Übung zur Vorlesung

Maschinelles Lernen 2

Sommersemester 2011

Abteilung Maschinelles Lernen
Institut für Softwaretechnik und
theoretische Informatik
Fakultät IV, Technische Universität Berlin
Prof. Dr. Klaus-Robert Müller
Email: klaus-robert.mueller@tu-berlin.de

Blatt 5 - Deep Learning

Abgabe bis Montag, **23. Mai 2011**, 12:00 Uhr, Briefkasten bei Raum FR6061

Programming a restricted Boltzmann machine (14 pts)

- Download the files `mnist.mat`, `rbm.m`, `sigmoid.m` from the course website. These files contain a reduced version of the MNIST handwritten digits dataset and an incomplete code for training a restricted Boltzmann machine on this dataset.
- Complete the code and run the learning algorithm for at least $n = 10000$ iterations (this may take a few minutes). The simulation should write in the same folder the files `w.png` and `x.png` containing a human-readable visualization of the trained weight matrix and a few reconstructed samples.
- Attach to your submission the completed code and the images that you obtained, and also send the code per e-mail (ml-ml2@lists.tu-berlin.de).

Analysis (6 pts)

Let $\theta = (W, \boldsymbol{a}, \boldsymbol{b})$ be the parameters of a restricted Boltzmann machine of N_v visible units and N_h hidden units. The energy $E(\boldsymbol{v}, \boldsymbol{h}) = -\boldsymbol{v}^\top W \boldsymbol{h} - \boldsymbol{v}^\top \boldsymbol{a} - \boldsymbol{h}^\top \boldsymbol{b}$ of the restricted Boltzmann machine determines the joint probability over the set of visible and hidden units as:

$$\mathrm{p}(\boldsymbol{v}, \boldsymbol{h}) = \frac{e^{-E(\boldsymbol{v},\boldsymbol{h})}}{\sum_{\boldsymbol{u},\boldsymbol{g}} e^{-E(\boldsymbol{u},\boldsymbol{g})}}$$

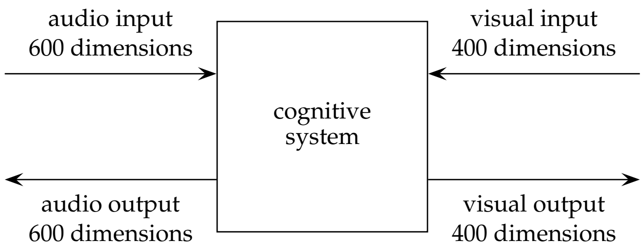
- **Question:** (3 pts) Compute analytically the probability $p(\boldsymbol{v}, \boldsymbol{h})$ when $\boldsymbol{a} = \boldsymbol{0}$, $\boldsymbol{b} = \boldsymbol{0}$ and $W = 0$.
- **Question:** (3 pts) Compute analytically the probability $p(\boldsymbol{v} = \boldsymbol{1}, \boldsymbol{h} = \boldsymbol{1})$ as a function of the parameters of the restricted Boltzmann machine when $W = 0$.
  Hint: Use the fact that disconnected units are independent.

Designing a cognitive system (10 pts)

We would like to build a multi-purpose cognitive system with duplex audio and visual pathways. The system must fulfill the following specifications:

1. Given an audio input (someone reading loud a digit), display an appropriate handwritten digit.
2. Given an visual input (someone writing a digit on an input device), play an appropriate spoken digit.

The system is depicted below:



We do not consider the problem of the acquisition and the synthesis of the audio and visual signal. The visual input/output channel are presented to the system as black and white images of size 20×20. The auditive input/output channel is presented to the system as black and white spectrograms of size 20×30 (20 mel-frequency coefficients sampled at 30 different time intervals). A dataset of 10000 pairs of spoken digits and handwritten digits has been collected from humans, leading to a binary data matrix of size 10000×1000 (10000 samples and $600 + 400$ dimensions). We would like our system to learn the bidirectional relation between spoken digits and handwritten digits in order to be able to mimic it. In this problem, you're asked to design a Boltzmann machine-based architecture that implements the functionality stated above.

- **Question:** (2 pts) Determine what type of Boltzmann machine is the most appropriate for implementing the system (standard restricted Boltzmann machine, third-order restricted Boltzmann machine, deep Boltzmann machine, ...) and motivate your choice.
- **Question:** (3 pts) Sketch a diagram of your Boltzmann machine showing your different units and how they are connected.
- **Question:** (2 pts) Describe the parameters θ of your Boltzmann machine and write the energy function E of your system.
- **Question:** (3 pts) Describe briefly what are the computations required for processing an audio input and deliver an adequate visual output.

Für Fragen zum Übungsblatt bitte in der Google Group http://groups.google.com/group/ml-tu registrieren und die Fragen an die Mailingliste stellen.