Homework10

Steven Black

4/24/2019

IS 677: Introduction to Data Science Spring 2019

Homework Assignment 10 (Due: April 28, 2019, midnight EST)

Use the CIFAR10 dataset that comes with Keras. It has 50,000 training and 10,000 testing images of 10 classes:

- airplane
- automobile
- bird
- cat
- deer
- dog
- frog
- horse
- ship
- truck

The image size is 32*32*3. The 3 indicates that these are color images.

```
library(keras)
k_clear_session()
```

1. Clear the session and load the CIFAR10 data into a variable called cifar. (5 points)

```
cifar <- dataset_cifar10()
```

2. Create a small training dataset using the first 1000 training images (and the corresponding labels) from CIFAR10. Similarly, create a small test dataset (and the corresponding labels) using the first test 500 images from CIFAR10. (5 points).

```
train_images <- cifar$train$x[1:1000, 1:32, 1:32, 1:3]
train_labels <- cifar$train$y[1:1000]

test_images <- cifar$test$x[1:500, 1:32, 1:32, 1:3]
test_labels <- cifar$test$y[1:500]
```

3. Create one-hot encoding for the labels for both train and test labels. (5 points)

```
train_labels_categorical <- to_categorical(train_labels)
test_labels_categorical <- to_categorical(test_labels)
```

4. Instantiate a VGG16 convolutional base without the top layer. (5 points)
5. Extract features from the CIFAR10 images so as to fit the conv_base. (40 points)
6. Flatten the features in order to feed them to a densely connected classifier. (5 points)
7. Build a model with one dense layer with 256 units and "relu" activation, one dropout alyer with 50% dropout rate, and a dense output layer with appropriate parameters. (15 points)
8. Compile the model with categorical_crossentropy as the loss function and optimizer_rmsprop with 0.01% learning rate (lr=0.0001). (5 points)

9. Fit the model using 30 epochs. Plot the loss and accuracies. (5 points)
10. Note that the model is likely to have low accuracy. Explain why. (10 points)