

## Vícevrstvé síť typu Perceptron

V této dvojevrstvé síti (neobsahuje skrytou vrstvu) tvoří základ model neuronu s lineárně váženou obvodovou funkcí a aktivační funkcí, kterou je **skoková funkce** (*threshold function*). Jedná se o McCulloch-Pittsův model. Vstup je tvořen vektorem vzorků  $\underline{x} = (x_1, x_2, \dots, x_n)$  a vektor požadovaných (cílových) hodnot  $\underline{c} = (c_1, c_2, \dots, c_n)$  v  $n$ -dimensionálním prostoru. Jedná se o učení s učitelem. Perceptrony se používají pro řešení jednoduchých klasifikačních úloh. Vstupní data musí být **lineárně separabilní**. Pouze jedna vrstva je trénována, a to pomocí **perceptronového trénovacího (učicího) algoritmu** (*Perceptron learning rule*). Tento algoritmus patří mezi chybová učení.

### Učící algoritmus Perceptronu.

- Pro obvodovou funkci  $u(\underline{x})$  a výstupní funkci platí (v jedné iteraci)

$$u(\underline{x}) = w_0 + \sum_{i=1}^n w_i x_i \quad (3.1)$$

$$y(\underline{x}) = \begin{cases} 1 & u(\underline{x}) \geq 0 \\ 0 & u(\underline{x}) < 0 \end{cases} \quad (3.2)$$

- Výstup je porovnán s cílovou hodnotou a je spočítána chyba  $e$ :

$$e = c - y \quad (3.3)$$

je-li  $y = c$  a chyba  $e = 0$ , váhy se nezmění; je-li  $y = 0$  a  $c = 1$ , chyba  $e = 1$  a k vahám se přičte vstupní hodnota, váhový vektor se přiblíží ke vstupnímu vektoru a v následující iteraci se zvýší pravděpodobnost správné klasifikace 1; je-li  $y = 1$  a  $c = 0$ , chyba  $e = -1$  a od vah se odečte vstupní hodnota, váhový vektor se vzdálí od vstupního vektoru a v následující iteraci se zvýší pravděpodobnost chybné klasifikace 0.

- Pro adaptaci vah a prahů platí rovnice

$$w_{ij}(t+1) = w_{ij}(t) + e_j x_i \quad (3.4)$$

$$b_j(t+1) = b_j + e_j \quad (3.5)$$

kde  $t$  označuje čas.

Z rovnic (3.3) a (3.4) plyne, že velikost chyby je závislá na velikosti vstupního vektoru. Jsou-li některé ze vstupních vektorů přiváděných do Perceptronu mnohem větší než ostatní, trvá učení déle (menší vektory musí být přiváděny mnohokrát, aby byl dosažen stejný efekt). Pro nerovnoměrné velikosti vstupních vektorů je vhodné je normalizovat.

Pomocí jednoho Perceptronu je možné klasifikovat pouze do dvou tříd. Pro klasifikaci do více tříd je třeba použít více Perceptronů nebo vícevrstvou síť s učením BPG (viz podkapitola 3.3).