Alex Ivensky
ECE 1395
Homework 7

0a.

1b.

```
1b. Accuracy:  0.7510666666666667
```

2b.

```
Cost for Lambda = 0.1: 1.1311995443418565
Cost for Lambda = 1: 1.1475583902938862
Cost for Lambda = 2: 1.1657348857961412
```

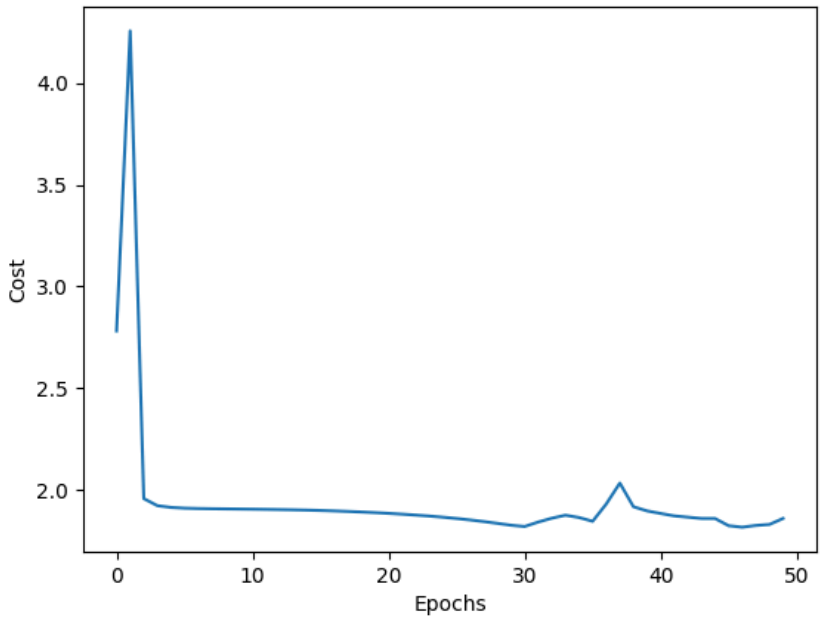
3.

```
Sigmoid Gradient Test:
 [[4.53958077e-05]
 [2.50000000e-01]
 [4.53958077e-05]]
```

4d.

Alpha = 0.0001. This didn't yield perfect results but it was the only reasonable value I could find that consistently converged without oscillating too much.

4e.



5.

	MaxEpochs = 50		**MaxEpochs = 300**	
	Training Acc.	**Testing Acc.**	**Training Acc.**	**Testing Acc.**
Lambda = 0.1	0.433	0.437	0.597	0.605
Lambda = 1	0.521	0.515	0.586	0.564
Lambda = 2	0.538	0.535	0.567	0.566

	MaxEpochs = 50		**MaxEpochs = 300**	
	Training Cost	**Testing Cost**	**Training Cost**	**Testing Cost**
Lambda = 0.1	1.849	1.845	1.599	1.605
Lambda = 1	1.763	1.764	1.593	1.597
Lambda = 2	1.758	1.757	1.620	1.633

6.
The most likely cause of the neural network's inaccuracy for this dataset is overfitting as a result of overdimensionality. Our dataset has 1024 features, but only 15000 examples. With roughly a 15:1 sample-to-feature ratio, there is a large possibility of noise being introduced and misinterpreted by the model. A good solution would be to preprocess the training set using PCA to perform dimensionality reduction. This would reduce the number of input units to the neural network as well, making NN-related computations faster.