

Hyperparameter	Default setting	Experiment1			Experiment
learning_rate	0.001	0.01	0.1	0.001	
batch_size	32	32	32	16	
epochs	10	10	10	10	
dropout_rate	0.2	0.2	0.2	0.2	
hidden_layer_size	64	64	64	64	
Test Accuracy	0.976400018	0.957099974	0.395099998	0.974799991	
Training Time	23.85433110001C	22.08665529999	22.01153550005	48.44654840009	

Hyperparameter	Default Setting	Experiment6			Experiment
learning_rate	0.001	0.01	0.1	0.001	
batch_size	32	32	32	128	
epochs	10	20	40	10	
dropout_rate	0.2	0.2	0.2	0.2	
hidden_layer_size	64	64	64	256	
Test Accuracy	0.976400018	0.957099974	0.499300003	0.980000019	
Training Time	23.85433110001C	35.05695370002	57.66288029996	7.980868	

Hyperparameter	Default Setting	Experiment10			Experiment
learning_rate	0.001	0.00001	0.00001	0.001	
batch_size	32	32	32	32	
epochs	10	10	100	10	
dropout_rate	0.2	0.2	0.2	0	
hidden_layer_size	64	64	64	64	
Test Accuracy	0.976400018	0.900399983	0.951099992	0.974900007	
Training Time	23.85433110001C	15.1367163	143.4568245	14.7178486	

ment2	Experiment3		Experiment4		Experiment	
0.001	0.001	0.001	0.001	0.001	0.001	0.001
64	32	32	32	32	32	32
10	5	20	10	10	10	10
0.2	0.2	0.2	0.4	0.8	0.95	0.2
64	64	64	64	64	64	16
0.973999977	0.968400002	0.976199985	0.9713	0.93730003	0.8791	0.974900007
16.075956000015	11.932146100094	44.134347399	16.261028	15.968310900	15.771401	16.016044100

ment7	Experiment8		Experiment9
0.001	0.001	0.001	0.001
256	32	1024	4096
10	2	2	20
0.2	0.2	0.2	0.2
512	64	64	4096
0.980099976	0.959200025	0.912899971	0.980000002
8.2417537	3.5206199	0.9924654	47.2603804

nent11

0.001

32

10

0

4096

0.985000014

278.3101474

ent5

0.001

32

10

0.2

128

0.9772

20.343311400036328 seconds