Data Collection and Preprocessing Phase

Section	Description
Project Overview	The machine learning project aims to classify lymphography data based on patient information. Using a dataset with features such as age, lymph nodes status, and other relevant attributes, the objective is to build a model that accurately classifies lymphography conditions (e.g., normal, metastases, malign lymph, fibrosis), facilitating efficient and informed decision-making in the medical diagnostic process.

Data Collection Plan	• Search for datasets related to lymphography, medical diagnostics, and patient details. • Prioritize datasets with diverse and comprehensive medical information

Date	09 JULY 2024
Team ID	740024
Project Title	Evolving efficient classification patterns in Lymphography
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.

Raw Data Sources Identified	The raw data source for this project includes the dataset obtained from UCI, a popular platform for data science repositories. The provided sample data represents a subset of the collected information, encompassing variables such as age, lymph nodes status, and other lymphography-related details for machine learning analysis

Source Name	Description	Location/URL	Format	Size	Access Permission

Raw Data Sources Template

Data set Link	The dataset comprises patient details (age, lymph nodes status), and lymphography conditions (normal, metastases, malign lymph, fibrosis).	https://archive.ics.uci.edu/ml/datasets/Lymphography	CSV	3.5 kB	Public