

Gyakorlás

Eddig kihagytuk a gyakorlás területét.

A hálózatot a gyakorláshoz mintaadatokkal tápláljuk.

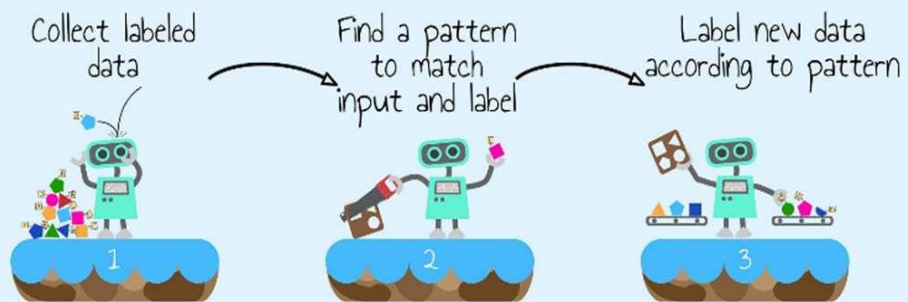
```
["f", 60, 165]    # Alice  
["m", 73, 183]    # Bob  
["m", 69, 178]    # Charlie  
["f", 54, 152]    # Diana
```

Az adatokat "m" (férfi) vagy "f" (nő) jelöléssel látjuk el.
Ezt nevezzük felügyelt tanulásnak.

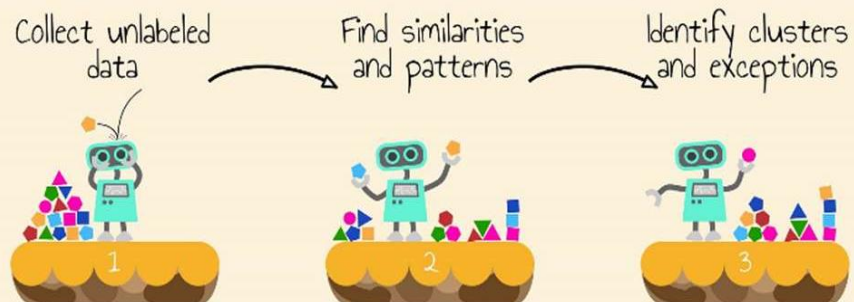
Összességében megkülönböztethetjük:

Types of Machine Learning

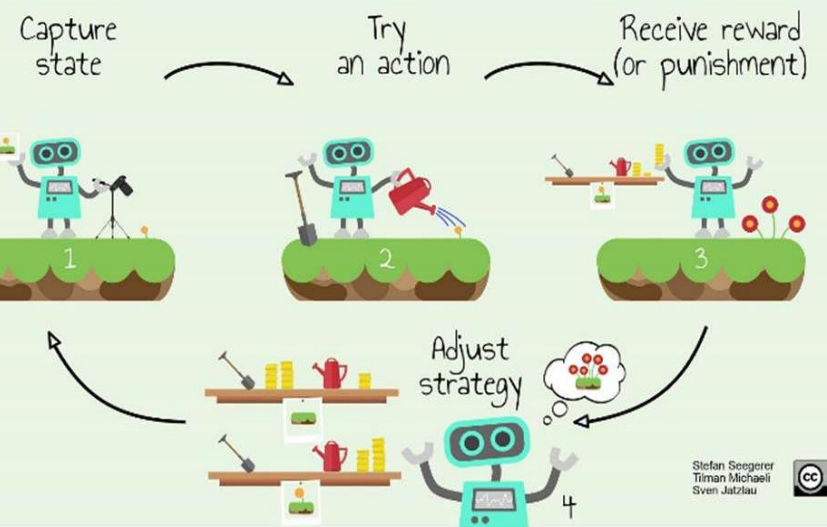
Supervised Learning



Unsupervised Learning



Reinforcement Learning



Megvalósítás

A technikai megvalósítást itt csak nagyon röviden és nem részletesen szeretnénk ismertetni:

Kezdetben minden súlyt és torzítást véletlenszerűen rendelünk hozzá.

Minden ilyen kombinációhoz hozzárendelhető egy hibaérték.

A hibaérték kiszámítása

Erre gyakran használják az átlagos négyzetes hibát (MSE).

$$\text{MSE} = \frac{1}{n} \sum_{i=1}^n (Y_i - \hat{Y}_i)^2$$

MSE = mean squared error

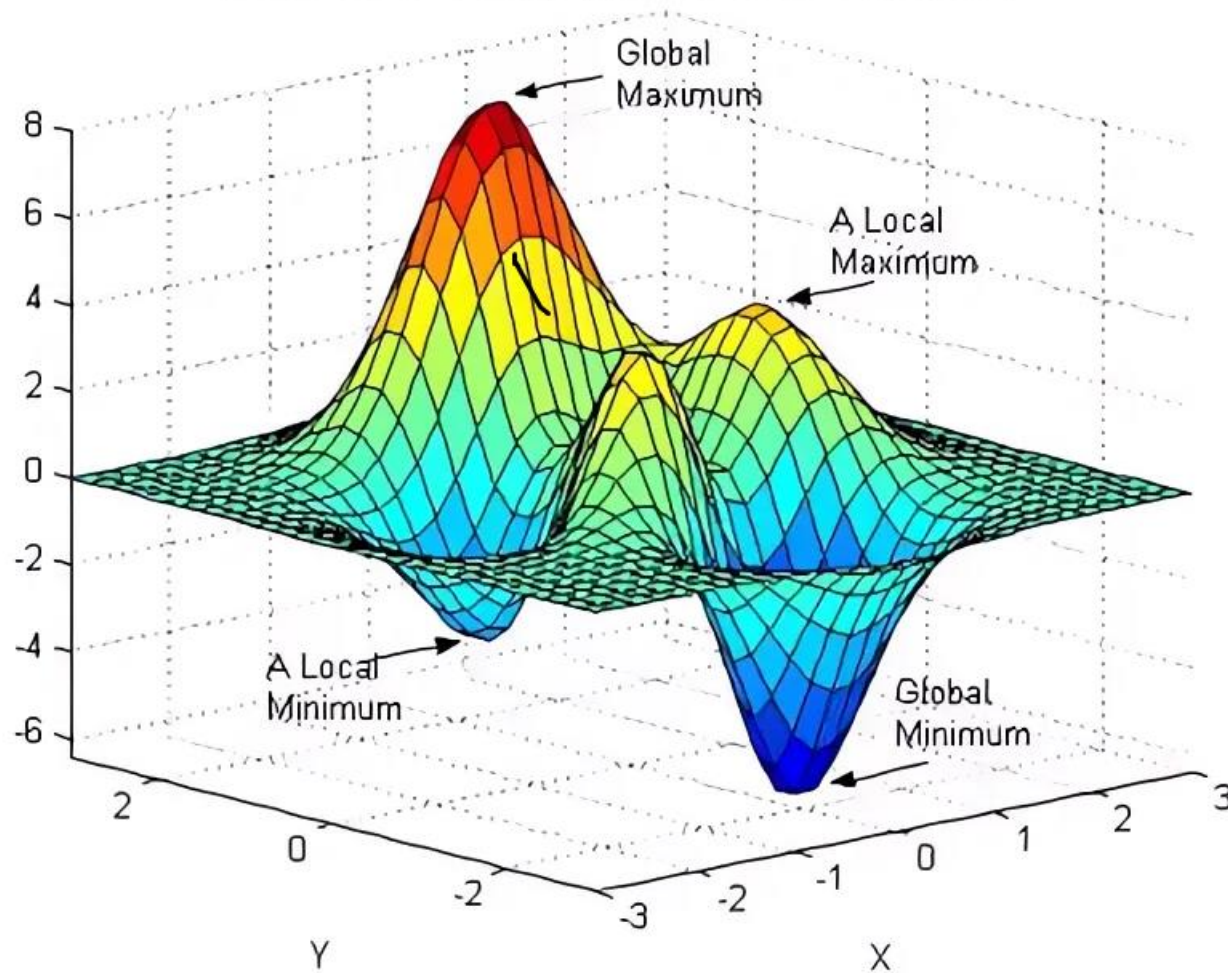
n = number of data points

Y_i = observed values

\hat{Y}_i = predicted values

n megfigyelésre került sor. Minden esetben a várható érték ("f"=1, "m"=0) és a megfigyelt érték közötti különbséget négyzetre kell emelni. Ezután az összes értéket összeadjuk, és az összeget elosztjuk n -nel.

A Non-Convex Combination of Gaussian Distributions



Zachary kaplan, CC BY-SA 4.0, via Wikimedia Commons

A hibaértékkel együtt a paraméterek egy n -dimenziós síkot fednek le. A szemléltetés érdekében ez az eset, amikor csak 2 paraméter van. A megoldás keresése itt megfelel a (lokális vagy globális) minimumok keresésének ezen a síkon.

Fokozatos süllyedés

Nagyon gyakran, de nem kizárólagosan, az úgynevezett "gradiens süllyedés módszerét" alkalmazzák. Ez meghatározza az aktuális ponton a meredekséget (gradiens), és megpróbál lefelé haladni a minimum felé.

A hálózatot így sikeresen kiképeztük és teszteltük.
A kapott súlyok és torzítások ezután kiolvashatók.
Próbáld meg magad elvégezni a számítást a megadott 2 értékpárra.

Férfi vagy nő?
Te is megpróbálhatsz egy értéket.

Jegyzetek

A számítás előtt a súly és a magasság átlagos értékeit kivonjuk az értékpárból. A hálózatot így képeztük ki. A szigmoid függvényt használjuk kényszerfüggvényként, lásd az alábbi sigmoid.pdf fájlt.