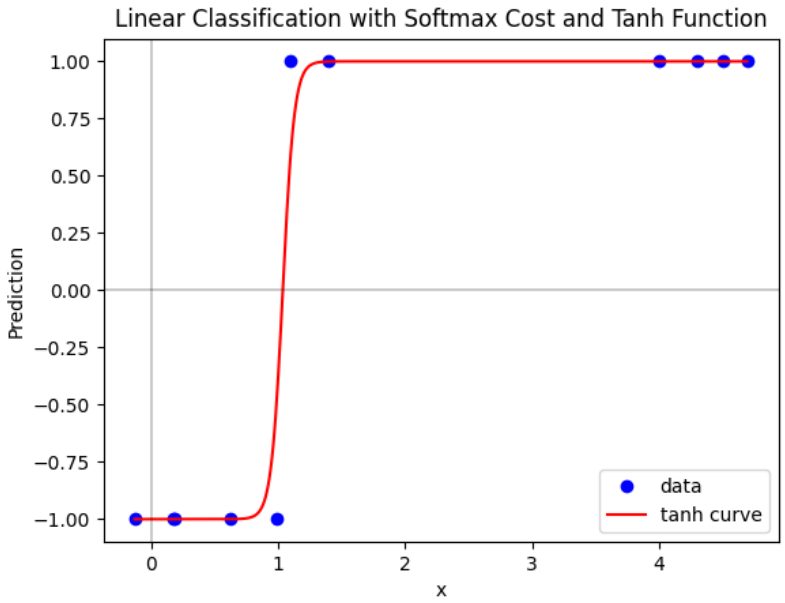
COMPSCI 589: MACHINE LEARNING

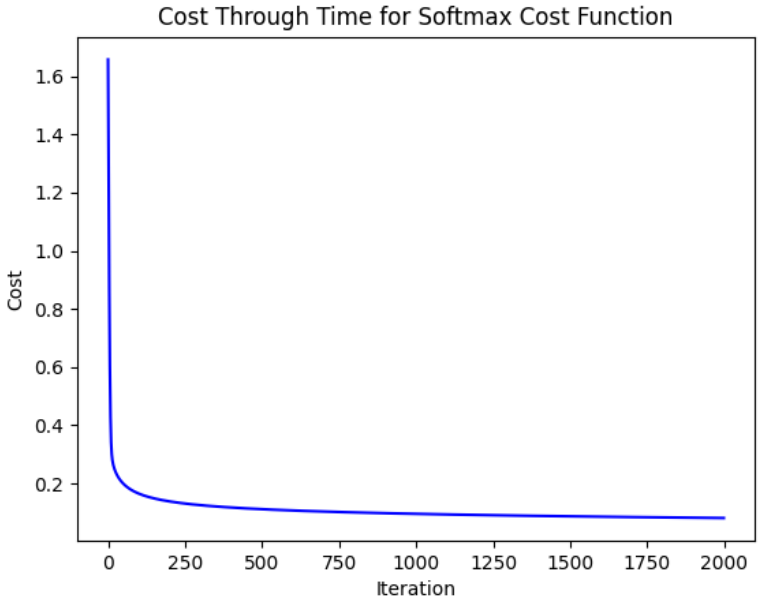
PROJECT 4 REPORT

Student: Trung Dang – 33858723

I. Task 1



- Accuracy: 100 %
- Misclassification of model = 0
- Figure shows cost history over 2000 iterations.



II. Task 2

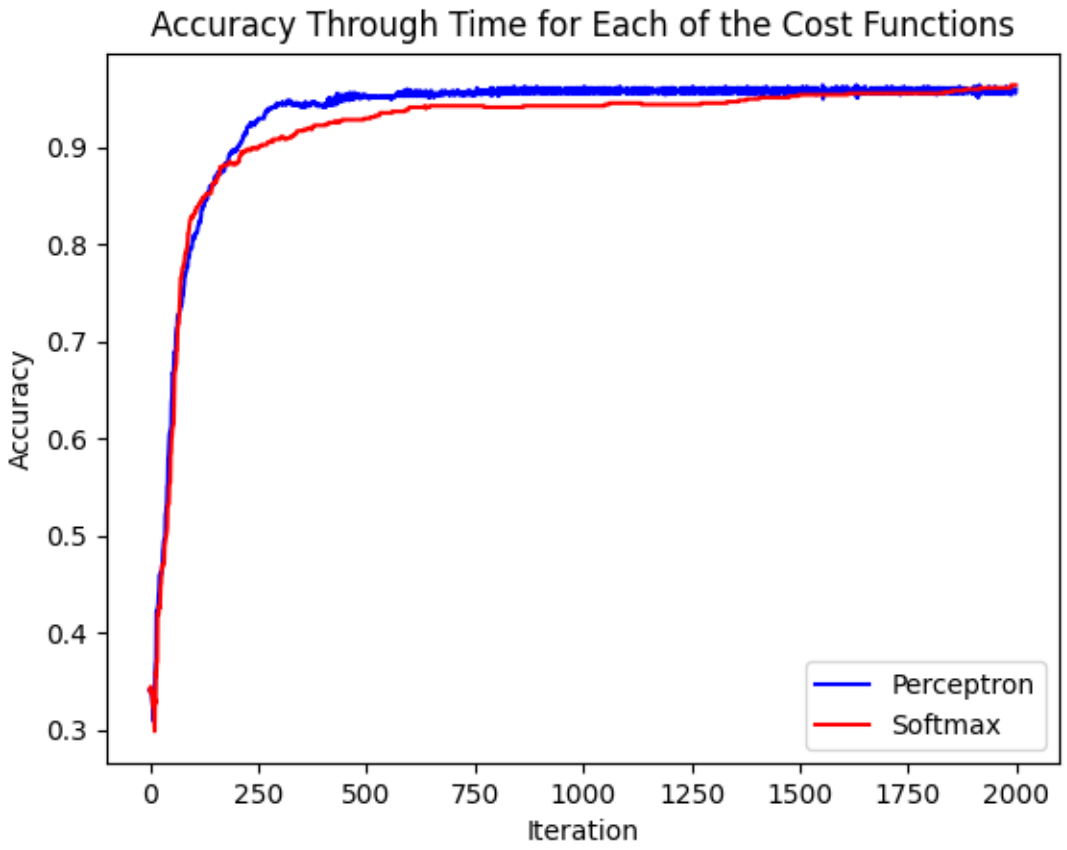
- Initial values are described by:

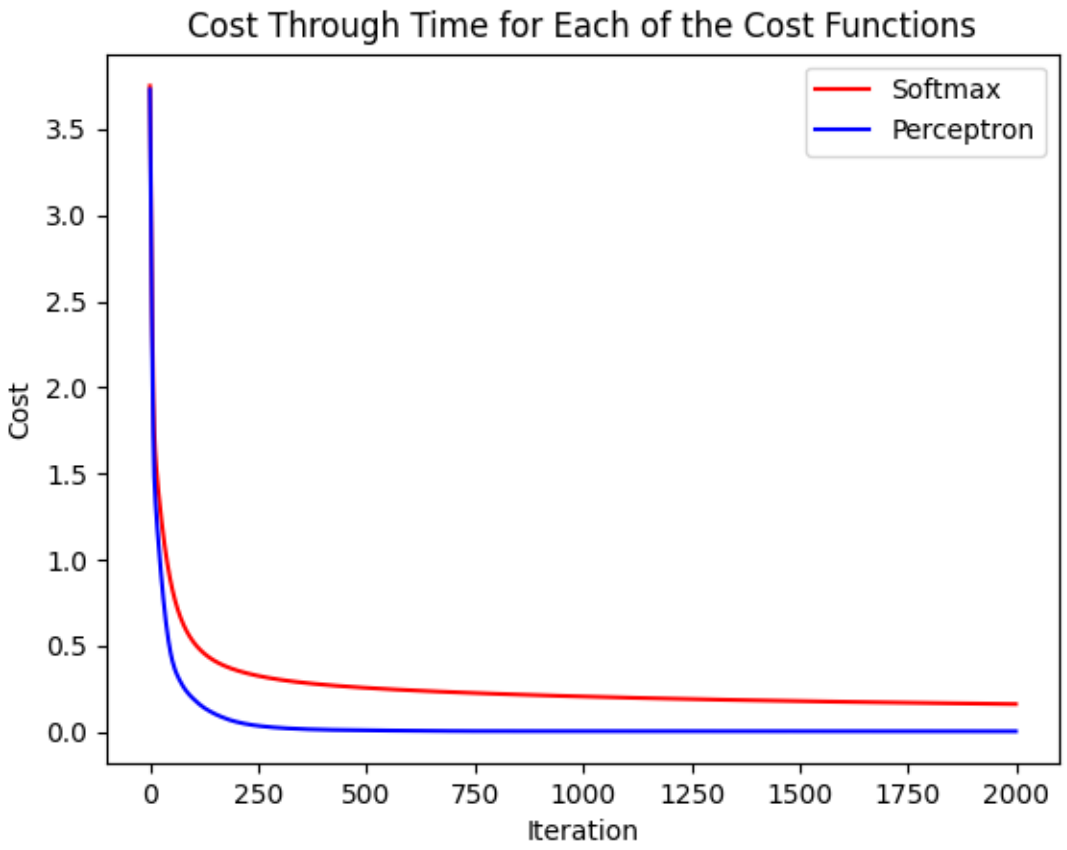
```
key = random.PRNGKey(40)
default_iteration = 2000
training_loop(random.normal(key, (9,)), 0.027, default_iteration, softmax, False)
```

- Here, the weights **(w)** are generated as a random array of shape (9,), (8 given parameters + 1 dimension for non-feature-touching weight). To ensure replication between different runs, the seed for the random number generator is set to 40. 40 is a personal preference. In fact, it can be replaced by any constant c.
- The learning rate is chosen to be 0.027. I found that the Perceptron model tends to diverge with large step length. In fact, a step length as small as 0.07 have already caused divergence in the Perceptron model. In **appendix A,** I have included the graph of the model when run with alpha = 0.07, as well as a table of misclassification of each model when run with different learning rate to demonstrate why 0.027 is approximately optimal
- number of epochs **(max_it)** = 2000. I have run the same program with 50, 300, 2000, and 5000 epochs. 2000 seems to be the most suitable max_it, as it allows the model to be refined while do not run into the same aforementioned divergence problem with Perceptron.
- I have explained why this happens in Campuswire post #193 at: https://campuswire.com/c/G460FD1BA/feed/193
- `diminishing_steplength` in the given example is False

- The accuracy and misclassifications from the two cost functions are reported as below:

```
Accuracy of Softmax is:  0.92846924
Accuracy of Perceptron is:  0.9198856
Misclassification of Softmax is:  25
Misclassification of Perceptron is:  28
```


Accuracy Through Time for Each of the Cost Functions
Accuracy
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0
250
500
750
1000
1250
1500
1750
2000
Iteration
Perceptron
Softmax


Cost Through Time for Each of the Cost Functions
Cost
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
0
250
500
750
1000
1250
1500
1750
2000
Iteration
Softmax
Perceptron

APPENDIX A: An experiment with learning rates

Alpha	Misclassification	
	Softmax	Perceptron
0.1	40	33
0.2	33	28
0.25	30	27
0.27	25	28
0.3	25	32
0.4	26	34

Table 1. Misclassification for several alpha

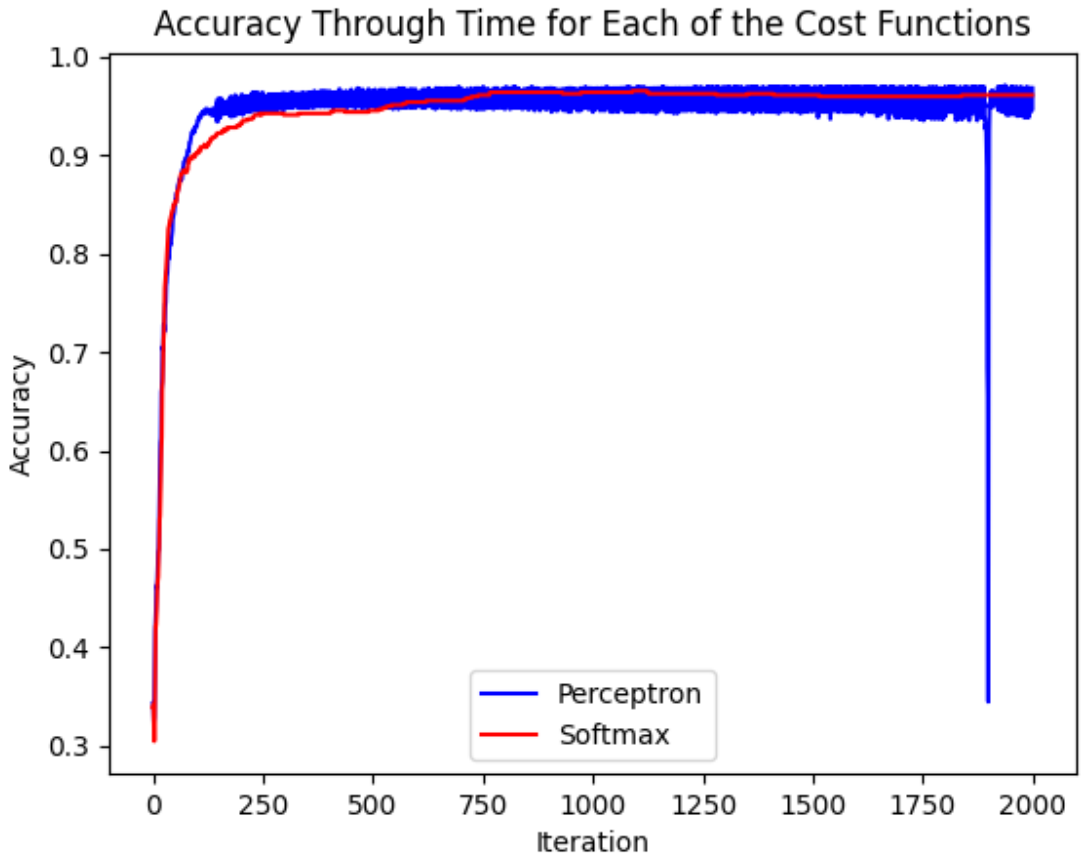


Fig 1. Accuracy through time for alpha = 0.07