Data Collection and Preprocessing Phase

Date	June 2024
Team ID	Team-739815
Project Title	Power Consumption Analysis For Households
Maximum Marks	6 Marks

PreparationTemplate

The images will be preprocessed by resizing, normalizing, augmenting, denoising, adjusting contrast, detecting edges, converting color space, cropping, batch normalizing, and whitening data. These steps will enhance data quality, promote model generalization, and improve convergence during neural network training, ensuring robust and efficient performance across various computer vision tasks.

Section	Description
Data Overview	There are many popular open sources for collecting the data. Eg: kaggle.com, UCI repository, etc. In this project we have used .csv data.
Data Preparation	These are the general steps of pre-processing the data before using it for machine learning
Handling missing values	We use Handling missing values For checking the null values
Data Preparation	
Collect the dataset	Please refer to the link given below to download the dataset. Link: household_power_consumption.txt
| Importing the libraries | ```python
import pandas as pd
import numpy as np
import seaborn as sns
import matplotlib.pyplot as plt
import warnings
``` |

| Loading Data | ```python
dataset = pd.read_csv("household_power_consumption.txt",delimiter=';',low_memory=False)
dataset.head()
``` |
| --- | --- |
| Handling missing values | ```
dataset.isnull().sum()

Global_active_power     25979
Global_reactive_power   25979
Voltage                 25979
Global_intensity        25979
Sub_metering_1          25979
Sub_metering_2          25979
Sub_metering_3          25979
Year                        0
Month                       0
Sub_metering_4          25979
dtype: int64
``` |