

ACT REPORT

‘WeRateDogs’ is a very popular twitter handle, with millions of followers famous for rating dogs in a humorous way. In this report we’re going to analyze its twitter data gathered from three different sources, and then assessed and cleaned to derive some interesting insights out of it.

Dog pictures from tweets were extracted and fed into the neural network to see if the algorithm can predict correctly the breed of the dog and whether or not it is a dog. The algorithm did a pretty neat job in identifying, however there were some cases where it went crazy. I wanted to study the cases where it was not identified as a dog and then check those images to find out if the algo was right. It was interesting to explore those cases where neural network failed to identify a dog and why.

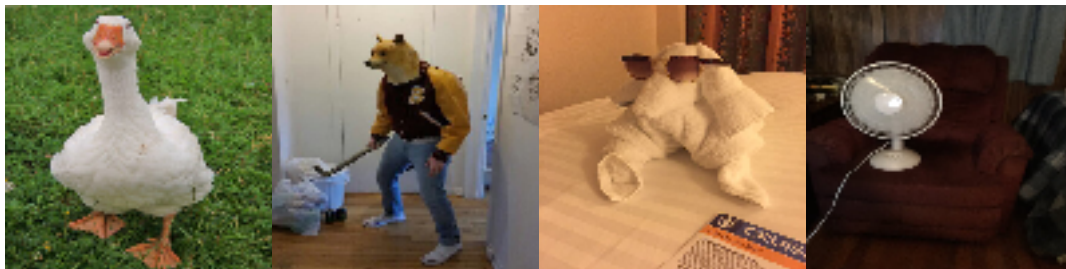
I studied two cases –

1. all the three predictions call it not a dog
2. the first prediction calls it not a dog but the second and third prediction does.

All the three predictions call it not a dog-the finding:

Cases where Algo got it right:

1-In most of these pictures it’s a look alike but clearly not a dog. Well done AI!!



2-People have presumably sent pictures of their toys and pets to rate. Who can call them a dog? Good Job AI!

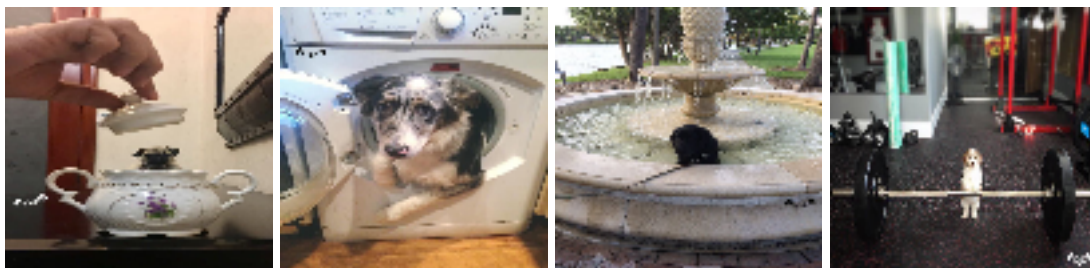


3- it was a cartoon image instead of an image of a real dog

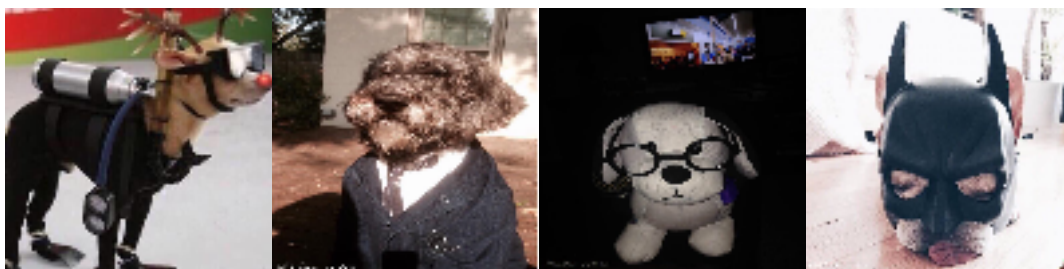


Cases where the Algo got it wrong:

1-some other animal or object dominated the picture



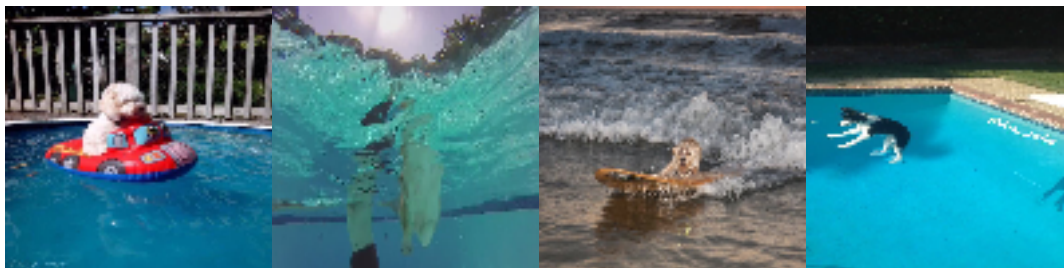
2-dog was wearing a mask or some fancy object which hid its face and body



3-dog which was sleeping or was camouflaged

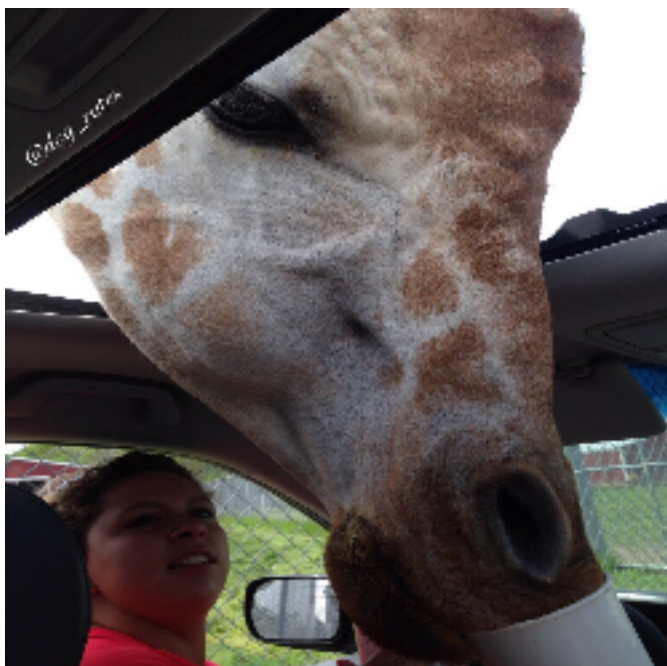


4-dog in water



The first prediction calls it not a dog but the second and third prediction does-the finding:

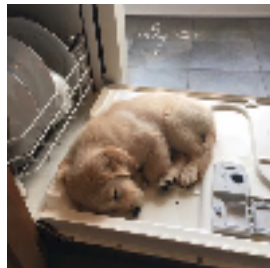
All of them are dog pictures except one. The algo failed to identify this giraffe! The first prediction calls it 'seat belt', the second 'Great Dane' and the third 'Weimaraner'



Why the algo got it wrong :

- some other object dominated the picture, so algorithm predicted that object in the first try and then predicted the dog in the subsequent tries.

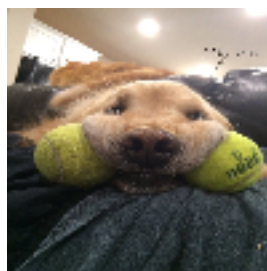
-1st predict: Dishwasher , 2nd predict: Golden Retriever , 3rd predict: Chow



-1st predict: Refrigerator , 2nd predict: Malinois , 3rd predict: Kelpie



- dogs are holding some object e.g. tennis ball in the mouth
-1st predict: Tennis ball , 2nd predict: Chesapeake Bay Retriever , 3rd predict: Labrador Retriever



Retweets, Favorited Tweets and Dog Ratings

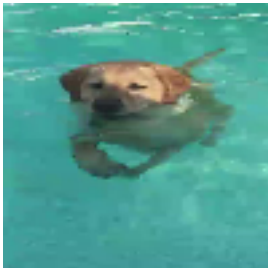
I compared the ratings of the 100 most retweeted tweets with the average ratings for all the dogs. The most retweeted had a significantly higher rating (1.65), almost 50% higher than the average ratings of dogs (1.1)

I compared the ratings of the 100 most favorited tweets with the average ratings for all the dogs. The most favorited had a slightly higher rating (1.26) compared to an average rating of dogs (1.16)

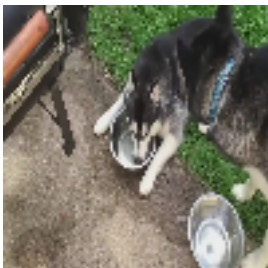
Fun Analysis

Here are the three most retweeted tweets:

1-Here's a doggo realizing you can stand in a pool. 13/10 enlightened



2-Here's a doggo blowing bubbles. It's downright legendary. 13/10 would watch on repeat forever

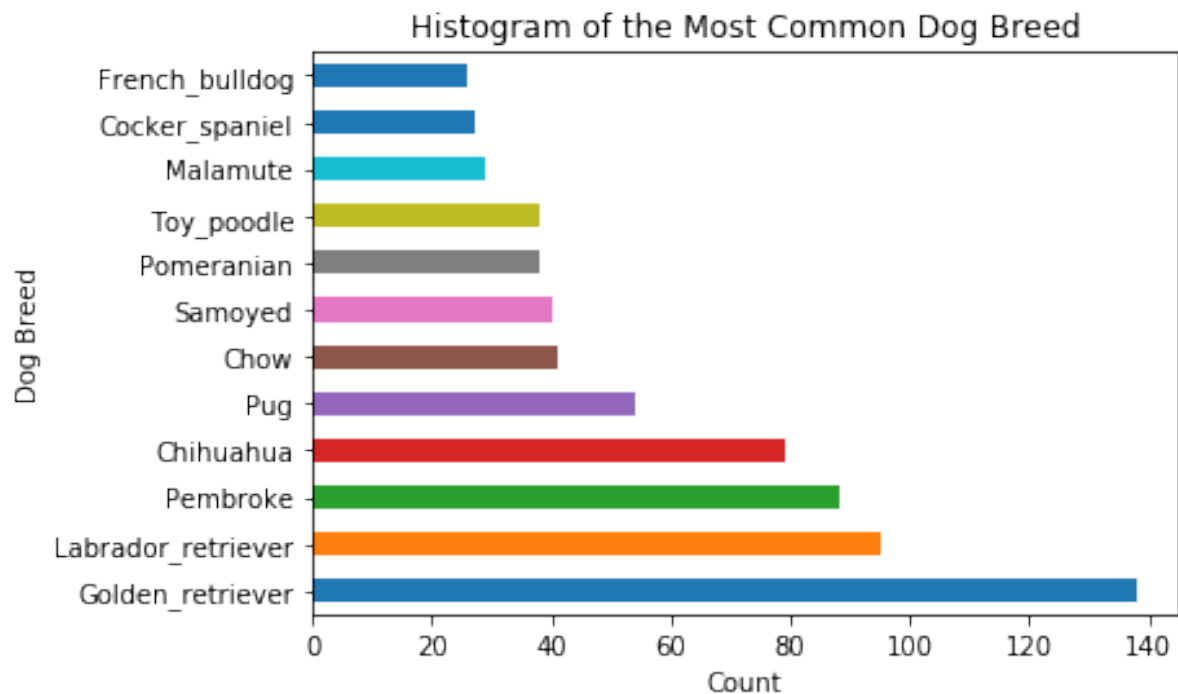


3-This is Stephan. He just wants to help. 13/10 such a good boy



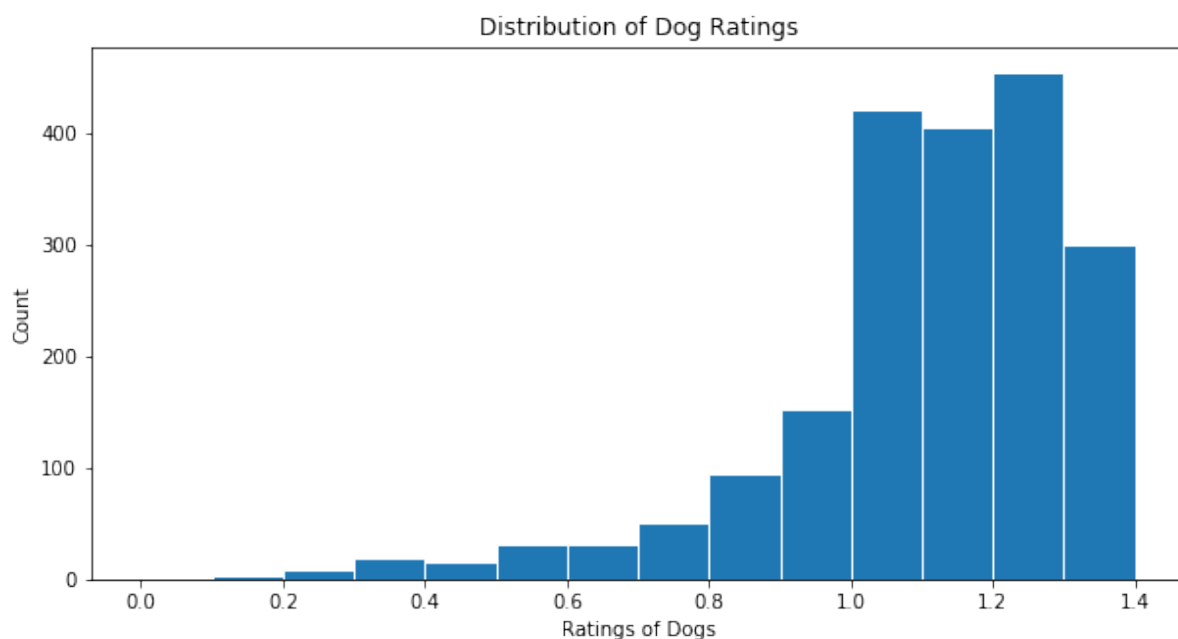
Most Common Dog Breeds

The most common dog breeds were studied through a histogram and Golden Retriever and Labrador Retriever were found to be most common.

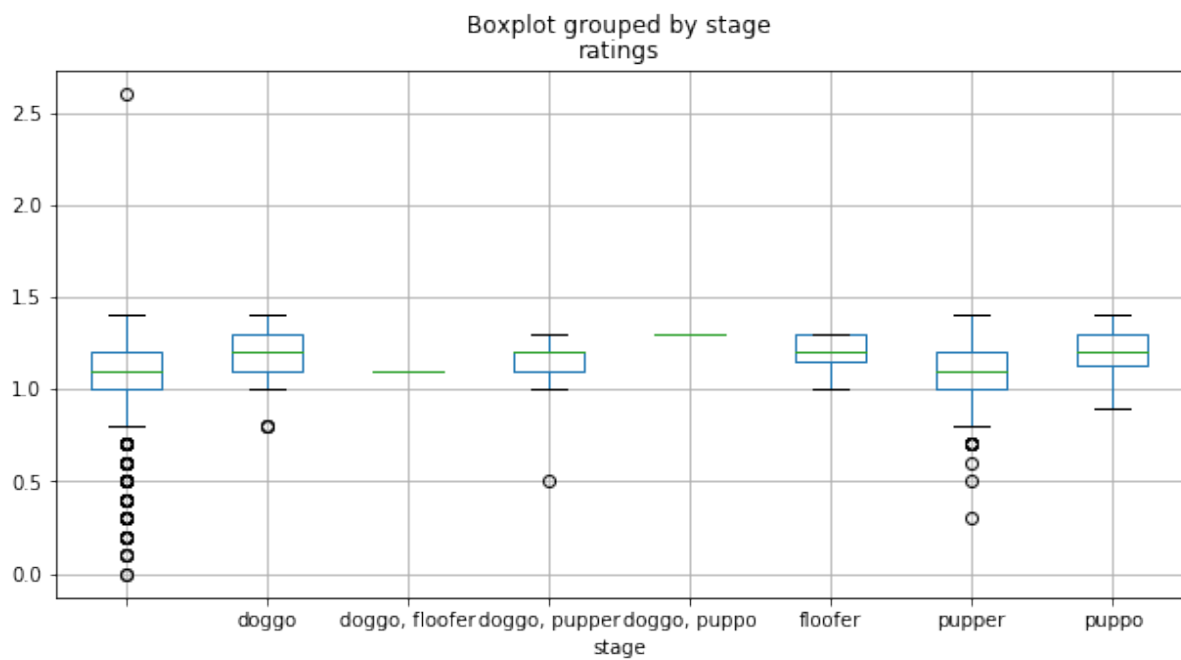


Distribution of Dog Ratings

Most dogs have been rated 10 and above (Why? Because “they’re good dogs Brent”) since most of the dogs have ratings above 1.



Boxplot Rating Distribution by dog stages



- Pupper has the lowest median rating (1.1) and doggo-pupper has the highest (1.3) also Pupper is the most common stage.
- Most dogs have a median rating of 1.2 and all the dogs have Q1s above 1 that means
- 75% of the dogs are rated more than 10
- Floofer has the least representation in terms of data, therefore its distribution is not very reliable
- Doggo-Floofer and Doggo-Puppo have just one data point so the box plot is not useful