Homework 8

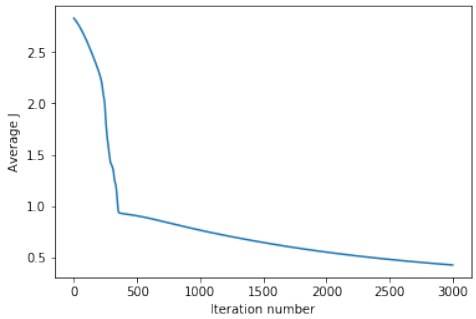
Name: Yuan Li
ID: N19728558 netID yl6606
Email: foxerlee1@gmail.com

Part II: Programming and Questions

11. (a)

Answer:

After adding regularization, the Average J curve:

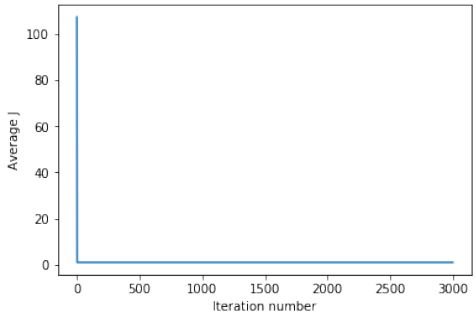


The Assessing accuracy: 91.098%

11. (b)

Answer:

When using ReLU, the Average J curve:

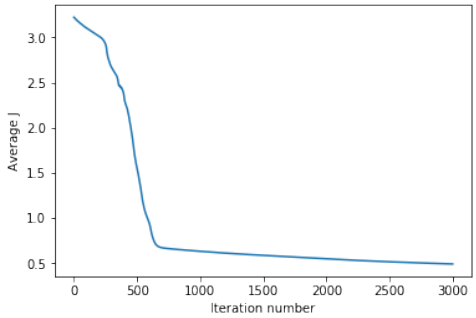


The Assessing accuracy: 10.152%

11. (c)

Answer:

When using ReLU, the Average J curve:



The Assessing accuracy: 89.568%

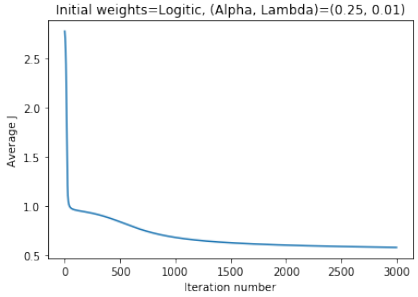
11. (d)

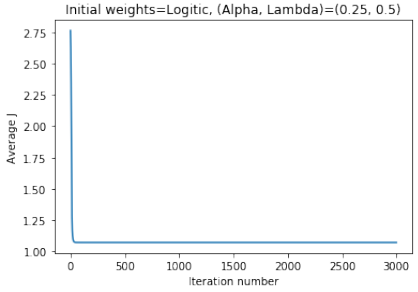
Answer:

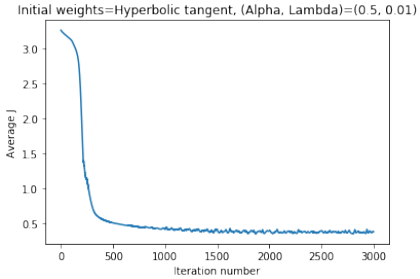
Try the different weight initializations given in the lecture notes:

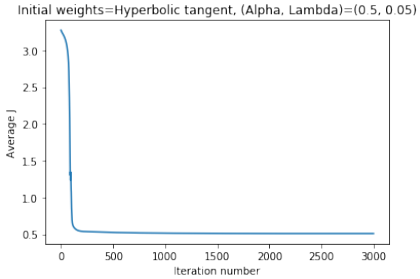
Initial weights	(Alpha, Lambda)	Activation	Result(ACC)
Logistic	(0.25, 0.01)	Sigmoid	90.681%
Logistic	(0.25, 0.5)	Sigmoid	8.623%
Hyperbolic tangent	(0.5, 0.01)	Tanh	97.913%
Hyperbolic tangent	(0.5, 0.05)	Tanh	93.602%
ReLU	(0.8, 0.01)	Tanh	95.549%

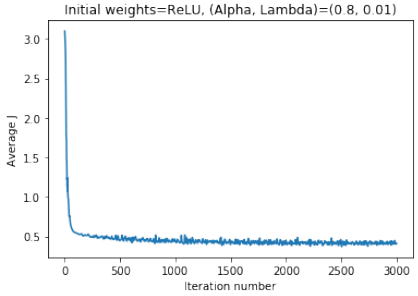
And the corresponding Average J curves are as follow:


Initial weights=Logitic, (Alpha, Lambda)=(0.25, 0.01)
Average J
Iteration number


Initial weights=Logitic, (Alpha, Lambda)=(0.25, 0.5)
Average J
Iteration number


Initial weights=Hyperbolic tangent, (Alpha, Lambda)=(0.5, 0.01)
Average J
Iteration number


Initial weights=Hyperbolic tangent, (Alpha, Lambda)=(0.5, 0.05)
Average J
Iteration number


Initial weights=ReLU, (Alpha, Lambda)=(0.8, 0.01)
Average J
Iteration number

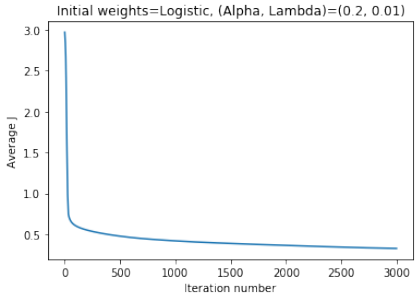
11. (e)

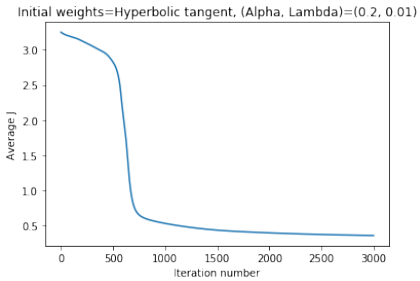
Answer:

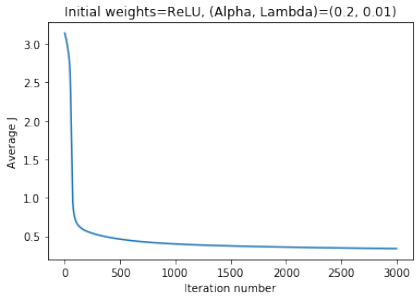
My own different hyper-parameters are:

Initial weights	(Alpha, Lambda)	Activation	Result(ACC)
Logistic	(0.2, 0.01)	Tanh	96.522%
Hyperbolic tangent	(0.2, 0.01)	Tanh	95.688%
ReLU	(0.2, 0.01)	Tanh	95.827%

And the corresponding Average J curves are as follow:







Others:

For my code, if you want to use different activations, you can uncommand the code in **block 8**. And the different weights initializations are **block 9**.