**Data Collection and Preprocessing Phase**

|  |  |
| --- | --- |
| Date | 29 April 2024 |
| Team ID | 737820 |
| Project Title | EcoForecast: AI-Powered Prediction of Carbon Monoxide Levels |
| 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 | EcoForecast is a machine learning project that uses AI algorithms to provide real-time and accurate predictions of carbon monoxide levels for environmental monitoring and health protection. |
| Data Collection Plan | EcoForecast collects data from sources such as air quality sensors, meteorological data, and satellite imagery. |
| Raw Data Sources Identified | EcoForecast uses AI algorithms to provide real-time, accurate predictions of carbon monoxide levels using raw data sources such as weather data, traffic information, industrial emissions, sensor data from monitoring stations, and historical air quality records. |

**Raw Data Sources Template**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Source Name** | **Description** | **Location/URL** | **Format** | **Size** | **Access Permissions** |
| Dataset 1 | Description of the data in this source. | https://www.kaggle.com/datasets/anubhav3242/carbon-monoxide-ppm-data-for-regression | CSV | 52KB | Public |
| Dataset 2 | Description of the data in this source. | Link of Dataset 2 | Excel | YY GB | Private (with access) |
| … | … | … | … | … | … |