Pornrapee Kajorndech

August 24th, 2020

Dog Breed Classifier

Domain background

The Dog breed classifier is a problem that can be solved with the ML model. The problem is identifying a breed of a dog with a dog image that is given as an input. This is a multi-class classification problem where we can use a supervised ML model to solve. After completing the ML model then I build a web application and Rest-API for users to upload an image and get results from the ML model. This project gives me the knowledge to build an end-to-end ML project, so I have chosen this project as my capstone project.

Problem statement

The end goal of the project is to build an end-to-end ML model that can be used by web applications that connect to API on AWS cloud platform to mimic real world use cases. For the ML part, the problem is to predict dog breed from a dog image as input.

Dataset and input

Dog images dataset: The dog images dataset has 8351 total image including train (6680 Images), validation (835 Images) and test (836 Images) and have 133 dog breed classes. The dog images dataset for this project is provided by Udacity.
An input has a type as an image, because we want to input an image to identify the breed of the dog.

Solution statement

To identify dog breeds from images we will use pretrained model like VGG16, RestNet50 and the other architecture (I will decide later). In training we will use augmentation technique to improve performance of model.

Benchmark model

The [VGG-16 bottleneck features](https://github.com/udacity/deep-learning-v2-pytorch/tree/master/project-dog-classification) from https://github.com/udacity/deep-learning-v2-pytorch/tree/master/project-dog-classification have 73% accuracy

Evaluation metrics

- Accuracy: The ratio of correct predictions to the total size of data
- Recall: The ratio of true positives to the true positive and false negative
- Precision The ratio of true positives to the true positive and false positive

Project design

- Import the necessary dataset and libraries.
- Pre-process the data and create train, validation, and test dataset. And use the augmentation technique on training data
- Create a neural network to classify dog breeds by using transfer learning with ResNet50. And then evaluation of the model accuracy.
- Write a function for serving to prepare to deploy the model
- Write an HTML file to serve in web application
- Deploy on AWS by using Sagemaker, Lambda, API Gateway