



EVROPSKÁ UNIE
Evropské strukturální a investiční fondy
Operační program Výzkum, vývoj a vzdělávání



MINISTERSTVO ŠKOLSTVÍ,
MLÁDEŽE A TĚLOVÝCHOVY

Název projektu	Rozvoj vzdělávání na Slezské univerzitě v Opavě
Registrační číslo projektu	CZ.02.2.69/0.0./0.0/16_015/0002400

Dolování dat

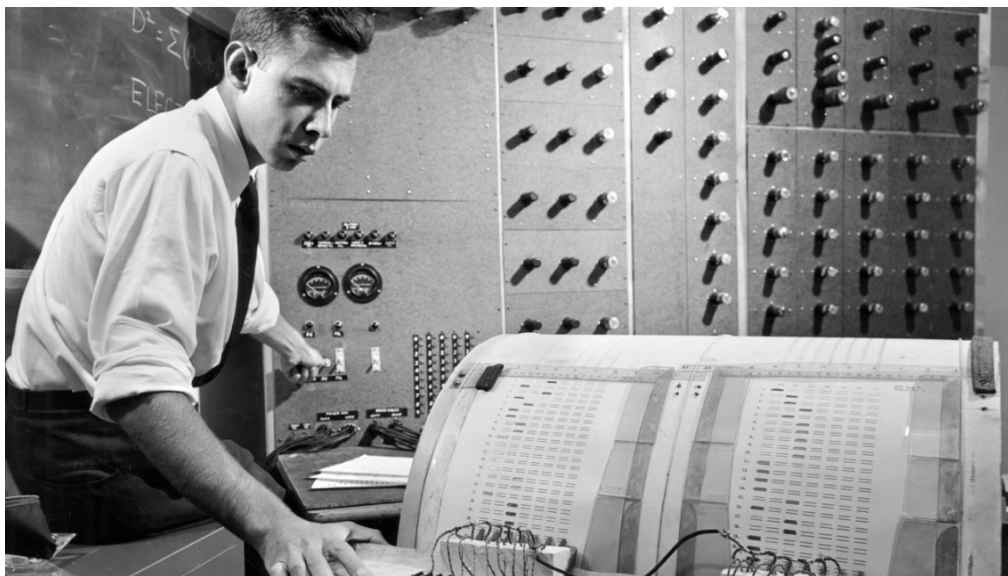
Neuronové sítě

Jan Górecki



**SLEZSKÁ
UNIVERZITA**
OBCHODNĚ PODNIKATELSKÁ
FAKULTA V KARVINĚ

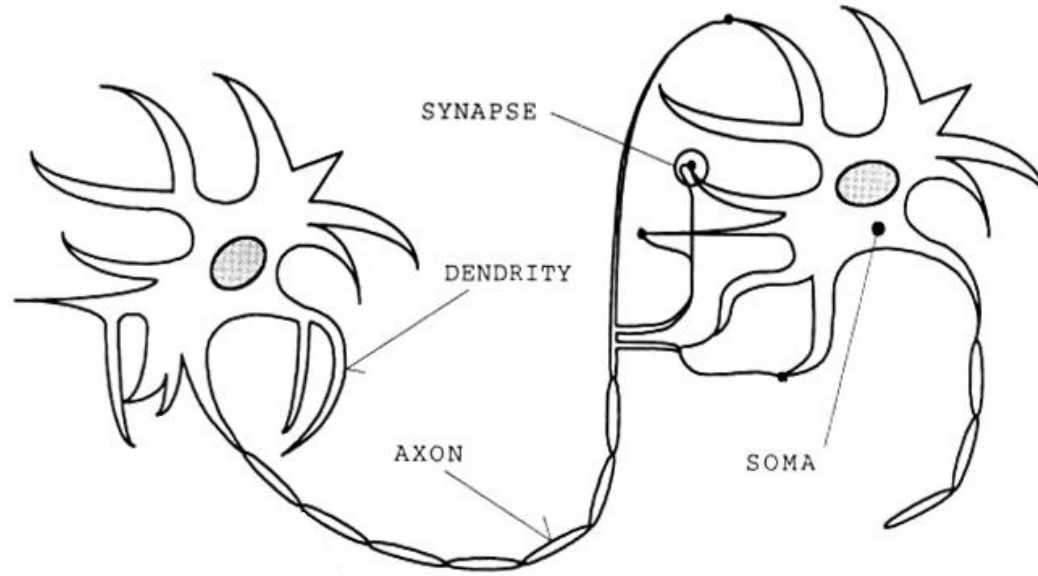
- Biologický neuron
- Modely neuronu
- Učení neuronu
- Vícevrstvý perceptron
- Backpropagation
- Příklady použití



Biologický neuron



**SLEZSKÁ
UNIVERZITA**
OBCHODNĚ PODNIKATELSKÁ
FAKULTA V KARVINĚ



1. Logický neuron (McCulloch, Pitts, 1943)

$$w \in \mathbb{R}, x, y \in \{0, 1\}$$

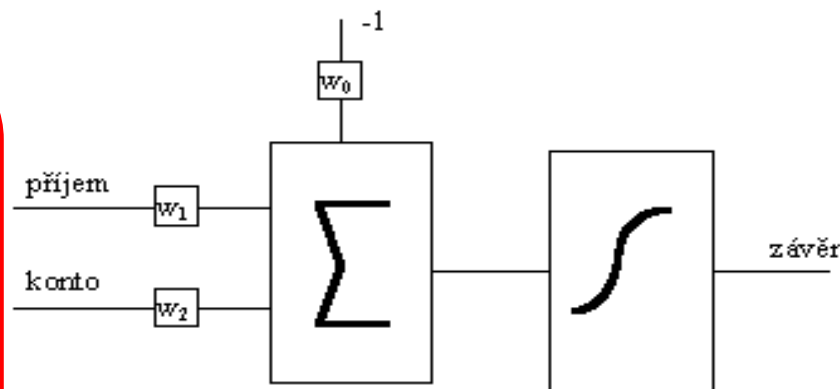
2. ADALINE (Widrow, 1960)

$$x, w \in \mathbb{R}, y \in \{0, 1\}$$

$$\hat{y} = 1 \text{ pro } \sum_{i=1}^d w_i x_i \geq w_0$$

$$\hat{y} = 0 \text{ pro } \sum_{i=1}^d w_i x_i < w_0,$$

kde d je počet vstupů



Příklad ADALINE



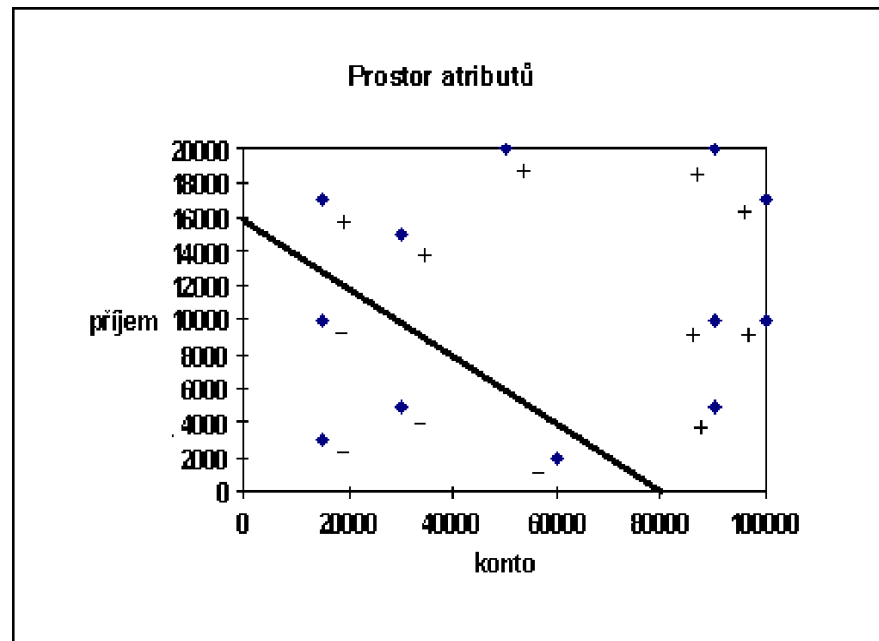
SLEZSKÁ
UNIVERZITA
OBCHODNĚ PODNIKATELSKÁ
FAKULTA V KARLOVĚ

příjem	konto	úvěr	úvěr(bin)
3000	15000	ne	0
10000	15000	ne	0
17000	15000	ano	1
5000	30000	ne	0
15000	30000	ano	1
20000	50000	ano	1
2000	60000	ne	0
5000	90000	ano	1
10000	90000	ano	1
20000	90000	ano	1
10000	100000	ano	1
17000	100000	ano	0

$$\begin{aligned}x_1 &= \text{příjem,} \\x_2 &= \text{konto,} \\y &= \text{úvěr (bin),} \\w_1 &= 1, \\w_2 &= 0.2, \\w_0 &= 16000\end{aligned}$$

Separující přímka:

$$\sum_{i=1}^d w_i x_i = w_0$$



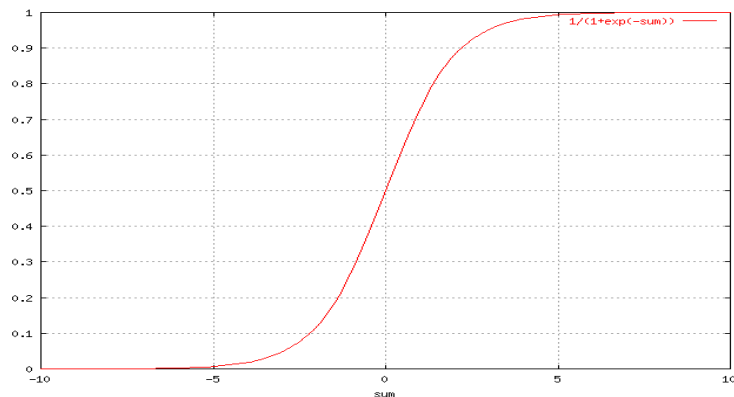
$$\text{příjem} + 0.2 \text{ konto} - 16000 = 0$$

Modely neuronu

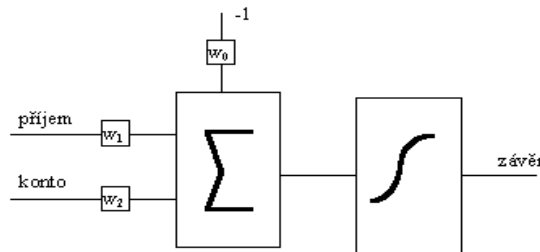
3. Současné modely

$x, w \in \mathbb{R}, y \in [0, 1]$ (nebo $[-1, 1]$)

Nelinearita v podobě skokové funkce, tak jak je použita v případě *Adaline* (dávající pouze dvě výstupní hodnoty neuronu), je ve složitějších modelech neuronu nahrazena hladkými funkcemi, které nabývají hodnot z celého intervalu (sigmoida, tang, ale taky identita).

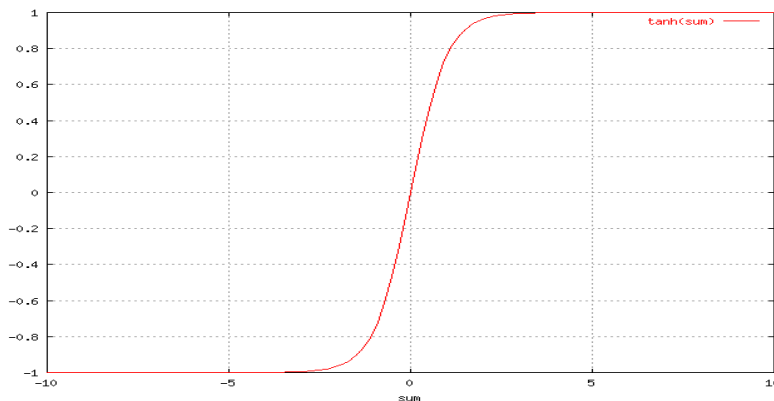


$$\text{Sigmoida } f(\text{SUM}) = \frac{1}{1 + e^{-\text{SUM}}}$$



$$\text{SUM} = \sum_{i=0}^d w_i x_i = \mathbf{w} \mathbf{x},$$

kde $\mathbf{x} = (-1, x_1, \dots, x_d)$ a $\mathbf{w} = (w_0, w_1, \dots, w_d)$



$$\text{hyperbolický tangens } f(\text{SUM}) = \tanh(\text{SUM})$$

Schopnost učení

- Modifikace vah \mathbf{w} na základě dat $[\mathbf{x}_i, y_i], i = 1, \dots, n$
- Učení jako aproximace – hledáme parametry dané funkce $f(\mathbf{x})$

Gradientní metoda

Střední kvadratická chyba

$$\text{Err}(\mathbf{w}) = \frac{1}{2} \sum_{i=1}^n (y_i - \hat{y}_i)^2$$

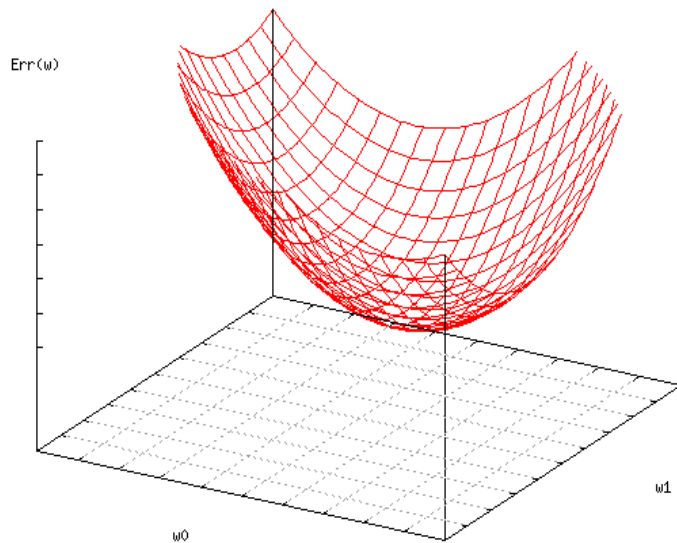
Minimalizace probíhá iterativně modifikací vah

$$\mathbf{w} \leftarrow \mathbf{w} + \Delta \mathbf{w},$$

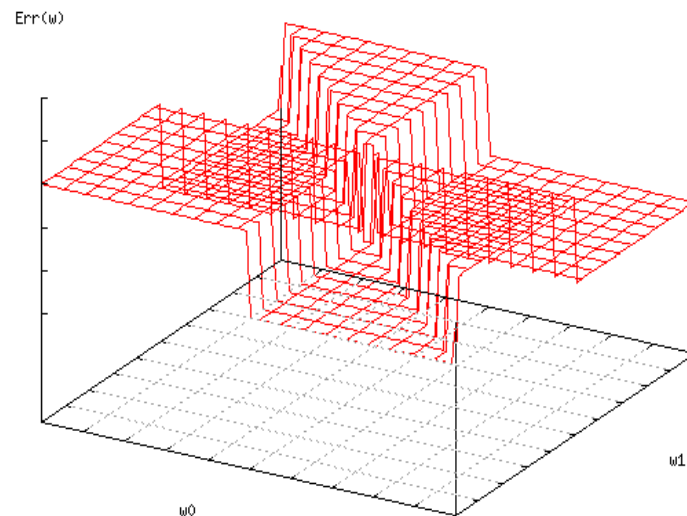
kde

$$\Delta \mathbf{w} = -\eta \frac{\partial \text{Err}}{\partial \mathbf{w}}$$

Příklady chybových funkcí



pro lineární aktivaci:
 $f(SUM) = SUM$



pro skokovou aktivaci (ADALINE):
 $f(SUM) = 1$ pro $SUM \geq 0$,
jinak $f(SUM) = 0$

Perceptron



**SLEZSKÁ
UNIVERZITA**
OBCHODNĚ PODNIKATELSKÁ
FAKULTA V KARVINĚ

Vol. VI, No. 2, Summer 1958

research trends
CORNELL AERONAUTICAL LABORATORY, INC., BUFFALO 21, NEW YORK

The Design of an
Intelligent AUTOMATON

by FRANK ROSENBLATT

Introducing the perceptron — A machine which senses, recognizes, remembers, and responds like the human mind.

STORIES about the creation of machines having human qualities have long been a fascinating province in the realm of science-fiction. Yet we are now about to witness the birth of such a machine — a machine capable of perceiving, recognizing, and identifying its surroundings without any human training or control.

Development of that machine has stemmed from a search for an understanding of the physical mechanisms which underlie human experience and intelligence. The question of the nature of these processes is at least as ancient as any other question in western science and philosophy, and, indeed, ranks as one of the greatest scientific challenges of our time.

Our understanding of this problem has gone perhaps as far as had the development of physics before Newton. We have some excellent descriptions of the phenomena to be explained, a number of interesting hypotheses, and a little detailed knowledge about events in the nervous system. But we lack agreement on any integrated set of principles by which the functioning of the nervous system can be understood.

We believe now that this ancient problem is about to yield to our theoretical investigation for three reasons:

First, in recent years our knowledge of the functioning of individual cells in the central nervous system has vastly increased.

Second, large numbers of engineers and mathematicians are, for the first time, undertaking serious study of the mathematical basis for thinking, perception, and the handling of information by the central nervous system, thus providing the hope that these problems may be within our intellectual grasp.

Third, recent developments in probability theory and in the mathematics of random processes provide new tools for the study of events in the nervous system, where only the gross statistical organization is known and the precise cell-by-cell "wiring diagram" may never be obtained.

Receives Navy Support

In July, 1957, Project PARA (Perceiving and Recognizing Automaton), an internal research program which had been in progress for over a year at Cornell Aeronautical Laboratory, received the support of the Office of Naval Research. The program had been concerned primarily with the application of probability theory to

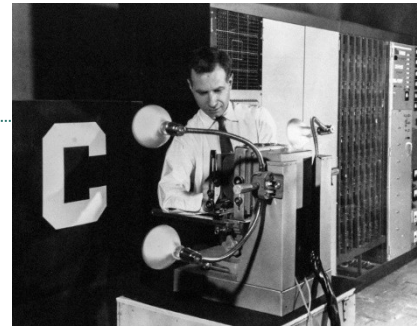
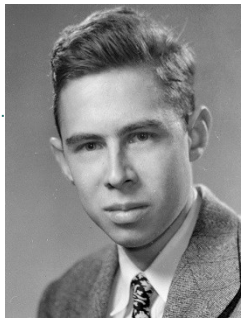


FIG. 1 — Organization of a biological brain. (Red areas indicate active cells, responding to the letter X.)

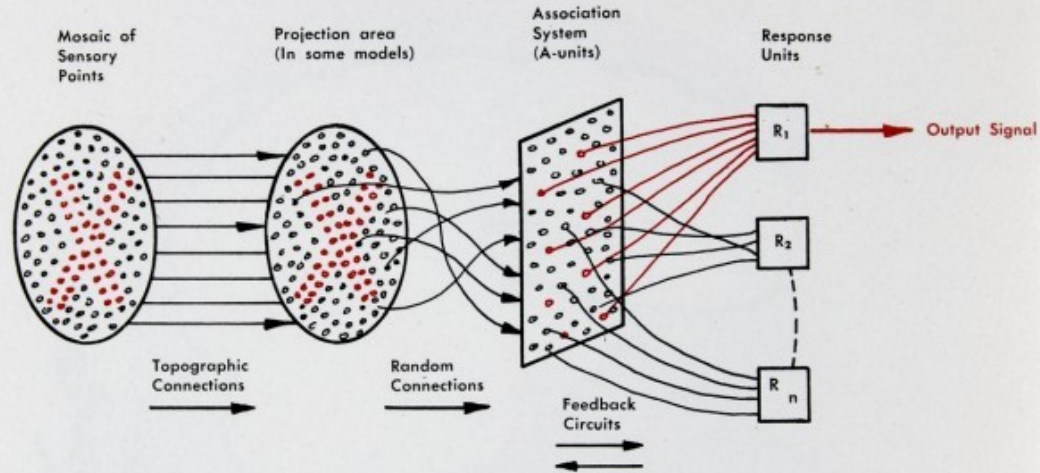


FIG. 2 — Organization of a perceptron.

Perceptron



SLEZSKÁ
UNIVERZITA
OBCHODNĚ PODNIKATELSKÁ
FAKULTA V KARVÍNĚ

Umí:

- Rozpoznat písmena z abecedy
- Rozpoznat tvary (pattern recognition)

Neumí:

- XOR problem
- Nelineárně separabilní problémy

Problém:

- Jak ho učit?

⇒ Celá oblast neuronových sítí
utihla na téměř 20 let!

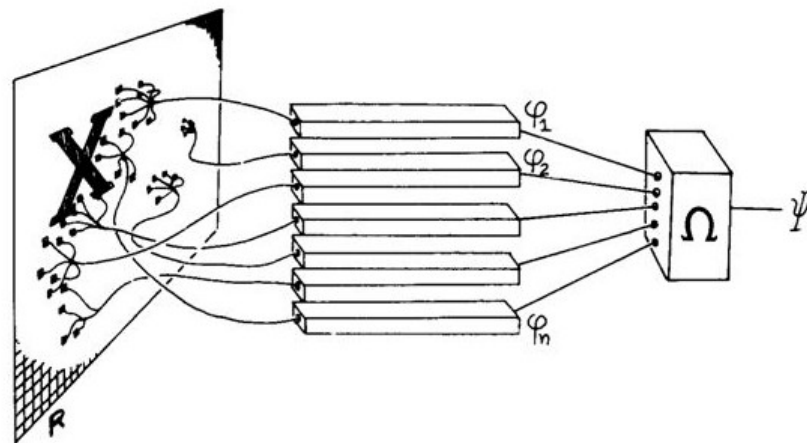


FIGURE 1. The one-layer perceptron analyzed by Minsky and Papert. (From *Perceptrons* by M. L. Minsky and S. Papert, 1969, Cambridge, MA: MIT Press. Copyright 1969 by MIT Press. Reprinted by permission.)

Vícevrstvý perceptron (MLP)



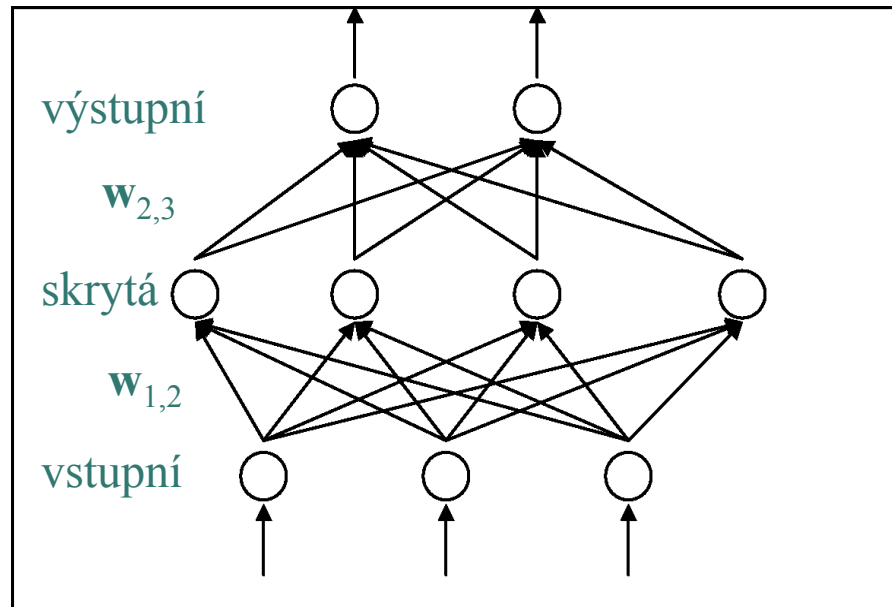
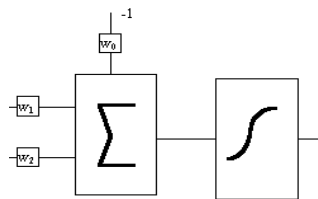
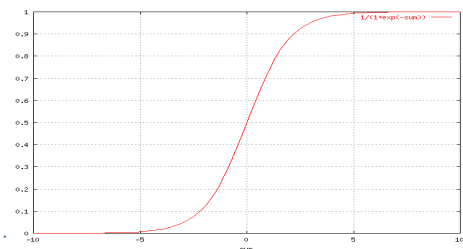
SLEZSKÁ
UNIVERZITA
OBCHODNĚ PODNIKATELSKÁ
FAKULTA V KARVÍNĚ

Síť používaná pro klasifikaci resp. predikci

3 vrstvy:

1. vstupní – přenáší vstupní data dál
2. skrytá
3. výstupní – ukazuje výsledek klasifikace resp. predikce

sigmoidální aktivační funkce pro neurony ve **skryté** a **výstupní** vrstvě

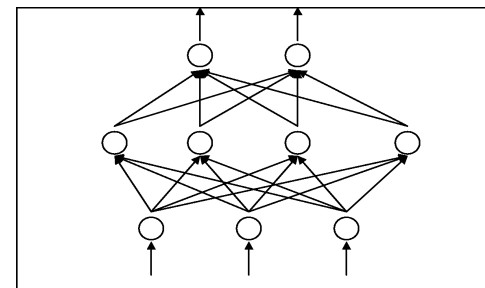


Vícevrstvý perceptron s jednou skrytou vrstvou

Backpropagation - algoritmus učení s učitelem

- cílem je minimalizovat chybu založenou na druhé mocnině rozdílu mezi skutečným ($y_{i,v}$) a očekávaným ($\hat{y}_{i,v}$) výstupem sítě pro příklad x_i

$$Err(\mathbf{w}^{(i)}) = \frac{1}{2} \sum_{v \in \text{výstupy}} (y_{i,v} - \hat{y}_{i,v})^2$$



Změna váhy vazby vedoucí od neuronu j k neuronu k se bude řídit gradientem funkce Err

$$\Delta w_{jk} = -\eta \frac{\partial Err}{\partial w_{jk}}$$

Kritéria zastavení



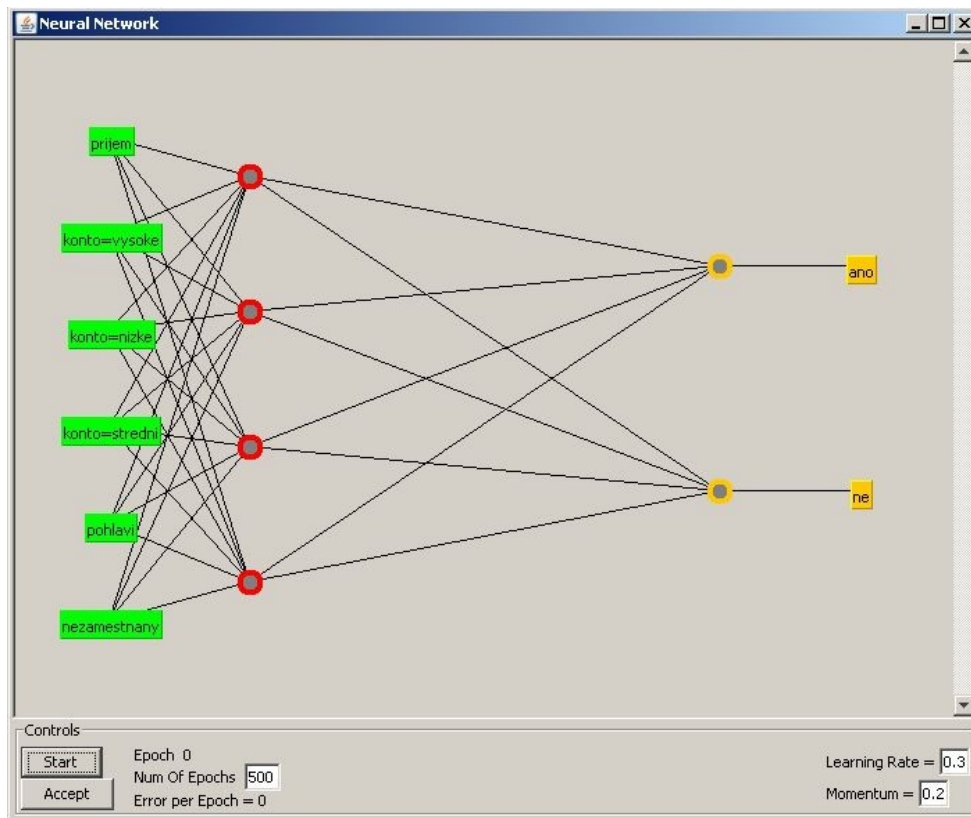
- počet iterací
- velikost chyby
- změna chyby mezi iteracemi



Implementace (Weka)



**SLEZSKÁ
UNIVERZITA**
OBCHODNĚ PODNIKATELSKÁ
FAKULTA V KARVINĚ



Způsob aplikace neuronové sítě

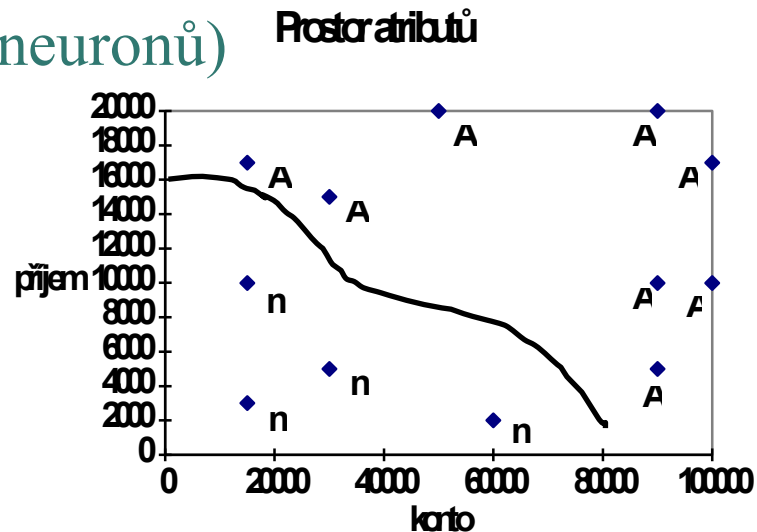


SLEZSKÁ
UNIVERZITA
OBCHODNĚ PODNIKATELSKÁ
FAKULTA V KARVÍNĚ

- volba vhodné sítě (topologie, počet neuronů)
- výběr trénovacích příkladů
- učení sítě

Typy úloh:

- klasifikace
- predikce



- neuronové sítě jsou vhodnější pro numerické atributy
- lze „popsat“ složité shluky v prostoru atributů
- nalezené znalosti jen obtížně interpretovatelné

Klasifikace



**SLEZSKÁ
UNIVERZITA**
OBCHODNĚ PODNIKATELSKÁ
FAKULTA V KARVINĚ

Úloha: vyhodnocování bonity klienta banky

Vstupy: údaje o žadateli

Výstupy: doporučení „poskytnout/neposkytnout úvěr“

Řešení:

Typ sítě: MLP

Vstupní vrstva: jeden neuron pro každý údaj o žadateli

Výstupní vrstva: jeden neuron

Skrytá vrstva: ??

Učení: informace o dřívějších klientech + závěr banky

Predikce



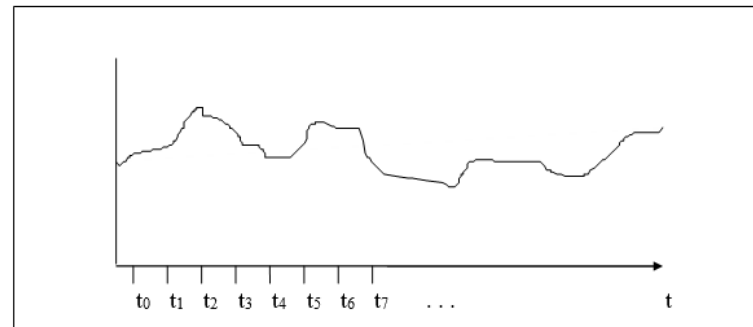
SLEZSKÁ
UNIVERZITA
OBCHODNĚ PODNIKATELSKÁ
FAKULTA V KARVINĚ

Na základě historických dat předpověď dalšího vývoje

Úloha: predikce vývoje směnných kursů

Vstupy: předchozí hodnoty kursu

Výstupy: budoucí hodnota kurzu



Řešení:

Typ sítě: MLP

Vstupní vrstva: jeden neuron pro každou hodnotu

Výstupní vrstva: jeden neuron pro každou hodnotu

Skrytá vrstva: ??

Učení: historické + aktuální údaje

vstupy				výstup	
<u>$y(t_0)$</u>	$y(t_1)$	$y(t_2)$	$y(t_3)$	$y(t_4)$	$[\text{sign}(y(t_4) - y(t_3))]$
<u>$y(t_1)$</u>	$y(t_2)$	$y(t_3)$	$y(t_4)$	$y(t_5)$	$[\text{sign}(y(t_5) - y(t_4))]$
<u>$y(t_2)$</u>	$y(t_3)$	$y(t_4)$	$y(t_5)$	$y(t_6)$	$[\text{sign}(y(t_6) - y(t_5))]$
...					

Děkuji za pozornost

Některé snímky převzaty od:
prof. Ing. Petr Berka, CSc. berka@vse.cz