

part_1b

Learning Rate	No. of hidden layers	Size of each hidden layer	regulariser	RMSE(train)	RMSE(dev)
0.001	1	64	0	52.05158	51.23241
0.001	1	64	5	25.13367	24.89866
0.001	1	128	0	12.63330	12.43023
0.001	1	128	5	11.65381	11.49076
0.001	2	64	0	10.02679	9.86091
0.001	2	64	5	10.66215	10.47763
0.001	2	128	0	9.94422	9.80324
0.001	2	128	5	10.63068	10.44362
0.01	1	64	0	9.88792	9.74939
0.01	1	64	5	11.54509	11.38096
0.01	1	128	0	9.91508	9.76971
0.01	1	128	5	11.58479	11.42427
0.01	2	64	0	9.77584	9.61950
0.01	2	64	5	10.62352	10.43472
0.01	2	128	0	9.71150	9.56050
0.01	2	128	5	11.21178	11.02068