Data Collection and Preprocessing Phase

Date	15 March 2024
Team ID	SWTID1720521440
Project Title	Dog Breed Identification 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	Dog breed identification is a challenging yet fascinating problem in the field of computer vision. This project aims to leverage transfer learning, a powerful technique where a pre-trained model on a large dataset is fine-tuned on a smaller, specific dataset, to identify different dog breeds from images.
Data Collection Plan	The dataset can be sourced from platforms like Kaggle (e.g., the "Dog Breed Identification" competition dataset)
Raw Data Sources Identified	The dataset can be sourced from platforms like Kaggle (e.g., the "Dog Breed Identification" competition dataset)

Raw Data Sources Template

Source Name	Description	Location/URL	Format	Size	Access Permissions
Train Dataset	The training data for the Kaggle Dog Breed Identification competition consists of labeled images of dogs belonging to various breeds, serving as the primary dataset for model training	https://www.kaggle.com/competitions/dog-breed-identification/data?select=train	ZIP(images)	10.2k	Public
Test Dataset	The test data for the Kaggle Dog Breed Identification competition comprises unlabeled images of dogs, used for	https://www.kaggle.com/competitions/dog-breed-	ZIP(images)	10.4k	Public

	evaluating the trained models' performance and generating predictions for breed identification.	identification/data?select=train			
Labels	The labels data for the Kaggle Dog Breed Identification competition provides breed annotations for the training images	https://www.kaggle.com/competitions/dog-breed-identification/data?select=train	.CSV	482KB	Public
Sample Submission	The sample submission data for the Kaggle Dog Breed Identification competition outlines the required format for result submissions	https://www.kaggle.com/competitions/dog-breed-identification/data?select=train	.CSV	25.2MB	Public