**Assignment 1 Executive Summary**

The goal of this analysis was to apply a convnet from scratch, and pre-trained to image data of cats and dogs, evaluate their performance based on sample sizes and any data augmentation used. Sample size and using a pre-trained model vs a model built from scratch greatly changed the effectiveness of the models on test data.

**Findings for Each Step in Model Development**

1. **Initial training sample of 1000, validation of 500, and test sample of 500, with dropout of 50%:**
   1. Using the parameters listed above as the starting point, I used batch sizes of 20, 10, and 10 for training, validation, and testing, while using 50 steps per epoch. The model achieved 64% accuracy and a loss of 63% before using any kind of data augmentation or dropout. Once I used data augmentation as well as dropout of 50%, the model achieved 65% accuracy and a loss of 65%. Not very encouraging.
2. **Increased training sample size (I chose 2000 for training):**
   1. Using the parameters listed above as the starting point, I used batch sizes of 20, 10, and 10 for training, validation, and testing, while using 50 steps per epoch. The model achieved 71% accuracy and a loss of 59% before using any kind of data augmentation or dropout. Once I used data augmentation as well as dropout of 50%, the model achieved 61% accuracy and a loss of 69%.
3. **Better performance sample size (trying 1750 training)**
   1. I achieved 66% accuracy and a loss of 63% with 1750 training samples, while keeping all else the same in the model. I can’t seem to break the mid 60’s for accuracy.
4. **Pre-Trained Model:**
   1. The most accurate Pre-trained model achieved an amazing 97% accuracy. No matter the samples chosen, this model seemed to achieve much higher accuracy right out of the gate.

**Recommendations:**

After running several scenarios of all models, the pre-trained model had a much larger effect on the accuracy and loss of the model than the sample size selection did. My recommendation would be to always use a pre-trained model while selecting as many training samples as possible – using data augmentation if needed as well. This will give you the most accurate results when using a convnet model.

**Appendix:**

Relevant charts

First try: 1000 training, 500 validation and testing samples:A graph of training and validation loss

AI-generated content may be incorrect. A graph of training and validation loss

AI-generated content may be incorrect.

Second Try: 2000 training, 1000 validation and testing samples:

A graph with blue dots and white text

AI-generated content may be incorrect. A graph with blue dots

AI-generated content may be incorrect.

Third Try:

A graph with blue lines and white text

AI-generated content may be incorrect. A graph with blue lines and dots

AI-generated content may be incorrect.

Pretrained Model:

A graph with blue dots

AI-generated content may be incorrect. A graph with blue dots

AI-generated content may be incorrect.