

Act report

This report will illustrate the analysis and visualisation steps of the project. Once again, the data used to produce the analysis and visualisation are gathered from three resources:

1. Enhanced Twitter Archive
2. Twitter API data
3. Image Predictions File

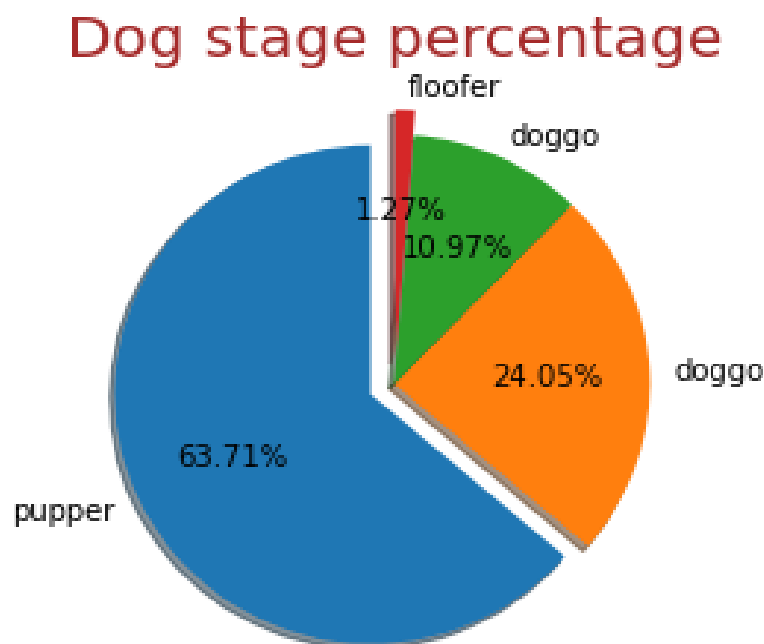
- **Analysing and Visualizing Data:**

In this part of the project, there are three insights produced after the analysis, along with some visualisation.

- 1- **Insight #1:** The goal here is to discover what the most popular dog stage is in the twitter_archive_master dataset. After conducting the analysis, the following results were found:

Dog Stage	Number of the dogs	percentage %
Popper	151	63.71
Doggo	57	24.05
Puppo	26	10.97
Floofer	3	1.27

The most popular dog stage has been found to be pupper which forms the majority with a high percentage of 63.71%. This is illustrated in the following pie chart:



2- Insight #2: The twitter_archive_master dataset includes three predictions sourced from the neural network. There are three columns for predictions of the dog breed. This insight aims to find the most common dog breed in each prediction. Therefore, the most common breed based on the three predictions is as follows:

- Prediction 1: Golden retriever (128)
- Prediction 2: Labrador retriever (90)
- Prediction 3: Labrador retriever (66)

3- Insight #3: For this insight, I will explore whether the three prediction confidences are accurate or not. Each prediction confidence column will be averaged. Also, the "true" and "false" will be counted to find their percentages. Therefore, if the percentage of the decision is higher and the average of the prediction is higher, then the prediction is accurate, and vice versa. After conducting the analysis, the following results have been found:

- Prediction 1 is accurate
- Prediction 2 is not accurate
- Prediction 3 is not accurate