Neuronové sítě

Vlasta Radová
Západočeská univerzita v Plzni
katedra kybernetiky

Motivace pro výzkum umělých neuronových sítí

- lidský mozek pracuje jiným způsobem než běžné číslicové počítače
- počítače přesně a rychle provádějí posloupnosti instrukcí, které pro ně byly formulovány
- člověk dokáže lépe řešit řadu výpočetně náročných úkolů (šachy, porozumění řeči, zpracování vizuální informace…)

jaké jsou ?

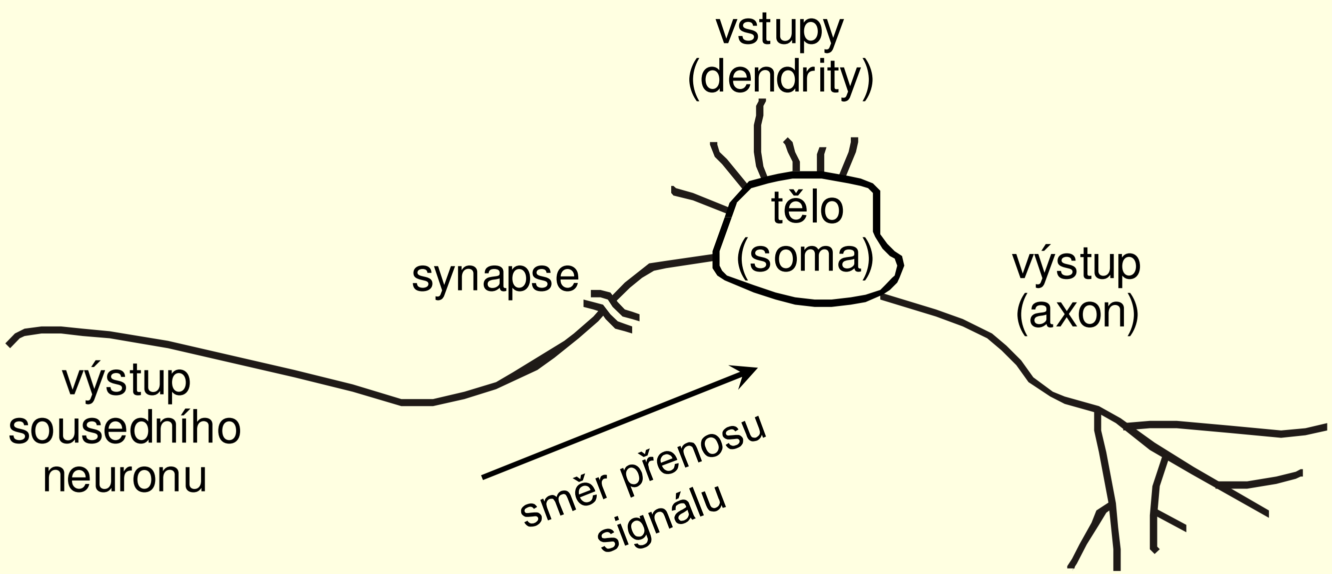
snaha napodobit schopnosti mozku a vytvořit umělou neuronovou síť

Lidský mozek

- skládá se asi ze 100 miliard výpočetních elementů - **neuronů**
- neurony mezi sebou komunikují prostřednictvím sítě vazeb
- vstup do sítě zajišťují tzv. **receptory** (zrak, sluch, čich, chuť, hmat)
- podnět z receptorů je ve formě elektrických impulsů přenášen do centrálního nervového systému, kde je zpracován
- podle výsledku zpracování jsou řízeny tzv. **efektory** (tj. výkonné orgány) a odezva člověka na daný podnět se projeví ve formě rozmanitých akcí

Biologický neuron

- biologický neuron je základní buňkou biologických neuronových sítí



Biologický neuron (pokračování)

- části biologického neuronu
  - tělo (soma)
  - výstup (axon) – jediný, poměrně dlouhý (až 60 cm)
  - vstupy (dendrity) – poměrně krátké (do 3 mm), je jich až několik tisíc
  - rozhraní (synapse) – jednosměrné brány umožňující přenos signálu pouze ve směru axon → dendrit, je jich asi 10 tisíc na každý neuron

Biologický neuron (pokračování)

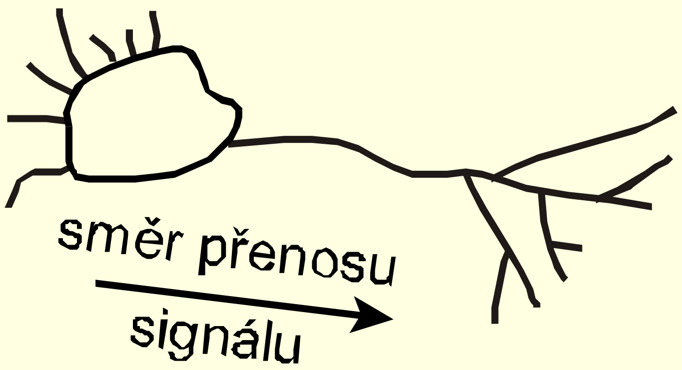
- přenášené signály jsou elektrické impulsy, jejich přenos je ovlivněn uvolňováním budicích a tlumicích látek v synapsích
  - **excitátory** (budicí látky) umožňují, aby neuron generoval impuls
  - **inhibitory** (tlumicí látky) snižují schopnost neuronu generovat impuls
- překročí-li hodnota budicích signálů hodnotu tlumicích signálů o určitý **práh**, nastává tzv. **aktivace neuronu** – na výstupu neuronu se objeví impuls

Biologický neuron (pokračování)

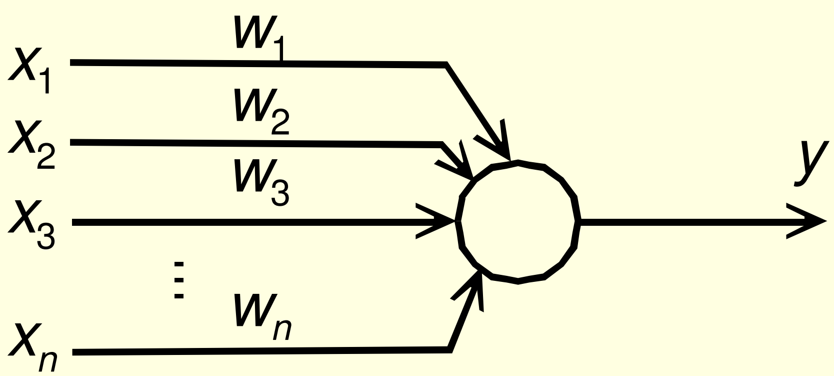
- po vygenerování impulsu se neuron na určitou dobu (tzv. **období pauzy**) dostane do stavu, kdy nereaguje na žádné podněty → chování neuronu lze popsat diskrétně v čase
- období pauzy není pro všechny neurony stejně dlouhé → neurony v mozku pracují asynchronně

Model biologického neuronu

biologický neuron



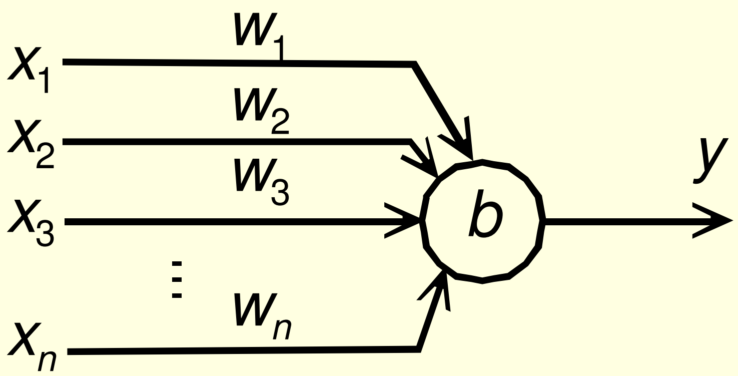
model neuronu



x_i … vstupy
w_i … váhy
y … výstup
n … počet vstupů neuronu

Model neuronu McCullochův – Pittsův model

- navržen v roce 1943



pro výstup platí:

$$y(k+1) = \begin{cases} 1, & \sum_{i=1}^{n} w_i\, x_i(k) \geq b \\ 0, & \sum_{i=1}^{n} w_i\, x_i(k) < b \end{cases}$$

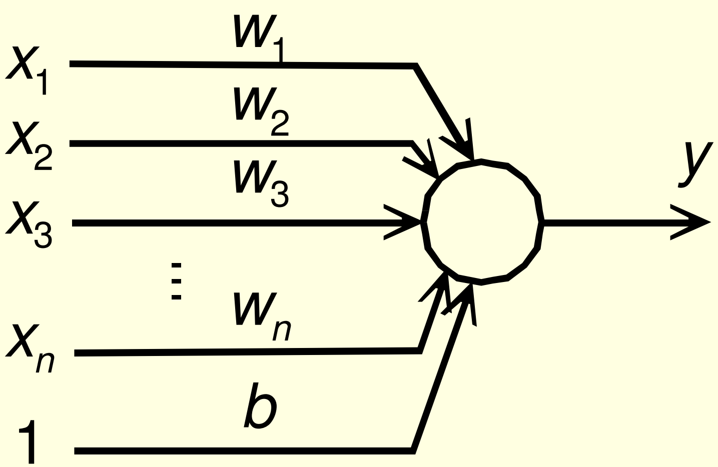
- vstupy x_i, $i = 1,\ldots,n$, mohou nabývat pouze hodnot 0 nebo 1 podle toho, zda je přítomen signál nebo ne
- váhy w_i, $i = 1,\ldots,n$, $w_i = +1$ pro budicí signály (excitátory), $w_i = -1$ pro tlumicí signály (inhibitory)
- b je práh

Vlastnosti McCullochova – Pittsova modelu

- jednoduchý
- má pouze 2 výstupy (stavy), a to 0 a 1 – jedná se o binární model
- váhy jsou pevně nastavené pouze na -1 nebo +1
- prahy jsou pevně stanoveny
- neumožňuje realizovat všechny logické funkce (umožňuje AND, OR, NOT, ale neumožňuje XOR)

Model neuronu – perceptron

- diskrétní verze poprvé uvedena v roce 1958 F. Rosenblatem



pro výstup platí:

$$y(k+1) = f\left(\sum_{i=1}^{n} w_i x_i(k) + b\right) =$$

$$= f(\boldsymbol{w}^{\mathrm{T}}\boldsymbol{x}(k) + b)$$

f je aktivační funkce

$\boldsymbol{w} = [w_1, w_2, \ldots, w_n]^{\mathrm{T}}$ je váhový vektor

$\boldsymbol{x} = [x_1, x_2, \ldots, x_n]^{\mathrm{T}}$ je vstupní vektor

b je práh

$\xi = \boldsymbol{w}^{\mathrm{T}}\boldsymbol{x} + b$ je aktivační hodnota

Model neuronu – perceptron

Práh *b* se někdy přidává jako (n+1). složka váhového vektoru. Pak je nutné přidat 1 jako (n+1). složku vstupního vektoru. Potom tedy

$\tilde{\boldsymbol{w}} = [w_1, w_2, \ldots, w_n, b]^{\mathrm{T}}$ je váhový vektor

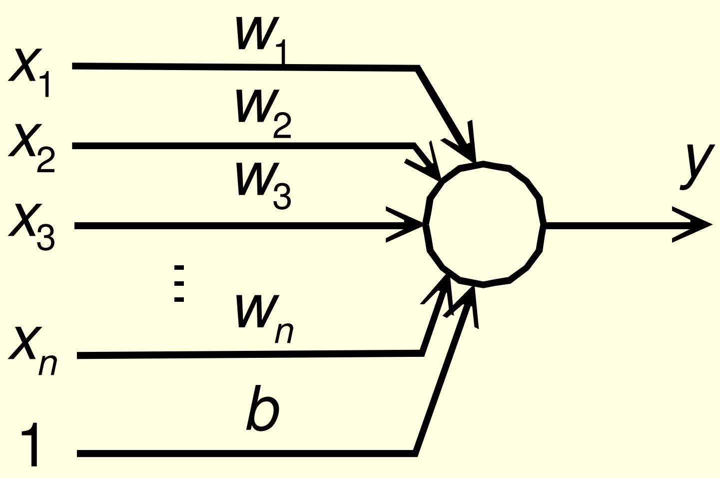
$\tilde{\boldsymbol{x}} = [x_1, x_2, \ldots, x_n, 1]^{\mathrm{T}}$ je vstupní vektor

$\xi = \tilde{\boldsymbol{w}}^{\mathrm{T}} \tilde{\boldsymbol{x}}$ je aktivační hodnota

Pro výstup perceptronu pak platí:

$$y(k+1) = f\left(\sum_{i=1}^{n+1} \tilde{w}_i \tilde{x}_i(k)\right) =$$

$$= f\left(\tilde{\boldsymbol{w}}^{\mathrm{T}} \tilde{\boldsymbol{x}}(k)\right)$$

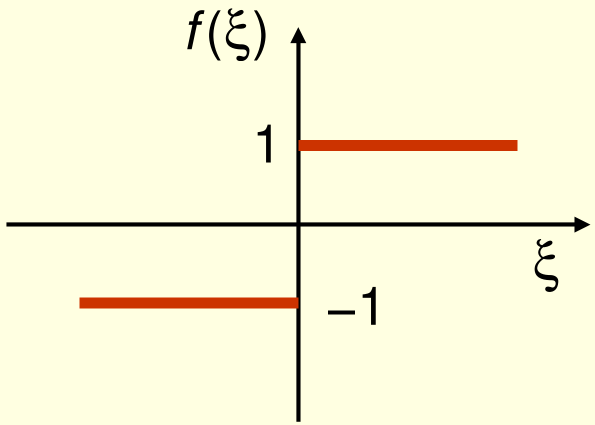


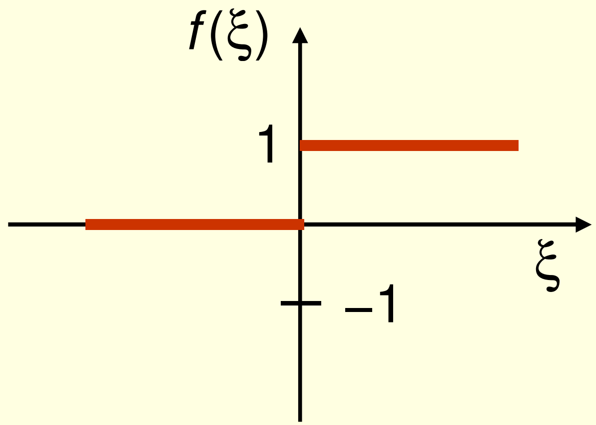
Model neuronu – perceptron (pokr.)

- Nejčastěji užívané aktivační funkce
  - bipolární binární
  - unipolární binární

$$f(\xi) = \mathrm{sgn}(\xi) = \begin{cases} +1 & \text{pro } \xi \geq 0 \\ -1 & \text{pro } \xi < 0 \end{cases}$$

$$f(\xi) = \begin{cases} +1 & \text{pro } \xi \geq 0 \\ 0 & \text{pro } \xi < 0 \end{cases}$$



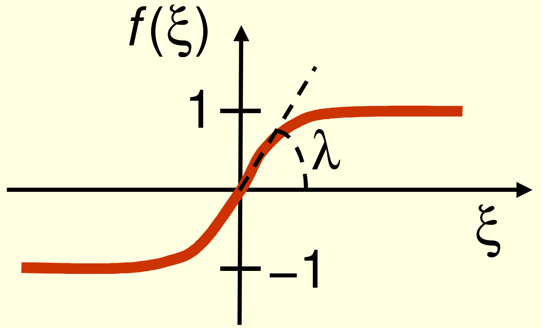


Model neuronu – perceptron (pokr.)

- Nejčastěji užívané aktivační funkce

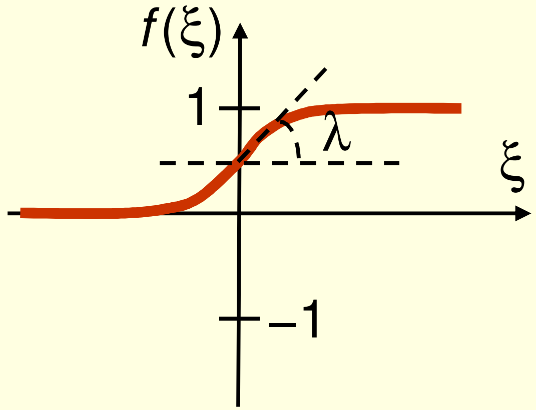
- bipolární spojitá

$$f(\xi) = \frac{2}{1 + \exp(-\lambda\,\xi)} - 1$$



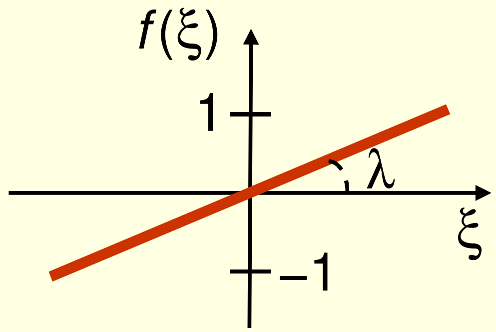
- unipolární spojitá

$$f(\xi) = \frac{1}{1 + \exp(-\lambda\,\xi)}$$



- lineární

$$f(\xi) = \lambda\,\xi$$

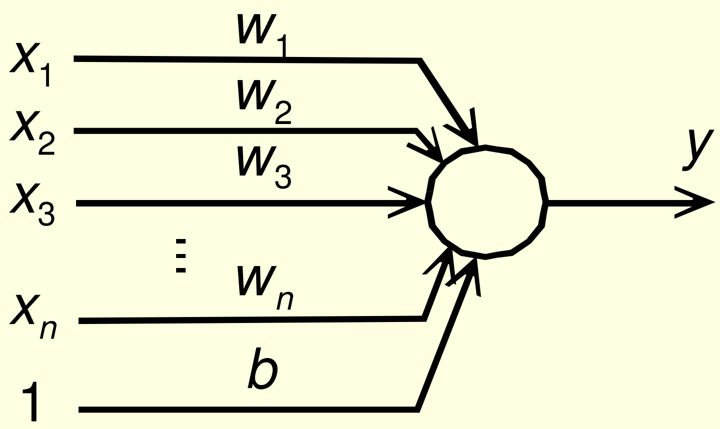


λ … strmost aktivační funkce

Model neuronu – perceptron (pokr.)

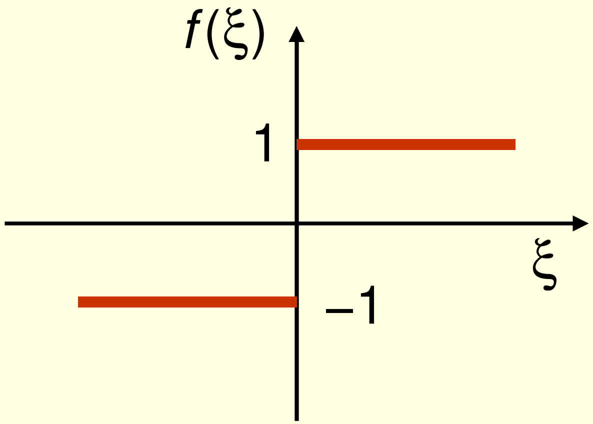
- binární aktivační funkce jsou speciálním případem spojitých pro $\lambda \rightarrow \infty$
- bipolární spojitá a unipolární spojitá funkce se též nazývají **sigmoidální funkce** (mají tvar písmene S)

Činnost perceptronu s bipolární binární aktivační funkcí



$$y(k+1) = f\left(\sum_{i=1}^{n} w_i x_i(k) + b\right) =$$

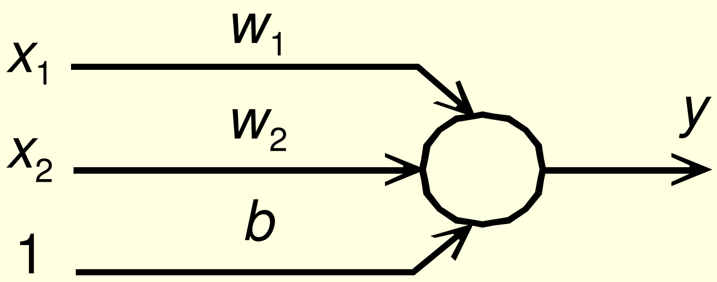
$$= f\left(\boldsymbol{w}^{\mathrm{T}}\boldsymbol{x}(k) + b\right)$$



$$f(\xi) = \mathrm{sgn}(\xi) = \begin{cases} +1 & \text{pro } \xi \geq 0 \\ -1 & \text{pro } \xi < 0 \end{cases}$$

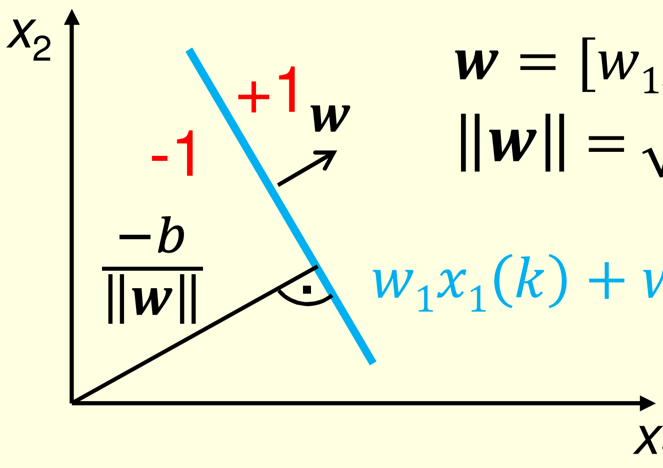
Činnost perceptronu s bipolární binární aktivační funkcí (pokr.)

pro n=2



$$y(k+1) = \operatorname{sgn}\left(w_1 x_1(k) + w_2 x_2(k) + b\right) =$$

$$= \begin{cases} +1, & w_1 x_1(k) + w_2 x_2(k) + b \geq 0 \\ -1, & w_1 x_1(k) + w_2 x_2(k) + b < 0 \end{cases}$$



$$\boldsymbol{w} = [w_1, w_2]^{\mathrm{T}}$$

$$\|\boldsymbol{w}\| = \sqrt{(w_1)^2 + (w_2)^2}$$

$$w_1 x_1(k) + w_2 x_2(k) + b = 0$$

perceptron rozdělí vstupní rovinu x_1 x_2 na dvě poloroviny, jedné přiřadí +1, druhé -1

Činnost perceptronu s bipolární binární aktivační funkcí (pokr.)

- Perceptron s bipolární binární aktivační funkcí rozdělí vstupní prostor $\boldsymbol{x} = [x_1, x_2, \ldots, x_n]^{\mathrm{T}}$ na dvě oblasti, jedné přiřadí +1, druhé -1.
- Nadplocha, která vstupní prostor rozděluje, je dána rovnicí $\boldsymbol{w}^{\mathrm{T}}\boldsymbol{x} + b = 0$, resp. $\widetilde{\boldsymbol{w}}^{\mathrm{T}}\widetilde{\boldsymbol{x}} = 0$.