

**Data Collection and Preprocessing Phase**

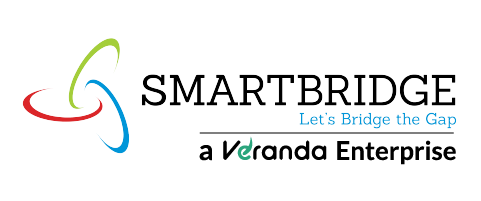
|  |  |
| --- | --- |
| Date | 15 June 2024 |
| Team ID | 739904 |
| Project Title | Income Activities Using Machine Learning |
| Maximum Marks | 2 Marks |

**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 | Machine learning (ML) has emerged as a pivotal technology reshaping income-generating activities across various sectors. By leveraging vast amounts of data, ML algorithms can extract valuable insights and drive actionable decisions, ultimately enhancing profitability and efficiency. Here’s an overview of how ML is transforming income activities.This overview outlines how machine learning technologies are transforming income activities across sectors, highlighting key benefits and applications. |
| Data Collection Plan | ● Search for datasets related to income predictions.  ● Prioritize datasets with diverse demographic information. |
|  |



|  |  |
| --- | --- |
| Raw Data Sources  Identified | The raw data sources for this project include datasets obtained from Kaggle , the popular platforms for data science competitions and repositories. The provided sample data represents a subset of the collected information, encompassing variables such as ambient pressure (AP), relativ |

**Raw Data Sources Report:**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Source Name** | **Description** | **Location/URL** | **Format** | **Size** | **Access Permissions** |
| Kaggle  Dataset | The dataset comprises details like income predictions,salary of people,collection of data. | https://www.kaggle.com/datasets/jocelyndumlao/income-activities | CSV | 51kB | Public |