

Óbudai Egyetem Alba Regia Kar			
Tantárgy neve és kódja: Adattudományi ismeretek AGXAD2HMLF Kreditérték: 5 Nappali tagozat 2024/25. tanév 2. félév			
Szakok, melyeken a tárgyat oktatják: Geoinformatika mesterképzési szak MSc			
Tantárgyfelelős oktató:	Piglerné dr. Lakner Rozália	Oktató:	Piglerné dr. Lakner Rozália
Féléves óraszámok: 20	Előadás: 10	Tantermi gyak.: 0	Laborgyakorlat: 10
Számonkérés módja:	vizsga		
A tananyag			
<p><i>Oktatási cél:</i> A tárgy célja az adattudomány alapfogalmainak, eszközeinek és módszereinek megismertetése, valós alkalmazási példákon keresztül. Főbb témakörök: Az adatok általános jellemzése és alapvető statisztikai jellemzői. Adatmodellek, strukturált, félig-strukturált és strukturálatlan adatok. Az adatok előfeldolgozásának lépései, adattisztítás, adatintegráció, adatredukció, adattanszformáció. Adatfeltárás, az adatvizualizáció eszközei. A gépi tanulás módszerei és eszközei: felügyelt és nemfelügyelt tanulás; osztályozás, regresszió és csoportosítás. Tanulási, validációs és tesztelési halmaz, cross-validáció. Osztályozók értékelése: pontosság, felidézés, F1-mérték, ROC görbe. Gyakorlatok: az adatfeldolgozás, gépi tanulás, megjelenítés feladatainak megvalósítása valódi adatokkal (pl. Kaggle adathalmazok), Knime adatelemző szoftver használatával.</p>			
<i>Tematika:</i>			
Témakör			Óraszám
1. Bevezetés. Áttekintés az adattudományról. Az adatok általános jellemzése és alapvető statisztikai jellemzői. Adatmodellek, strukturált, félig-strukturált és strukturálatlan adatok. Az adatok előfeldolgozásának lépései, adattisztítás, adatintegráció, adatredukció, adattanszformáció. Adatfeltárás, az adatvizualizáció eszközei. / KNIME megismerése, adathalmazok importálása és alapvető adattípusok felfedezése. Adatok előfeldolgozása KNIME segítségével - adatok tisztítása, hiányzó értékek kezelése, leíró statisztikák generálása. Adatvizualizáció.			4 + 3
2. A gépi tanulás módszerei és eszközei: felügyelt és nemfelügyelt tanulás. Az osztályozás alapjai és módszerei. Tanulási, validációs és tesztelési halmaz, cross-validáció. Osztályozók értékelése: keveredési mátrix, pontosság, felidézés, F1-mérték, ROC görbe. / Egyszerű osztályozó modellek készítése KNIME-ban és a modellek teljesítményének értékelése metrikák segítségével.			3 + 4
3. A regresszió alapjai és módszerei. Regressziós értékelési metrikák. Csoportosítás alapjai, csoportosítási algoritmusok. / Egyszerű regressziós modellek készítése KNIME-ban és a modellek teljesítményének értékelése metrikák segítségével. Csoportosítási algoritmusok implementálása KNIME-ban és az eredmények vizualizálása.			3 + 3
Félévközi követelmények			
Aláírás feltétele:	Egyéni feladatok minimum 50%-os teljesítése.		
Pótlás módja:	a TVSZ előírása szerint.		
A vizsga módja (írásbeli, szóbeli, teszt, stb): szóbeli			

Irodalom:	
Kötelező:	Az egyetem e-Learning rendszerén közzétett tananyag
Ajánlott:	Jiawei Han, Micheline Kamber and Jian Pei, Data Mining: Concepts and Techniques, 3rd ed. The Morgan Kaufmann Series in Data Management Systems, Morgan Kaufmann Publishers, July 2011. ISBN 978-0123814791 Abonyi János (szerk): Adatbányászat - A hatékonyság eszköze, Computerbooks, 2006. ISBN: 9789636183424 Fogarassyné Wathy Ágnes, Starkné Werner Ágnes: Intelligens adatelemzés, Typotex, 2011.