**Data Collection and Preprocessing Phase**

|  |  |
| --- | --- |
| Date | 9 July 2024 |
| Team ID | SWTID1720162737 |
| Project Title | Predicting Compressive Strength Of Concrete Using Machine Learning |
| 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 machine learning project aims to predict compressive strength of concrete based on applicant information. Using a dataset with features such as cement, blast\_furnace\_slag, fly\_ash, water, superplasticizer, coarse\_aggregate, fine\_aggregate, age, concrete\_compressive\_strength ,the objective is to build a model that accurately classifies, facilitating efficient and informed decision-making in the process. |
| Data Collection Plan | * Search for datasets related to concrete in internet. * Collect the datasets from the Construction companies, concrete material suppliers or research laboratories.   ● Prioritize datasets with Data Characteristics. |
| 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 cement, blast furnace slag, fly ash, water, superplasticizer, coarse aggregate, fine aggregate, age and concrete compressive strength, details for machine learning analysis . |

**Raw Data Sources Template**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Source Name** | **Description** | **Location/URL** | **Format** | **Size** | **Access Permissions** |
| Kaggle Dataset | Data set contains attributes such as cement, blast furnace slag, fly ash, water, superplasticizer, coarse aggregate, fine aggregate, age and concrete compressive strength | https://www.kaggle.com/datasets/elikplim/concrete-compressive-strength-data-set | CSV | 11 GB | Public |
| UCI | Data set contains attributes such as cement, blast furnace slag, fly ash, water, superplasticizer, coarse aggregate, fine aggregate, age and concrete compressive strength. | https://archive.ics.uci.edu/dataset/165/concrete+compressive+strength | Excel | 32 GB | Private (with access) |