liver disease** where healthy liver tissue is replaced with scar tissue (fibrosis).
* This scarring **blocks blood flow** through the liver and **impairs its function**.
* It often results from long-term damage due to:
  + Chronic **alcohol abuse**
  + **Hepatitis B/C infections**
  + **Non-alcoholic fatty liver disease** (NAFLD)
* Symptoms may include fatigue, jaundice, abdominal swelling, and confusion.
* Cirrhosis is a **progressive disease** and can lead to **liver failure** or **liver cancer** if untreated.

**Objective:**

* The objective of this project is to develop a machine learning model capable of predicting the **stage of liver cirrhosis** using patient data comprising **clinical symptoms and biochemical test results**. The dataset, sourced from Cirrhosis.csv, contains multiple features such as age (converted from days to years), ascites, bilirubin, albumin, copper, platelet count, SGOT, and cholesterol, among others. The target variable is stage, which represents the severity of cirrhosis on an ordinal scale.

**Data preprocessing** involved:

* Dropping irrelevant or incomplete entries,
* Median imputation for missing values in key features (triglicerides, cholesterol, platelets, copper),
* Feature normalization and selection based on correlation heatmap analysis.

The feature matrix X and target vector y were split into an **80-20 train-test set**, ensuring a representative distribution of cirrhosis stages across sets.

A **custom neural classifier** was implemented using:

* **ReLU activation** in the output layer for multiclass classification,
* **Cross-entropy loss** as the objective function,
* **One-hot encoding** for target labels to match ReLU outputs,
* Accuracy and Precision as evaluation metrics.

**Result:**

Achieved 40% accuracy on test data and 42% accuracy on train data.

**Limitations and Future scope**

While the model achieved satisfactory performance in classifying most stages, limitations such as a relatively small sample size and imputed data pose risks of overfitting and bias. Future enhancements could include integrating ensemble models (e.g., Random Forest, XGBoost) or fine-tuned deep learning architectures with regularization and hyperparameter tuning.

This study illustrates the feasibility of using machine learning techniques to assist in **automated liver disease staging**, offering valuable clinical insights and decision support.