|  |  |
| --- | --- |
| **Section** | **Description** |
| Project Overview | The project aims to develop a system that utilizes biosignals obtained from wearable devices to accurately estimate the presence or absence of smoking behavior in individuals. Smoking is a significant public health concern globally, and traditional methods of monitoring smoking behavior often rely on self-reporting, which can be unreliable due to biases and inaccuracies. |
| 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, |

**Data Collection and Preprocessing Phase**

|  |  |
| --- | --- |
| Date | 21 June 2024 |
| Team ID | 740037 |
| Project Title | Estimating Presence or Absence of smoking Through Bio Signals |
| 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:**

**Raw Data Sources Report:**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Source**  **Name** | **Description** | **Location/URL** | **Format** | **Size** | **Access Permissions** |
| Kaggle  Dataset | The dataset comprises applicant details (gender,age,height,waist,  weight,eyesight) | [Body signal of smoking (kaggle.com)](https://www.kaggle.com/datasets/kukuroo3/body-signal-of-smoking) | CSV | 4 MB | Public |