

Choosing Hyper-parameter for Neural Networks:

Hyper-parameters are variables that are set prior to optimizing the model's parameters. For the project, since Neural Networks can have many hyper-parameters we had to focus on choosing the correct ones which determined how the network was trained.

How we chose the number of hidden units was through a process of elimination. Regularization was used in the Neural Network to avoid overfitting problem. We chose by picking values that we felt would be easier to show correlation in the graphs.

lambda	hidden layer	Training set accuracy	validation set acc.	test set acc.
10	4	0.468300	0.444800	0.465500
10	8	0.899240	0.889800	0.895900
10	12	0.923200	0.917700	0.921400
10	16	0.935560	0.924400	0.935100
10	20	0.936860	0.929900	0.934400

lambda	hidden layer	Training set accuracy	validation set acc.	test set acc.
20	4	0.610540	0.601200	0.609600
20	8	0.870700	0.864300	0.871100
20	12	0.921280	0.908300	0.918900
20	16	0.923900	0.918400	0.922400
20	20	0.923620	0.918400	0.920200

lambda	hidden layer	Training set accuracy	validation set acc.	test set acc.
30	4	0.786240	0.779990	0.784400
30	8	0.858000	0.854400	0.861500
30	12	0.919720	0.915200	0.914800
30	16	0.933220	0.920000	0.929700
30	20	0.933300	0.926200	0.931000

lambda	hidden layer	Training set accuracy	validation set acc.	test set acc.
40	4	0.732880	0.728100	0.725800
40	8	0.879820	0.874400	0.883800
40	12	0.918020	0.909700	0.915900
40	16	0.925280	0.915000	0.922500
40	20	0.932520	0.927700	0.932100

lambda	hidden layer	Training set accuracy	validation set acc.	test set acc.
50	4	0.708080	0.693900	0.709000
50	8	0.894020	0.891300	0.893800
50	12	0.919140	0.916300	0.916300

50	16	0.925480	0.922400	0.925200
50	20	0.931440	0.926200	0.931500
validation set				
lambda	hidden layer	Training set accuracy	acc.	test set acc.
60	4	0.641440	0.631700	0.639500
60	8	0.892160	0.888700	0.892300
60	12	0.912140	0.905200	0.909300
60	16	0.921060	0.914400	0.919900
60	20	0.927720	0.919300	0.928000
0	50	0.955860	0.944900	0.951800





