

- **Dr. Geoffrey E. Hinton**, University of Toronto, CANADA

Deep belief nets are probabilistic generative models that are composed of multiple layers of stochastic, latent variables. The latent variables typically have binary values and are often called *hidden units* or *feature detectors*. The top two layers have undirected, symmetric connections between them and form an associative memory. The lower layers receive top-down, directed connections from the layer above. The states of the units in the lowest layer represent a data vector.

The two most significant properties of deep belief nets are:

- There is an efficient, layer-by-layer procedure for learning the top-down, generative weights that determine how the variables in one layer depend on the variables in the layer above.
- After learning, the values of the latent variables in every layer can be inferred by a single, bottom-up pass that starts with an observed data vector in the bottom layer and uses the generative weights in the reverse direction.

Deep belief nets are learned one layer at a time by treating the values of the latent variables in one layer, when they are being inferred from data, as the data for training the next layer. This efficient, greedy learning can be followed by, or combined with, other learning procedures that fine-tune all of the weights to improve the generative or discriminative performance of the whole network.

Discriminative fine-tuning can be performed by adding a final layer of variables that represent the desired outputs and backpropagating error derivatives. When networks with many hidden layers are applied to highly-structured input data, such as images, backpropagation works much better if the feature detectors in the hidden layers are initialized by learning a deep belief net that models the structure in the input data (Hinton & Salakhutdinov, 2006).

Contents

- 1 Deep Belief Nets as Compositions of Simple Learning Modules
- 2 The Theoretical Justification of the Learning Procedure
- 3 Deep Belief Nets with Other Types of Variable
- 4 Using Autoencoders as the Learning Module
- 5 Applications of Deep Belief Nets
- 6 References
- 7 See also

Deep Belief Nets as Compositions of Simple Learning Modules

A deep belief net can be viewed as a composition of simple learning modules each of which is a restricted type of Boltzmann machine that contains a layer of *visible units* that represent the data and a layer of *hidden units* that learn to represent features that capture higher-order correlations in the data. The two layers are connected by a matrix of symmetrically weighted connections, W , and there are no connections within a layer. Given a vector of

activities v for the visible units, the hidden units are all conditionally independent so it is easy to sample a vector, h , from the factorial posterior distribution over hidden vectors, $p(h|v, W)$. It is also easy to sample from $p(v|h, W)$. By starting with an observed data vector on the visible units and alternating several times between sampling from $p(h|v, W)$ and $p(v|h, W)$, it is easy to get a learning signal. This signal is simply the difference between the pairwise correlations of the visible and hidden units at the beginning and end of the sampling (see Boltzmann machine for details).

The Theoretical Justification of the Learning Procedure

The key idea behind deep belief nets is that the weights, W , learned by a restricted Boltzmann machine define both $p(v|h, W)$ and the prior distribution over hidden vectors, $p(h|W)$, so the probability of generating a visible vector, v , can be written as:

$$p(v) = \sum_h p(h|W)p(v|h, W)$$

After learning W , we keep $p(v|h, W)$ but we replace $p(h|W)$ by a better model of the *aggregated* posterior distribution over hidden vectors – i.e. the non-factorial distribution produced by averaging the factorial posterior distributions produced by the individual data vectors. The better model is learned by treating the hidden activity vectors produced from the training data as the training data for the next learning module. Hinton, Osindero and Teh (2006) show that this replacement, if performed in the right way, improves a variational lower bound on the probability of the training data under the composite model.

Deep Belief Nets with Other Types of Variable

Deep belief nets typically use a logistic function of the weighted input received from above or below to determine the probability that a binary latent variable has a value of 1 during top-down generation or bottom-up inference, but other types of variable can be used (Welling et. al. 2005) and the variational bound still applies, provided the variables are all in the exponential family (i.e. the log probability is linear in the parameters).

Using Autoencoders as the Learning Module

A closely related approach, that is also called a deep belief net, uses the same type of greedy, layer-by-layer learning with a different kind of learning module -- an *autoencoder* that simply tries to reproduce each data vector from the feature activations that it causes (Bengio et.al., 2007; LeCun et. al. 2007). However, the variational bound no longer applies and an autoencoder module is less good at ignoring random noise in its training data (Larochelle et.al., 2007).

Applications of Deep Belief Nets

Deep belief nets have been used for generating and recognizing images (Hinton, Osindero & Teh 2006, Ranzato et. al. 2007, Bengio et.al., 2007), video sequences (Sutskever and Hinton, 2007), and motion-capture data (Taylor et. al. 2007). If the number of units in the highest layer is small, deep belief nets perform non-linear dimensionality

reduction and they can learn short binary codes that allow very fast retrieval of documents or images (Hinton & Salakhutdinov,2006; Salakhutdinov and Hinton,2007).

References

- Bengio, Y., Lamblin, P., Popovici, P., Larochelle, H. (2007) Greedy Layer-Wise Training of Deep Networks, Advances in Neural Information Processing Systems 19, MIT Press, Cambridge, MA.
- Hinton, G. E, Osindero, S., and Teh, Y. W. (2006). A fast learning algorithm for deep belief nets. Neural Computation, 18:1527-1554.
- Hinton, G. E. and Salakhutdinov, R. R. (2006). Reducing the dimensionality of data with neural networks. Science, 313:504-507.
- Larochelle, H., Erhan, D., Courville, A., Bergstra, J., Bengio, Y. (2007) An Empirical Evaluation of Deep Architectures on Problems with Many Factors of Variation. International Conference on Machine Learning.
- LeCun, Y. and Bengio, Y. (2007) Scaling Learning Algorithms Towards AI. In Bottou et al. (Eds.) Large-Scale Kernel Machines, MIT Press.
- M. Ranzato, F.J. Huang, Y. Boureau, Y. LeCun (2007) Unsupervised Learning of Invariant Feature Hierarchies with Applications to Object Recognition. Proc. of Computer Vision and Pattern Recognition Conference (CVPR 2007), Minneapolis, Minnesota, 2007
- Salakhutdinov, R. R. and Hinton, G. E. (2007) Semantic Hashing. In Proceedings of the SIGIR Workshop on Information Retrieval and Applications of Graphical Models, Amsterdam.
- Sutskever, I. and Hinton, G. E. (2007) Learning multilevel distributed representations for high-dimensional sequences. AI and Statistics, 2007, Puerto Rico.
- Taylor, G. W., Hinton, G. E. and Roweis, S. (2007) Modeling human motion using binary latent variables. Advances in Neural Information Processing Systems 19, MIT Press, Cambridge, MA
- Welling, M., Rosen-Zvi, M., and Hinton, G. E. (2005). Exponential family harmoniums with an application to information retrieval. Advances in Neural Information Processing Systems 17, pages 1481-1488. MIT Press, Cambridge, MA.

Internal references

- Jan A. Sanders (2006) Averaging. Scholarpedia, 1(11):1760.
- Geoffrey E. Hinton (2007) Boltzmann machine. Scholarpedia, 2(5):1668.

See also

Sponsored by: Prof. Ke CHEN, School of Computer Science, The University of Manchester, U.K.

Reviewed by (http://scholarpedia.org/w/index.php?title=Deep_belief_networks&oldid=61111) : Anonymous

Reviewed by (http://scholarpedia.org/w/index.php?title=Deep_belief_networks&oldid=61111) : Dr. Max Welling, School of Information and Computer Science, University of California, Irvine, CA

Reviewed by (http://scholarpedia.org/w/index.php?title=Deep_belief_networks&oldid=61111) : Dr. Yoshua

Category: Pattern Recognition

*This page was last modified on 21
October 2011, at 04:07.*



*This page has been accessed 206,673
times.*

*"Deep belief networks" by Geoffrey E.
Hinton is licensed under a Creative
Commons Attribution-
NonCommercial-ShareAlike 3.0
Unported License. Permissions
beyond the scope of this license are
described in the Terms of Use*