

Kiberfizikai Rendszerek Intézet			Mintatanterv szerinti 1. félév			
Tantárgy neve:	Kódja:	Kredit:	Óraszám			
				ea	tg	lab
Bevezetés az adattudományba	NKWBA1HBNF	5	heti	2	0	2
Tárgyfelelős: Dr. Kail Eszter			Beosztás: nappali tagozat			
Oktató(k): Dr. Dineva Adrienn, Németh Ernő Nándor						
Előtanulmányi feltételek:	Adatbázisok NIXAB0HBNE					
Számonkérés módja:	évközi jegy					
A tananyag						
Oktatási cél:	A tárgy célja az adattudomány alapfogalmainak és folyamatainak gyakorlati megközelítésű megismertetése. A hallgatók a gyakorlati életből vett valós alkalmazási példákon keresztül az ismereteket megtapasztalva, egyre mélyebbre haladva precíz elméleti és egyúttal praktikus gyakorlati ismeretekhez jutnak. Az elméleti ismeretek gerincét a gépi tanulás algoritmusai adják, a gyakorlati feladatok pedig az elmélet gyakorlati megjelenítését szolgálják Python nyelv használatán keresztül					
Tematika:	Előadás: Történet, példák, esettanulmányok, az adattudományba sorolható területek. CRISP DM és a 6 fázisa. Adattisztítás. Felügyelt, felügyelet nélküli és megerősített tanulás. Túltanulás, alultanulás. Modellek validálása, tanulási/validációs/tesztelési halmaz, cross-validáció, Bias-Variance, ROC görbe. Legkisebb négyzetek módszere. Lineáris Regresszió. Gradiens módszer, maximum-likelihood becslés. Logisztikus regresszió, Tanulási/validációs/tesztelési halmaz, cross-validáció, Bias-Variance tradeoff, Precision-Recall, F1-score, ROC görbe. SVM. Neurális hálók. Döntési fák. Véletlen erdők. Boosting. Nem Ellenőrzött tanulás. Klaszterezés. K-means klaszterezés. Megerősített tanulás. Nagyobb esettanulmányok, kitekintés. Gyakorlat: Az adatmanipulálás, prediktív analízis, megjelenítés lépései valódi adatokkal elsősorban Python-csomagok (pandas, scikit-learn, matplotlib, ggplot) használatával.					

Féléves ütemezés	
Oktatási hét (konzultáció)	Témakör
1.	Bevezetés, fogalmak, csoportosítás, klasszikus adatbányászati technikák
2.	Adatelemzés folyamata, statisztikai alapok, adattípusok átvizsgálása
3.	Adatelőkészítés, modellértékelés, modellválasztás technikái
4.	Döntési fák, Bagging, Random Forest
5.	Statisztikai modellek alapjai, regresszió
6.	SVM
7.	rektori szünet
8.	Neurális hálók
9.	Klaszterezés
10.	Megerősítő tanulás
11.	AtML. esettanulmányok
12.	Konzultáció, zh felkészülés
13.	ZH (labor és elmélet)
14.	Pót ZH (labor és előadás)
Félévközi követelmények	
Évközi jegy / aláírás megszerzésének feltételei:	évközi jegy

Zárthelyi dolgozatok	
Oktatási hét	Témakör
13.	elméleti zh, labor zh
14.	pótló zárthelyi elmélet és labor
Az évközi jegy kialakításának módszere (csak évközi jegyes tárgyak esetében töltendő ki)	
Pótlás módja	
A ZH / évközi jegy / aláírás pótlásának módja:	Az utolsó gyakorlaton az egyik ZH pótolható. Évközi jegy pótláson mindkét ZH pótolható. Itt is érvényes, hogy minimum 50%-ot kell teljesíteni mindkét ZH-n.)
Vizsga módja (csak vizsgás tantárgy esetében töltendő ki)	
Vizsgajegy kialakítása (csak vizsgás tantárgy esetében töltendő ki)	
Az egyes érdemjegyek ponthatárai:	
A félévzáró érdemjegy a ZH-k pontszáma alapján áll össze, amikre maximum 100 pont szerezhető. Az elégséges jegyhez 52, a közepeshez 63, a jóhoz 74, a jeleshez 85 pontot kell elérni.	
Irodalom	
Kötelező:	A Moodle rendszerben közzétett jegyzet.
Ajánlott:	Richard S. Sutton and Andrew G. Barto: Reinforcement Learning: An Introduction J. Leskovec, A. Rajaraman, J. D. Ullman: Mining of Massive Datasets J. Gareth et al.: An Introduction to Statistical Learning
Egyéb segédletek:	