Seminario rivolto agli studenti del III anno di INGEGNERIA BIOMEDICA, CIBERNETICA ed ELETTRONICA

9 Aprile 2019, ore 15:00-19:00, Aula Savagnone ex DEIM

Introduction to Neural Networks for Image and Signal Processing

Speakers: Simona Etinski & Ivan Lazić, university of Novi Sad, Serbia

Abstract:

Living in the era of information technologies provides us with an opportunity to obtain large volumes of data and utilize it in order to predict future events, make decisions and gain deeper understanding of the data by discovering hidden relations within it. In the last five decades different data driven machine learning methods were successful in solving a whole range of the tasks. However, traditional machine learning methods, such as naive Bayes classifier, k-nearest neighbors algorithm (k-NN) or support vector machines, require a lot of domain expertise and even so, some of the tasks remains too hard to be approached and the difficulty increases as more convoluted data is used. It was in 2009, when powered by hardware advances (graphical processing units) and large data volumes, research and use of deep neural networks has flourished resulting in a superhuman performance on many challenging tasks.

In this lecture we are going to present theoretical foundations of neural networks, as well as their practical implementation in image processing, biomedical engineering and audio signal processing. The first part of the lecture introduces neural networks, explains basic theoretical concepts behind them and underlines their advantages over some other machine learning methods. In the second part an implementation of neural networks is illustrated on several real-world problems.

The first thing that we point out is that neural networks perform well on both *regression tasks*, predicting data values, as well as in *classification tasks* of predicting data labels. Such a good performance is achieved via deep data analysis, performed by neural networks, which circumvents complex feature engineering required in classical machine learning approaches. After that, we give a toy example of a simple classification problem in order to describe a mathematical model of the basic neural network unit, called perceptron, as well as to illustrate an intuitive decision-making performed by a perceptron. The perceptron learning algorithm is then introduced as a method that enables us to approach linearly separable classes.

We then introduce a nonlinear problem that reveals limitations of the perceptron, which proves to deal well with linearly separable classes but fails to do so for those which are not linearly separable. Some of nonlinear problems (such as the XOR problem) can be circumvented using a multilayer perceptron approach, but most of them remain unsolvable until nonlinearities are introduced in the network. The neuron unit, as an extension of the perceptron unit, introduces the nonlinearity, in the form of an activation function, and manages to create nonlinear regions for class separation. Throughout the evolution of neural networks different activation functions have emerged, addressing some issues that could arise during the training process of the network. We touch upon some of the activation functions which are most commonly used, such as the sigmoid activation function, tanh, relu & leaky relu.

Following that, we present concepts of shallow and deep networks with the benefits and flaws of each and the learning process of neural networks. The learning process is illustrated by an example of data flow from an input to an output and using the backpropagation algorithm for weight adjustment. As in traditional machine learning, different problems can be met in a training process, such as underfitting and overfitting. To overcome these issues, there are number of practical advices on selection of network architecture, number of units in a layer, a proper network initialization, and more advanced learning techniques that include regularization as well as usage of dropout, early stopping and data augmentation.

In the final section of the lecture more complex network architectures and their use in the engineering fields are described. Some of the topics that we cover in this section are convolutional networks, recurrent networks and generative adversarial networks.