UMĚLÁ INTELIGENCE

2018/2019

0. VÝVOJ, (PRE)HISTORIE, OBSAH

viz slajdy

INTRO

ANNs = matematické modely (výpočetní systémy) inspirované biologickými NN

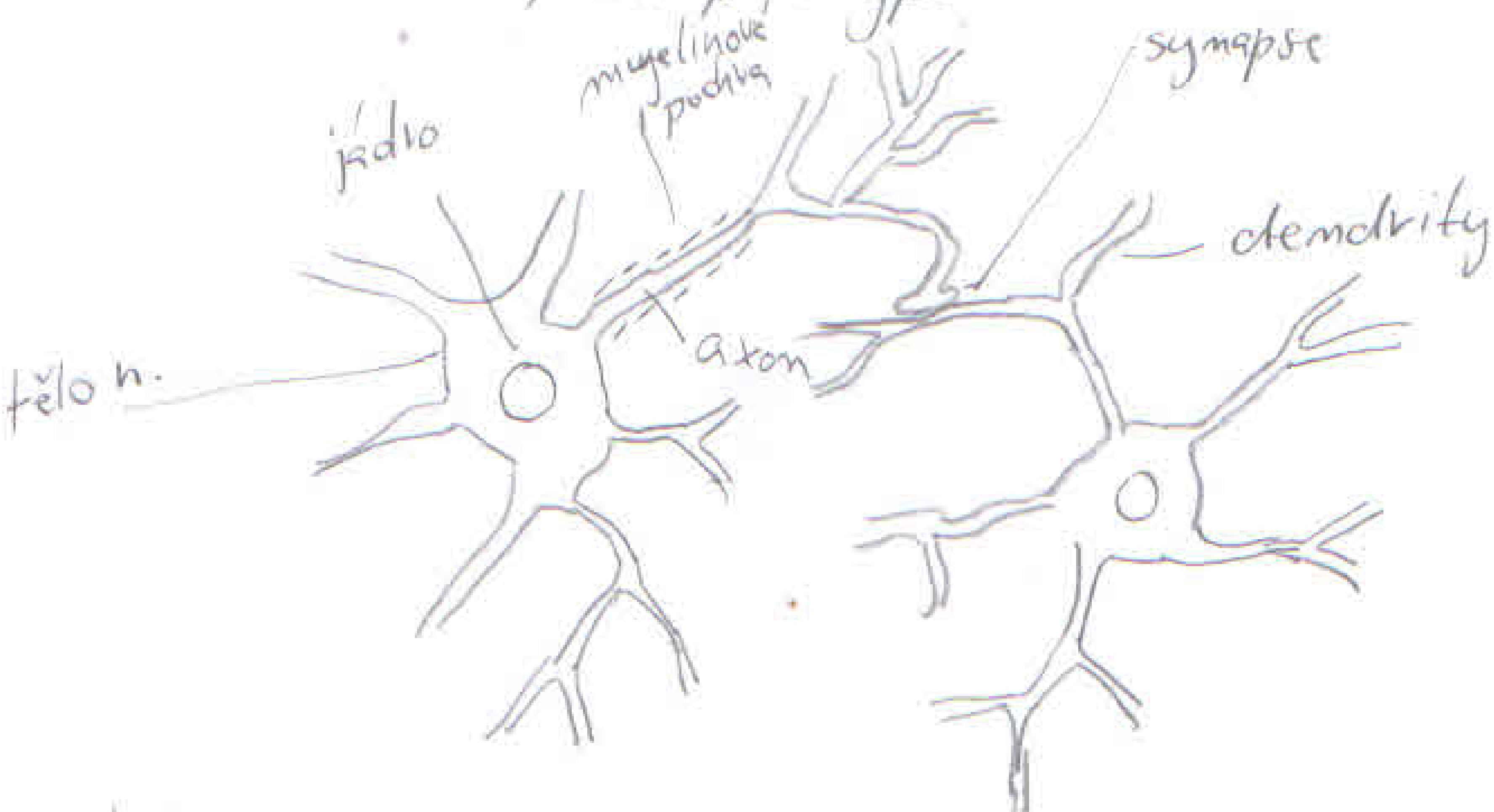
- skládají se z umělých neuronů, které si posílají signály po tzv. spojeních
- ANN jsou součástí směru v AI zvaném connectionism
- význam nejen pro AI, ale i zpětně pro neurovědy; některé fyzikální modely mají podobnou strukturu
- zaměříme se na aspekty pro AI
- alternativa: physical symbol system hypothesis (Newell & Simon 1976):
  A physical symbol system has the necessary and sufficient means for general intelligent action.
  (AI logic-based, symbolic vs. ANNs sub-symbolic)
- ANNs are not easily interpretable - disadvantage
  ↓
  knowledge is encoded in many neurons and connections
  ↓
  black box effect
- ANNs outperform symbolic AI systems in many tasks:
  - pattern recognition, prediction
  - easy-to-develop tool, many existing applications

BIOLOGICAL NNs - SOURCE OF INSPIRATION

NNs - jsou součástí centrální nervové soustavy : = mozek + senzorické systémy (vizuální, sluchový, teplota, hmat) + motorický systém

- skládají se z neuronů navzájem propojených (další části pomocné buňky)
- velký počet, odhad 10^{11} neuronů v lidském mozku
- pracují paralelně (tj. masivní paralelismus)

<u>neuron</u> - buňka, která přijímá a vysílá el. signály
- mnoho druhů, ale prototyp:



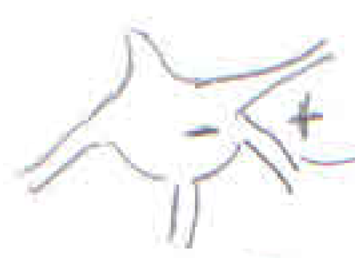
tělo ≈ 5-100 μm

dendrity = krátká spojení, axon = dlouhé spojení mm až 1 m

↓ signál do neuronu; ↓ signál z neuronu do dendritů ostatních neuronů

<u>synapse</u> = místo připojení axonu na dendrit (šířka 10-50 nm)

přenos signálu: - axon v synapsi uvolňuje chem. látky, tzv. neurotransmitery
- neurotransmit. mění el. potenciál dendritu



rozdíl ≈ 70 mV

- ... převládají negativní ionty
+ ... -//- kladné -//-

synapse zmenšující rozdíl el. potenciálu dendritu → <u>excitační</u>
zvětšující -//- → <u>inhibiční</u>

v lidském mozku: 1 neuron má 1000 - 10 000 spojení k dalším neuronům

šíření signálu
excitační a inhibiční efekty synapsí (pro daný) neuron se akumulují ("sčítají")
při dostatečně vysoké excitaci (počtes rozdílu potenciálu uvnitř a vně neuronu), dochází až k depolarizaci počátku axonu tím, že kladné ionty sodíku (Na^+) vstupují do neuronu;
na krátký okamžik (1 ms) má vnitřek neuronu více kladných náboj než vnějšek; vede k uvolnění iontů draslíku (K^+) a dochází k obnovení rozdílu potenciálu uvnitř a vně neuronu (pomocí iontových pump v membráně mezi těla neuronu)
zmíněná náhlá změna v potenciálu se šíří podél axonu rychlostí 0.5 – 130 m/s → "šíření signálu" mezi neurony
(rychlost závisí mj. na tloušťce myelinové pochvy axonu)
signál dojde na konec axonu, způsobí uvolnění neurotransmitterů a dojde k přenosu do dendritů dalšího neuronu atd.

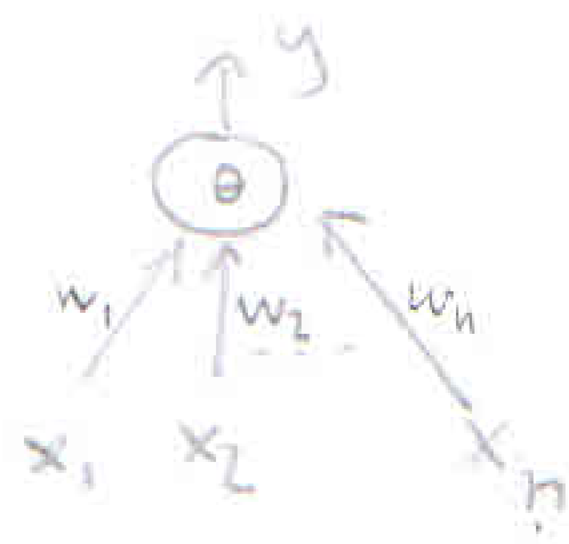
tedy schematicky:
signály (změny v el. potenciálu) se v těle neuronu akumulují, a pokud tato dojde k překročení jisté prahové hodnoty, signál se začne šířit dál po axonu k synapsím a dendritům dalších neuronů

firing rate = počet vyslaných signálů/sec., v lidském mozku ≈ i 100 a více

→ z této představy vychází ANNs

Perceptron = jednoduchý model neuronu (McCulloch & Pitts 1943)
také nazýván McCulloch-Pitts neuron

model :



$y \in \{0,1\}$

$w_i \in \mathbb{R}$

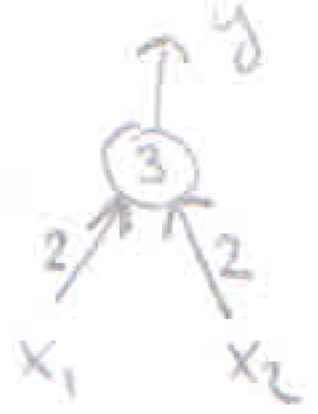
$x_i \in \mathbb{R}$

$$y = \begin{cases} 1 & \text{pokud } \sum w_i x_i \geq \theta \\ 0 & \text{jinak} \end{cases}$$

váha (≈ vlastnosti synapse
$w_i > 0$ ... excitace
$w_i < 0$ ... inhibice)

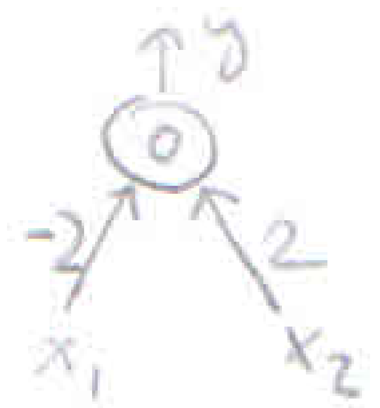
vstupní signál
(z okolí, např. okolních
neuronů)

Př.: - perceptron modelující konjunkci



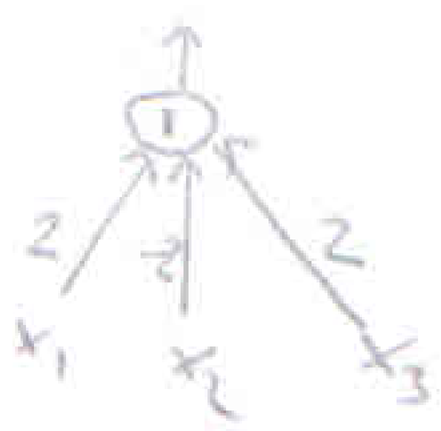
pro $x_1, x_2 \in \{0,1\}$ je $y = x_1 \wedge x_2$

- implikaci



$y = x_1 \rightarrow x_2$

- uvaz.



pravd. tabulka:

x_1	x_2	x_3	$\sum w_i x_i$	y
0	0	0	0	0
0	0	1	2	1
0	1	0	-2	0
0	1	1	0	0
1	0	0	2	1
1	0	1	4	1
1	1	0	0	0
1	1	1	2	1

$$y = (x_1 \wedge \neg x_2) \vee (x_1 \wedge x_3) \vee (\bar{x}_2 \wedge x_3)$$

Učení perceptronu

idea: dána trénovací množina $T = \{\langle x_p, y_p \rangle \mid p \in P\}$

kde $x_p \in \mathbb{R}^n$, $y_p \in \{0,1\}$
$x_p = \langle x_{p1}, \dots, x_{pn} \rangle$

Cílem je najít perceptron, který bude realizovat fci y takže $\forall p \in P$ je $y(x_p) = y_p$;

Pozn. - Perceptron se má chovat tak, jak říká/předepisuje T.
- Chceme "naučit" perceptron tomu, co říká trénovací množina.
- Chování perceptronu (tj. funkce realizovaná perceptronem) je dáno jeho parametry $w_1, \dots, w_n, \Theta$. Cílem je tedy najít tyto parametry.
- Hledání parametrů = učení (adaptace)
  algoritmus, který parametry hledá = učící algoritmus (adaptační)
- učení zpravidla iterativní proces

2 základní způsoby učení (úpravy parametrů)
- online learning - podíváme se na vzor $\langle x_p, y_p \rangle$, upravíme váhy, atd.
- batch learning - podíváme se na více vzorů, pak upravíme váhy

u obou způsobů provádíme iterativně, dokud nesplníme ukončující podmínku, např. celá T byla naučena (popř. stokrát celé T)

Základem pro učení perceptronu je tzv. delta rule (Widrow & Hoff 1960, na základě biol. vzoru), nebo také Widrow-Hoff rule:

Je-li vzor $\langle x_p, y_p \rangle$, tj $x_p = \langle x_{p1}, \dots, x_{pn} \rangle \in \mathbb{R}^n$ vstup, y_p požadovaný výstup, a $y \in \{0,1\}$ skutečný výstup perceptronu upraví Θ a w následovně:

$\Theta^{(t+1)} := \Theta^{(t)} + \Delta\Theta$, kde $\Delta\Theta = -\eta(y_p - y)$ — $\Theta^{(t)}$ je Θ v čase t, ...

$w_i^{(t+1)} := w_i^{(t)} + \Delta w_i$, kde $\Delta w_i = \eta(y_p - y)x_i$ — $\eta \in \mathbb{R}$ ... learning rate (rychlost učení) delta rule, přiblížení zdůvodnění

Pozn.: S Θ lze zacházet jako s vahou w_0 pro konstantní vstup $x_0 = -1$:

$$\sum_{i=1}^{n} w_i x_i \geq \Theta \iff \sum_{i=1}^{n} w_i x_i - \Theta \geq 0 \iff \sum_{i=0}^{n} w_i x_i \geq 0$$

Pak přejde pravidlo pro adaptaci Θ v pravidlo pro adaptaci w_0.

Algoritmus učení perceptronu (online)

T... trén. množina
$\mathbf{w}, \Theta$... poč. hodnoty vah a prahu, např. náhodné nastavené
η... learning rate (např. $=1$)

```
Perceptron-online-training(T, w, Θ, η)
  repeat
    e ← 0
    for each p ∈ P do
      if Σ_{i=1}^{n} w_i x_pi ≥ Θ then y=1
                                  else y=0
      if y ≠ y_p then
        Θ ← Θ - η(y_p - y)
        for i=1 to n  w_i ← w_i + η(y_p - y) x_pi
        e ← e + |y_p - y|
  until e = 0
```

Algoritmus učení perceptronu (batch)

analogicky, změny w_i a Θ se provádí po průchodu celou T, kumulují se

Pozn.: Pro $x_i = 0$ nedochází k adaptaci; řešením je $x_i = -1$ místo $x_i = 0$; učení probíhá stejně, ale často je rychlejší. Původně použito v systému ADALINE.

Věta (konvergence algoritmu učení perceptronu)

Nechť $T, w, \Theta, n > 0$ jsou libovolné vstupy. Jestliže existují $v_1, \dots, v_n, \varphi$ t.ž. $\forall p \in P$ je $\sum_{i=1}^{n} v_i x_{pi} \geq \varphi$ pro $y_p = 1$ a $\sum_{i=1}^{n} v_i x_{pi} < \varphi$ pro $y_p = 0$ (tj. T je naučitelná), pak algoritmus učení, online i batch, skončí (a tedy naučí se množinu T).

Důkaz:

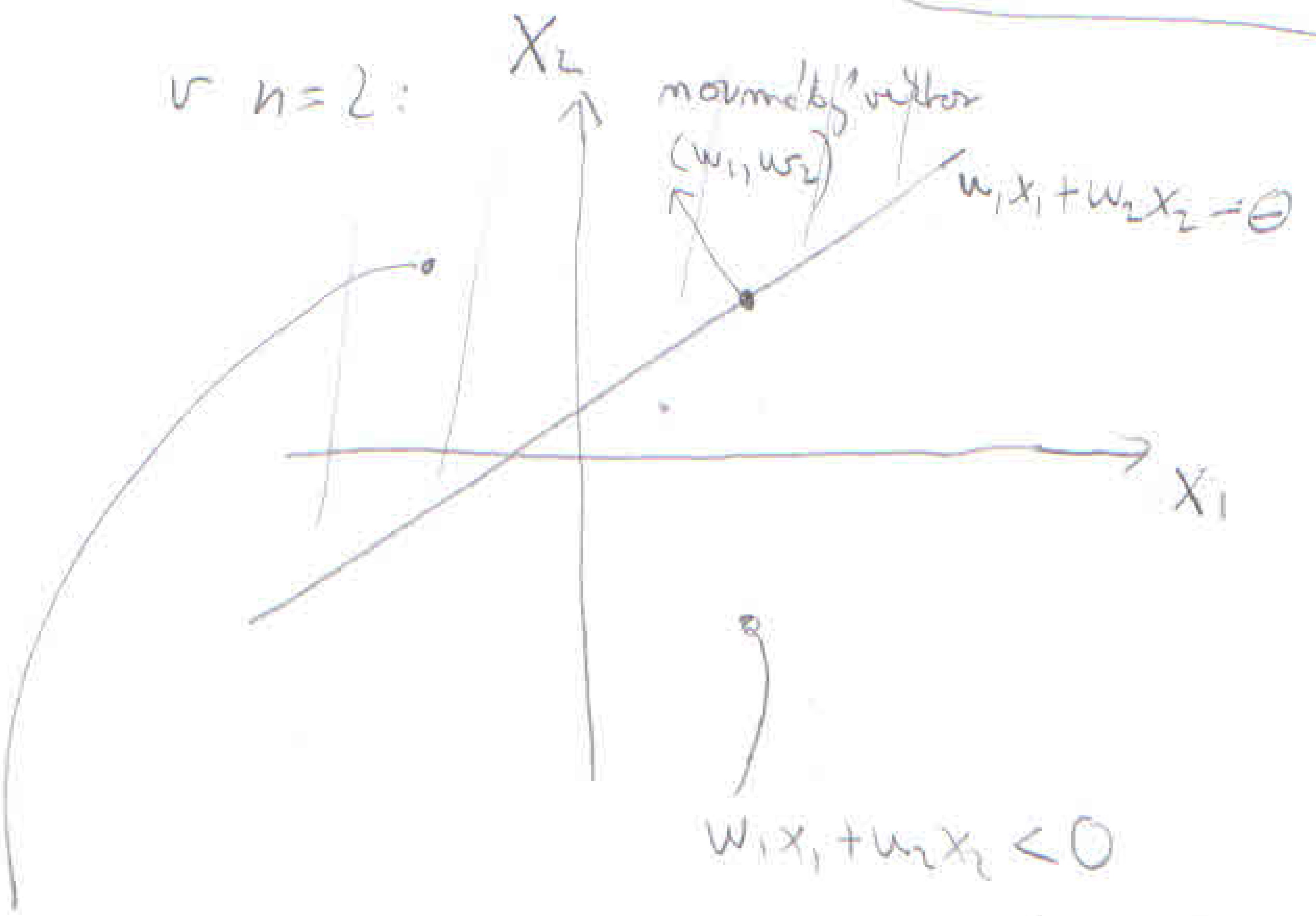
Meze perceptronu

umožňuje naučit se pouze lineárně separabilní množiny.

Geometrický pohled na $\sum_{i=1}^{n} w_i x_i \geq \Theta$ neboli

$$\underbrace{\overbrace{w_1 x_1 + \dots + w_n x_n - \Theta}^{\text{nadrovina}} \geq 0}_{\text{poloprostor}}$$

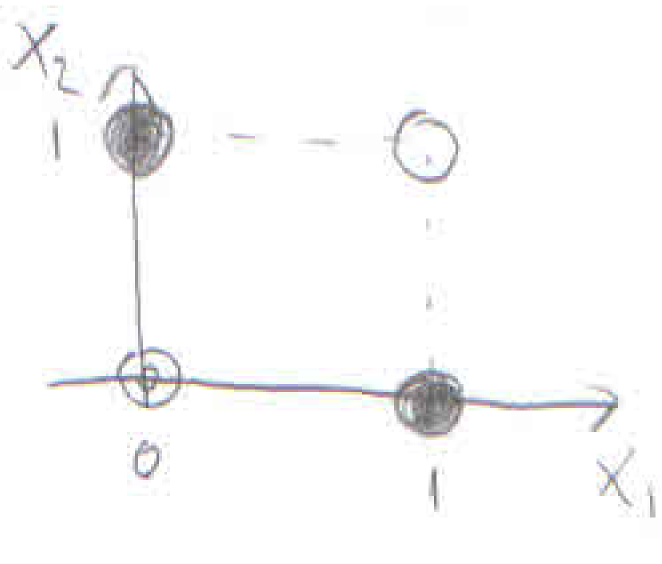
v $n=2$:



Pokud T není lineárně separabilní, nelze perceptron naučit.
Pokud T je — // ——, lze — // —— (+ teorém o konvergenci zaručuje, že se to podaří).

Př.: Funkce XOR

	0	1
0	0	1
1	1	0



není lin. separabilní

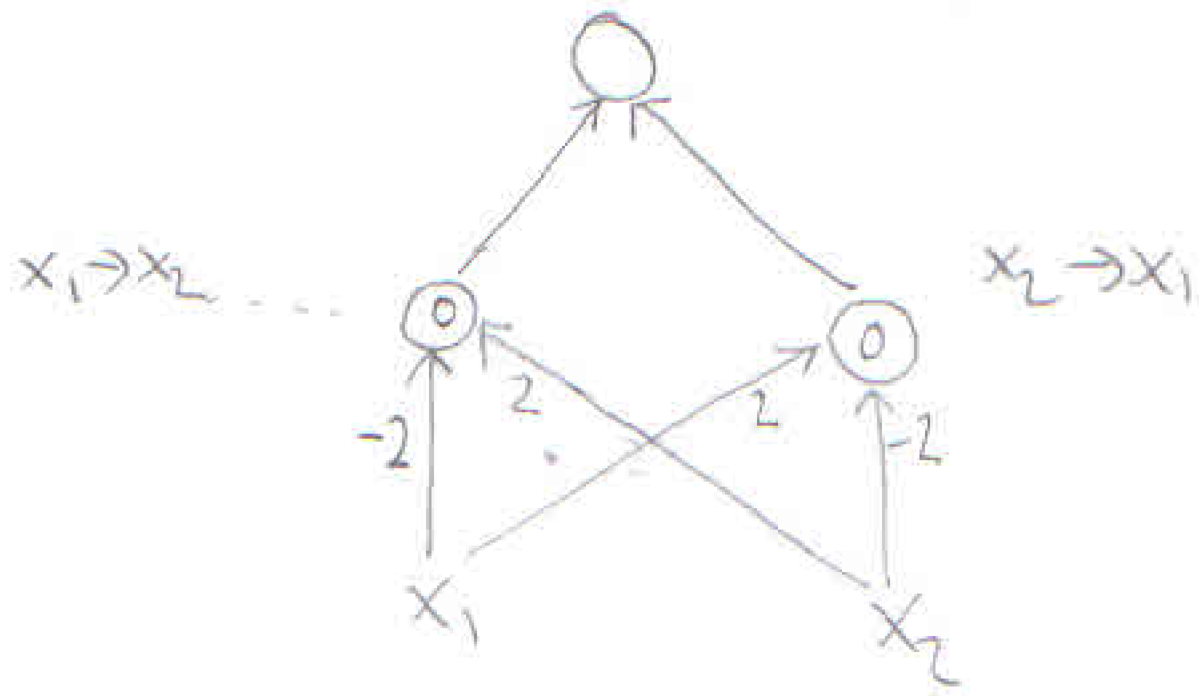
je lin. separ.

+ Kruse str 20 tabulka 3.1

Sítě perceptronů jako řešení problému lin. neseparability

- výstupy neuronů přivedeny na vstupy dalších neuronů

Př.:



Pozn. Sítěmi z perceptronů lze reprezentovat i složité funkce.
ROZVÉST

Problém: Zatímco pro jednoduchý perceptron existuje jednoduchý učící algoritmus se zaručenou konvergencí, pro sítě perceptronů pravidlo (učící algoritmus) nebylo objeveno.

Minsky & Papert (1969): Perceptrons: An introduction to computational geometry
kniha zastavila vývoj v neuronových sítích na 20 let