Machine Learning Engineer Nanodegree

Capstone Proposal

Varghese Simon
April 10, 2018

DOMAIN BACKGROUND

Dogs are always human's best friend and, they always show us what true love and loyalty is. When we reach home, they will be the ones who invite us with all their joy, waiting eagerly at the door.

There are about 340 breeds recognized by the Fédération Cynologique Internationale (FCI), the world governing body of dog breeds, sometimes known as the World Canine Organization. Each dog breed has its charactericts, like color and patterns of fur, height, so on and so forth. Even for a dog lover, it may seem impossible to distinguish all dog breeds. Some characteristics of one breed may be very much similar to other breed, and also there a large number of breeds now.

As being a dog lover, I have been part of developing an iOS application which provides a means to post the availability and requirements of puppies for sale/adoption in Bangalore, India. However, as the development carry forward, one thing became clear, the app was not the difficult task, but filtering dogs and segregating it into different breeds where one of the significant challenges we have faced. It felt like there are zillions of dog breeds out there (pun intended). Hence, we have to hire 2 employees just for segregating dogs based on their breed.

Recently I came across a competition in Kaggle, and I was surprised to see the results people have uploaded. The competition was to classify dogs based on their breed. There are submissions which have close to 100% accuracy. Even though I have moved out of the company I mentioned earlier, I see a need for to create a solution for the challenge I have faced early in my career. In this project, I aim to create a model which can predict dogs' breed from the pictures with significant accuracy.

After doing some research, I came across paper on AlexNet (2012) (Link) which explains about Image recognition using Convolutional Neural Networks. Authored by Alex Krizhevsky, Ilya Sutskever, and Geoffrey E. Hinton, this 2012 paper won the ImageNet Large Scale Visual Recognition Challenge with a 15.4% error rate, which was a milestone in case of CNN.

PROBLEM STATEMENT

As mentioned earlier, there exists a certain level of difficulty in identifing dogs' breed from an image, even for human eyes. Hence, this project aims to create a model to predict dog's breed based on the image of a dog with acceptable **Accuracy** (above 90%). To measure the performance we can use set of testing samples and calculate the percentage of the images labeled correctly. Also, we can use **Log Loss** for measuring losses.

$$-\log P(yt|yp) = -(yt \log(yp) + (1 - yt) \log(1 - yp))$$

The solution for the problem here is to create a classifier which take images of dogs and print out the dogs' breed on console. The aim is to limit the input to have only one dog per image so that to avoid detection of multiple breeds from an image.

DATASET AND INPUTS

The dataset for this project is acquired from a competition that was held by Kaggle. In the competition, they have provided a collection of images of dogs, and a csv file which has labels for each image. The images that have been shared by Kaggle in this competition is taken from "**Stanford Dogs Dataset**" which in turn is a subset of ImageNet competition held by Stanford University.

The images in this dataset has color, and are having different resolution. There are 10222 images with labels for each image representing breed of the dog in the image. There are images of 120 different dog breeds with almost 100 images for each breed, but some breeds are having more images as high as 120 and some are having low as 66 images.

We can use these images to train a **Convolutional Neural Network** and use that model to predict the breeds. To evaluate the models performance I will take 20% of dataset.

Link: https://www.kaggle.com/c/dog-breed-identification

SOLUTION STATEMENT

Convolutional Neural Networks (CNN) are one of the best algorithms for Computer Vision problems. Hence, I aim to model and train a CNN to solve the problem at hand. The program will be able to take images of dogs and classify the images based on their breeds. Using “Accuracy” and “Log Loss” we can measure the performance of the system.

BENCHMARK MODEL

For benchmark model, I will create a basic CNN, with one Convolutional Layer, One Flattening layer and One Fully Connected layer. Then the performance of this model will be measure with the same set of training images and test images. The performance will be kept as a benchmark for my model, so that the performance of the model should be higher than that of the benchmark model.

EVALUATION MATRIX

Logarithmic Loss (**Log Loss**) is a measure of performance of classifier which predicts a probability value between 0 and 1. When Log Loss is ‘0’, the model is said to a perfect model, and as Log Loss increases, the predicted probability diverges from the actual label.

PROJECT DESIGN

Project creation requires us to follow different steps. Firstly, I have to download the required dataset (training and test) from Kaggle. Once it is downloaded, extract both compressed folders into different locations, for training and test. Next step is to load these images and preprocess the images. This step includes resizing, normalizing and cropping. Along with this I could do Data Augmentation to create more images for training. Along with this, I will load the labels (Breed names) and associate it with corresponding images.

Third step is to create a Convolutional Neural Network (CNN), as they are the best way forwards in case of Computer Vision. It will consist of Convolution layers, Pooling Layers, Dropout Layers and Fully Connected Layer. The number of layer needed will be decided based on the accuracy and log loss of neural network.

Next step is to train the neural network using the images and associated labels. While training I have to define some hyperparameters. Later on, I have to adjust these parameters to increase the accuracy and reduce the Log loss of the system. To measure the accuracy at this stage I will be using a small subset of training set as a Validation set.

Once the CNN acquires an acceptable accuracy, I will do testing using test dataset as a final step to note down the performance of my system. After this, the system can be used to predict breed of a dog in an image. For this