

part_1b

Learning Rate	No. of hidden layers	Size of each hidden layer	λ (regulariser)	RMSE(train)	RMSE(dev)
0.001	1	64	0	128.46	11.16574842
0.001	1	64	5	123.52	10.717357
0.001	1	128	0	127.93	11.107376
0.001	1	128	5	123.52	10.717357
0.001	2	64	0	nan	nan
0.001	2	64	5	nan	nan
0.001	2	128	0	nan	nan
0.001	2	128	5	nan	nan
0.01	1	64	0	128.254	11.14391
0.01	1	64	5	123.52	10.718
0.01	1	128	0	123.52	10.718
0.01	1	128	5	123.526	10.718
0.01	2	64	0	nan	nan
0.01	2	64	5	123.526	10.718
0.01	2	128	0	nan	nan
0.01	2	128	5	nan	nan