Data Collection and Preprocessing Phase

Date	12 July 2024
Team ID	SWTID1719935963
Project Title	Automated Weather Classification using Transfer Learning
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 uses deep learning to classify weather images (sunny, cloudy, rainy, foggy, sunrise). Pre-trained models improve accuracy and efficiency. This benefits farmers (water usage), environmental agencies (fog advisories), and disaster management (early warnings). Overall, it improves weather understanding and decision-making.
Data Collection Plan	Search for datasets related to weather images and choose the suitable dataset
Raw Data Sources Identified	The raw data sources for this project include datasets obtained from Kaggle. the popular platform for data science competitions and repositories

Raw Data Sources Template

Source Name	Description	Location/URL	Format	Size	Access Permissions
Weather Classification	The dataset features 5 different classes of weather collected from the above said different sources, however it's real life data so any system for weather classification must be able to handle this sort of images. The dataset contains about 1500 labelled images including the validation images. Images are not of fixed dimensions and the photos are of different sizes. Each image has only one weather category and are saved in separate folder as of the labelled class.	https://www.kaggle.com/datasets/vijaygiitk/multiclass-weather-dataset	ZIP	141 MB	Public