Data Collection and Preprocessing Phase

Date	7 July 2024
Team ID	SWTID1720426301
Project Title	Cognitive Care: Early Intervention for Alzheimer's Disease
Maximum Marks	2 Marks

Data Collection Plan & Raw Data Sources Identification Report:

The Data Collection Plan and Raw Data Sources Identification are crucial for effective data training strategies. These components involve strategic planning to identify and gather relevant data sources systematically. By ensuring the integrity and comprehensive coverage of data, organizations can enhance the accuracy and reliability of their training datasets, leading to more informed decision-making and robust model performance in data-driven analyses and applications.

Data Collection Plan

Section	Description
Project Overview	The machine learning project aims to develop robust Alzheimer's Disease diagnostic models using MRI scans. Its objective is to leverage comprehensive data collection and rigorous training strategies to enhance model accuracy and support informed medical decision-making.
Data Collection Plan	The data for this project was collected from Kaggle, a popular platform for datasets and machine learning competitions.

Raw Data Sources Identified	The raw data sources for this project include JPEG images sourced from Kaggle, categorized into four classes. These images represent various stages of Alzheimer's Disease progression.

Raw Data Sources Report:

Source Name	Description	Location/URL	Format	Size	Access Permissions
Kaggle Dataset	The data source comprises JPEG images from Kaggle representing MRI scans categorized into four classes: MildDemented, ModerateDemented, NonDemented, and VeryMildDemented, each depicting different stages of Alzheimer's Disease progression.	https://www.kaggle.com/datasets/tourist55/alzheimers-dataset-4-class-of-images	JPEG	33.0 MB	Public