Data Collection and Preprocessing Phase

Date	20 October 2024
Team ID	739755
Project Title	Bird Species Classification
Maximum Marks	2 Marks

Data Collection Plan & Raw Data Sources Identification Template

For this project, the data will be collected from the publicly available 200 Bird Species with 11,788 Images dataset on Kaggle. This dataset includes 11,788 labeled images of 200 different bird species, which will serve as the primary source for training the bird species classification model. The dataset can be downloaded in JPEG image format along with a CSV file containing the corresponding species labels for each image.

Data Collection Plan Template

Section	Description
Project Overview	This project focuses on Bird Species Classification using a Convolutional Neural Network (CNN). The primary objective is to classify bird species from images using deep learning techniques. The model will be trained on a dataset of bird species and will aim to accurately identify and classify different species based on their images.
Data Collection Plan	The data for this project will be collected from a publicly available dataset on Kaggle, specifically the 200 Bird Species with 11,788 Images dataset. The images will be used to train and evaluate a deep learning model for bird species classification. The dataset includes labeled images of birds from various species and will serve as the primary data source.
Raw Data Sources Identified	This dataset includes 11,788 labeled images of 200 different bird species. It is hosted on Kaggle and contains diverse images of birds taken in different environments and angles, which will help train a robust model. The images are labeled with the species name, and the dataset also provides metadata about each image (such as the bird species and the image filename).

Raw Data Sources Template

Source Name	Description	Location/URL	Format	Size	Access Permissions
Dataset 1	Description of the data in this source.	https://www.kaggle.com/datasets/veeralakrishna/200-bird-species-with-11788-images	CSV	**1.2** GB	Public