1) **Nodes:** number of nodes in input layer + number of nodes in each hidden layer + number of nodes in output layer
Edges: number of edges = (number of nodes in input layer * number of nodes in the first hidden layer) + (number of nodes in first hidden layer * number of nodes in second hidden layer) + ... + (number of nodes in (number of hidden layers-1)th hidden layer) * number of nodes in (number of hidden layers)th hidden layer) + (number of nodes in nth hidden layer * number of nodes in output layer)

2a) 7 Segment LED

Configuration	Error Rate (%)	Bit-level error
1 Hidden Layer, 10 Nodes	37.5	0.411374234569
1 Hidden Layer, 100 Nodes	92.96875	0.509298757696
1 Hidden Layer, 200 Nodes	92.96875	0.509250181182
2 Hidden Layers, 10 Nodes	92.96875	0.658477828083
2 Hidden Layers, 30 Nodes	92.96875	0.655555321964
2 Hidden Layers, 50 Nodes	92.96875	0.654918308387
3 Hidden Layers, 10 Nodes	92.96875	0.501948993054
3 Hidden Layers, 20 Nodes	100.0	0.501860208338
3 Hidden Layers, 30 Nodes	100.0	0.501707775994

9 Segment LED

Configuration	Error Rate (%)	Bit-level error
1 Hidden Layer, 10 Nodes	97.65625	0.500829641012
1 Hidden Layer, 100 Nodes	97.65625	0.506584834339
1 Hidden Layer, 200 Nodes	97.65625	0.506818911938
2 Hidden Layers, 10 Nodes	97.65625	0.695970525305
2 Hidden Layers, 30 Nodes	97.65625	0.693997518627
2 Hidden Layers, 50 Nodes	97.65625	0.692390100541
3 Hidden Layers, 10 Nodes	97.65625	0.500972651075

3 Hidden Layers, 20 Nodes	100.0	0.500975578459
3 Hidden Layers, 30 Nodes	100.0	0.500999186835

14 Segment LED

Configuration	Error Rate (%)	Bit-level error
1 Hidden Layer, 10 Nodes	99.2258972555	0.500233559913
1 Hidden Layer, 100 Nodes	98.9913206662	0.50083746341
1 Hidden Layer, 200 Nodes	99.038235984	0.500960120317
2 Hidden Layers, 10 Nodes	98.5221674877	0.866093114153
2 Hidden Layers, 30 Nodes	98.5221674877	0.866093105811
2 Hidden Layers, 50 Nodes	98.5221674877	0.865955403407
3 Hidden Layers, 10 Nodes	98.5221674877	0.500136820463
3 Hidden Layers, 20 Nodes	100.0	0.500136652215
3 Hidden Layers, 30 Nodes	100.0	0.500136884486

2b) Analyze how each neural net's results compare. Attempt to draw more precise conclusions than the general ones previously described. You can also discuss the more subjective correctness measure. **Be thorough, yet concise, in finding and describing interesting relationships in the data.**

As the number of hidden layers and nodes increases, the error rate increases as well. In almost all cases, when there are at least 20 nodes and 3 hidden layers, the error rate is 100%. The more segments in the LED, the higher the error rate. As seen in the charts above, a 7 segment LED maintains an error rate around 92%, 9 segment around 97%, and 14 segment around 99%. The 2 hidden layers always have the highest bit-level error regardless of the number of segment LED. When there is one or three hidden layers, the bit rate error seems to hover around 0.5. However, for 2 hidden layers, the bit-level error increases. For7, 9, and 14 segment LEDs respectively, 0.65, 0.69 , 0.87 are the bit-level errors. The lowest error rate occurred with the lowest of all parameters. It occurred when there was the lowest segment LED (7 segment LED), the lowest number of hidden layers (1 hidden layer), and the lowest number of nodes (10 nodes). This resulted in an error rate of 37.5%. This also, predictably, resulted in the lowest bit-level error from the test runs, producing a bit-level error of 0.41. Among a constant level of hidden layers, the bit-level error and the error rate

percentages are always very similar. For example, For each 3 hidden layer test, there are error rates of 100% and bit-level errors of 0.5.

3a) One Hidden Layer, 10 Nodes

Number of training runs	Error Rate (%)	Bit-level error
5	97.65625	0.500894377651
10	97.65625	0.500703201379
15	97.65625	0.500459789892
25	97.65625	0.500605317696
30	97.65625	0.500232580617
35	97.65625	0.500845954153
40	19.140625	0.394126504363
45	19.53125	0.382691189481
50	97.65625	0.500307587953

Three Hidden Layers, Ten Nodes Each

Number of training runs	Error Rate (%)	Bit-level error
5	97.65625	0.500975586398
10	97.65625	0.500977135573
15	97.65625	0.500975551536
25	97.65625	0.500974645825
30	97.65625	0.500968779164
35	97.65625	0.500810803429
40	97.65625	0.500975586336
45	97.65625	0.500975529053
50	97.65625	0.500975708728

3b) Analyze how the results depend on the amount of training data. Attempt to draw more precise conclusions than the general ones previously described. You can also discuss the more subjective correctness measure. **Be thorough, yet concise, in finding and describing interesting relationships in the data.**

The data for three hidden layers was a lot more consistent that the data for one hidden layer. For example, when given one hidden layer with 10 nodes, the training data produced an error rate of around 98% for every number of training runs except 40 and 45 were significantly lower, and around 20%. This is reflected concurrently in the bit-level error: the error is 0.5 for the majority of the runs, but 0.39 for 40 and 45 runs. However for three hidden layers with 10 nodes, the data is consistent across the board. Every training run has an exact error rate of 97.65625%. All of the bit-level errors are around 0.5 as well. An interesting but not necessarily significant relationship is that both of the anomalies in the set of training data from 10 nodes one hidden layer are in the 40s. They also closely resemble each other (~19%).

4a) One Hidden Layer, 10 Nodes

Trial	Error Rate (%)	Bit-level error
1	97.65625	0.50085454711
2	97.65625	0.500581747486
3	97.65625	0.500697533927
4	97.65625	0.50051480214
5	97.65625	0.50061244665

Three Hidden Layers, 10 Nodes Each

Trial	Error Rate (%)	Bit-level error
1	97.65625	0.500975514455
2	97.65625	0.500975877023
3	97.65625	0.501031321536
4	97.65625	0.500975642246
5	97.65625	0.500975498555

4b) Analyze how consistent the results are. E.g., how much does the randomness effect the results compared to varying the number of training runs or hidden layer configuration? Which of these configurations gives the overall best results? You can also discuss the subjective correctness measure. **Be thorough, yet concise, in finding and describing interesting relationships in the data.**

The training data in this case could be very flawed, so it is hard to draw accurate conclusions. The information that *can* be concluded from this data is that for each trial, the bit-level error and error rate % are almost identical. For all 5 trials on both one hidden layer - 10 nodes and three hidden layers - 10 nodes the error rate is 97.65625%. The same can be said about the bit-level error. They are all nearly identical, around 0.5. Seeing that the data is basically identical, it could be (whether correctly or incorrectly) that the number of hidden layers has no effect on error rate or bit rate error. This doesn't seem like a very plausible conclusion, but it is the extent of what I can conclude based on this limited data.

Data shared among: Benjamin Klimko, Jacob Barrios, Swapnav Deka. Ragib Mostofa, Ashwin Leo