PERFORMANCE COMPARISON OF DEEP BELIEF NETS AND NEURAL NETWORKS

EECS 700 – COMPUTATIONAL NEUROSCIENCE

Sai Nivedita Chandrasekaran

May 3, 2012

OVERVIEW

- Biological Neural Network
- Artificial Neural Network
- History of NNs
- Restricted Boltzmann Machines
- Deep Belief Nets
- Dataset
- Problem Statement and Hypothesis
- Experimental Setup for DBN & NN
- Results
- Conclusion
- Acknowledgement

BIOLOGICAL NEURAL NETWORK

- Human brain has about 10^{11} neurons and 10^{14} synapses.
- Neuron consists of a soma, axons and dendrites.
- Synapse links an axon to a dendrite.
- For a given stimulus:

  Synapse might increase – excite

  Synapse might decrease – inhibit
- Learning happens by changes to synapses.

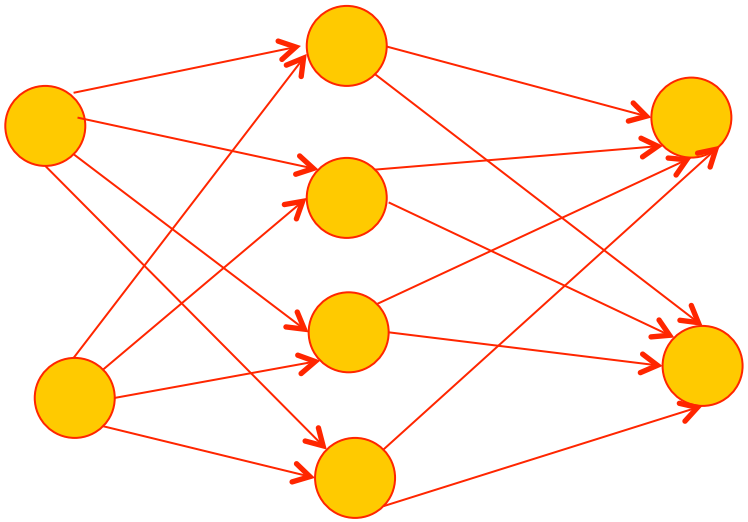
ARTIFICIAL NEURAL NETWORKS

- An artificial neural network is inspired by the biological neural network.

Biological	Artificial
Soma	Unit
Axon, Dendrite	Connection
Synapse	Weight
Potential	Weighted sum
Threshold	Bias weight
Signal	Activation

HISTORY OF NNs

- First generation
  - Perceptrons used one layer of hand coded features and recognized targets by learning to weigh the features.
  - Perceptrons are limited to what they can learn.

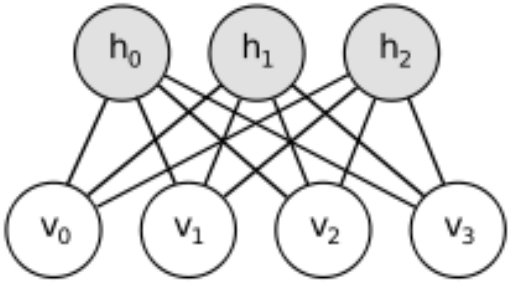


HISTORY Contd…

- Output is compared with the expected target and the error is calculated.
- The error is back propagated to determine the derivatives.
- The network might have multiple hidden layers.
- Disadvantage
  - Requires labeled training data.
  - Time consuming for multiple hidden layers.
  - Possibility of converging in local optima.

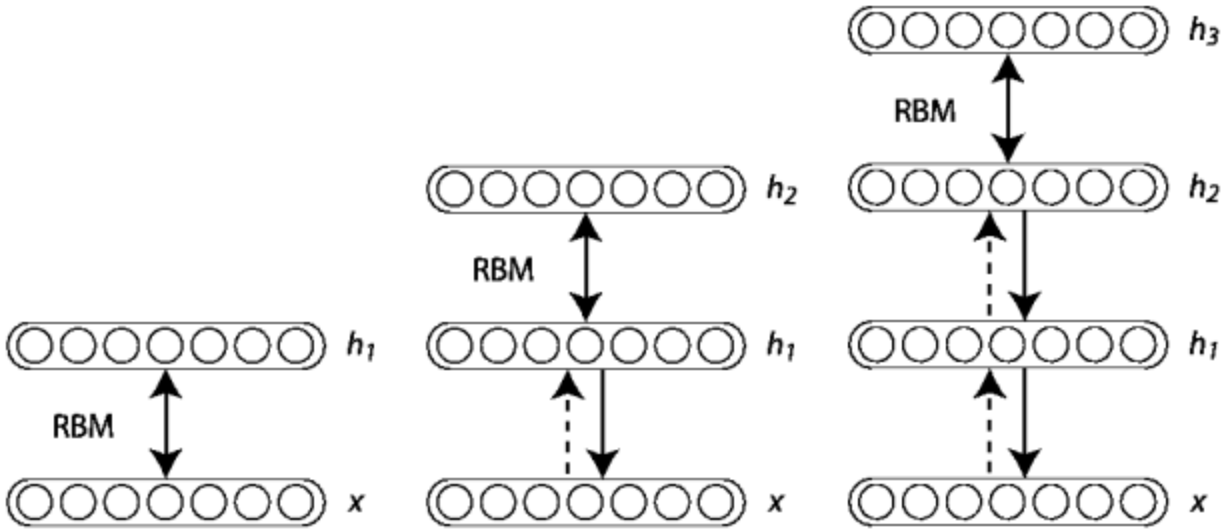
Restricted Boltzmann Machines

- RBMs is a stochastic neural network.
- Each neuron have random behavior when activated.
- It has one layer of visible units and one layer of hidden units.
- The connections are bidirectional.
- The learning is unsupervised until the last layer.



DEEP BELIEF NETS

- RBMs are stacked and trained in a greedy way to form deep belief nets.
- Model the joint distribution between observed output and the hidden layers.



DATASET

- Wisconsin Breast Cancer dataset. (UCI repository)
- Number of instances = 699 (250-B , 449-M)
- Number of features = 10
- Example of attributes: Clump thickness, marginal adhesion, uniformity of cell shape, etc.
- Class:

  - 2 for benign

  - 4 for malignant
- Ignored the sample ID.

PROBLEM STATEMENT

- Binary classification problem.
- To predict if the sample is benign or malignant.

HYPOTHESIS

- Deep belief nets will outperform traditional neural networks.

EXPERIMENTAL SETUP FOR DBN

- Considered 600 samples.
- Training data: 20 batches containing 20 samples each.
- Testing data: 20 batches containing 10 samples each.
- Relabeled benign to be 0 and malignant to be 1.
- Used 3 layers.
- For DBNs, modified the code given for the MNIST data.

EXPERIMENTAL SETUP FOR NN

- Considered 600 samples
- Training data: 400 samples
- Test data : 200 samples
- Learning rate : Varied it between 0.5 and 1.0
- Performed model selection to pick the best learning rate
- Used 2 layers
- Implemented NN in matlab

RESULTS

Test Result

Measure	Definition	DBN	NN
Precision	TP/TP+FP	90%	95%
Recall	TP/TP+FN	45%	88%
Accuracy	TP+TN/ (TP+TN+FP+FN)	87%	90%

- DBN had zero error in training and had an accuracy of 100%
- NN had a training accuracy of 95.48%

HYPOTHESIS FAILED!!

- Low sample size ~ 600.
- DBN is effective for large sample sizes.
- Low sample size doesn't help in capturing the effects of features.
- DBN is capable of over fitting! (My training error was zero)
- Bad generalization performance. (Ineffective classification of new samples)

FUTURE WORK

- Experiment with large datasets.
- The most exciting work according to me would be to exploit multi-task learning using DBNs.

ACKNOWLEDGEMENT

- I would like to thank Dr.Brian for suggesting deep belief nets for my project as it opened a whole new direction for me in machine learning.

REFERENCES

- http://www.cs.utsa.edu/~bylander/cs6243/neural-networks
- http://www.cs.toronto.edu/~hinton/nipstutorial/nipstut3
- Deeplearning.net
- http://www.cs.toronto.edu/~hinton/MatlabForSciencePaper.html
- Wolberg, W.H., & Mangasarian, O.L. (1990). *Multisurface method of pattern separation for medical diagnosis applied to breast cytology*. In Proceedings of the National Academy of Sciences, 87, 9193--9196.
- Zhang, J. (1992). *Selecting typical instances in instance-based learning*. In Proceedings of the Ninth International Machine Learning Conference (pp. 470--479). Aberdeen, Scotland: Morgan Kaufmann.

Thanks!!!

Questions ?!?!