



# A neurális hálózatok tanításának alapjai I.: Felépítés, működés, tanulási ráta, torzítás

Szoldán Péter

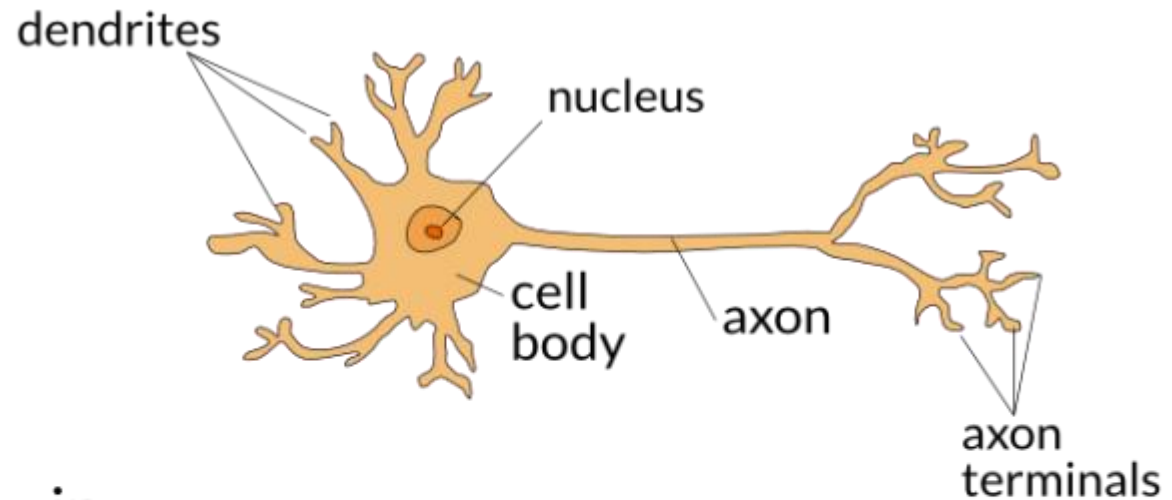
# Adminisztratív információ

- Az előadásokon látott prezentációk a következő weboldalon érhetők el:

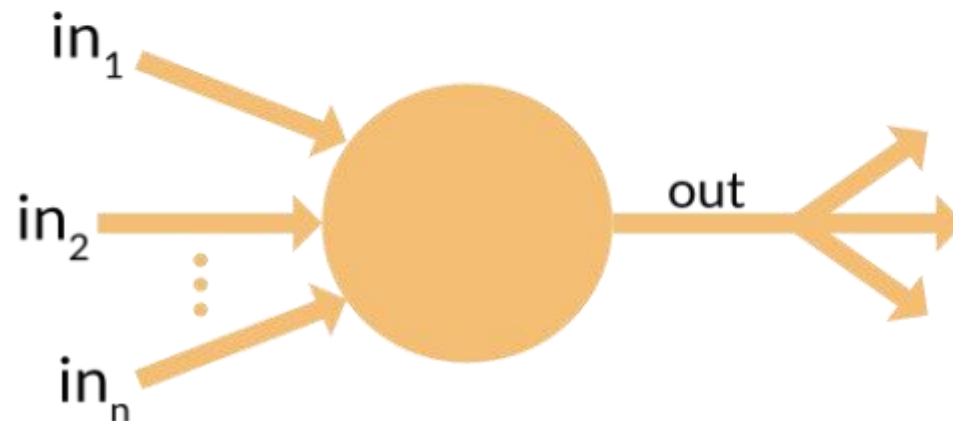
<http://semmelweis.hu/radiologia/oktatas/eloadas-kivonatok/>

# A virtuális neuron

Biológiai  
neuron:



Virtuális  
neuron:



A bemenetek súlyozott összegéből áll a kimenet

# Akciós potenciál, tüzelési ráta

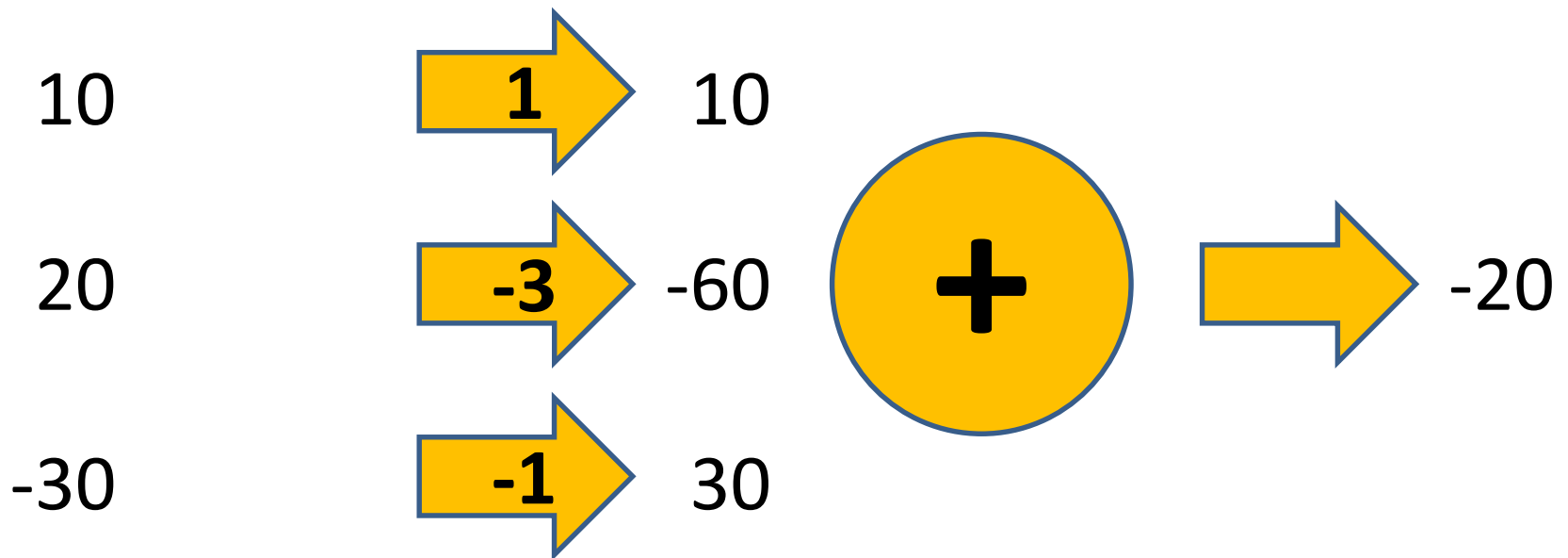
- A természetes idegsejt eléri az akciós potenciált, „tüzel” egyet
- Ha erősen van ingerelve, folyamatosan tüzel
- Tüzelési ráta jelzi, hogy mennyire aktív
- A mesterséges „idegsejt” a tüzelési rátával, mint számmal dolgozik, nem tüzeléseket szimulálunk, mert így sokkal gyorsabban lehet számolni

# Súlyozott összegzés

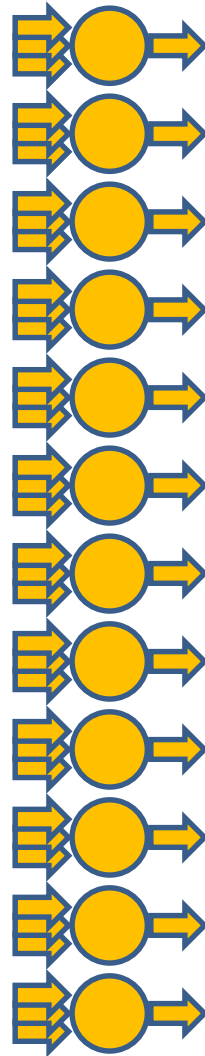
Be

Súlyok

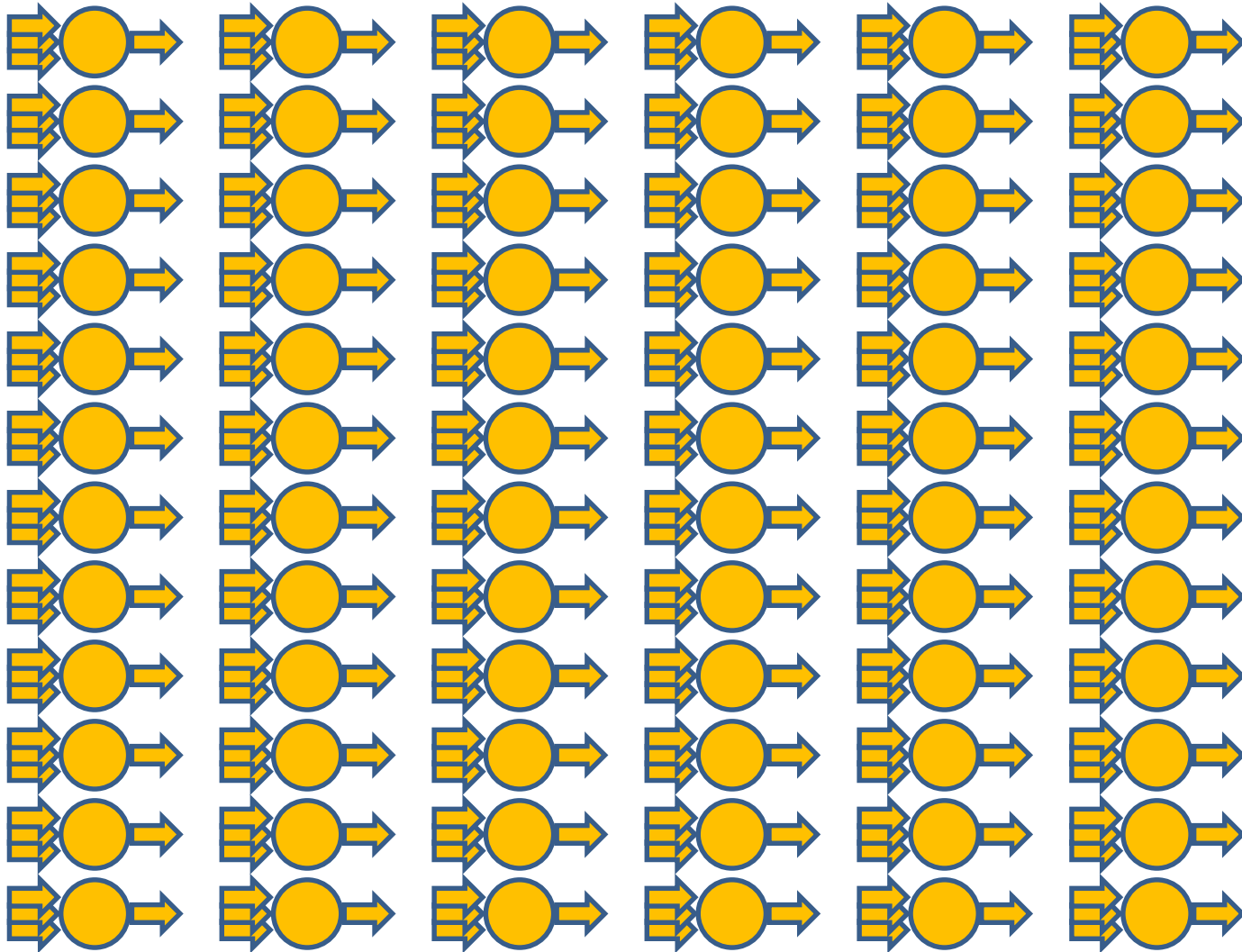
Ki



***Réteg***: sok neuron egymás mellett

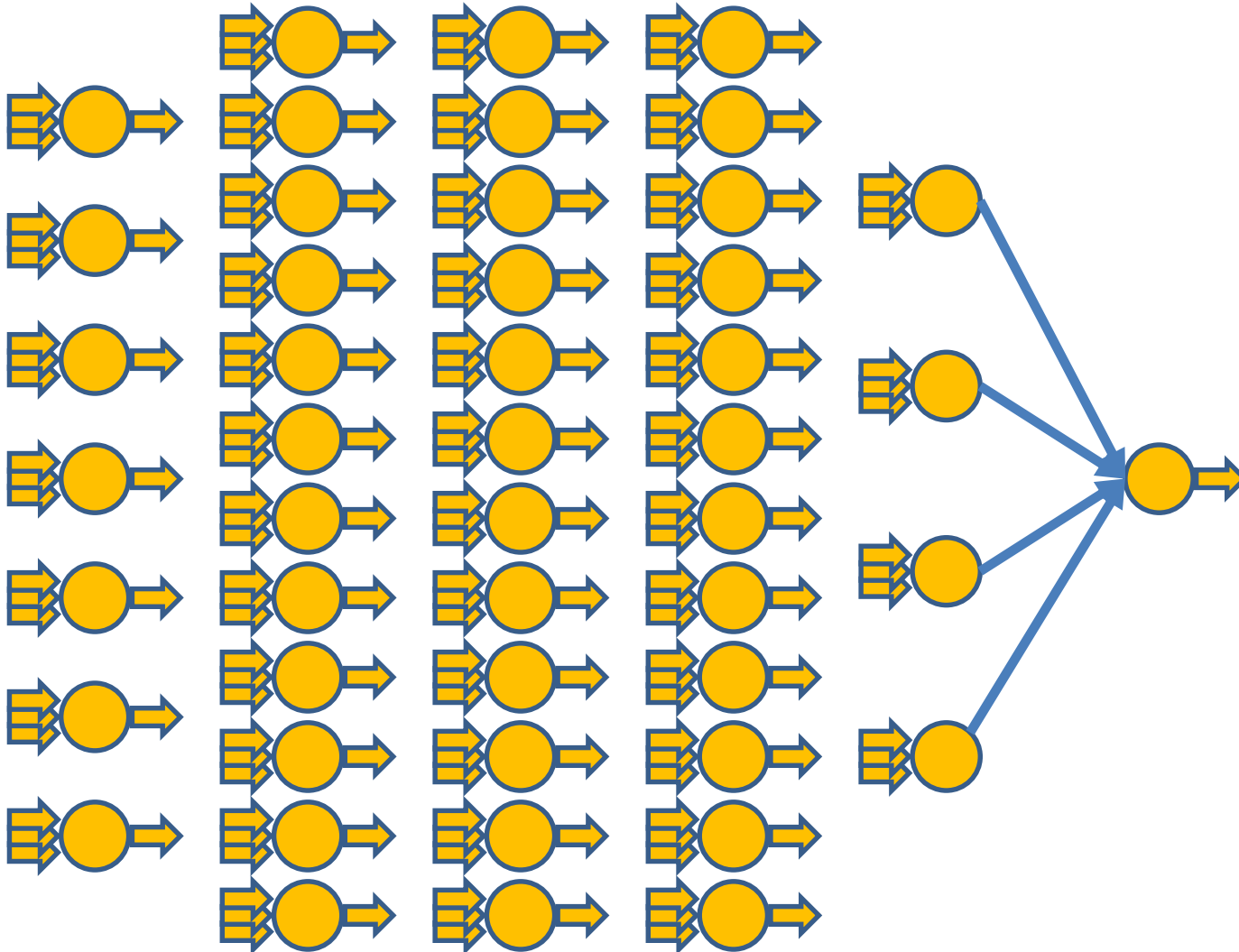


# *Mély hálózat: sok réteg*

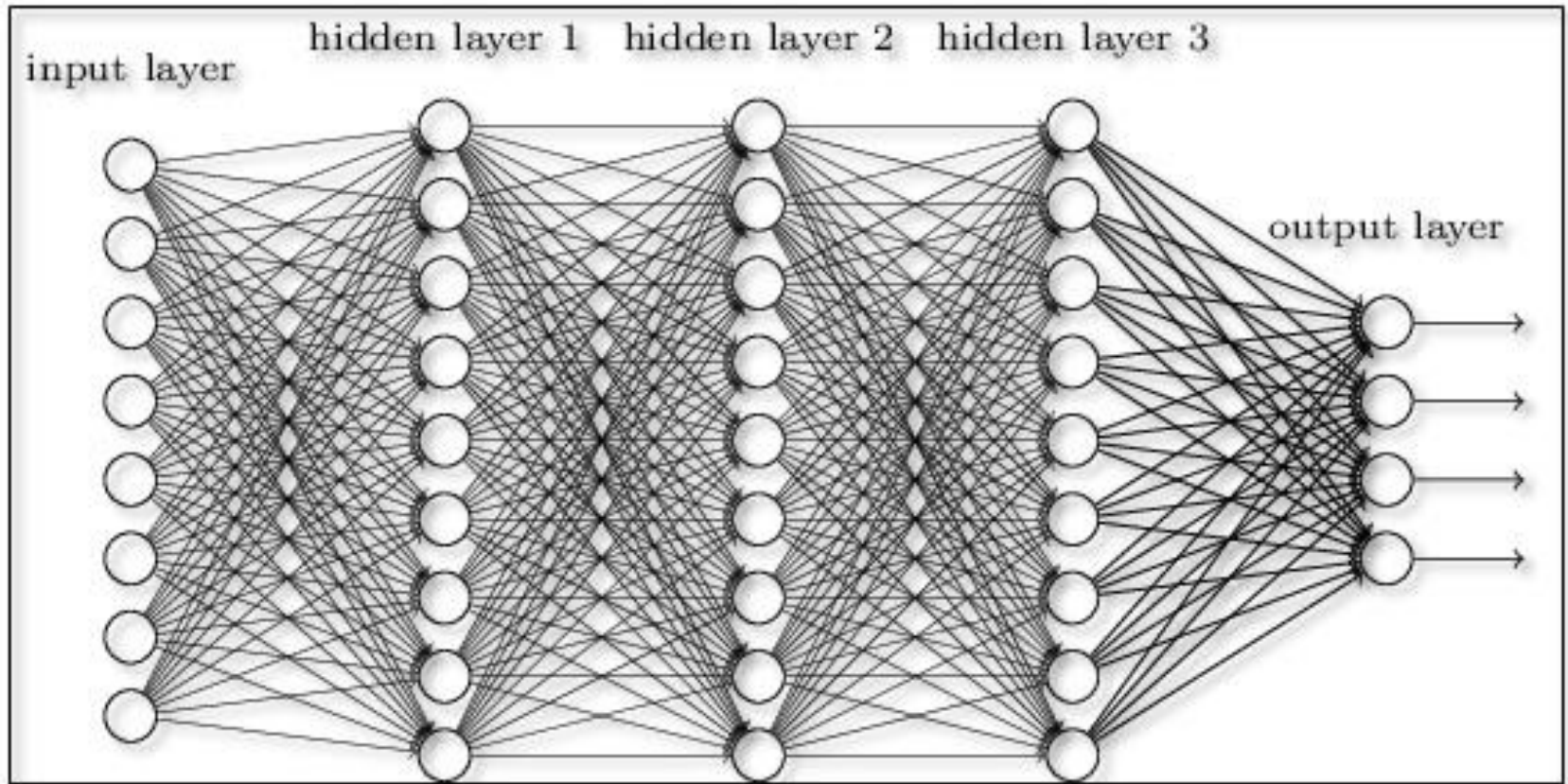




# Különböző számú neuron rétegenként

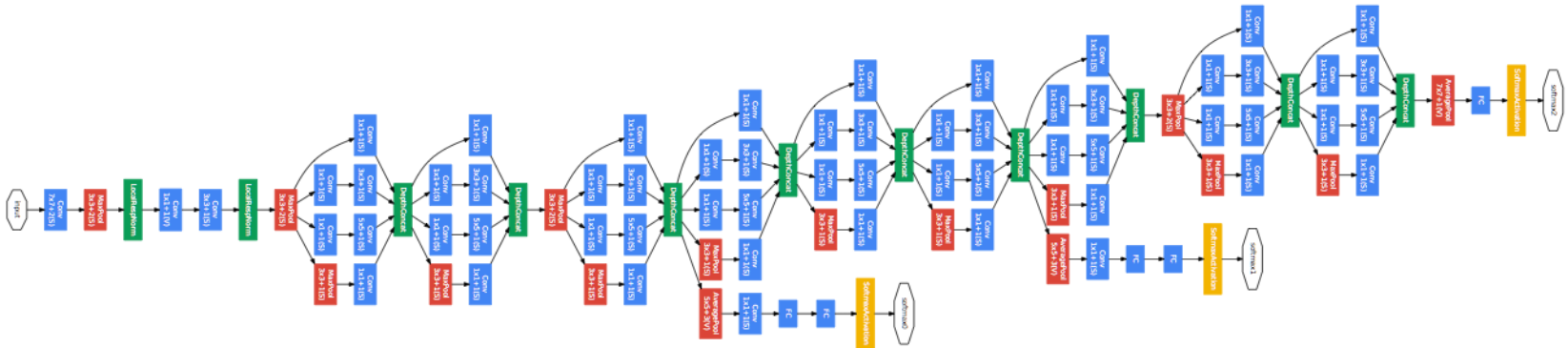


# Minden nyíllal berajzolva...



Mindig csak előre megy nyíl, nincs visszacsatolás,  
nem úgy, mint az agyban

# Architektúra: hogyan kapcsolódnak?



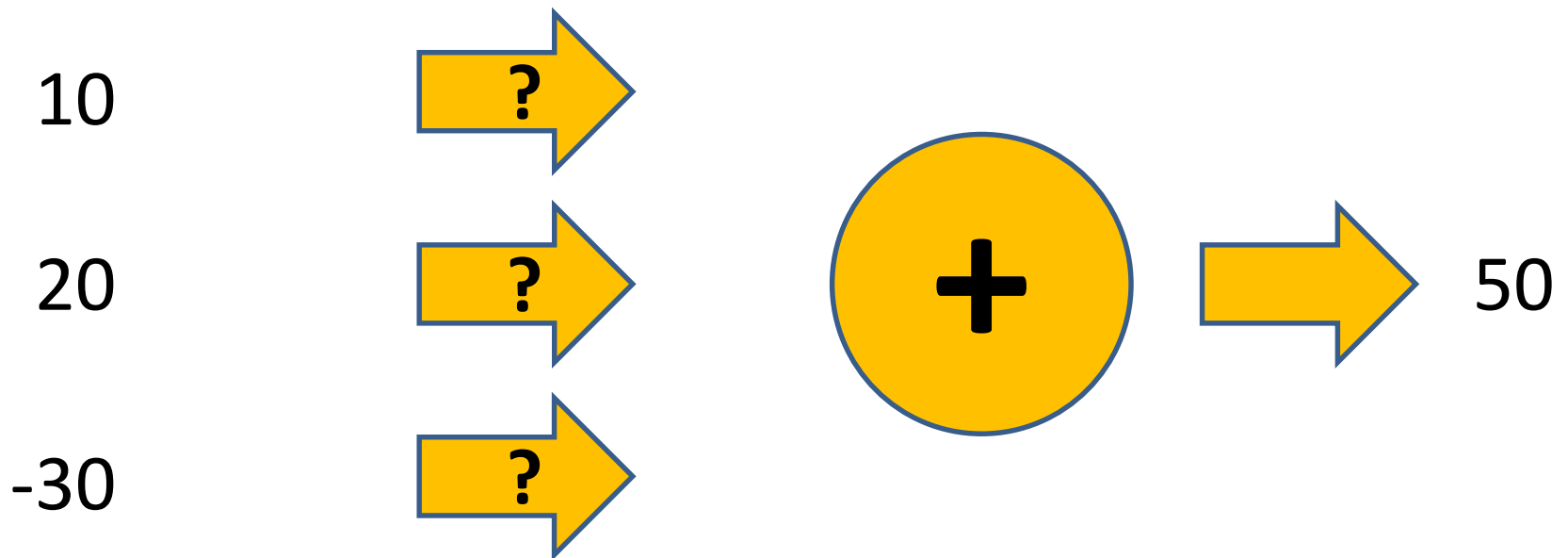
- Több millió neuron is lehet egy-egy rétegben, az egész hálózatban meg több száz millió is.
- Ez még mindig az emberi agy százmilliárd neuronja alatt van három nagyságrenddel.

# A súlyokat változtatjuk: ez a tanítás

Be

Súlyok

Cél

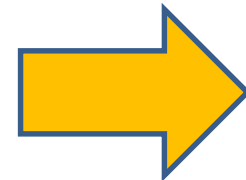
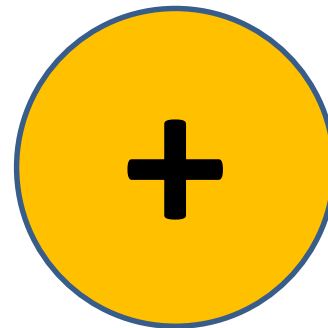
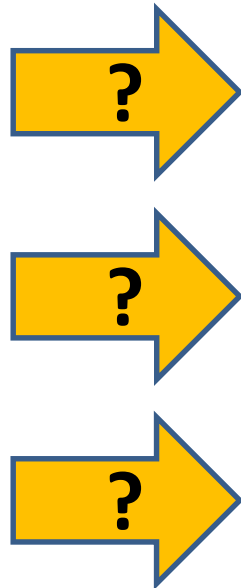


# Felügyelt tanítás: adott bemenethez adott kimenetet várunk

Be

Súlyok

Cél  
(kutya?)



1  
(igen)

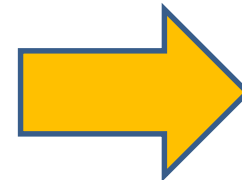
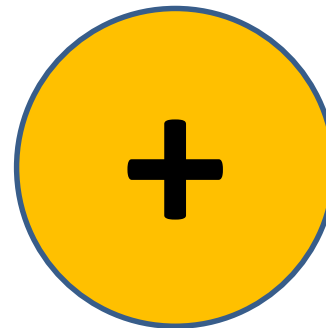
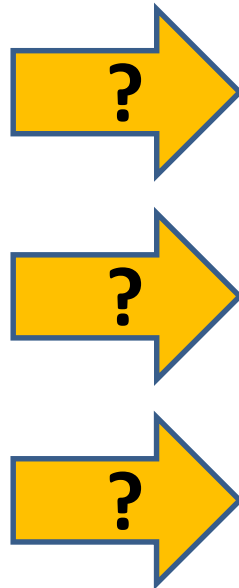
Milyen súlyokra kapjuk azt, amit szeretnénk?

# Felügyelt tanítás: adott bemenethez adott kimenetet várunk

Be

Súlyok

Cél  
(kutya?)



0  
(nem)

Milyen súlyokra kapjuk azt, amit szeretnénk?

# A kép, mint egy csomó szám

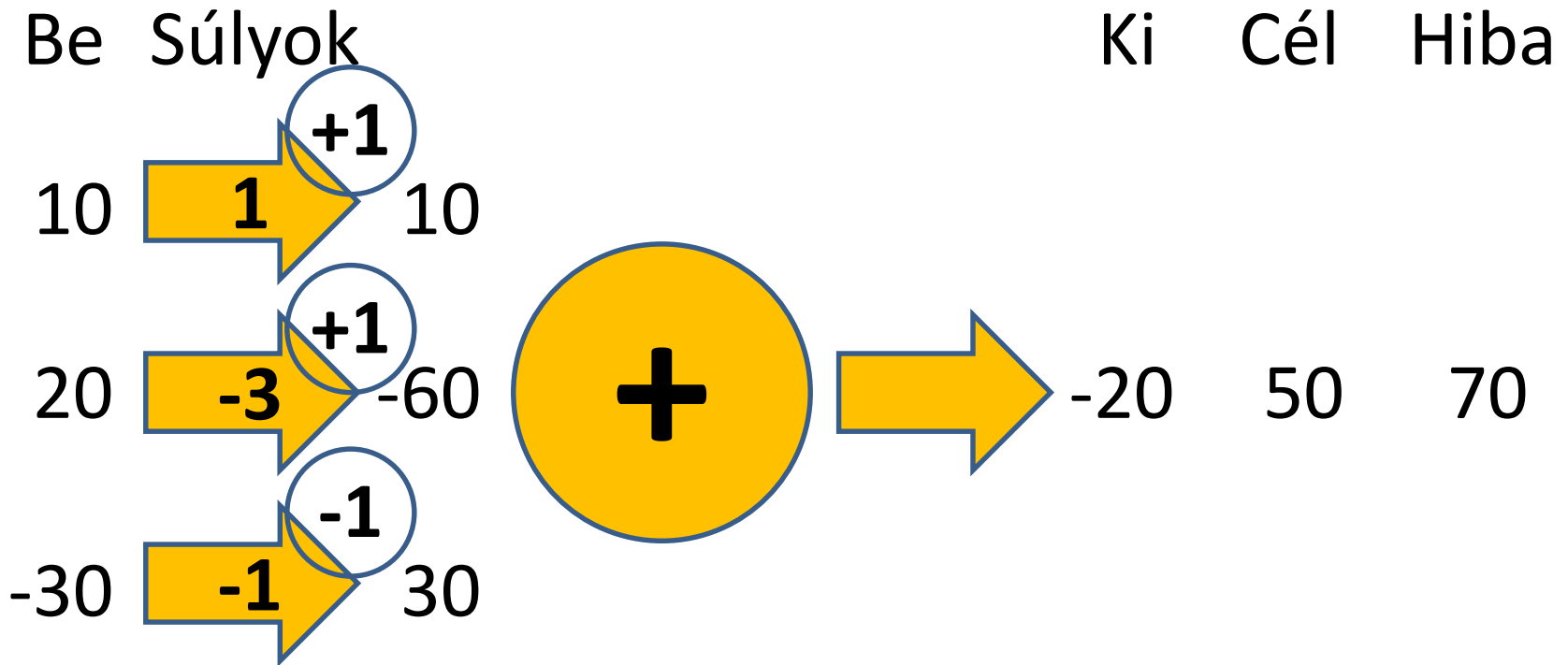
- Az egyes pixelek színe számként van tárolva, RGB formátumban
- Például kutya kép bal felső sarka: halványkék, 202 piros, 237 zöld, 255 kék
- Pl. 10 megapixeles kép (nem ez): 30,000,000 byte (szám 0-255 között), 30MB



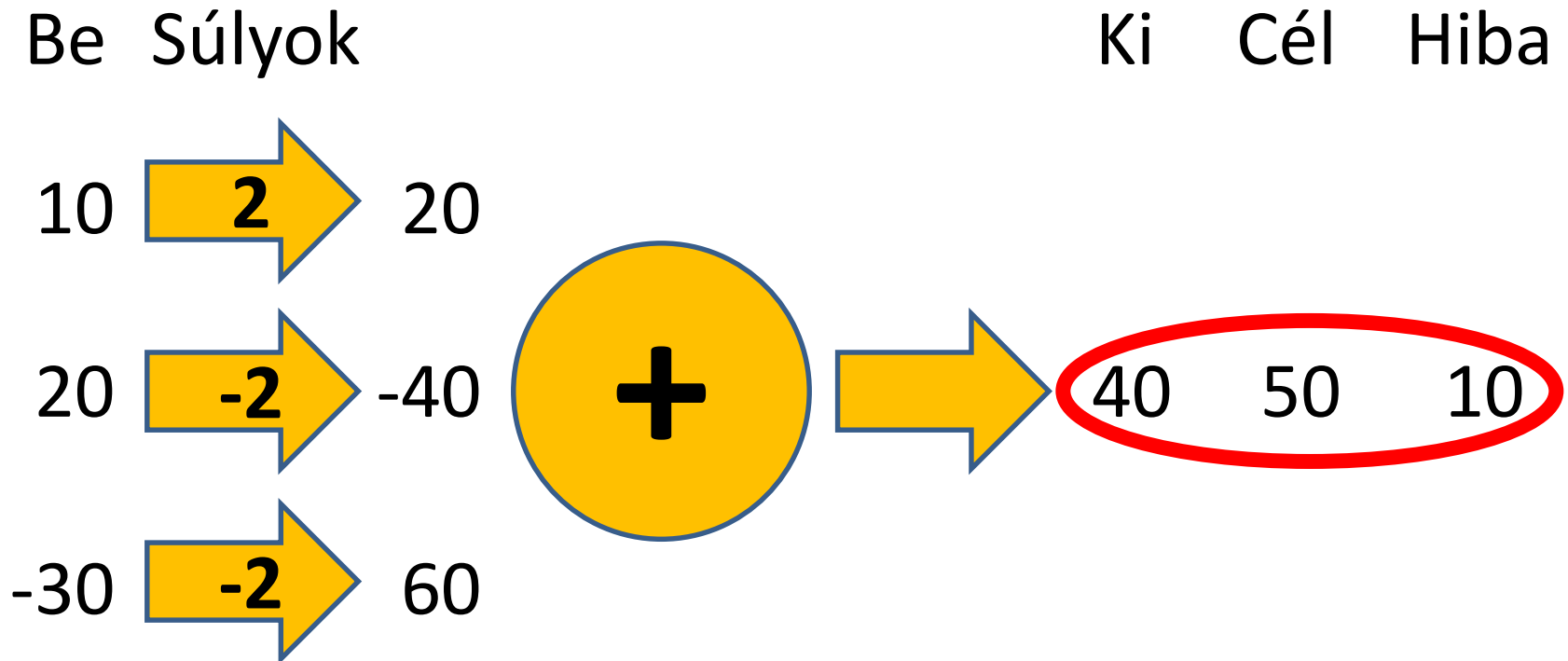




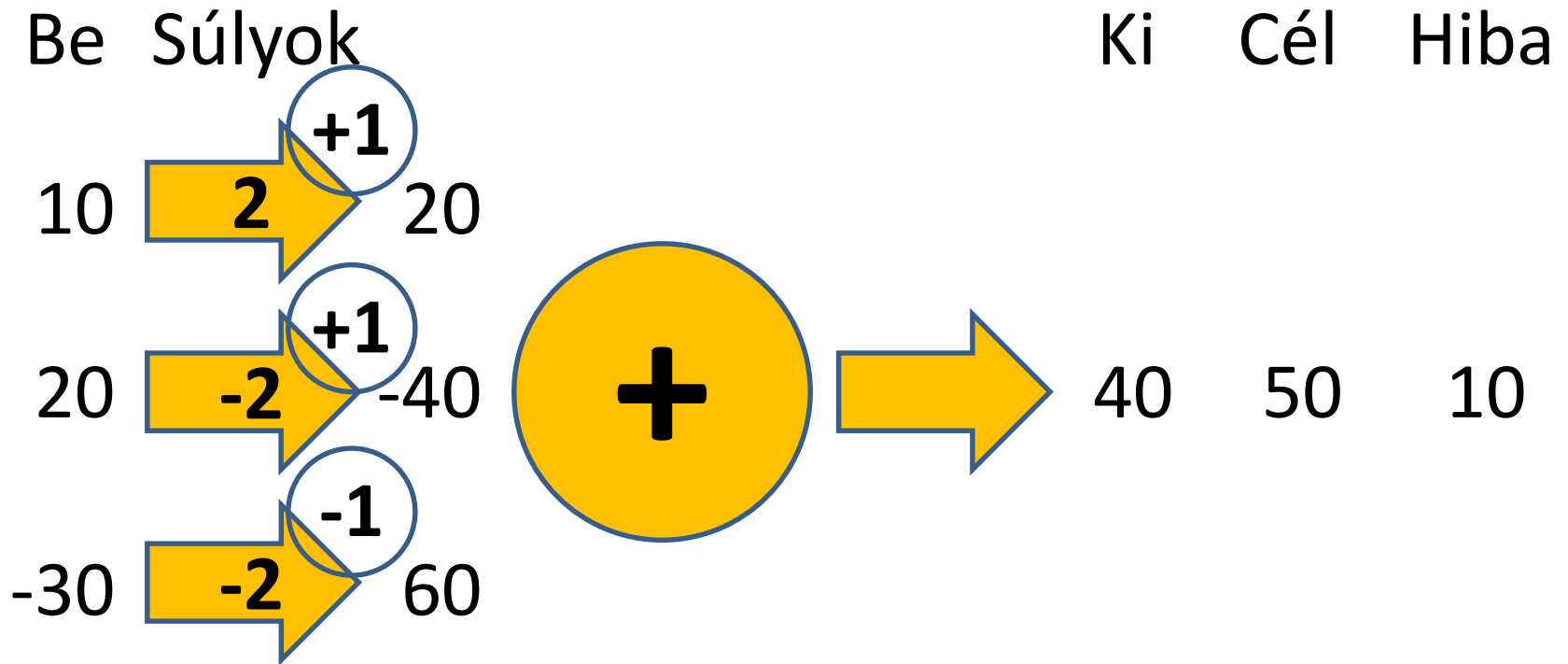
Mindent súlyt abba az irányba változtassunk, hogy a hiba csökkenjen!



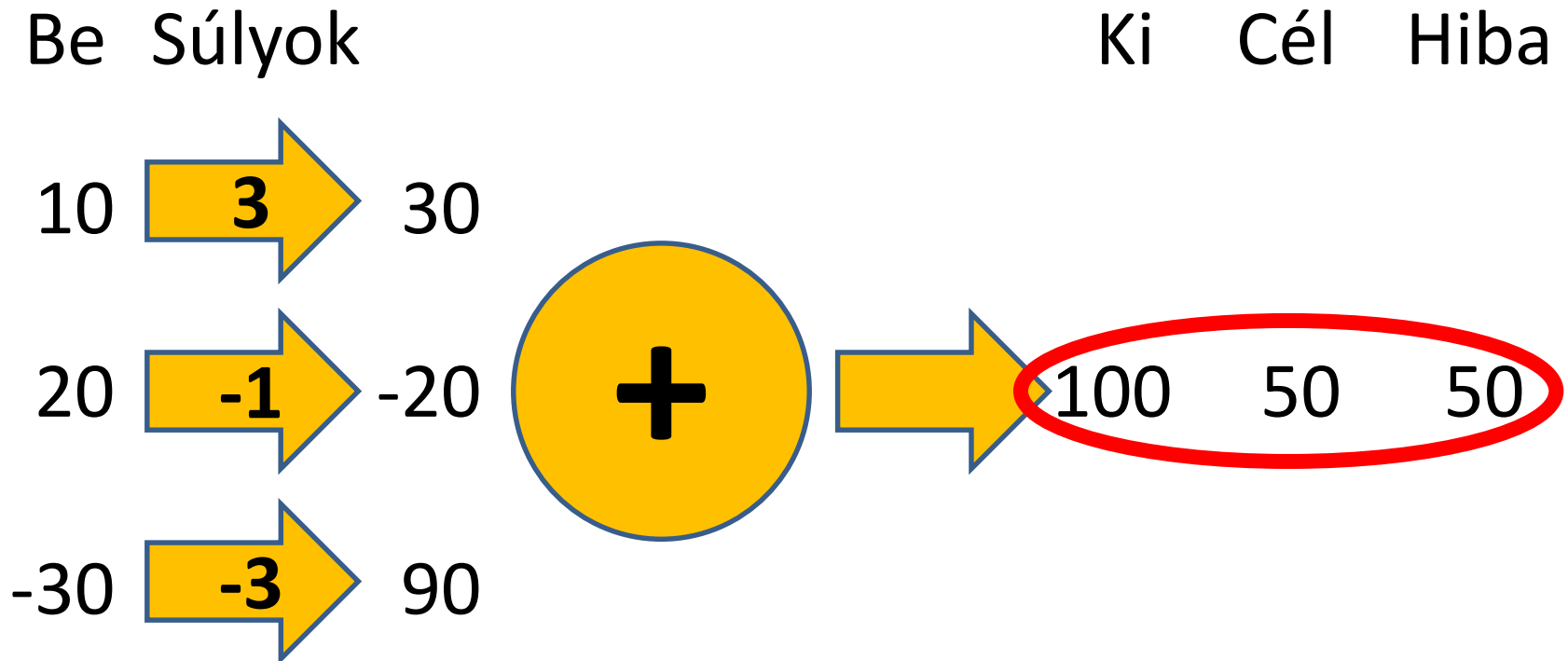
# Már majdnem jó...



# Na még egyszer...



# Túllőttünk a célon...



Nem lehet túl nagy lépésekben haladni, úgy nem fog működni, *divergál*.

# Tanulási ráta („learning rate”)

- Azt határozza meg, hogy mekkora lépésekben haladunk a súlyok változtatásánál.
- Ha túl nagy, nem tudjuk jól eltalálni a célt, a modell *divergál*.
- Ha túl kicsi, nagyon sokáig tart a tanítás.
- Nehéz dolog a jó tanulási ráta megtalálása.
- Gyakran nagyobb értékkel indulunk, és a tanulás folyamán folyamatosan csökkentjük.

# Előre- és visszapropagálás

- Előre propagálásnak („forward propagation”) hívjuk azt, amikor a bemenetből és a pillanatnyi súlyokból kiszámoljuk a kimenetet.
- Visszapropagálásnak („backpropagation”) hívjuk azt, amikor a hiba alapján kiszámoljuk, hogy hogyan változzanak a súlyok, és aszerint meg is változtatjuk őket.
- Egy tanítási lépés egy előre- és visszapropagálásból áll. Ezt ismételtetjük.

# Súlyok változtatása nem konstans

- Itt most +1 vagy -1-gyel változtattuk a súlyokat az egyszerűség kedvéért.
- A valóságban ez nem így történik, jobban változtatjuk azokat, amik jobban hozzájárultak a hibához.
- Erre is nagyon sok féle eljárás van, az egyik alapvetőt úgy hívjuk, hogy Stochastic Gradient Descent, rövidítve SGD.

# Melyik ruhát vegyem fel?

Szeretnénk a buliban egy kis figyelmet kelteni magunk iránt. Melyik ruhát vegyük fel, a pirosat vagy a feketét?





# Jóbarátok tanácsot adnak...

Andrea: Pirosat!

Bernadett: Nem tudom...

Cecília: Feketét!



# Jóbarátok tanácsot adnak...

Andira hallgat, pirosat vesz fel. Egész este megvan az érdeklődés!

Legközelebb kire hallgat?



# Az este tanulságai

- Piros ruha jól áll
- Andi ért a divathoz!
- Bernadett használhatatlan
- Cecília nem ért a divathoz

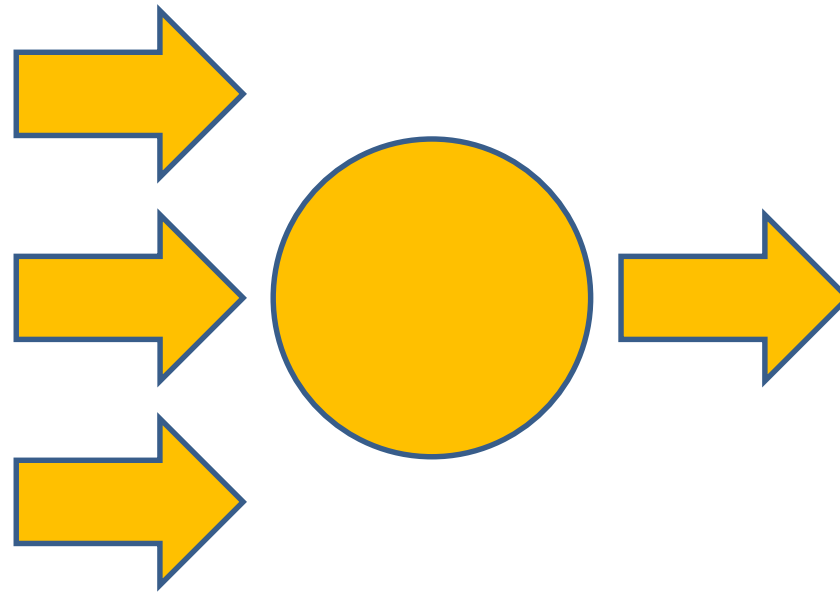
# Súlyok a neurális hálózatban

- Legközelebb nagyobb súllyal veszi figyelembe Andrea véleményét, mert már „bizonyított”.
- Ha valaki nem mond semmi használhatót, mint Bernadett, kár kérdezni, nem éri meg a fáradságot, a véleményének adott súly a nullához közelít, akár meg se kérdezi.
- Aki nem jót mond, mint Cecília, arra egyre kevésbé hallgatunk, a súly csökken, sőt, negatív is lehet, pont az ellenkezőjét tesszük...

# Visszapropagálás, saját memória

- Andrea kérdésére elmondja, hogy jól sikerült az este. Ezáltal ő magabiztosabb lesz a saját döntésében, ő is erősíti a magában, hogy azok a faktorok, ami alapján mondta a piros ruhát, a jó faktorok voltak.
- A „piros ruha jól áll”, önálló vélemény csak azokban a neuron fajtákban van, amiknek van belső memóriájuk, ez nem általános.

# ***Minibatching***: Egyszerre több példával tanítjuk



1	0
0	0
1	1
1	0

# Minibatching nagyon gyors

- Nagyon sok időt lehet megspórolni, mert ha ugyanazokat a súlyokat alkalmazom sok képre, akkor a gép nagyon hatékonyan tud szorozni és összeadni
- Visszapropagálás pedig csak egyszer van az egész batchre, a hibák összege alapján
- Viszont nagyon sok memóriát igényel a videó kártyán – ezért a sok kártya

