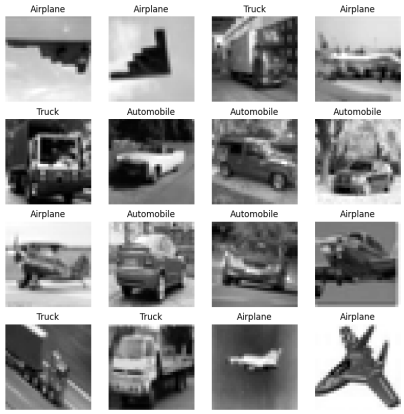
Keshav Shankar
ECE 1395
Assignment 7

0a)



Note:
Directions are misleading.
- 1 part says "16 images" while another part says "25 images", so I just picked 16

1b)

```
Training accuracy: 75.14%
```

2b)

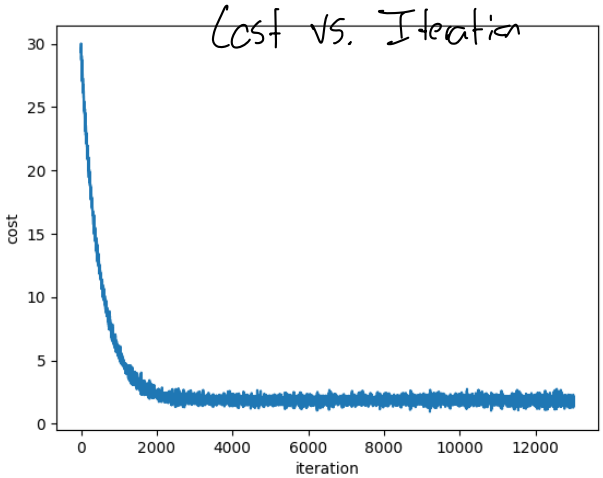
```
Cost when lambda = 0.1: 1.13
Cost when lambda = 1: 1.15
Cost when lambda = 2: 1.17
```

3)

```
Training accuracy: 75.14%
Sigmoid gradient: [4.53958077e-05 2.50000000e-01 4.53958077e-05]
```

4d) $\alpha = 0.01$

4e)



5)

	MaxEpochs = 50				MaxEpochs = 300			
	Training data accuracy	Testing data accuracy	Training data cost	Testing data cost	Training data accuracy	Testing data accuracy	Training data cost	Testing data cost
λ = 0.1	33.22%	34.05%	6.05	5.98	33.22%	34.05%	6.17	6.09
λ = 1	33.22%	34.05%	4.15	4.09	33.22%	34.05%	4.16	4.11
λ = 2	33.22%	34.05%	3.91	3.85	33.22%	34.05%	3.92	3.86

It seems that the number of epochs does not affect the network. This does not seem right, as with more epochs, the model should learn more from the data, thus reduce error in training, at the very least. However, too many epochs can also overfit the data.

A greater global minimum may have not been found since the accuracies are all the same.

6) I think the model in general is poor. There could be better algorithms to fit the set of data. Additionally, the data quality may be poor, failing to find good features. It also seems like there are too many features, so some dimensionality technique may be needed.