

A félév első felében vett anyag (a zárthelyi anyagának) vázlatos ismertetése

Gépi tanulás (VIMIM136)

2012/2013 2. (tavaszi) félév

Ez csupán segítség kíván lenni, nem a tananyag tételes, kimerítő felsorolása! Senki se hivatkozzon erre, ha a zh-ban olyasmit kérdezek, ami előadáson szerepelt, de itt nem szerepel!

Még átnézem, és esetleg módosítom!

A zárthelyin mind elméleti kérdések, mind egyszerű feladatok lesznek/lehetnek.

Az anyag *legnagyobb része* megtalálható a Russel-Norvig könyv 17. és 18. fejezetében (2005-ös kiadás), továbbá a Neurális hálók könyv (2006-os kiadás) 2.3.1, illetve 9. fejezetében. (Pl. a hibaarány-komplexitás alapon történő metszés tudtommal egyikben sincs benne úgy, ahogy az előadáson vettük.) Igyekeztem az egyes kérdéskörökhöz odaírni, hogy hol található meg róluk a legtöbb infó, remélem nem rontottam el. (Az irodalmak felsorolását lásd a Gépi tanulás/Jegyzetek alatt.)

Általában is az a szokásom, most is ehhez ragaszkodom: a zh két részre bomlik (***** a felsorolásnál), és mindkét részből külön-külön el kell érni a megfelelt szintet!

Ahol nem hangsúlyozom külön, ott az előadáson elhangzott anyag elegendő. Jelenleg ez alól a Neurális hálók könyv 2.3.1 fejezete a kivétel, itt a könyvben több van.

- A tanulás három alapvető fajtája (ellenőrzött, nemellenőrzött, megerősítéses). Az ellenőrzött tanulás alapvető blokkvázlata. Mit nevezünk induktív tanulásnak? Mikor realizálható egy tanulási probléma?
- Döntési fa tanulása példák alapján. Osztályozásra, illetve regresszióra használt döntési fák felépítése. A döntési fa tanulás pszeudókódja. Az attribútum-teszt választás az információnyereség kritérium alapján.
- A tanuló algoritmus teljesítményének becslése: a tanulási görbe kétféle értelmezése. A tanítás során elkövetett tipikus hiba: a kukucskálás. A zaj és túlilleszkedés.
- A döntési fák metszése szignifikancia teszt, illetve hibaarány-komplexitás kompromisszum alapján.

■

- Hiányzó adatok problémájának kezelési lehetőségei. Nagyon sok értékkel rendelkező attribútumok okozta nehézség – az információ nyereség arány. Folytonos bemeneti attribútumok felhasználása.
- A tanulás számítási elmélete. A szükséges példamennyiségre adható alsó korlát. (Russel-Norvig 18. fejezet)
- Az ERM konzisztenciája, a szükséges mintaszámra adható másik korlát (Neurális hálózatok könyv 2.3.1 fejezetében) – a könyvben több van, mint ami előadáson elhangzott!
- Moduláris eszközök kialakítása a feladat dekompozíciója alapján. (Neurális hálózatok könyv 9.1).
- Hipotézishalmaz együttes tanulása, a pontos, de különböző megoldások integrálásának követelménye. (ld. a Neurális hálózatok könyv 9.2 fejezetében)