09 Mély neuronhálók

Fogalomtár

A *mesterséges* *neuronok* a mesterséges neuronhálók számítási egységei. A továbbiakban olyan neuronokról lesz szó, amelyek *rétegekbe* szerveződnek, legalább egy *bemenettel* és *kimenettel* rendelkeznek; az *egyesített* bemeneti értékeikre alkalmazzuk az *aktivációs függvényt*, amelynek az értéke a mesterséges neuron kimeneti értéke, ezt a kimeneti értéket adja tovább minden kimenetén.

A továbbiakban a mesterséges neuronok bemeneti értékeit azok *lineáris kombinációinak egy adott konstanssal való eltoltjával* egyesítjük. Ha az eltoláshoz (biashoz) tartozó bemeneti értéket jelöli, ahol , az *n* darab, a megelőző neuronok által átadott bemeneti érétket , és az ezekhez tartozó súlyokat rendre , akkor a bemeneti értékek lineáris kombinációjának konstanssal eltolt értéke . A lineáris kombinációban szereplő súlyokat a háló *paramétereinek* nevezzük.

Gyakran használt *aktivációs függvények*:

*Identikus függvény*:

*Egység lépcsős függvény*, vagy ugrásfüggvény:

*Szigmoid függvény*, vagy logisztikus függvény:

*Tangens hiperbolikusz függvény*:

*ReLU* (rectified linear unit):

*Sima* (smooth) *ReLU*, vagy softplus függvény:

A mesterséges neuronhálókban a mesterséges neuronok *rétegekbe* szerveződhetnek. A *bemeneti réteg* a bemeneti adatokat változatlanul továbbítja, a *kimeneti rétegen* a regresszió vagy osztályozás eredménye jelenik meg. A bemeneti és kimeneti rétegek közti rétegeket *rejtett rétegnek* nevezzük.

*Előrecsatolt neuronhálónak* nevezzük az olyan neuronhálókat, amelyekben a kimeneti neuronok kivételével minden neuron csak utána következő neuronnak ad át információt, az információátadásban nem alakulnak ki hurkok és körök.

A *sekély neuronhálók* olyan mesterséges neuronhálók, amelyek legfeljebb egy rejtett réteget tartalmazhatnak.

A *mély neuronhálók* olyan mesterséges neuronhálók, amelyek legalább két rejtett réteget tartalmaznak.

A bemeneti adatokhoz tartozó elvárt értékek () és a mesterséges neuronhálóval becsült értékek () eltérését a *hibafüggvénnyel* mérjük. Az általunk használt hibafüggvény: (a hibatagok négyzetösszegének fele).

A háló paramétereit úgy változtatjuk meg, tehát a paramétereket úgy *frissítjük*, hogy a hibafüggvény értéke csökkenjen. Erre a továbbiakban a gradiensereszkedésen alapuló *hiba visszaterjesztés* módszerét alkalmazzuk.

A *tanulási ráta* gradiensereszkedés mértékét meghatározó tényező.

A háló paramétereinek egyszeri frissítését *epochnak* nevezzük.