Feed Forward Neural Networks: Hidden Layer Analysis

Stuart Eiffert,
stuarteiffert@gmail.com

Abstract

Blah Blah Blah

1 Introduction

The choice of parameters regarding hidden layers in feedforward neural network can have significant impact on the classification and training performance, leading to the question in architecture selection of how many hidden units to use. Standard approaches suggest choice of neurons in a single layer to be ‘somehwre between the input layer size and output layer size [1], however other suggest these rules are ‘total nonsense’ [2].

This report performs a comparison of results of an FNN with varying numbers of hidden layers, and varying numbers of neurons in each hidden layers for a number of datasets.

2 Method

Using a feed-forward neural network with a single hidden layer, 30 experimental runs have been performed uwing stochastic method on a number of datasets to determine the mean and std dev of performance for training and test datasets.

The number of hidden layers was increased and this procedure was repeated for 2, 3 ,4 and 5 hidden layers, all of the same number of neurons.

For a selection of the used datasets, further testing was carried out wherein the number of neurons in each hidden layer was varied also.

The data sets used for our experiments are taken from the UCI machine learning repository [3] :

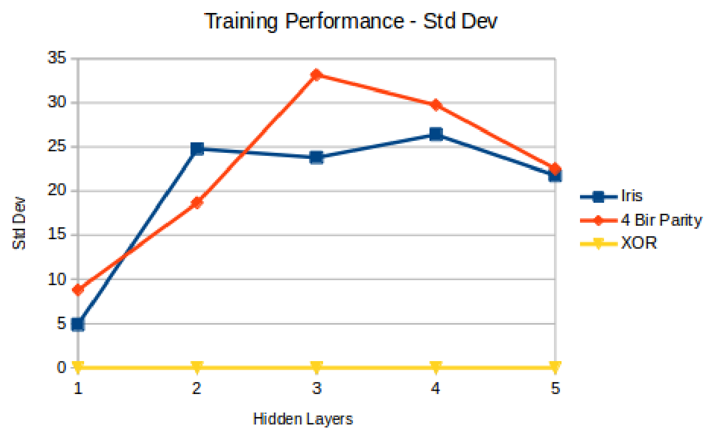
- Iris
- Breast Cancer Wisconsin (diagnostic)
- Cervical Cancer (risk faactors)
- Fertility
- Image Segmentation

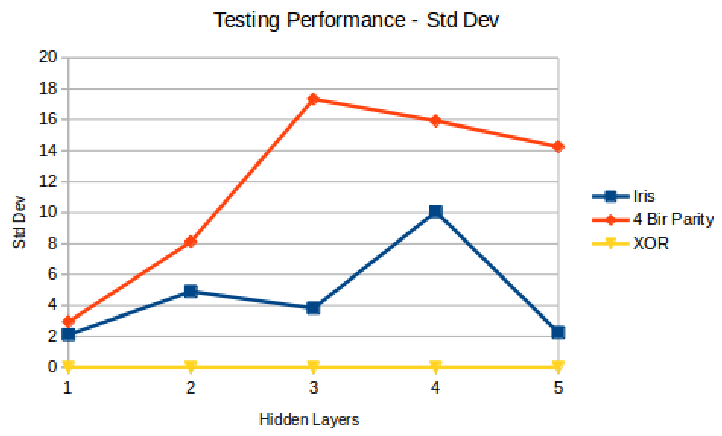
as well as,

- XOR Gate
- 4 bit parity

3 Results

3.1 Hidden Layer Number Comparison

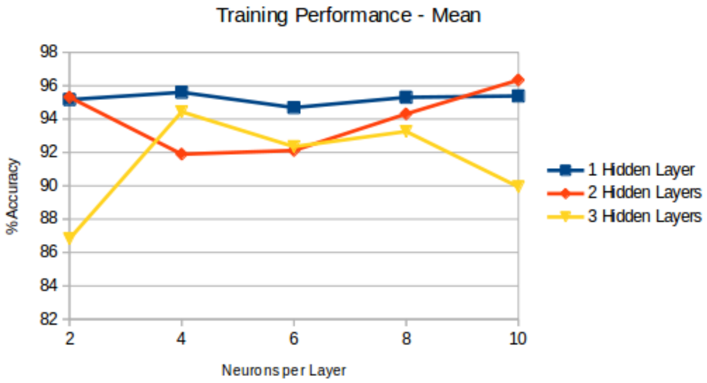


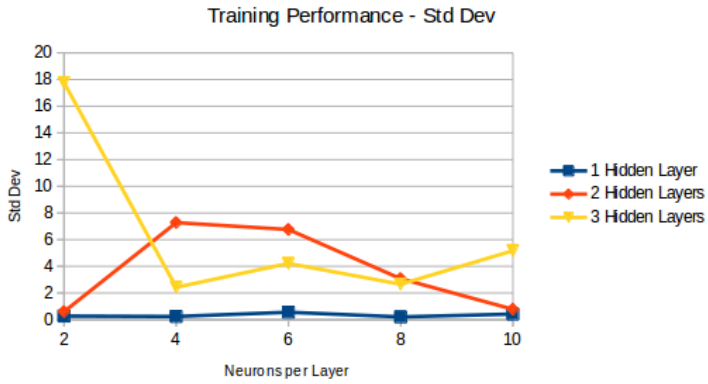


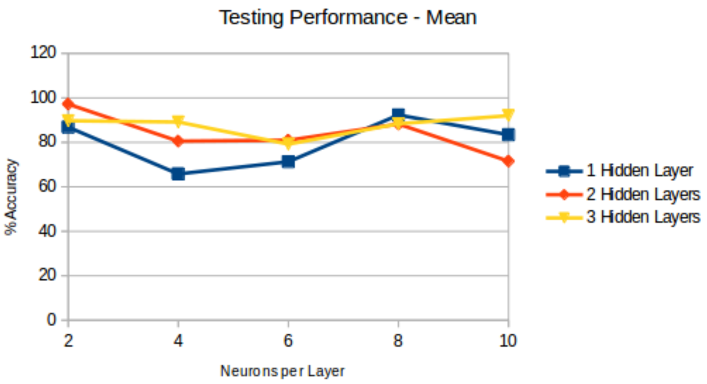
3.2 Hidden Layer Neurons Comparison

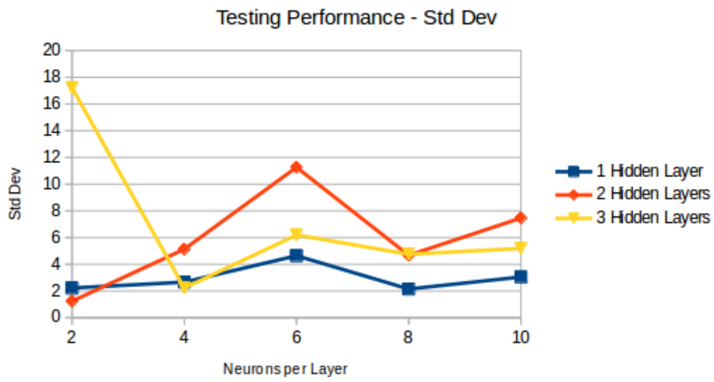
Note that results have been computed for all architectures for the Breast Cancer and Fertility datasets only due to computation time.

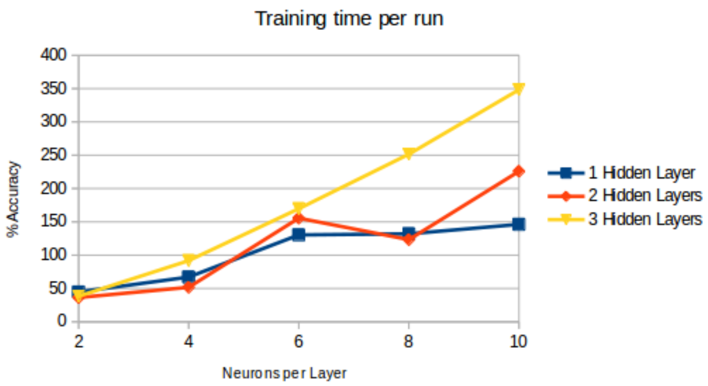
Breast Cancer Dataset



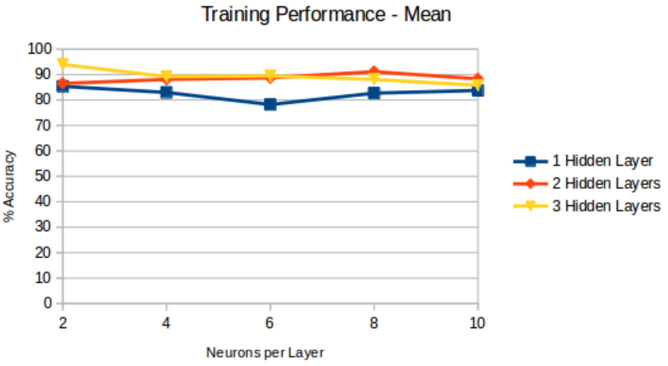


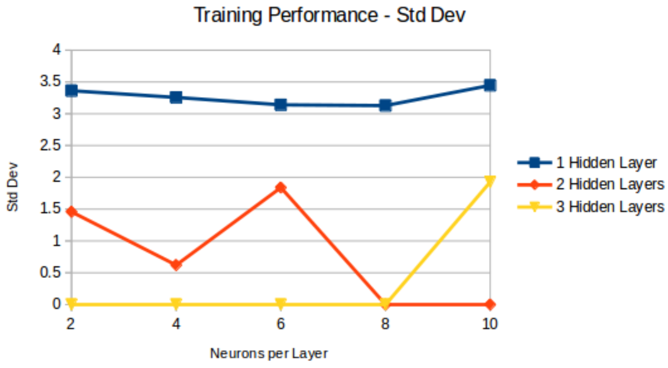


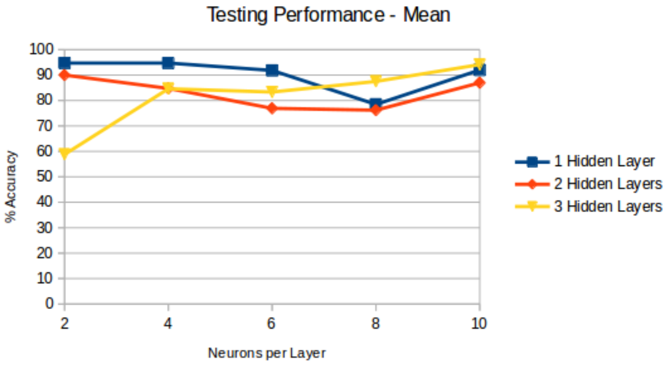


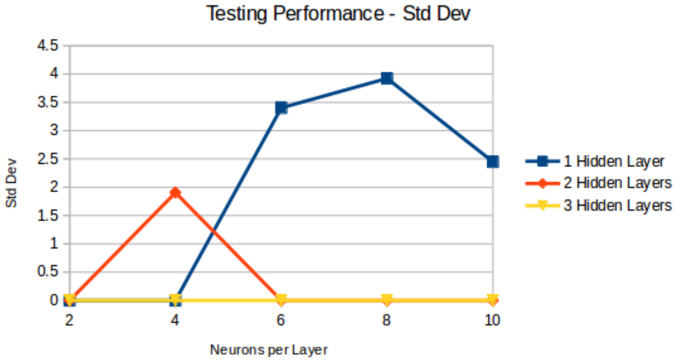


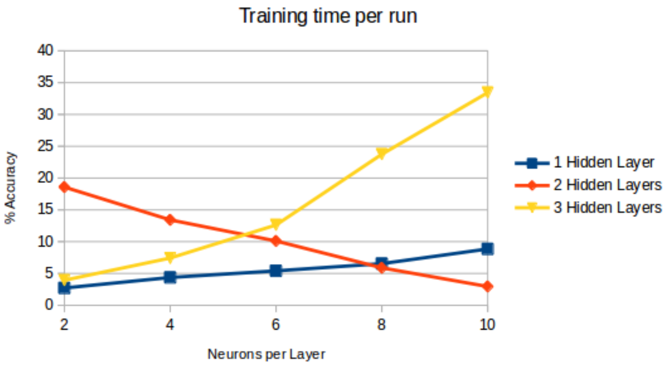
Fertility Dataset











4 Conclusion

From our results, it is not apparent that an increase in number of hidden layers always relates to increase in performance, for either the training or testing datsets.

Regrding the iris, fertility and breast cancer datasets only, an increase in hidden layers does correlate to what appears to be a significant increase in performance accuracy. An increase in the number of neurons per layer also appears to correlate to a significant increase in performance accuracy.

Greater repetition and an increase in experimental runs would be required to confirm these observations.

The performance of a given system architecture can also be seen to differ between datasets, leading to the conclusion that choice of hidden layer and neuron number is dependent on the distribution of input data and required output format.

In summary, there appears to be a small level of benefit to using a greater number of hidden layers, as opposed to a shallower, yet wider network. One hidden layer with a large number of neurons “suffices for the ‘universal approximation’ property” [2][5]. It shouold also be noted that additional hidden layers increases the training time significatly. As such, it is recommended to use a one or two hidden layers and only increase the number of layers if performance cannot be satisfactorily achieved by increasing the number of neurons per layer.

5 References

[1] A. Blum, Neural Networks in C++, Wiley, 1992.

[2] D. Svozil, V. Kvasnicka and J. Posp´ıchal, Introduction to multi-layer feed-forward neural networks , Chemom. Intell. Lab. Syst., 1997, 39, 43.

[3] Lichman, M. (2013). UCI Machine Learning Repository [http://archive.ics.uci.edu/ml]. Irvine, CA: University of California, School of Information and Computer Science.

[4] Backpropagation with one and two neural nets

[5] K.Hornik, Approximation capabilities of multi-layer neural networks, Neural Networks 4 (2) (1991) 251-257.

6 Appendices

OUTPUT DATA

				Training		Testing		Time	
Problem	Experimental Runs	No. Hid	H_layers size	mean	std_dev	mean	std_dev	mean	std_dev
Iris	30	1	6	89.151515152	4.8726669678	96.333333333	2.1147629234	7.634077239	1.4191655839
Iris	30	2	6,6	59.484848485	24.77744835	96.083333333	4.9025220267	14.668650556	2.7893972152
Iris	30	3	6,6,6	67.333333333	23.774576478	95.166666667	3.8151743808	15.008756272	2.8874896605
Iris	30	4	6,6,6,6	13.515151515	26.397818314	96.833333333	10.060925515	27.941106804	5.3064713809
Iris	30	5	6,6,6,6,6	13.181818182	21.762238886	99.583333333	2.2438186696	29.534411542	5.7378049708
4 Bit Parity	30	1	6	94.583333333	8.7994633675	97.9166666666	2.9462782549	2.8050020615	1.2538881268
4 Bit Parity	30	2	6,6	78.125	18.66299257	91.666666667	8.1223286207	6.8211506446	0.4817007971
4 Bit Parity	30	3	6,6,6	56.666666667	33.147921839	76.25	17.335536719	9.8347765923	0.2871531131
4 Bit Parity	30	4	6,6,6,6	23.541666667	29.742091619	56.875	15.922174004	12.839442118	0.6269601808
4 Bit Parity	30	5	6,6,6,6,6	11.25	22.5	50	14.252192814	16.17000045	0.5717422375
XOR	30	1	6	100	0	100	0	0.0043167353	0.002570154
XOR	30	2	6,6	100	0	100	0	0.0043167353	0.002570154
XOR	30	3	6,6,6	100	0	100	0	0.0050983906	0.003046937
XOR	30	4	6,6,6,6	100	0	100	0	0.0058403095	0.006022957
XOR	30	5	6,6,6,6,6	100	0	100	0	0.0079867681	0.008734891
Breastcancer	5	1	2	95.169300226	0.2708803612	86.825396825	2.2222222222	44.788816977	7.4526215296
Breastcancer	5	1	4	95.604395604	0.2407571681	65.789473684	2.6606580505	67.103379059	10.60887956
Breastcancer	5	1	6	94.690265487	0.5596951611	71.282051282	4.6374905868	130.36547914	15.226529212
Breastcancer	5	1	8	95.306553911	0.20714501	92.291666667	2.144922946	131.95559478	7.7630060395
Breastcancer	5	1	10	95.387931034	0.4223258177	83.428571429	3.0476190476	146.19833765	44.750535063
Breastcancer	5	2	2,2	95.3125	0.5989467797	97.190082645	1.2369115328	36.366285753	11.613964115
Breastcancer	5	2	4,4	91.895424837	7.2773407141	80.545454545	5.1297221744	51.719136429	35.098623323
Breastcancer	5	2	6,6	92.117117117	6.7582580915	80.96	11.261545187	155.52389216	45.453290912
Breastcancer	5	2	8,8	94.310722101	3.0728209508	88.214285714	4.6702488681	123.29287505	66.991609759
Breastcancer	5	2		96.344086022	0.7813291236	71.538461538	7.46292612	225.82780356	68.916026778
Breastcancer	5	3	2,2,2	86.824034335	17.771868345	89.708737864	17.199543602	38.304488564	19.501490827
Breastcancer	5	3	4,4,4	94.437086093	2.4278546728	89.137931034	2.2213963322	91.800189543	31.495460064
Breastcancer	5	3	6,6,6	92.348993289	4.2172247581	79.180327869	6.1905598651	169.69346929	46.363140361
Breastcancer	5	3	8,8,8	93.259911894	2.6600051347	88.347826087	4.7500870552	251.57583084	82.859628101
Breastcancer	5	3		89.934640523	5.1710688276	92	5.1873973155	348.37580481	78.515013493
Fertility	5	1	2	85.365853659	3.3619631103	94.736842105	0	2.7078350544	0.1363234087
Fertility	5	1	4	82.962962963	3.2570138169	94.736842105	0	4.3668388367	0.0397814486
Fertility	5	1	6	78.205128205	3.1403714651	91.818181818	3.4015067152	5.4049056053	0.0692086546
Fertility	5	1	8	82.702702703	3.1286045683	78.461538462	3.9223227028	6.5400352478	0.0934385724
Fertility	5	1	10	83.75	3.446012188	92	2.4494897428	8.8466826916	0.063808494
Fertility	5	2	2,2	86.5	1.4577379737	90	0	18.578540945	0.1131336477
Fertility	5	2	4,4	88.101265823	0.6201239855	84.761904762	1.9047619048	13.419491529	0.0324536671
Fertility	5	2	6,6	88.735632184	1.8390804598	76.923076923	0	10.105801249	0.0897440965
Fertility	5	2	8,8	91.139240506	0	76.19047619	0	5.8837672234	0.0947533404
Fertility	5	2	10,10	88.311688312	0	86.956521739	0	2.9401092052	0.0246012339
Fertility	5	3	2,2,2	93.975903614	0	58.823529412	0	3.930153656	0.034406463
Fertility	5	3	4,4,4	89.189189189	0	84.615384615	0	7.4243725777	0.1088714179
Fertility	5	3	6,6,6	89.473684211	0	83.333333333	0	12.624750185	0.047673394
Fertility	5	3	8,8,8	88.095238095	0	87.5	0	23.744304562	0.2042625672
Fertility	5	2	10,10,10	85.78313253	1.9277108434	94.117647059	0	33.432540131	0.1720947393