Machine Learning Engineer Nanodegree

Capstone Proposal

Aadam
June 14, 2017

Domain Background

The project is taken from the kaggle competition, dogs-vs-cats-redux-kernels-edition, in which we need to predict whether a given picture if of a dog or a cat. For humans, this task seems quite easy and intuitive but for computers, the task of correctly recognizing objects in a picture seems quite difficult and daunting. For many years, researchers in computer vision have struggled with this problem, making computer understand a picture and what's in it. But with the advent of new techniques and algorithms, now the computers can be trained to predict what's in a picture quite effectively.

Problem Statement

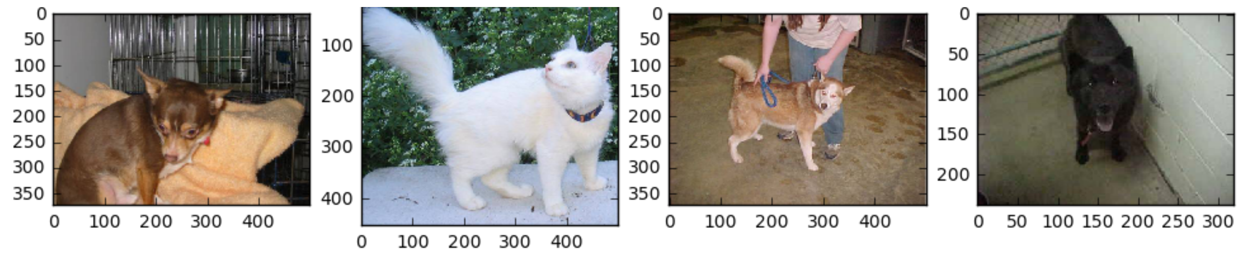
In this problem, our task is to identify whether a picture contains a dog or a cat. If it's a dog then we need to correctly identify him and vice versa. We need to analyze the pictures and then calculate the probabilities of every category, dog or cat, showing how confident we are in our predictions that it is a dog or a cat. So, we can summize that this is a classification problem. One possible solution for this problem is to train CNN models on the dataset and use those to classify whether a picture is of a dog or a cat.

Datasets and Inputs

The dataset, provided here, contains labeled pictures of dogs and cats. The train folder contains 25,000 images of dogs and cats. Each image in this folder has the label as part of the filename. The test folder contains 12,500 images, named according to a numeric id. For each image in the test set, you should predict a probability that the image is a dog (1 = dog, 0 = cat).

We will use these pictures to train our CNN model and based on which we will generate predictions. Then we'll use the test dataset to validate our prediction model and see how well it performs.

Here are some of the sample images from the dataset:



Solution Statement

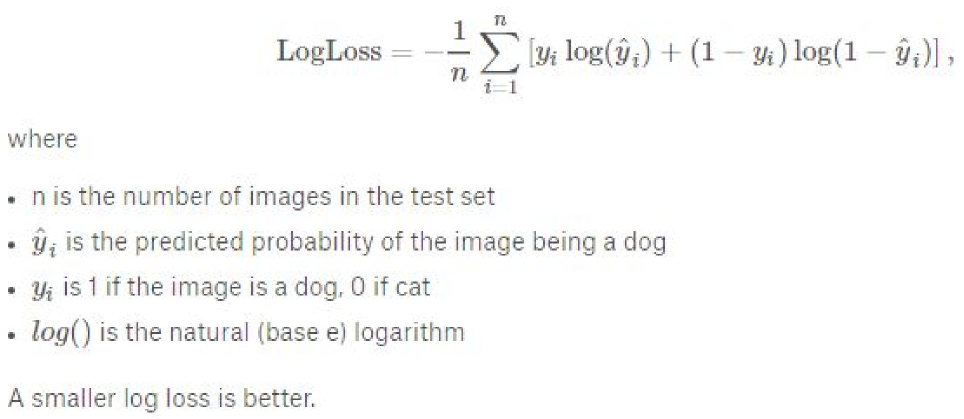
We can use the Convolutional Neural Network (CNN) which have been know to perform quite well on the Imagenet database. We will train our CNN model on the pictures that are provided of the dogs and cats and then we will validate it using the test dataset.

Benchmark Model

According to this kaggle competition, the state of the art, in this kind of task 3 years ago, was considered to have an accuracy score above 80%. Now, by looking at the leaderboard on the kaggle competition, this accuracy level has increased quite a bit, in some cases, reaching to 97%, meaning that 97% of the time, the prediction that the model produces is the correct one.

Evaluation Metrics

Instead of using the accuracy of predictions as our evaluation metric, the model will be evaluated based on its **LogLoss** value, which is defined to be:



$$\text{LogLoss} = -\frac{1}{n}\sum_{i=1}^{n}\left[y_i \log(\hat{y}_i) + (1 - y_i)\log(1 - \hat{y}_i)\right],$$

where

- n is the number of images in the test set
- $\hat{y}_i$ is the predicted probability of the image being a dog
- y_i is 1 if the image is a dog, 0 if cat
- $log()$ is the natural (base e) logarithm

A smaller log loss is better.

Project Design

We will use **Keras** library to train our CNN model on this dataset and it will use **Theano** as the backend.

We will first extract some sample and validation datasets from the train dataset. We will use validation dataset to validate how our model is doing and we'll use sample dataset so that we can test our model quickly as it will contain smaller amount of data as compared to the train dataset. We will first divide the pictures in the train dataset into separate categories, namely dogs and cats and then we'll load those as batches and then train our CNN model on those batches. After we have fine tuned our model and trained it on the train dataset, we will then validate our model using the test dataset and see how well it performs.

So, our action plan will be something like this:

1. Create Validation and Sample sets
2. Rearrange image files into their respective directories
3. Finetune and Train model
4. Generate predictions
5. Validate predictions