

Trial #	# of hidden layers	Neuron in each layer	Learning Rate	Epoch	Train MSE error	Val MSE error
1	1	1024	0.001			
2	1	512	0.001			
3	1	128	0.001	35	0.0106	0.0107
4	1	64	0.001			
5	2	256,128	0.001	21	0.0106	0.0106
6	2	128,256	0.001	4	0.0112	0.0108
7	2	128,128	0.001	12	0.0108	0.0106
8	2	128,512	0.001	9	0.0109	0.0107
9	3	128,256,128	0.001			
10	3	64,128,64	0.001			
11	3	128,256,512	0.001	17	0.0107	0.0106
12	2	256,128	0.0002	13	0.0105	0.0105
13	2	128,256	0.0002	15	0.0104	0.0103
14	2	128,512	0.0002	8	0.0106	0.0104
15	2	256,128	Adaptive, init=0.0002 After 5 epochs, decay: Init/(epoch-4)	35	0.0101	0.0103
16	2	128,256	Adaptive, init=0.0002 After 5 epochs, decay: Init/(epoch-4)	35	0.0100	0.0102
17	2	128,512	Adaptive, init=0.0002 After 5 epochs, decay: Init/(epoch-4)	35	0.0099	0.0101

```
accuracy_score(ytestnotnorm, test_pred)
```

```
0.0631027457068012
```



