Multilayer Perceptron

Ilja Livenson

University of Tartu

16.10.2006
Single layer perceptron

- Structure
1 Single layer perceptron
   - Structure

2 Multilayer perceptron
   - Details
   - Single element
   - Composite network
1. Single layer perceptron
   - Structure

2. Multilayer perceptron
   - Details
   - Single element
   - Composite network

3. Perceptron function
   - MLP with 1 Hidden Layer
1. Single layer perceptron
   - Structure

2. Multilayer perceptron
   - Details
   - Single element
   - Composite network

3. Perceptron function
   - MLP with 1 Hidden Layer

4. More on MLP
   - Training
   - Overfitting
   - Applications
Overview
Single layer perceptron
Multilayer perceptron
Perceptron function
More on MLP

Structure

\[ y = \phi(W^T x + b) \]
• Considered to be the classical neural network
• Neurons are organized into layers (input - \( n \) hidden- output)
• The signal is transported only forwards starting from the input layer
$y = \phi(\hat{W}\hat{x})$
Overview
Single layer perceptron
Multilayer perceptron
Perceptron function
More on MLP

Details
Single element
Composite network

\[ x_0 = 1 \]
\[ x_1 \]
\[ x_2 \]
\[ \vdots \]
\[ x_m \]

Input layer

First hidden layer

Second hidden layer

Output layer

\[ y_1 \]
\[ y_2 \]
Taking simple MLP with only one hidden layer...

\[
F(x_1, x_2, \ldots, x_{m_0}) = \sum_{i=1}^{m_1} \alpha_i \phi\left(\sum_{j=1}^{m_0} w_{ij} x_j + b_j\right)
\]

where

- \(m_0\) - number of input neurons
- \(m_1\) - number of neurons in the hidden layer
- \(\alpha_i\) - weight of \(i\) neuron
There are a lot of algorithms for training MLP:

- Back-propagation (*classical*)
- Back-propagation with momentum
- Conjugate-gradient
- Levenberg-Marquardt
Overfitting - “when too many neurons is bad”

- Loss of generalization
- Complete fit of training set
- Motivates the usage of validation set/cross-validation
Applications

- Non-linear regression
- Classification
- Feature extraction
Demo

XOR-problem
Thank you!