Matthew Oakley

Deep Learning w/ Tensorflow

Dr. Rivas

11/26/17

The best combination came from the first layer having 1024 and the second layer having 256. The reason I believe that the first layer is the largest is due to the necessity of transitioning from a larger size to smaller size when using an autoencoder. This is because when encoding you go from bigger to smaller sizes. The more neurons helps because some of these images are more complex and needs more neurons to help process it. Some data as seen in Figure 2 was overfit due to overtraining and the training error was 0 while other error rates were not as close to zero.

The best combination which was 1024 and 256 saw an improvement when the learning rate was changed to .0001 but not when it was changed to .001.As seen in Figure 5 the lowest testing error was when the learning rate was .0001 which means that it did not go too far in the training. The problem with different learning rates is that the network could overtrain and experience explosive gradient and not be able to recover from it. I believe that this learning rate was just right in that in improved error but did not go too far.

Training Loss					
Neurons	64	128	256	512	1024
32	1.23	1.11	1.12	1.16	1.15

Neurons	64	128	256	512	1024
64	1.12	1.1	1.1	1.18	
128		1.17	1.06	1.06	
256				1.05	1.05
512					1.05

Figure 1. This is the training loss at the end of each test.

Training Error					
Neurons	64	128	256	512	1024
32	0.0833	0.0833	0.0000	0.1000	0.1000
64		0.1000	0.0667	0.0167	0.3333
128			0.0500	0.0000	0.0000
256				0.0333	0.0167
512					0.0167

Figure 2.This is the training error at the end of each test.

Validation Error					
Neurons	64	128	256	512	1024
32	0.5833	0.6833	0.6500	0.6500	0.7167

64		0.6000	0.6500	0.6833	0.6333
128			0.6333	0.7333	0.6500
256				0.7500	0.7000
512					0.7000

Figure 3. This is the validation error at the end of each test.

Testing Error					
Neurons	64	128	256	512	1024
32	0.7000	0.6833	0.7500	0.7000	0.7000
64		0.7667	0.6833	0.7500	0.7000
128			0.7667	0.7500	0.7000
256				0.7667	0.6667
512					0.7167

Figure 4. This is the testing error at the end of each test

	1.08, .0333, .7667,	
1.05, .0167, .7000, .6667	.6333	1.20, .2833, .6667, .6667

Figure 5. This is the loss rate, Training, Validation, Testing error with learning rates of .00001, .0001, .001 respectively.