Act report
Data Analysis and Visualization

Introduction

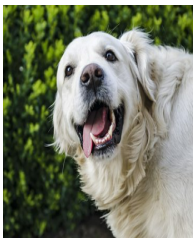
That document include the insights and displays the visualizations produced from the data.

dog breeds predictions

Using a neural network, the 'Image predictions' table stores classification results for dog breeds. How does this model work? Is this model accurate? In order to visualize and analyze the results, I have presented them below:

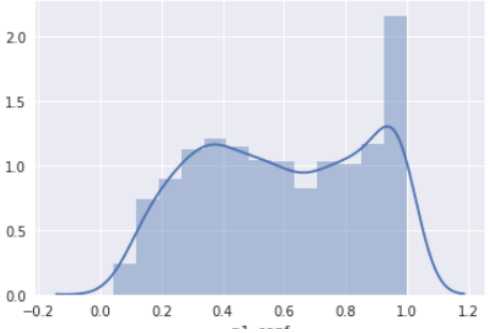
golden_retriever	150
Labrador_retriever	100
Pembroke	89
Chihuahua	83
pug	57
chow	44
Samoyed	43
toy_poodle	39
Pomeranian	38
cocker_spaniel	30

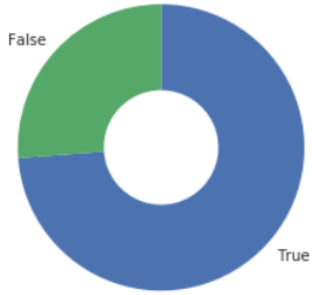
These are the top 10 dog breeds this model predicted.
Top 2 predictions are Golden retrievers and Labrador retrievers.



Here they are!! So cute!!

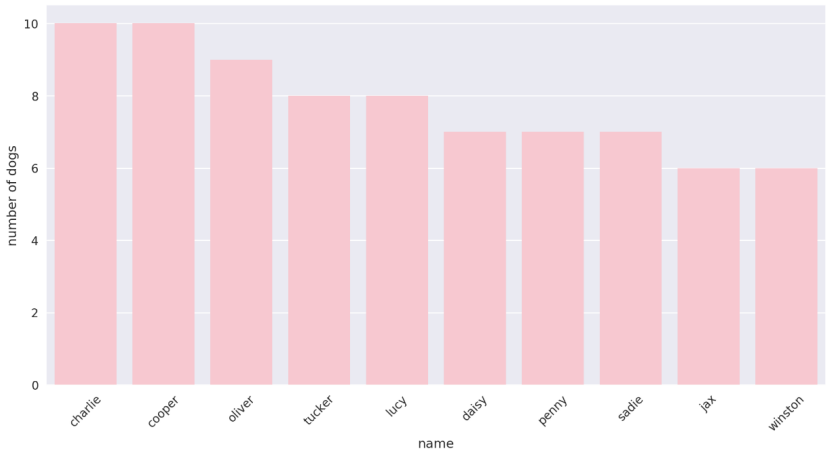
There may be a reason for this. Those are the two most common breeds in the United States. Because we have more images of those breeds, we were able to train a better model.



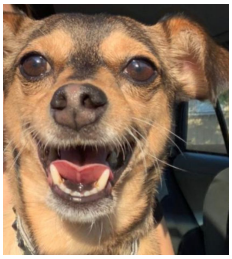


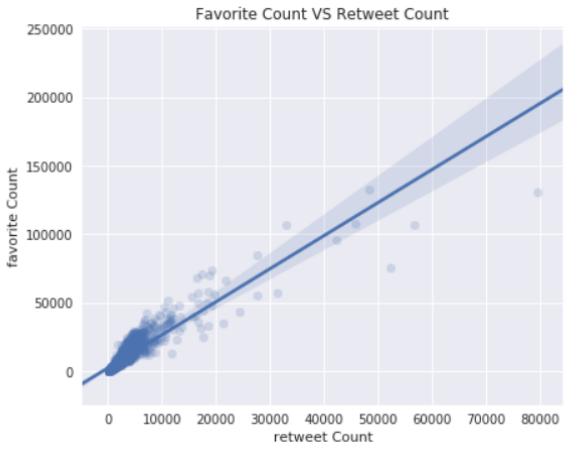
The left plot shows how confident the algorithm is in its first prediction.
There are more than 100 cases, but the distribution is centered around 0.1 to 0.8. This could also indicate that the model is not accurate.
The pie chart shows the prediction success rate of whether or not first prediction is a breed of dog.
The pie chart shows that almost 2/3 predictions are correct, but this result is not good enough for a deep learning model.

Most common dog names



We can see that Charlie and Cooper are the most popular dog names.
MY dog name is Happy , because he is happy all the time, look at him !!





Favorite Count and Retweet Count

There is a reasonable hypothesis that most popular tweets get a large number of retweets and favorites.
I test the correlation between ‘retweet_count’ and ‘favorite_count’ and I discovered that there is a positive correlation between favorite counts and retweet counts.