Tensorflow Practice 3

Hanock Kwak
2016-11-03
Biointelligence Lab
Computer Science and Engineering, Seoul National University
DBNs are stacks of restricted Boltzmann machines forming deep (multi-layer) architecture.

2000 top-level neurons
500 neurons
500 neurons
28x28 pixel image (784 neurons)
Restricted Boltzmann machines (RBM)

- We restrict the connectivity to make learning easier.
  - Only one layer of hidden units.
    - We will deal with more layers later.
  - No connections between hidden units.

Energy($v, h$) = $-b'v - c'h - h'Wv$

- In an RBM, the hidden units are conditionally independent given the visible states.
  - So we can quickly get an unbiased sample from the posterior distribution when given a data-vector.
  - This is a big advantage over directed belief nets.

Approximation of the log-likelihood gradient:
- Contrastive Divergence

$$\frac{\partial E(v, h)}{\partial w_{ij}} = -v_i h_j$$
Deep Belief Networks

- **Stacking RBMs** to form Deep architecture
- DBN with \( l \) layers of models the joint distribution between observed vector \( x \) and \( l \) hidden layers \( h \).
- Learning DBN: **fast greedy learning algorithm** for constructing multi-layer directed networks on layer at a time
A neural model of digit recognition

When training the top layer of weights, the labels were provided as part of the input.

The labels were represented by turning on one unit in a ‘softmax’ group of 10 units:

\[
p_i = \frac{\exp(x_i)}{\sum_j \exp(x_j)}
\]
Looking into the ‘mind’ of the machine
Code

• http://solarisailab.com/archives/374
Guide

• First, copy and paste codes to the IPython Notebook