



## **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	739734					
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.





## Collection Plan Template



Raw Data	a Sources	The raw data source for this project includes the dataset obt UCI, a popular platform for data science repositories. The prosample data represents a subset of the collected information encompassing variables such as age, lymph nodes status, and lymphography-related details for machine learning analysis	ovided 1,	1	
Source Name					Access Permission
	Description	Location/URL	Format	Size	

**Raw Data Sources Template** 





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



