

decay	1	0.1	0.01	0.001	0.0001	0.00001
training accuracy	92.45%	95.69%	96.20%	96.19%	96.25%	96.06%
testing accuracy	92.72%	95.20%	95.63%	95.66%	95.86%	95.60%
Result:	Add a small decay increases the accuracy slightly. Increasing decay beyond 0.0001 reduces the training and testing accuracy slightly. Decay reduces the chances of overfitting.					
2.2.4						
ReLU+ReLU+Softmax						
decay	0.0001	0.001	0.01	0.1	1	
training accuracy	98.70%	98.33%	98.22%	97.78%	94.19%	
testing accuracy	97.51%	97.12%	97.06%	97.25%	94.42%	
Result:	the decay rate 0.0001 is the best as it gives slightly higher testing and training accuracy of 98.70% and 97.51%					
3.1						
3.1.1						
learning rate	0.1	0.01	0.001	keeping decay as 0		
training accuracy	21.55%	46.06%	40.57%			
testing accuracy	20.91%	44.38%	40.75%			
decay	0.00001	0.0001	0.001	0.01	keeping learning rate as 0.01	
training accuracy	47.35%	46.01%	45.17%	44.83%		
testing accuracy	45.73%	44.46%	43.79%	43.33%		
Result:	Keeping the decay 0, I increased the learning rate from 0.1 to 0.001. On getting the best results for learning rate as 0.01, I kept that as constant and reduced the decay starting from 0.01 to 0.00001. Increasing decay above 0.00001 and increasing learning rate beyond 0.01 resulted in poor training and testing accuracy. The best training and testing accuracy achieved was 47.35% and 45.73% for learning rate=0.01 and decay=0.00001.					