



Data Collection and Preprocessing Phase

| Date | 25 June 2024 | |
|---------------|---|--|
| Team ID | 739776 | |
| Project Title | Prediction Of Full Load Electrical Power Output Of A Base Load Operated Combined Cycle Power Plant Using Machine. | |
| Maximum Marks | 2 Marks | |

Data Collection Plan & Raw Data Sources Identification Report:

Elevate your data strategy with the Data Collection plan and the Rw 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 | | | | |
|------------------|---|--|--|--|--|
| | | | | | |
| | The machine learning project aims to prediction of full load electrical power | | | | |
| Project Overview | output of a base load operated combined cycle power plant using Machine. | | | | |
| | Using a dataset with features such as ambient pressure, relative humidity, | | | | |
| | exhaust vaccum, ambient temperature and other variables. The objective is to | | | | |
| | develop a machine learning model that accurately predicts. The objective of this | | | | |
| | project is to develop a machine learning model capable of accurately predicting the | | | | |
| | full load electrical power output of a base load operated combined cycle power | | | | |
| | plant. The prediction model should help in optimizing the plant's performance, | | | | |
| | reducing operational costs, and improving reliability. | | | | |

| Data Collection Plan | Search for datasets related to hospital readmission prediction. | | |
|----------------------|---|--|--|
| | 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), relative humidity (Rh), exhaust vaccum(V), ambient temperature(T) and other variables. |
|-----------------------------|---|
| | |

Raw Data Sources Report:

| Sour ce Na me | Descriptio n | Location/URL | Form at | Si ze | Access Permissi ons |
|------------------------|-----------------|--|------------|----------|---------------------------|
| | The dataset | https://archive.ics.uci.edu/ml/datasets/combined+cycle | | | |
| | comprises | +power+plant | | | |
| | details like | | | | |
| Kaggle | ambient | | COLL | | D 111 |
| Dataset | pressure(A | | CSV | 15 | Public |
| | P), relative | | | kB | |
| | humidity(R | | | | |
| | h), exhaust | | | | |
| | vaccum(v) | | | | |
| | , ambient | | | | |
| | temperature | | | | |
| | (T) and | | | | |
| | other | | | | |
| | variables. | | | | |

| | • | |
|--|---|--|