

Témaleírás

## **Klasszifikációs algoritmusok bizonytalanságának eloszlás-független, nem-aszimptotikus vizsgálata**

Gépi tanulás témakörben a klasszifikáció az egyik legalaposabban vizsgált probléma. Az osztályozás számos alkalmazott területen vált elemi fontosságúvá (egészségügy, pénzügy, informatika stb.). Ennek a problémának a megoldására több jól működő tanuló algoritmus létezik. Az egyik legelterjedtebb módszer a szupport vektor gépek használata. Különösen jól használhatóak ezek kernelizált változatai.

A klasszifikáló algoritmusok többsége determinisztikus úton működik és nem szolgáltat garanciát az osztályozás helyességét illetően. Habár léteznek olyan módszerek, amik valószínűséget is rendelnek a klasszifikált elemhez ezek nagy része valamilyen eloszlást feltételez a mintáról vagy aszimptotikus eredményekre épít.

Az önálló projekt során a diák feladata, hogy minél gyengébb feltételek mellett vizsgálja a szupport vektor klasszifikáció bizonytalanságát különös figyelmet szentelve a kernelizált esetekre. A kutatás során eloszlás-független, nem-aszimptotikus, fél-parametrikus statisztikai módszerek tervezése a cél melyre az újramintavételezés vagy „adat perturbáció” szolgálhat alapul. Lehetőség szerint egzakt, de mindenképpen garantált konfidencia halmazok konstruálása értékes eredménynek számítana ezen a területen.