Data Collection And Preprocessing Phase

Date	15 November 2024
Team ID	739909
Project Name	Unlocking the Minds: Analyzing Mental Health with NLP
Maximum Marks	2 Marks

Data Collection Plan &Raw Data Sources Identification Report:

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:

Section	Description
Project Overview	Develop a Natural Language Processing (NLP)-based system to analyze and extract insights from textual data related to mental health, enabling enhanced understanding, support, and intervention.
Data Collection Plan	- Search for datasets related to mental_health_analysis details. - Prioritize datasets with diverse demographic information.
Raw Data Sources Identified	The raw data sources for this project include datasets obtained from Kaggle, the popular platforms for data science competitions and repositories, details for machine learning Analysis and NLP text_processing.

Raw Data Sources Report:

Source Name	Description	Location/URL	Format	Size	Access Permissions
Kaggle Dataset	The dataset comprises details like (text, label classes)	https://www.kaggle.com/code/swathiunnikrishnan/mental-health-analysis-using-nlp/input	CSV	1.28 MB	Public