

part_1b

Learning Rate	No. of hidden layers	Size of each hidden layer	λ (regulariser)	RMSE(train)	RMSE(dev)
0.001	1	64	0	10.91828	10.71737
0.001	1	64	5	10.91828	10.71737
0.001	1	128	0	10.91828	10.71737
0.001	1	128	5	10.91828	10.71737
0.001	2	64	0	10.91828	10.71737
0.001	2	64	5	10.91828	10.71737
0.001	2	128	0	10.91828	10.71737
0.001	2	128	5	10.91828	10.71737
0.01	1	64	0	10.91894	10.71856
0.01	1	64	5	10.91894	10.71856
0.01	1	128	0	10.91894	10.71856
0.01	1	128	5	10.91894	10.71856
0.01	2	64	0	10.91894	10.71856
0.01	2	64	5	10.91894	10.71856
0.01	2	128	0	10.91894	10.71856
0.01	2	128	5	10.91894	10.71856