Regularized neural networks with applications in classification problems

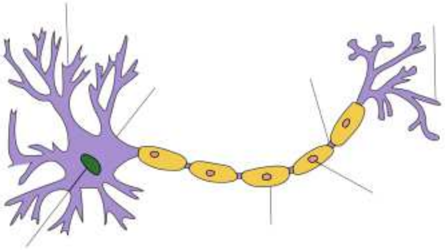


Figure: Structure of a typical neuron (Wikipedia).

Machine learning is a field of artificial intelligence which gives computer systems the ability to “learn” using available data. One of the important parts of machine learning are **neural networks**. **Neural networks** are computing systems that are inspired by biological neural networks and "learn" using available data to perform tasks by considering examples, generally without being programmed with task-specific rules.

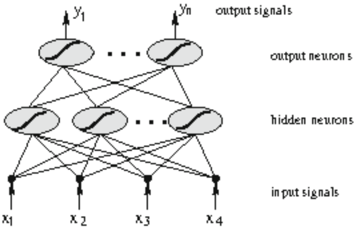


Figure: Example of neural network which contains two interconnected layers (M. Kurbat, *An Introduction to machine learning*, Springer, 2017.)

- The goal of this project is numerical studies and comparison of regularized and non-regularized neural networks in the problem of detection of inter-class boundaries, or classification problem. You should experiment with different number of hidden neurons and 1-3 hidden layers and study how they affect the learning of the neural network. We will study the case when there are no communication between neurons of the same layer and adjasted layer are fully interconnected, see Figure.

Classification problem

- Suppose that we have data points $(x_i, y_i), i = 1, ..., m$. These points are separated into two classes A and B. Assume that these classes are linearly separable.

  Definition

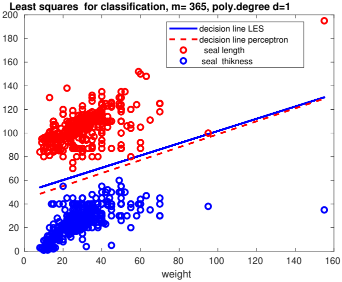
  Let A and B are two data sets of points in an n-dimensional Euclidean space. Then A and B are linearly separable if there exist $n+1$ real numbers $\omega_1, ..., \omega_n, l$ such that every point $x \in A$ satisfies $\sum_{i=1}^{n} \omega_i x_i > l$ and every point $x \in B$ satisfies $\sum_{i=1}^{n} \omega_i x_i < -l$.

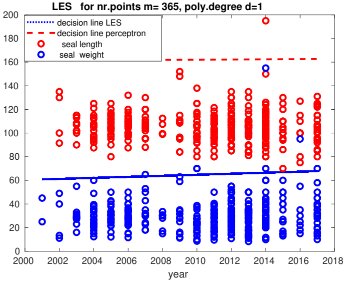
- Our goal is to find the decision line which will separate these two classes. This line will also predict in which class will the new point fall.

Example: classification of grey seals

Experimental data is provided by the DEPARTMENT OF BIOLOGICAL AND ENVIRONMENTAL SCIENCES, University of Gothenburg, and is available for download from

`www.waves24/download`





Our goal is to find optimal weights $\omega_{ji}^{(1)}$ and $\omega_{kj}^{(2)}$ in forward propagation

$$y_i = f(\Sigma_j \omega_{ji}^{(1)} x_j) = f(\Sigma_j \omega_{ji}^{(1)} \underbrace{f(\Sigma_k \omega_{kj}^{(2)} x_k)}_{x_j}). \tag{1}$$

To do this we introduce functional

$$F(\omega_{ji}^{(1)}, \omega_{kj}^{(2)}) = \frac{1}{2}\|t_i - y_i\|^2 = \frac{1}{2}\sum_{i=1}^{m}(t_i - y_i)^2. \tag{2}$$

Here, $t = t(x)$ is the target vector which depends on the concrete example x. In the domain with m classes the target vector $t = (t_1(x), ..., t_m(x))$ consists of m binary numbers such that

$$t_i(x) = \begin{cases} 1, & \text{example } x \text{ belongs to } i\text{-th class,} \\ 0, & \text{otherwise.} \end{cases} \tag{3}$$

- Non-regularized neural network

$$F(w) = \frac{1}{2}\|t_i - y_i(w)\|^2 = \frac{1}{2}\sum_{i=1}^{m}(t_i - y_i(w))^2. \tag{4}$$

- Regularized neural network

$$F(w) = \frac{1}{2}\|t_i - y_i(w)\|^2 + \frac{1}{2}\gamma\|w\|^2 = \frac{1}{2}\sum_{i=1}^{m}(t_i - y_i(w))^2 + \frac{1}{2}\gamma\sum_{j=1}^{M}|w_j|^2 \tag{5}$$

Here, γ is reg.parameter, $\|w\|^2 = w^T w = w_1^2 + ... + w_M^2$, M is number of weights.

Project exercises

- Take experimental data from the link

  `https://archive.ics.uci.edu/ml/datasets.html`

  Or download *.xlsx file with data for grey seals and classify time-dependent data in this file for length, thinkness and length of seals depending on season (winter, spring, summer, autumn). Use perceptron learning algorithm and least squares for classification. You can download Matlab-programs with these two algorithms at

  `www.waves24/download`

- Use regularized and non-regularized neural networks for classification. In this project we will restrict ourself to the specific class of neural network called the multilayer perceptron.
- Discover convergence of studied algorithms and compare their performance with respect to applicability, reliability, accuracy, and efficiency. Programs written in Matlab should demonstrate performance for every algorithm.

Project literature

The proposed project books:

- AI: Miroslav Kurbat, *An Introduction to Machine Learning*, Springer, 2017.
- AI: Christopher M. Bishop, *Pattern recognition and machine learning*, Springer, 2009.
- Solution of least squares problems :

  L. Beilina, E. Karchevskii, M. Karchevskii, *Numerical Linear Algebra: Theory and Applications, Springer, 2017*. Book is available at Cremona.

  Matlab and C++ programs for examples in this book are available for download from the course homepage: go to the link of the book and click to

  `GitHub Page with MATLAB Source  Codes`

  on the bottom of this page.