INFO 7374 Spring 2019
Assignment 1
Group 6

Part 1

Small epochs will cause under fitting. When the epochs get bigger, the accuracy will become better and the loss will decrease. Large epochs need more time to train the model. Usually, small epochs can't attach the limit of a model.

Small batch size will cost less time per step, but if the batch is too small, we need more epochs to find the convergence. If the batch size is too big, we may run out of memory or need to do more computing to get the same level of accuracy. So we need to choose a moderate batch size.

Appropriate learning rate will find the limit quickly and accurately. Small learning rate takes more computing cost, while bigger learning rate can't reach the limit.

Activation functions are really important for a neural network to learn and make sense of something really complicated and non-linear complex functional mappings between the inputs and response variable. Appropriate activation function can help us to improve accuracy.

Dropout rate help us solve over-fitting. The model with lower dropout rate may over fitting, and the model with higher dropout rate may under fitting.

You are expected to provide a recommendation for the best model you would recommend for classification. Which model (with parameter values) would you choose and why?

Here is my recommend model:

```
batch_size = 128
num_classes = 10
epochs = 100
learning_rate = 0.1
dropout_rate = 0.5
lr_decay = 0.0002
model = Sequential()
model.add(Dense(2048, activation='relu', input_shape=(3072,)))
```

```
model.add(Dropout(0.3))
model.add(Dense(1024, activation='relu'))
model.add(Dense(512, activation='relu'))
model.add(Dropout(0.3))
model.add(Dense(256, activation='relu'))
model.add(Dropout(0.3))
model.add(Dense(128, activation='relu'))
model.add(Dense(num_classes, activation='softmax'))
```

I tried several parameters, and this one got the best accuracy 0.4976. The number of epochs is 100, because if it too small, the model will under-fitting, and if it too large, the model may over-fitting. The learning rate is 0.1, because the training speed is slow with small learning rate, and the accuracy can't reach the top with large learning rate. The dropout rate is 0.2, I had tried small dropout rate before, but that will cause overfitting.

Comment on how good your model is ? Does it overfit/underfit data ? What could you do to improve the model?

Output of my model:
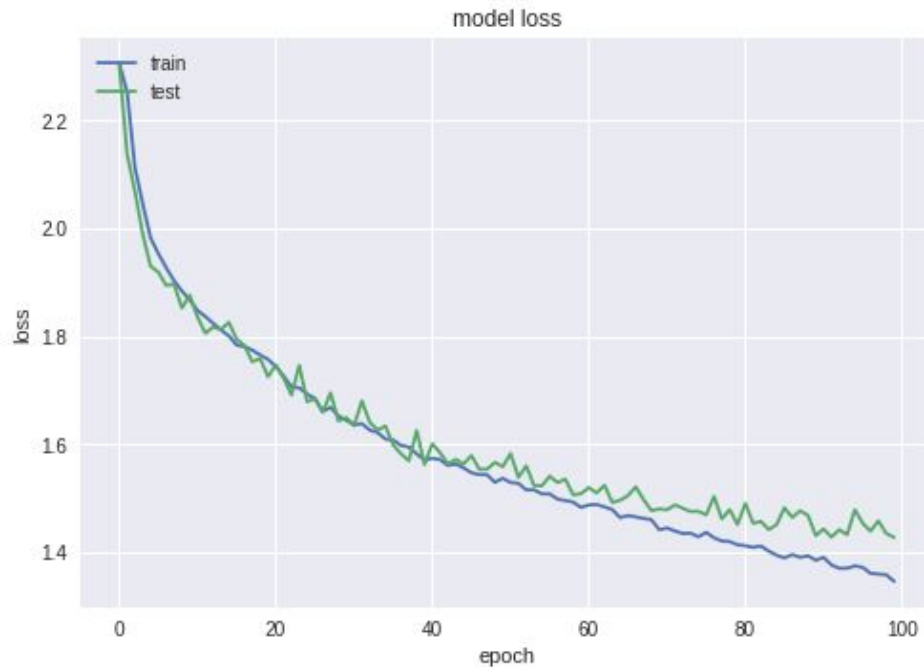
From the graphs we can see, the model is a little overfitting around after 70 epochs. The accuracy is about 0.49. I tried to add more layers to improve the accuracy, and used larger drop rate to prevent overfitting but the results didn't get better. In order to get more accurate result, we may need more data or choosing other model with higher efficiency.