



Data Collection and Preprocessing Phase

Data Concetion and Treprocessing Thase				
Date	10 July 2024			
Team ID	739851			
Project Title				
	Beyond The Veil Of Wellness: Machine Learning's Unique Journey In Animal Health Classification			
Maximum Marks	2 Marks			

Data Quality Report:

The Data Quality Report will summarize data quality issues from the selected source, including severity levels and resolution plans. It will aid in systematically identifying and rectifying data discrepancies.

Data Quality Report:

Data Source			
	Data Quality Issue	Severity	Resolution Plan

Kaggle Dataset	The Dangerous column has 2 missing values (869 non-null out of 871 entries).	Moderate	Use mean/median imputation.
Kaggle Dataset	Categorical Data: AnimalName, symptoms1, symptoms2, symptoms3, symptoms4, symptoms5, and Dangerous.	Moderate	Encoding has to be done in the data.