

(3)

After we make two improvements :

```
validation set accuracy: 0.7079565340107254  
test set accuracy: 0.6944482077335591
```

The final validation accuracy and test set accuracy of our new model are almost same as baseline models.

The old model with 0.691645 validation accuracy and 0.688117 test accuracy.

Although the accuracy rate has not improved much, we still know that weight can help us get better Optimization.

Then for two regularization term in the cost function, we find that help us get convergent faster. (Achieve same effect in less time)

```
The old version runtime:  
0:00:35.425259
```

```
The new version(part b) runtime:  
0:00:06.529537
```

The old version runtime is almost six times the new version runtime.

Regularization makes each update more appropriate.

#### 4. Limitation 1.

When we get weight improvement, we assume the difficulty of each question are follow normal distribution. And the difficulty of each question are independent. However, in real life situation, the difficulty of each question may not follow normal distribution or independent.

One possible extension:

We can use more data and test more distribution. Then we may find a better distribution for our model.

#### Limitation 2

We apply a new weight. And we use data set to get average correctness. That may be overfit this data set.