Data Collection and Preprocessing Phase

Date	25 July 2024
Team ID	739716
Project Title	Predicting Baseline Histological staging in HCVpatients using Machine Learning
Maximum Marks	2 Marks

Data Collection Plan & Raw Data Sources Identification Template

Elevate your data strategy with the Data Collection plan and the Raw Data Sources report, ensuring meticulous data curation and integrity for informed decision-making in every analysis and decision-making endeavor.

Section	Description
Project Overview	In summary, machine learning empowers us to predict HCV stage, aiding early diagnosis and personalized treatment decisions. As we continue to refine these models, patient outcomes will improve globally. If you have any specific questions or want to explore further, feel free to ask
Data Collection Plan	surveys and interviews HCV paitents
Raw Data Sources Identified	• **Objective**: Accurately predict the severity of HCV. • **Features**: Thirteen different blood biomarkers were used.

i

urce ame	Description	Location/URL	Format	Size	Access Permission
aggle taset	The data set comprises patients details	https://www.kaggle.com/datasets/mohamedzaghloula/hepatitis-c-virus-egyptian-patients?select=HCV-Egypt-Data.csv	xlss	csv	public