**Data Collection and Preprocessing Phase**

|  |  |
| --- | --- |
| Date | 9 July 2024 |
| Team ID | SWTID1720075414 |
| Project Title | Panic Disorder Detection |
| Maximum Marks | 2 Marks |

**Data Collection Plan**

|  |  |
| --- | --- |
| **Section** | **Description** |
| Project Overview | The goal of this project is to develop a reliable and automated detection system using machine learning techniques. This system will analyse patterns in physiological data (e.g., heart rate, electrodermal activity), behavioural indicators (e.g., activity levels, sleep patterns), and psychological assessments (e.g., standardised questionnaires) to identify  individuals with panic disorder. |
| Data Collection Plan | * Search and obtain dataset related to Panic Disorders * Choose the dataset that is more balanced and divergent |
| Raw Data Sources Identified | The raw datasets for this project includes dataset that was taken from Kaggle. The dataset consists of 20,000 entries with 17 columns, including demographic details, medical and psychiatric history, current stressors, symptoms, and coping mechanisms. It primarily aims to identify factors related to panic disorder diagnosis, which is represented as a binary column. |

**Raw Data Sources Template**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Source Name** | **Description** | **Location/URL** | **Format** | **Size** | **Access Permissions** |
| Kaggle Dataset | The dataset contains various information about medical history(Family History, Personal History, Medical history, Psychiatric history), Current Conditions(Current Stressors, symptoms, Impact on life, severity), Lifestyle and Support(Substance Use, Coping Mechanisms, Social Support, Lifestyle Factors) and participant information(Participant ID, age, gender, demographics). | <https://www.kaggle.com/datasets/muhammadshahidazeem/panic-disorder-detection-dataset> | CSV | 12.4 MB | Public |
| Kaggle Dataset | The dataset contains various information about medical history(Family History, Personal History, Medical history, Psychiatric history), Current Conditions(Current Stressors, symptoms, Impact on life, severity), Lifestyle and Support(Substance Use, Coping Mechanisms, Social Support, Lifestyle Factors) and participant information(Participant ID, age, gender, demographics). | <https://www.kaggle.com/datasets/muhammadshahidazeem/panic-disorder-detection-dataset> | CSV | 2.5 MB | Public |