Data Collection and Preprocessing Phase

Date	15 July 2024
Team ID	SWTID1720151584
Project Title	Early Prediction of Chronic kidney Disease
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.

Data Collection Plan Template
Raw Data Sources Template

Raw Data Sources Identified	For raw data on CKD early prediction, use the following sources: 1. **Electronic Health Records (EHRs) from hospitals and clinics** 2. **Public health databases like NHANES** 3. **Clinical trials data** 4. **Patient surveys and questionnaires**

Source Name	Description	Location/URL	Format	Size	Access Permissions
Dataset 1	This dataset is designed for early prediction of chronic kidney disease (CKD) and includes patient information such as age, gender, diabetes and hypertension status, as well as clinical measurements like blood pressure, blood glucose, blood urea, serum creatinine, glomerular filtration rate (GFR), and urine protein levels. It aims to provide comprehensive data for developing and validating machine learning models to identify CKD at an early stage.	https://drive.google.com/file/d/1mPl4yaTKuKZ3017YfYC19Ni7Y964eCNI/view	CSV	15KB	Public