Data Collection and Preprocessing Phase

Date	12 July 2024
Team ID	SWTID1720420728
Project Title	Dog Breed Identification Using Transfer Learning
Maximum Marks	2 Marks

Data Quality Report Template

The Data Quality Report Template will summarize data quality issues from the selected source, including severity levels and resolution plans. It will aid in systematically identifying and rectifying data discrepancies.

Data Source	Data Quality Issue	Severity	Resolution Plan
Dataset	class imbalance, data quality issues, breed variability, annotation errors, overfitting, augmentation limitations, and transfer learning suitability, impacting model robustness and performance.	Moderate	Addressing issues in dog breed identification involves balancing class distribution, improving image quality, simulating breed-specific variations, ensuring accurate labeling, preventing overfitting, enhancing augmentation strategies, and aligning with dataset characteristics for transfer learning effectiveness.

| Dataset | Insufficient representation of certain breed variations, such as coat patterns or body shapes, may hinder the model's ability to differentiate between closely related breeds with similar appearances. | Moderate | Augment underrepresented breed features, generate diverse variations synthetically, adapt learning focus, combine models emphasizing variations, and transfer insights from related domains for breed sensitivity. |