MAI4CAREU

Master programmes in Artificial Intelligence 4 Careers in Europe

University of Ruse

INTELLIGENT COMPUTER SYSTEMS

Svetlana Stefanova

September, 2022

Co-financed by the European Union
Connecting Europe Facility

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

LECTURE 13

NEURAL NETWORK TRAINING

1. Introduction
2. Artificial intelligence learning
3. Learning algorithms
4. Training with a trainer
5. Training without a trainer

Co-financed by the European Union
Connecting Europe Facility

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

CONTENT 1

Learning

The question of learning is in the field of neuroscience and psychology.

Learning is the opposite of instinct:

- **instinct** is genetically determined behavior;
- **learning** is behavior that is changed by experience.

This Master is run under the context of Action
No 2020-EU-IA-0087, co-financed by the EU CEF Telecom
under GA nr. INEA/CEF/ICT/A2020/2267423

CONTENT 1

Why are some things learned and others instinctive?

- In 1896 James Mark Baldwin, an unknown American theorist, writes a philosophical survey article on the subject.
- In the late 1980s, a group of computer scientists decided that his reasoning had a lot to do with their task of making computers learn.

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

CONTENT 1

Plasticity

By learning something, we place ourselves in a selective environment, which contributes to the emergence of a new instinct in the future. Learning gradually gives way to instinct.

Brain plasticity - the ability to adjust the nervous system to its environment.

Co-financed by the European Union
Connecting Europe Facility

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

CONTENT 1

Learning through experience

Experiential learning is a fundamental ability of the human brain.

Learning in a biological system - takes place by correcting the synaptic connections between neurons or forming new connections.

During this process, the brain shows continuous improvement.

Co-financed by the European Union
Connecting Europe Facility

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

CONTENT 1

Cumulative learning

Evolution - in a sense, the human genome can be looked at as a cumulative learning spanning for four billion years.

Co-financed by the European Union
Connecting Europe Facility

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

Learning and AI

AI researchers put learning on a pedestal and their goal became a general purpose learning machine.

Attention is directed towards the synapses between neurons. Learning seems to be a change in their properties.

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

NN and the brain - similarities

- Knowledge enters NN from the environment and is used in the learning process;
- To accumulate knowledge, connections between neurons - synaptic weights - are applied.

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

NN training

One of the most important characteristics of NN is the ability to learn.

Training - adapting the network to better handle a task, given sample observations. It is done for better result by minimizing observed errors. Technically, it is a process of processing the input/output data while repeatedly using the training algorithm and adjusting the weights of the synapses until similarity is achieved.

NNs can also change their own topology - this is due to the fact that neurons in the human brain are constantly dying and new synaptic connections are constantly being created.

Co-financed by the European Union
Connecting Europe Facility

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

CONTENT 3

Learning algorithms

The rule by which connection weights are changed is called the **network's learning rule or algorithm**.

Most often, it starts with random values of the required weights, and then they iteratively change so that at the next step the new output of the network is better than the old one.

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

CONTENT 3

Completed training

- Training is completed when additional observations do not reduce the percentage of the errors.
- Even after training, the error rate is usually not 0.
- If, after training, the error rate is too high, the network must be redesigned.

This Master is run under the context of Action
No 2020-EU-IA-0087, co-financed by the EU CEF Telecom
under GA nr. INEA/CEF/ICT/A2020/2267423

CONTENT 3

NN according to the training method

- **trained with a trainer** - the difference between the obtained and the expected output of the network is monitored (the teacher sets the expected output) and iterative adjustments are made to the weights of the connections according to the selected training rule.
- **trained without a trainer** – there is missing data on the correct output of the network. The weights of the connections are adjusted so that the data representation in the network is the best according to the specified representation quality criterion.

Co-financed by the European Union
Connecting Europe Facility

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

CONTENT 4

NNs trained with a trainer

The task of training with a trainer consists of choosing a specific function *a* that optimally approximates the expected response *y.*

The selection is based on the set of *n* independent, uniformly distributed training examples.

Let *C(y,a(x,w))* be a **measure of the losses** or dissimilarities between the desired output *y* corresponding to an input vector *x* and the response *a* generated by the learning machine.

Co-financed by the European Union
Connecting Europe Facility

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

CONTENT 4

NNs trained with a trainer - problems

Making small changes to the weights *w* and biases *b* often does not lead to a change in the elements of the training data.

This makes it difficult to understand how to change the weights and biases to improve performance.

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

CONTENT 4

More famous learning rules

- perceptron learning rule (fixed-increment learning algorithm);
- error backpropagation learning algorithm;
- competitive learning rule;
- etc.

This Master is run under the context of Action
No 2020-EU-IA-0087, co-financed by the EU CEF Telecom
under GA nr. INEA/CEF/ICT/A2020/2267423

CONTENT 4

Smooth loss function

If we use a **smooth loss function** (such as the quadratic loss function), it turns out to be easy to figure out how to make small changes to the weights and biases so as to achieve an improvement in loss.

We first focus on quadratic loss minimization and then investigate the classification accuracy.

Co-financed by the European Union
Connecting Europe Facility

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

CONTENT 4

Smooth loss function - quadratic type

- ***w*** - all weights in the network;
- ***b*** – all biases;
- ***n*** - the total number of learning inputs;
- ***a*** - the vector of outputs from the network when ***x*** is input, and the sum is of all learning inputs ***x***. The output ***a*** depends on ***x***, ***w*** and ***b***.
- ***C*** – quadratic loss function, also known as the ***mean squared error*** or ***MSE***.

$$C(w,b) \equiv \frac{1}{2n} \sum_x \|y(x) - a\|^2.$$

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

CONTENT 4

Gradient learning

Designing and training an NN is not much different from training any other gradient descent machine learning model.

The biggest difference between linear models and NNs is that the nonlinearity of NNs makes the most interesting loss functions non-convex. his means that NNs are typically trained not by solving linear equations, but by using iterative gradient-based optimizers that drive the loss function to a very low value.

This Master is run under the context of Action
No 2020-EU-IA-0087, co-financed by the EU CEF Telecom
under GA nr. INEA/CEF/ICT/A2020/2267423

CONTENT 4

Learning through generalization

Generalization - the ability to obtain a reasoned result based on data that was not encountered in the learning process.

This property allows NNs to solve complex tasks, which at the given moment are considered difficult to solve.

In reality, the autonomous work of the NN cannot provide ready-made solutions. It is necessary to integrate them into complex systems. The complex task is divided into relatively simple sequences, some of which can be solved by NN.

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

CONTENT 5

NNs trained without a trainer (learning without a trainer)

The name itself emphasizes the absence of a manager controlling the process of setting the weighting coefficients. There are also no marked examples on which the training is conducted.

- We can distinguish two methods:

➢ reinforcement learning - the formation of reflected input signals into output signals is performed in the process of interaction with the external environment, with the aim of minimizing the scalar performance index.

➢ learning based on self-organization (self-organized) - takes place without the intervention of an external trainer or corrector controlling the training process. There is only a task-independent quality measure of performance that the NN needs to learn, and the free parameters of the network are optimized in regard to this measure.

Co-financed by the European Union
Connecting Europe Facility

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

MAI4CAREU

Master programmes in Artificial Intelligence 4 Careers in Europe

Co-financed by the European Union
Connecting Europe Facility

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423