Neural networks - Problem solving sessions II

FER

January, 2023

Administrative information about the final exam

- Thursday, 26. of January, 2023 at 5.30P.M.
- B1, B4
- 150 minutes
- 30 points
- the whole curriculum with an emphasis on the second part of the lectures (since *Recursive networks*)

Exam preparation resources

- all lecture notes (except *Committee machines*)
- the chapters of the textbook "Neural Networks - A Comprehensive Foundation" by Simon A. Haykin available on the Neural Net website
- all Jupyter notebooks (lab1-lab8)
- old Problem Solving Sessions
- old exams

Hopfield neural networks

General

- works in the same way as associative memory
- has ability to reconstruct the whole memorized pattern when only its fragment or its noisy version is presented to the network

Properties

- recurrent network consisting of only one layer of input-output nodes
- states of discrete Hopfield networks can achieve values $\{-1, 1\}$ or $\{0, 1\}$
- connection strengths or weights are symmetrical
- no self-connections
- weights can be updated **asynchronously** and **synchronously**

Hopfield neural networks

Learning phase

Let us assume we want to store a set of p N-dimensional vectors: $\{\xi_m | m = 1, ..., p\}$

$$W = \frac{1}{N} \sum_{m=1}^{p} \xi_m \xi_m^T - \frac{p}{N} I \tag{1}$$

Retrieval phase

- A probe vector x is imposed on the discrete HN as its state.
- The network operates using asynchronous updating rule
- Eventually, the network will converge to the stable state which satisfies the following alignment (stability) condition:

$$y = sgn(Wy - \Theta) \tag{2}$$

Hopfield neural networks

Consider a Hopfield network made up of four neurons which is required to store the following three fundamental memories:

$$s_1 = [1, 1, -1, -1]^T$$
$$s_2 = [1, -1, 1, 1]^T$$
$$s_3 = [1, -1, 1, -1]^T$$

Let threshold vector Θ be zero.

1. Determine the synaptic weight matrix W.
2. Find the state which follows the state $[-1, -1, 1, 1]^T$ after a synchronous update.
3. Calculate the energy of the state $[1, -1, 1, 1]^T$

Principal component analysis - PCA

Let there be a training dataset $T = \{(x_i, d_i), i = 1...N\}$ where x_i is a feature vector and d_i is a class label.
$T = \{([0,0],0),([1,1],0),([0,2],0),([2,0],0),$
$([7,5],1),([6,6],1),([5,7],1)\}$.

1. Determine the covariance matrix of the input data.
2. Reduce the dimensionality of the input data to one dimension by applying the PCA algorithm.
3. There is a 1D input data projection which has a smaller error than the one shown above (YES/NO).

Convolutional neural networks - CNNs

The input to the network is an image of size 64×64 with 3 channels. The first layer contains 4 filters of size 3×1 which are applied to each channel with **no** dimensionality reduction. The second layer contains 4 filters of size 1×3 which are applied to each channel with **no** dimensionality reduction. Next, the is a max pooling layer with a 2×2 size and stride 2 in each dimension. Finally, the last layer is a fully connected layer with 128 output neurons, where the activation function contains 2 extra parameters, alongside bias.

1. Sketch the network architecture.
2. How many parameters does the network contain?