Introduction to Machine Learning Perceptron

Barnabás Póczos



Contents

❑ History of Artificial Neural Networks ❑ Definitions: Perceptron, Multi-Layer Perceptron ❑ Perceptron algorithm

2

Short History of Artificial Neural Networks3

Short History

❑ **Progression (1943-1960)**

• First mathematical model of neurons

▪ Pitts & McCulloch (1943)

• Beginning of artificial neural networks

• Perceptron, Rosenblatt (1958)

▪ A single neuron for classification

▪ Perceptron learning rule

▪ Perceptron convergence theorem

❑ **Degression (1960-1980)**

• Perceptron can’t even learn the XOR function

• We don’t know how to train MLP

• 1963 Backpropagation… but not much attention…

Bryson, A.E.; W.F. Denham; S.E. Dreyfus. Optimal programming problems with inequality constraints. I: Necessary conditions for extremal solutions. AIAA J. 1, 11 (1963) 2544-2550

4

Short History

❑ **Progression (1980-)**

• 1986 Backpropagation reinvented:

▪ Rumelhart, Hinton, Williams:

Learning representations by back-propagating errors. **Nature**, 323, 533—536, 1986

• Successful applications:

▪ Character recognition, autonomous cars,…

• **Open questions**: Overfitting? Network structure? Neuron number? Layer number? Bad local minimum points? When to stop training?

• Hopfield nets (1982), Boltzmann machines,…

5

Short History

❑ **Degression (1993-)**

• SVM: Vapnik and his co-workers developed the Support Vector Machine (1993). It is a shallow architecture.

• SVM and Graphical models almost kill the ANN research. • Training deeper networks consistently yields poor results.

• Exception: deep convolutional neural networks, Yann LeCun 1998. (discriminative model)

6

Short History

**Progression (2006-)**

**Deep Belief Networks (DBN)**

• Hinton, G. E, Osindero, S., and Teh, Y. W. (2006).

A fast learning algorithm for deep belief nets.

Neural Computation, 18:1527-1554.

• Generative graphical model

• Based on restrictive Boltzmann machines

• Can be trained efficiently

**Deep Autoencoder based networks**

Bengio, Y., Lamblin, P., Popovici, P., Larochelle, H. (2007).

Greedy Layer-Wise Training of Deep Networks,

Advances in Neural Information Processing Systems 19

**Convolutional neural networks running on GPUs**

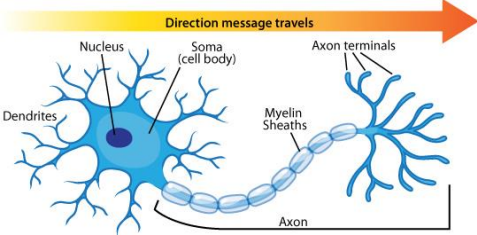
Alex Krizhevsky, Ilya Sutskever, Geoffrey Hinton, Advances in Neural Information Processing Systems 2012

7

The Neuron

8

The Neuron

– Each neuron has a body, axon, and many dendrites

– A neuron can fire or rest

– If the sum of weighted inputs larger than a threshold, then the neuron fires.

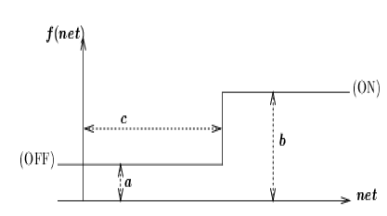
– Synapses: The gap between the axon and other neuron’s dendrites. It determines the weights in the sum.

9

The Mathematical Model of a Neuron



10

Typical activation functions• **Identity function **

• **Threshold function**

(perceptron)

• **Ramp function**

11

Typical activation functions • **Logistic function**

****

• **Hyperbolic tangent function**

****12

Typical activation functions

• **Rectified Linear Unit (ReLU)** 

****

• **Softplus function**

(This is a smooth approximation of ReLU)



• **Leaky ReLU**

****

• **Exponential Linear Unit**

****13

14

15

Structure of Neural Networks

16

Fully Connected Neural Network**Input neurons, Hidden neurons, Output neurons**

17

Layers, Feedforward neural networks **Convention: The input layer is Layer 0.**

18

Multilayer Perceptron

• **Multilayer perceptron:** Connections only between Layer i and Layer i+1

• The most popular architecture.

19

20

Recurrent Neural Networks



**Recurrent NN**: there are connections backwards too.21

The Perceptron

22

The Training Set







23

**The Perceptron**

****

****

****

****

****

****24

The Perceptron





1



-1





25

**Matlab:** opengl hardwarebasic**, nnd4pr**

**Matlab demos: nnd3pc**

****27

The Perceptron Algorithm

28

**The Perceptron algorithm**

The perceptron learning algorithm







29

The perceptron algorithm

**Observation**

****30

**The Perceptron Algorithm **

**How can we remember this rule?**

****An interesting property:

we do not require the learning rate to go to zero!

31

**The Perceptron Algorithm**

****32

**Perceptron Convergence**

****

33

**Perceptron Convergence**

****

****

****

****34

**Perceptron Convergence**

****

****

**Lemma** Using this notation, the update rule can be written as **Proof**

****

****

35

**Lemma**

**Perceptron Convergence**

****36

**Perceptron Convergence**

****37

**Lower bound**

****38

**Upper bound**

****

****

****

Therefore,



39

**Upper bound**

**Therefore,**

****40

**The Perceptron Algorithm**

****

****41

**Take me home!**

❑ **History of Neural Networks**

❑ **Mathematical model of the neuron**

❑ **Activation Functions**

❑ **Perceptron definition**

❑ **Perceptron algorithm**

❑ **Perceptron Convergence Theorem**

42