# Model Development Phase

| | |
|---|---|
| Date | 17th June 2025 |
| Team ID | SWTID1749820017 |
| Project Name | Dog Breed Identification using Transfer Learning |
| Maximum Marks | 3 Marks |

## Feature Selection Report

Feature selection is the process of identifying relevant inputs that contribute to the accuracy and efficiency of a model. This feature selection report is for a convolutional neural network that has been trained to recognize various dog breeds.

Usually in traditional structured datasets, features are numerical or categorical. But in our dog classification model we work with unstructured image data. Here the features of an image are directly derived from the pixel values of the image. However, other metadata such as file names, folder structures, image formats are also part of the dataset and their relevance must be evaluated for their relevance to the classification task.

This report lists out all the available features and indicates whether it was selected for model training and provides a concise explanation for its inclusion or exclusion.

| Feature | Description | Selected (Yes/No) | Reasoning |
|---|---|---|---|
| Folder Name | Each folder name includes the dog breed. (example: n02093256-Staffordshire_bullterrier) | Yes | Used as the label or target class during training and validation |
| Image File Name | Unique file name for each image. (example: n02093256_7577.jpg) | No | Used only for image identification. Not relevant to model performance |

| | | | |
|---|---|---|---|
| Image Pixel Data | The actual pixel values extracted from the images (resized to 331x331) | Yes | Core feature used as input to the CNN model |
| Image Size (KB) | The file size in kilobytes | No | Not a reliable indicator of image quality or class. Ignored during training |
| Image Format | All images are in .jpg format | No | Unfirm format – no variation to learn from. Not useful as a feature |
| Breed Name | Dog breed extracted from folder name (Example: Staffordshire Bullterrier) | Yes | Used for human-readable class mapping and reporting, though not directly fed into the model. |