Data Collection and Preprocessing Phase

Date	14 July 2024
Team ID	740073
Project Title	Exploratory Analysis of Rain Fall Data in India for Agriculture machine learning project
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	The machine learning project aims to analyse rainfall data in India and understand its impact on agriculture. By examining historical rainfall patterns, we aim to identify trends, anomalies, and correlations with agricultural outputs. This analysis will serve as a foundation for developing machine learning models that can predict rainfall and its potential effects on crop yields, thereby aiding farmers and policymakers in decision-making processes.
Data Collection Plan	• Search for datasets related to rain fall , Identify and gather initial data sources. • Implement data collection methods (API integration, web scraping, manual download)

Raw Data Sources Identified	The raw data sources for this project include datasets obtained from Kaggle & UCI, the popular platforms for data science competitions and repositories. The provided sample data represents a subset of the collected information, encompassing variables such as gender,

Raw Data Sources Report

Source Name	Description	Location/URL
UCI	This data concerns credit card applications; a good mix of attributes	https://docs.google.com/spreadsheets/d/1RA2OO0LZTeQykI_mvnensAjp6LM4YzWI1Tz0SUG5Ao/edit#gid=121883362

.