Intro to ML

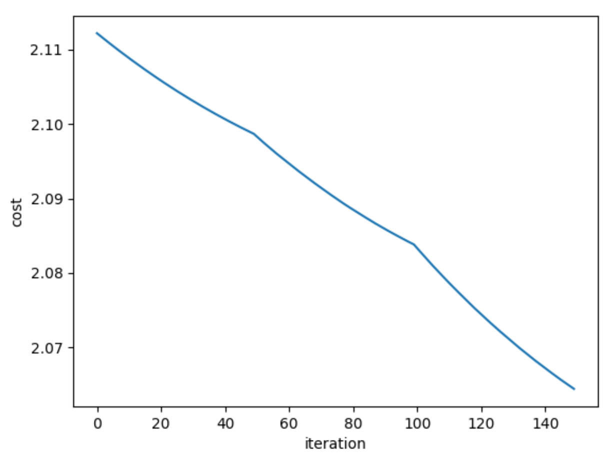
PS4 Report

Name: Timothy Horrell

ID: 4475990

Questions:

0) Dimensions verified.
1) .
    a. .
    b. 98%
2) .
    a. .
    b. Cost lambda(0): 0.203
       Cost lambda(1): 1.075
       Cost lambda(2): 1.958
3) Sigmoid gradient: [.000045, 0, .000045]
4) .
    a. .
    b. .
    c. .
    d. Alpha = 0.01
    e. Image below:



ps7-4-e-1.png

5) .

	Max_epochs = 50		Max_epochs = 100	
	Training data accuracy	Testing data accuracy	Training data accuracy	Testing data accuracy
Lambda = 0	43%	34%	69%	52%
Lambda = 0.01	69%	52%	69%	52%
Lambda = 0.1	36%	17%	69%	52%
Lambda = 1	36%	17%	36%	17%

	Max_epochs = 50		Max_epochs = 100	
	Training data cost	Testing data cost	Training data cost	Testing data cost
Lambda = 0	1.81	1.88	1.42	1.52
Lambda = 0.01	1.77	1.83	1.60	1.69
Lambda = 0.1	1.88	1.93	1.82	1.88
Lambda = 1	1.92	1.95	1.91	1.97

It appears that there are evident local minima that the gradient descent function falls into. Given the common accuracies of 69% and 36%, it seems that the gradient descent algorithm commonly falls into these local minimums. There may be a greater global minimum cost that was not found through these iterations. The number of epochs has no effect on getting stuck in local minimums.

Additionally, a lambda value of 0.01 appears to be best, as it provided the most accurate solution in each case and has relatively low cost. This means there is a small regularization coefficient.