# CS 521- Final project- Summary Dec-19-2020 [bouyang@bu.edu](mailto:bouyang@bu.edu) Beining Ouyang

## What the program does?

This program classifying Cats vs Dogs images with a Convolutional Neural Network (CNN).

This project topic is a very famous task in Kaggle, I found it on this link:

<https://www.kaggle.com/c/dogs-vs-cats-redux-kernels-edition/overview>

All the images (training and testing images) I used were also downloaded from Kaggle. The images are either of dog(s) or cat(s).There are 25,000 training images (labled) of dogs and cats. There are also 12,500 testing images. For each image in the test set, we need to write a program to predict a probability that the image is a dog (1 = dog, 0 = cat).

The code I submitted as the final project will ask user to input a number: “We have 12500 unlabeled pictures from testing set, how many of them do you want to classify?” Then it will feed these number of images through a CNN for classification and return the results of how many dogs and cats.

For example: if the user want to classify 34 pictures in the testing set, the program will process and return “We detected 16 dogs and 18 cats.”

To make it whole thing work, first, we need a CNN. I create my model followed this tutorial:

<https://pythonprogramming.net/convolutional-neural-network-kats-vs-dogs-machine-learning-tutorial/>

The model is based on TensorFlow and TFLearn.

After training my model, I saved it in the submittion folder.

In the final project code, I loaded the CNN model, ask for user’s input of how many images they want to classify, then return the results.

## Why it is useful?

I won’t say it is useful because we wanted to use it to detect dogs and cats --- I view this intro-level project as training for me. The "usefulness" behind it is that I want to learn image recognition. I work in the microbiology field: using neural networks to process/classify images is the future. For example, image recognition can be used for bacterial colony counting, classification of microscopic images, and more.

To be honest, the time I spent on the project was way over my estimate. Most of the time I was just googling “what is this error???”. I thought it matched my level (beginner), however the CNN model part...so many tears, it just didn't work even though I followed the tutorials. Eventually, it worked with a loss: 0.34108 - acc: 0.8514. It's satisfying to me because it runs! So another useful thing for me is: Always estimate more time than you think for a programming project --- you will need it.