

Data Collection and Preprocessing Phase

Date	13 JULY 2024
Team ID	XXXXXXXXXX
Project Title	Detection of Autistic Spectrum Disorder: Classification
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

Section	Description
Project Overview	The project aims to develop a classification model to detect Autistic Spectrum Disorder (ASD) based on various features and indicators. ASD is a neurodevelopmental disorder that affects communication, behavior, and social interaction.
Data Collection Plan	Gather a comprehensive dataset containing features such as demographic information, behavioral traits, medical history, and possibly genetic factors related to ASD..
Raw Data Sources Identified	To effectively detect Autistic Spectrum Disorder (ASD), several raw data sources can be utilized. Clinical databases from specialized healthcare institutions provide valuable clinical records, diagnostic assessments, and medical histories of individuals with ASD.

	Publicly available research datasets, such as those from Kaggle or the UCI Machine Learning Repository, offer diverse datasets encompassing demographic information
--	---

Raw Data Sources Template

Source Name	Description	Location/URL	Format	Size	Access Permissions
Kaggle Dataset	Age Number Age in years Gender String Male or Female Ethnicity String List of common ethnicities in text format	https://www.kaggle.com/code/faizunnabi/autism-screening-classification/input	CSV	58.58 KB	Public