

Neurális hálók redundáns kimeneti réteggel történő tanítása

Kis-Szabó Norbert

I. évf. programtervező informatika MSc.

Témavezető: Dr. Berend Gábor

SZTE TTIK Számítógépes Algoritmusok és Mesterséges Intelligencia Tanszék

A neurális hálók ma a legjobb eredményt elérni képes osztályozó algoritmusok között vannak. Rengeteg felhasználási területük van, főleg olyan feladatokra, amelyek az emberek számára nem okoznak problémát, de nehezen programozhatóak. Ilyen feladatok például a képfeldolgozás, természetesnyelv-feldolgozás valamint a beszédfelismerés területekhez tartozó problémák megoldása.

A hálók a tanuló halmaz elemeit felhasználva képesek egy általános reprezentációt létrehozni az adott osztályozási feladatról, így megbecsülni, hogy egy korábban még nem látott adat mely osztályba tartozik. Rengeteg hiperparaméter befolyásolja a tanulás menetét, és ebből kifolyólag a modellek hajlamosak a túltanulásra, azaz nagy pontossággal osztályozni a tanító halmazra, viszont minden egyéb adatra gyenge az osztályozó képessége. A hiperparaméterektől függ a konvergencia gyorsasága is, vagy hogy a modell egyáltalán képes-e konvergálni.

Dolgozatom célja annak vizsgálata, hogy milyen hatása van annak, ha több redundáns (ugyanazt a célfeladatot ellátó) predikciót végző kimenet réteget építünk a hálóba. Az ilyen módon létrehozott neurális háló tanítása sokkal kevésbé erőforrásigényes, mint az alkalmazott redundáns kimeneti rétegek számával megegyező mennyiségű egyetlen kimenettel rendelkező neurális háló független tanítása. A dolgozatban javasolt, redundáns kimeneti réteggel történő tanítás a manapság sokat használt multitask learning egy változatának is tekinthető, azzal a különbséggel, hogy esetünkben az összes output réteg ugyanazt a feladatot sajátítja el.

A javasolt módszer kiértékelésére a kísérletekhez egy szófaji kódsorozatokat meghatározó osztályozót használtam fel, a kísérleteket több (eltérő nyelvcsaládba tartozó) nyelven is végre hajtva, a megközelítés minél átfogóbb tanulmányozása érdekében. A vizsgálatok során észrevehető volt, hogy a hagyományos (egy kimenettel rendelkező) modellek általánosítóképessége gyakran elmaradt a redundáns kimenettel rendelkező hálótól.