Final Project Report

Lindsey Schaeffer

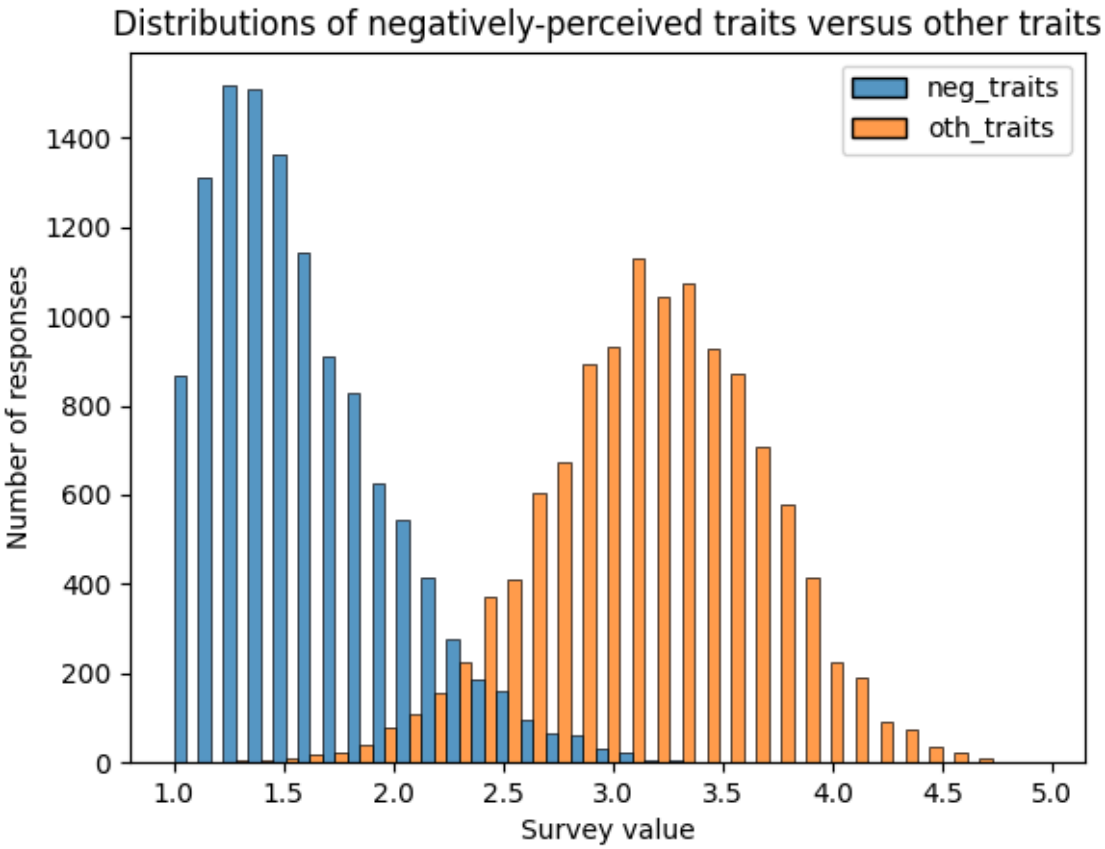
Is a pitbull more likely to bite than a labrador?

With so many dogs around us, of so many breeds, it is a question on the minds of many. Behavior is a difficult thing to quantify or predict, whether it be human or canine. In the name of safety, many apartment complexes impose breed restrictions on their tenants' dogs. Browsing through an animal shelter, one will find countless pit bulls and similarly-proportioned dogs. But should a dog's breed condemn them to a life in a cage?

The CBARQ (Canine Behavioral Assessment and Research Questionnaire) survey seeks to uncover some truths behind canine behavior. The dataset contains owner-provided information on around 12,000 dogs regarding 14 behaviors. These include aggressive (stranger-directed, owner-directed, dog-directed, and dog rivalry), fearful (dog-directed, stranger-directed, nonsocial, and separation-related), and other tendencies (trainability, chasing, touch sensitivity, excitability, attachment and attention-seeking behavior, and energy level). The survey was filled out by owners using a 1-5 scale.

Since the questionnaire was filled out by the dogs' owners, it unfortunately shows a lot of information bias. Negatively-associated traits (such as aggressive and fearful behaviors) are highly right-skewed, with values overwhelmingly on the low side of the scale. While it is possible that most dogs do not exhibit these behaviors, a low standard deviation and the extremity of these distribution differences hints at owner bias. It is

hard to admit that one's dog acts aggressively!



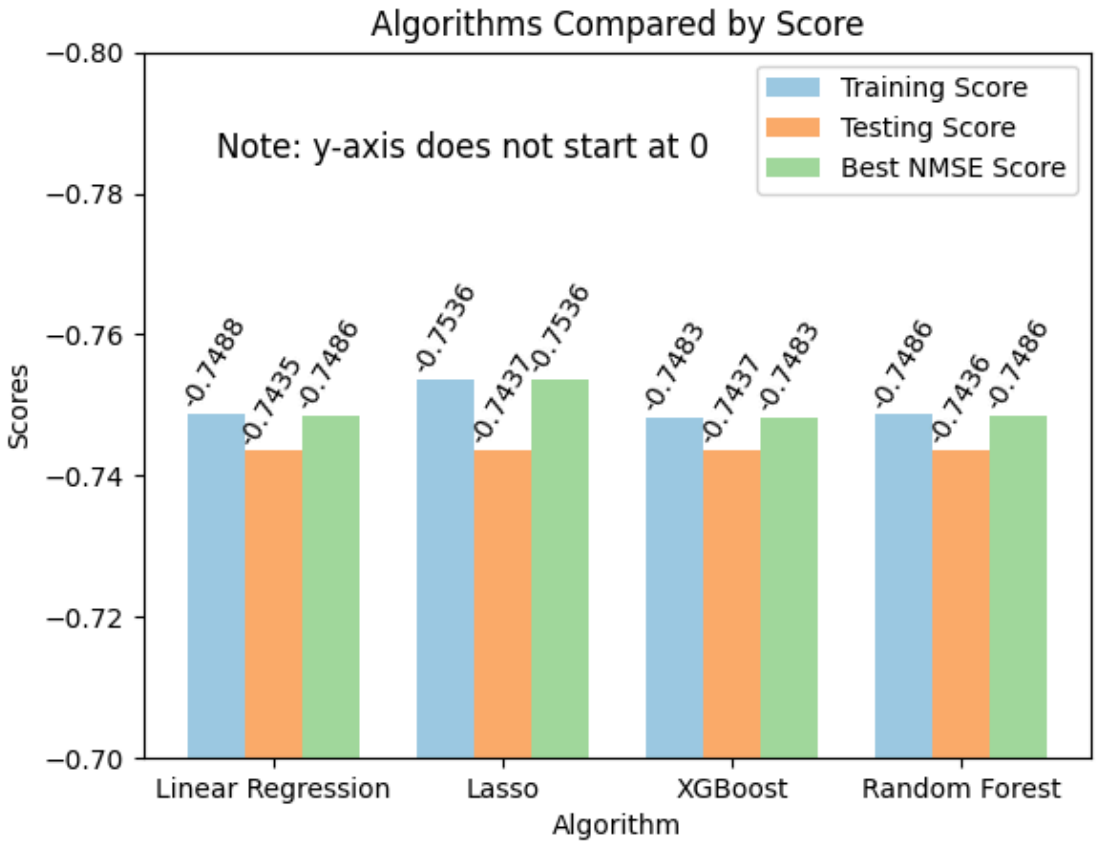
Therefore, unfortunately, this survey is not the answer to the question of breed-related behavior. However, it still can tell us something interesting! Looking at the "oth_traits" (other traits, as opposed to those that are negatively-associated) distribution, we see a nice, relatively even bell curve. This means that those traits are still viable to be examined.

So, after determining that the negative-based traits were too biased to provide any real benefit, we were left with a dataset examining breeds and 5 behaviors: chasing, excitability, energy, trainability, and attachment-related or attention-seeking behaviors. Of these, trainability still had a bias, although to the opposite end of the negative traits: it was left-skewed. After all, who doesn't want to believe their dog is well-trained? It reminded me of the quote "Everyone thinks they are an average driver". However, this bias was not nearly as significant than the right biases, so we kept the category.

After some more examination and some algorithmic testing, it seemed as though there could be a correlation found between breed and these behaviors! For simplicity's sake, I limited the dataset to the 10 most common breeds: Australian Shepherd, Border Collie, Doberman Pinscher, German Shepherd, Golden Retriever, Labrador Retriever, Mixed Breed/Unknown, Standard Poodle, Rottweiler, and Soft-Coated Wheaten Terrier.

These differences, however, were small enough that they were certainly not obvious in a graph and were rather only detectable through computational testing.

Finally, we created a machine learning model to help predict these differences. Several algorithms were created for comparison, including a Linear Regression, a Lasso Regression, an XGBoost, and a Random Forest Regressor. In the end, the Linear Regression model won, with the lowest difference between training and testing values and one of the highest negative mean-squared errors overall.



So, it does appear that there is some correlation between breed and behavior–at least, for these five traits. However, in order to get to the relationship between breed

and aggression, we will require a far more scientific approach, one conducted by professionals and not (albeit loving) owners.

In addition, I would add in other categories: for traits, barking, and for dogs, I would add background (whether they were rescued, bred, or acquired some other way), age, sex, and fixed status (whether or not they are neutered/spayed). These, I think, have a significant impact on a dog's behavior and are worth exploring, to avoid a "correlation without causation" situation.

Even though the CBARQ survey is not the answer we may want it to be, it does provide some hope for uncovering the mysteries behind canine behavior in the future, for more responsible breeding, dog ownership, and canine management.