

Data Collecting and Preprocessing Phase

| | |
|---------------|--|
| Date | 15 July 2024 |
| Team ID | 739703 |
| Project Title | Octagon Oracle: Machine Learning-Powered UFC FIGHT FORECAST |
| 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 | This project focuses on predicting the winner of UFC fights using machine learning. The goal is to develop a model with an accuracy greater than 90% by leveraging various features related to fighter statistics and fight outcomes. |
| Data Collection Plan | Kaggle |
| Raw Data Sources Identified | UFC Stats, Sherdog, Fight Metric |

Raw Data Sources Template

| Source Name | Description | Location/URL | Format | Size | Access Permissions |
|-------------|--|---|--------|------|--------------------|
| Dataset 1 | Fight-by-fight list of all UFC fights from 2013. | https://www.kaggle.com/datasets/calmdownkarm/ufcdataset | CSV | 4 MB | Public |