Data Collection and Preprocessing Phase

Date	4 June 2024
Team ID	SWTID1720076203
Project Title	Anemia Sense: Leveraging Machine Learning for Precise Anemia Recognitions
Maximum Marks	2 Marks

Data Collection Plan Template

Section	Description
Project Overview	Anaemia - sense utilizes machine learning models trained on vast datasets of blood parameters and patient profiles to detect early signs of anaemia. By analysing key indicators such as haemoglobin levels, red blood cell counts, and other relevant biomarkers, the system can flag potential cases for further investigation by healthcare professionals. Early detection enables timely interventions and treatment plans, improving patient outcomes.
Data Collection Plan	Kaggle dataset
Raw Data Sources Identified	The raw data sources for this project include datasets obtained from Kaggle, a popular platform for data science competitions and repositories. The provided sample data represents a subset of the collected information, encompassing variables such as gender, Hemoglobin, MCH, MCV and MCHC details for machine learning analysis

Raw Data Sources Template

Source Name	Description	Location/URL	Format	Size	Access Permissions
Kaggle dataset	The dataset comprises Gender, MCV, MCH, MCHC and the overall result	https://www.kaggle.com/datasets/biswaranjanrao/anemia-dataset	CSV	37KB	Public