Project: Traffic_Sign_Classifier
Name: Dongdong Wu

1.Data Set Summary & Exploration

training data / test data is provided. the result of data summary as below:

"Number of training examples = 34799 # size of training data
Number of testing examples = 12630 # size of test data
Image data shape = (32, 32, 3) # shape of data
Number of classes = 43 # labels number in training data"

the code is on "Provide a Basic Summary of the Data Set Using Python, Numpy and/or Pandas" part

2.Provide a Basic Summary of the Data Set Using Python, Numpy and/or Pandas

a. since the data is already 32*32 format, the preprocess only do shuffle for training and test data set

b. i split shuffled training images set with batch_size = 128 which is the same for validation and test set.the code is on 9th part cell

I pass images into model one batch by one batch contiuously.

the final test image is from GTSRB data set. the final test image is not always 32*32 shapen and ppm format.

I need load them by skimage lib, and resize them to 32*32 by numpy. the code is on 11th part cell.

c. based on what i learned, i use lenet architechture to build my CNN model. since this project is just for classification without location identification, and test image size is also quite small, so lenet should be good enough

[Layer	Description]
Input	32x32x3 RGB image
Convolution	5x5,1x1 stride, valid padding, outputs 28*28*6
RELU	YES
Max pooling	2x2 stride, outputs 14x14x6
Convolution_2	5x5,1x1 stride, valid padding, outputs 10*10*16
RELU	YES
Max pooling	2x2 stride, outputs 5x5x16
Fully connected	mean = 0, sigma = 0.1, input 400, output 120
Fully connected_2	mean = 0, sigma = 0.1, input 120, output 84
Fully connected_3	mean = 0, sigma = 0.1, input 84, output 10

Softmax output = 10

Optimizer = AdamOptimizer, Learning rate = 4*1e-4

Batch_Size = 128

Epoch = 100

d. training set accuracy = unrecorded
validation set accuracy = 0.899
test set accuracy = 0.887

3. final test images is from GTSRB data set in which images format is ppm, and the size is not 32 * 32.
before I start test, i need load ppm file by skimage and resize them to 32*32 by numpy
I pass five images to model

 label = 4  lable = 4  lable =31

 label = 13  lab3l = 8

and get results as below:

TopKV2(values=array([0.29774481, 0.07255951, 0.05254821, 0.04730011, 0.04410096],

[0.26439881, 0.11065474, 0.04948991, 0.04643316, 0.04405318],

[0.4554905 , 0.04427201, 0.0392456 , 0.03168693, 0.02614531],

[0.16883691, 0.10111474, 0.09836857, 0.07256201, 0.06892448],

[0.21717823, 0.0776408 , 0.07398154, 0.05089254, 0.04290496]),

indices=array([31, 19, 21, 23, 29],

[4, 1, 2, 0, 7], [13, 38, 41, 36, 1], [4, 7, 1, 5, 2],[8, 15, 9, 7, 3], dtype=int32))

Total accuracy = 1.0