Dalya Adams
MSDS 422 section 59
Cat vs Dog Classification

As humans, the ability to distinguish between different objects is easy. Telling the difference between shoes and sandals, trees and flowers, cats and dogs is not a challenge for humans. Despite the ease with which humans can complete this task, machines have a more difficult time. In its current state, machine learning is incapable of viewing a photo in the same way as a human and classifying it into categories. This does not mean image classification is an impossible task, it just must be approached in a different way. Specifically, through the use of neural nets.

The problem being addressed in this project is the ability of different neural nets to automatically label and classify images. In providing image classification to different types of customers, we must balance accuracy and training time/compute power necessary to train a model. For this project, we assume that the model with the highest accuracy should be prioritized over a balance of accuracy and compute time. As such, neural networks present the best solution to classifying images with increased accuracy. We must provide management with information on which machine learning models work the best, recommendations on what type of images to utilize, as well as recommendations for future data collection and model training.

The data provided for this classification project is 1000 images of cats and 1000 images of dogs. In the appendix, a sample of images of cat and dogs used in this analysis can be found. The images are 128x128 and in color. Prior to training the neural network, the images much be "one hot" encoded, or converted into a 0 or 1 value for each pixel value. These encoded values are then fed to the neural network for training. In selecting a deep learning algorithm for image classification, we attempt a multilayer perceptron (MLP) with a decreasing number of hidden layers as well as multiple convolution neural networks (CNN). These CNNs possess differing hyperparameters, such as learning rates and pooling layers, and are compared on accuracy. The accuracy and time results from these different models can be found in the appendix.

In programming this project, preprocessing was required prior to using the CNN versus the MLP. The CNN required the classification output to be in a 2-character format, as in (0,1) for a dog and (1,0) for a cat. The MLP processed based on a single character output. Once data was formatted for the corresponding model, data was split into a training and test set, with 20% of the data being randomly split into the test set. The different deep learning algorithms were trained on the training data and tested for accuracy on the test set. In building the CNNs, different hyperparameters were adjusted. Some of the parameters that were adjusted were: learning rate, the number of pooling layers, the drop out percent, as well as filter size and kernel size on each 2-dimensional convolution layer. All deep learning was done using TensorFlow. For the MLP, the model from assignment 6 with the highest accuracy on the MNIST dataset was selected to compare accuracy results.

With accuracy as the prime goal of this analysis, a convolution neural network is superior to the multilayer perceptron. The accuracy of the multilayer perceptron was approximately 62% whereas the best CNN is approximately 75%. Within the multiple CNNs that were tested, the CNN with 5 2-dimensional convolution layers, with filter sizes of 128, 256, 512, 256 and 128, 5 pooling layers, 2 fully connected layers, 1 drop out layer, with a dropout rate of 0.5 and a learning rate of 0.001 was the most accurate. This model achieved a training accuracy of 0.9156 and a test accuracy of 0.7475, with a processing time of ___. While this model was the most accurate, accuracy can likely be increased with more images to train the model on. Collecting additional images of cats and dogs and processing them to ensure noise in the image is minimized will likely ensure increased accuracy when classifying new images.

Appendix

Model	2D Layers	Pooling Layers	Fully Connected Layers	Drop Outs Layers	Learning Rate	Drop Out Percent	Test Accuracy	Training Accuracy	Time (sec)
CNN	5	5	2	1	0.001	0.8	0.969	0.74	599.7
CNN	5	5	2	1	0.001	0.8	0.88	0.7425	2771
CNN	1	2	3	2	0.01	0.8	0.50	0.5125	2240
CNN	5	5	2	1	0.001	0.5	0.9156	0.7475	2473
CNN	5	5	2	1	0.001	0.5	0.92	0.7400	1468
CNN	5	5	3	2	0.001	0.5, 0.25	0.5681	0.515	1462
CNN	3	3	2	1	0.001	0.25	0.503	0.4875	1445

Model	**Number of Layers**	**Nodes per Layer**	**Processing Time**	**Training Set Accuracy**	**Test Set Accuracy**
DNN	3	300, 200, 100	4941	0.6786	0.6175

Sample of Cat and Dog images