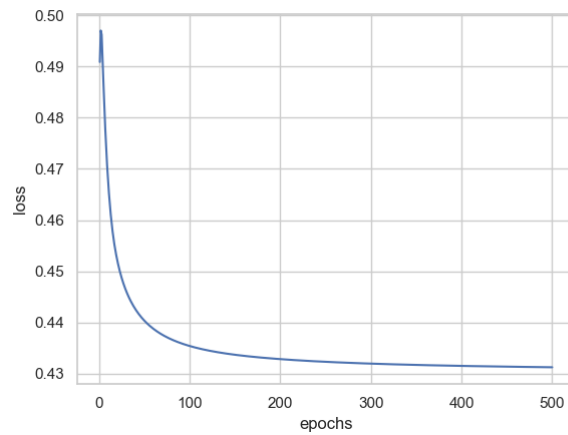
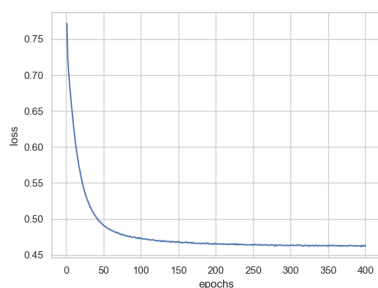
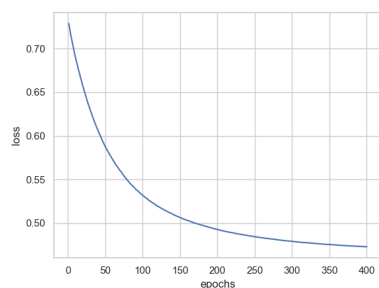
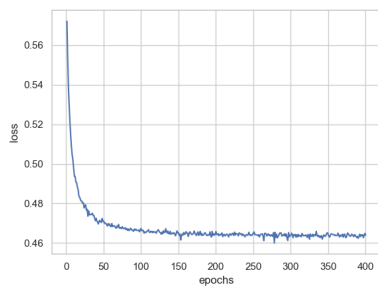
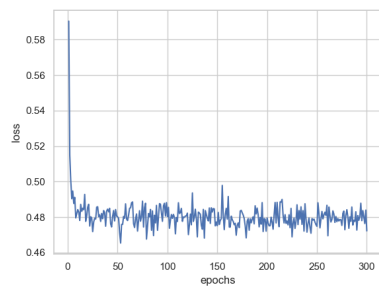


1.Loss curve of training process



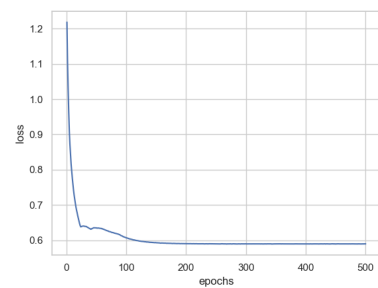
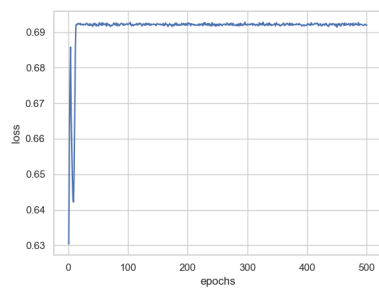
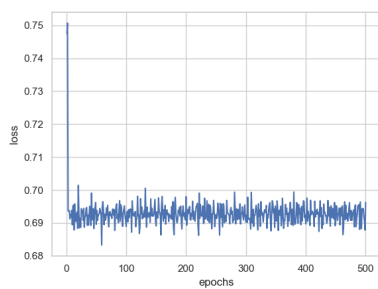
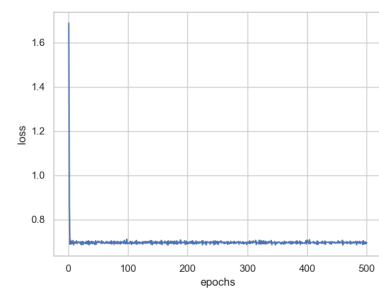
2.The comparison table of different parameters

learning rate	batch_size	γ	penalty	loss	prediction accuracy
0.1	16	None	None	0.472-0.484 横跳	0.8225806451612904
0.01	16	None	None	0.464	0.8145161290322581
0.001	16	None	None	0.473	0.8387096774193549
0.005	16	None	None	0.462	0.8225806451612904
0.005	5	None	None	0.463	0.8145161290322581
0.005	32	None	None	0.463	0.8145161290322581
0.005	16	1	l1	0.695	0.3225806451612903
0.005	16	0.5	l1	0.696	0.6048387096774194
0.005	16	0.1	l1	0.692	0.6774193548387096
0.005	16	0.01	l1	0.590	0.782258064516129
0.005	16	<i>1</i>	<i>l2</i>	0.679	0.6854838709677419
0.005	16	<i>0.5</i>	<i>l2</i>	0.667	0.6854838709677419
0.005	16	<i>0.1</i>	<i>l2</i>	0.626	0.6854838709677419
0.005	16	<i>0.01</i>	<i>l2</i>	0.569	0.6854838709677419
0.005	16	<i>0.001</i>	<i>l2</i>	0.501	0.8225806451612904



综合考虑选择 学习率 0.005

正则化 l1 时不同 γ 的 loss



The best accuracy of test data

The best : 0.8387096774193549