

Óbudai Egyetem				
Alba Regia Műszaki Kar				
Tantárgy neve és kódja: <i>Adattudományi ismeretek AGXAD2HMLF</i> Kreditérték: 5				
Levelező tagozat 2023/24 tanév 2 félév				
Szakok, melyeken a tárgyat oktatják: Geoinformatika mesterképzési szak MSc				
Tantárgyfelelős oktató:	Dr. Ediboglu Bartos Gaye	Oktatók:	Dr. Ediboglu Bartos Gaye	
Előtanulmányi feltételek: (kóddal)	-			
Heti óraszámok:	Előadás: 10	Tantermi gyak.: 0	Laborgyakorlat: 10	Konzultáció: 0
Számonkérés módja (s,v,f):	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, adattranszformáció. 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), adatelemző szoftverek (pl. RapidMiner, Knime) használatával.</p>				
<i>Tematika:</i>				
Témakör				Óraszám
Előadások				
Áttekintés az adattudományról. Az adatok alapjai és típusai. Az adatok megértése, az adattípusok.				1
Adatok előfeldolgozása és leíró statisztikák. Adatok tisztítása. Hiányzó értékek kezelése. Alapvető leíró statisztikák.				1
Bevezetés a gépi tanulásba. A gépi tanulás alapjai. A gépi tanulás típusai.				1
Felügyelt vs. Felügyelet nélküli tanulás.				1
A tanuló, validációs és tesztelési adathalmazok megértése, valamint a keresztvalidációs technikák bemutatása.				1
Az értékelési metrikák megértése, mint például a pontosság, a felidézés, az F1-pontszám, az ROC görbe és a zavaros mátrix megértése.				1
Klaszterezési fogalmak. Bevezetés a klaszterezésbe. Klaszterezési algoritmusok típusai.				1
Klasszifikáció alapok. Bevezetés a klasszifikációba. Egyszerű osztályozó modell készítése. Modell értékelési metrikák.				1
Bevezetés a neurális hálózatokba és alkalmazásaiba.				1
Bevezetés a regresszióba. Egyszerű regressziós modell készítése. Regressziós értékelési metrikák.				1

Gyakorlatok	
KNIME bevezetése, adathalmazok importálása és alapvető adattípusok felfedezése. Adatok előfeldolgozása KNIME segítségével - adatok tisztítása, hiányzó értékek kezelése, és leíró statisztikák generálása. Egy egyszerű gépi tanulási modell készítése KNIME-ban és a különböző gépi tanulás típusok megértése.	2
Gyakorlat a felügyelt és a felügyelet nélküli tanulás megkülönböztetésében KNIME használatával. Klaszterezési algoritmusok implementálása KNIME-ban és az eredmények vizualizálása. Egy egyszerű osztályozó modell létrehozása KNIME-ban és a modell teljesítményének értékelése metrikák segítségével.	4
Egy alapvető neurális hálózat modell létrehozása KNIME-ban és alkalmazásainak felfedezése. Egy egyszerű regressziós modell építése KNIME-ban és teljesítményének értékelése. Gyakorlatok KNIME-ban különböző adathalmazok megértése és kezelése a tanulóhoz és teszteléshez. Az alapvető fogalmak összefoglalása KNIME segítségével és gyakorlati alkalmazások egy példaadathalmazon.	4
Félévközi követelmények	
AZ ELŐADÁSOK LÁTOGATÁSA KÖTELEZŐ!	
13. hét	PótZH
A pótlás módja:	A HF aláíráspótló vizsga jelleggel pótolható
Aláírás feltétele:	Részvétel a gyakorlatokon, beadandó HF elkészítése elfogadott szinten, ZH legalább 50%-os teljesítése
A vizsga módja (írásbeli, szóbeli, teszt, stb): vizsga	

Irodalom:	
Kötelező:	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
Ajánlott:	Fogarassyné Wathy Ágnes, Starkné Werner Ágnes: Intelligens adatelemzés, Typotex, 2011