

Learning Rate	No. of hidden la	Size of each h ⁱ »(regulariser	RMSE(train)	RMSE(dev)
0.001	1	64	0 10.84017895	10.64049728
0.001	1	64	5 13.04568337	12.9072019
0.001	1	128	0 10.82649006	10.63692993
0.001	1	128	5 13.04541733	12.90701897
0.001	2	64	0 10.55274366	10.36053034
0.001	2	64	5 10.71739595	10.52654492
0.001	2	128	0 10.71319212	10.52569767
0.001	2	128	5 10.71992964	10.52928352
0.01	1	64	0 10.56222642	10.36860245
0.01	1	64	5 12.78968501	12.64277039
0.01	1	128	0 10.54229013	10.35313886
0.01	1	128	5 12.78968501	12.64277039
0.01	2	64	0 10.62976014	10.45068198
0.01	2	64	5 48.79561515	48.80976525
0.01	2	128	0 10.59454065	10.42500738
0.01	2	128	5 32.69644532	32.57568394