

A TANTÁRGY ADATLAPJA

1. A képzési program adatai

1.1 Felsőoktatási intézmény	Babeş–Bolyai Tudományegyetem
1.2 Kar	Matematika és Informatika Kar
1.3 Intézet	Magyar Matematika és Informatika Intézet
1.4 Szakterület	Informatika
1.5 Képzési szint	Mesteri
1.6 Szak / Képesítés	Adatelemzés és modellezés / Analiza datelor și modelare / Data analysis and modelling

2. A tantárgy adatai

2.1 A tantárgy neve	Bevezetés a gépi tanulásba / Introducere în învățarea automată a mașinilor / Introduction to machine learning						
2.2 Az előadásért felelős tanár neve	Prof. dr. CSATÓ Lehel						
2.3 A szemináriumért felelős tanár neve	Prof. dr. CSATÓ Lehel						
2.4 Tanul- mányi év	1	2.5 Félév	1	2.6. Értékelés módja	vizsga	2.7 Tantárgy típusa	kötelező – alap

3. Teljes becsült idő (az oktatási tevékenység féléves óraszama)

3.1 Heti óraszám	5	melyből: 3.2 előadás	2	3.3 szeminárium/labor/praktika	3
3.4 Tantervben szereplő össz-óraszám	70	melyből: 3.5 előadás	28	3.6 szeminárium/labor	42
A tanulmányi idő elosztása:					Óra
A tankönyv, a jegyzet, a szakirodalom vagy saját jegyzetek tanulmányozása					50
Könyvtárban, elektronikus adatbázisokban vagy terepen való további tájékozódás					24
Szemináriumok / laborok, házi feladatok, portfóliók, referátumok, esszék kidolgozása					36
Egyéni készségfejlesztés (tutorálás)					14
Vizsgák					6
Más tevékenységek:					
3.7 Egyéni munka össz-óraszama					130
3.8 A félév össz-óraszama					200
3.9 Kreditszám					8

4. Előfeltételek (ha vannak)

4.1 Tantervi	Nincs
4.2 Kompetenciabeli	Programozási készségek, matematikai alapismeretek (valószínűségi számítás).

5. Feltételek (ha vannak)

5.1 Az előadás lebonyolításának feltételei	<ul style="list-style-type: none"> • Az előadásokhoz video-projektor szükséges. • A példák kifejtéséhez és az ábrák számára tábla szükséges.
5.2 A szeminárium / labor	<ul style="list-style-type: none"> • A laboratóriumi órák alatt a diákok a számítógépet, az oktató a táblát

lebonyolításának feltételei	használja.
-----------------------------	------------

6. Elsajátítandó jellemző kompetenciák

Szakmai kompetenciák	<ul style="list-style-type: none"> • Adatok elemzésének módszertana, a gépi tanulás alapjai • Megismerkedés a gépi tanulási modellek fogalmával, a modellek adatokhoz történő illesztésével, illetve az adatokból kinyert információ értelmezéséhez. • Algoritmusok elemzése és fejlesztése, a valószínűségi modellek alkalmazása valós adatokon, a gépi tanulás módszertanának ismerete és alkalmazása.
Transzverzális kompetenciák	<ul style="list-style-type: none"> • Feladatmegoldó készségek fejlesztése. • A kritikus gondolkodás elsajátítása, elemzési kérdések megfogalmazása és a kérdésekre történő argumentált válasz megadása.

7. A tantárgy célkitűzései (az elsajátítandó jellemző kompetenciák alapján)

7.1 A tantárgy általános célkitűzése	<ul style="list-style-type: none"> • A cél, hogy napjaink eseményeit rögzítő, cselekvéseinket jellemző adatok feldolgozásának módszertanával megismerkedni.
7.2 A tantárgy sajátos célkitűzései	<ul style="list-style-type: none"> • Az adatok típusai és a hozzájuk illeszkedő algoritmusok megismerése; • A többdimenziós normális eloszlás használata adatok jellemzésénél, • A logisztikus regresszió és a GLM modell használata, • A valószínűség-alapú modellek esetén a Bayes-képlet alkalmazása. • A gépi tanulás fogalmának a megértése, a felügyelt tanulás alkalmazása valós feladatoknál.

8. A tantárgy tartalma

8.1 Előadás	Didaktikai módszerek	Megjegyzések
1. Bevezető fogalmak, definíciók, gépi tanulási modellek és sikeres alkalmazások, paraméterek, paraméterterek.	tanári magyarázat, rávezetés, programkódok tesztelése	
2. Valószínűség-számítási alapfogalmak, alkalmazások a gépi tanulásban, a többdimenziós normál-eloszlás,		
3-4. Paraméterek becslése, alkalmazás a lineáris modellnél. A „likelihood” fogalma, a „maximum likelihood” módszer, annak illusztrálása.		
5. Bayes tétele és a MAP módszer – a „maximum a-posteriori” és a Bayes becslés kapcsolata.		
6. Bayes-becslések és a valószínűségi becslések alkalmazásai. Alkalmazás az MLII-ben: egy modell hiper-paramétereinek a becslése.		
7. Információ-kinyerés Bayes-becslésekből. Egzakt, és		

mintavételezési modellek. A-poszteriori eloszlások kirajzolása.		
8. Közelítő Bayes-becslési módszerek: kvadratúra és variációs egyenlőtlenségek, a Kullback-Leibler divergencia és alkalmazásai. Alkalmazások különböző regressziós modelleknél és osztályozási feladatoknál.		
9. A PCA adatmodell: felépítés a többdimenziós normál eloszlás alapján; a saját-vektorok és saját-értékek szerepe az adatoknál és a rekonstrukciós hiba számításánál. Optimális lineáris vetítési modell.		
10. Az ICA algoritmus: a függetlenség fogalma, a független komponensek és a koktél-parti feladat. Függetlenítő algoritmusok. Alkalmazások EEG, MEG, gazdasági idősoroknál.		
11. Komponens-detektáló algoritmusok alkalmazása valós feladatokhoz.		
12. A „committee”-modellek: gyenge alap-osztályozókból generált „erős” együttesek. Elméleti alapok és alkalmazások.		
13. Keverék-modellek és az EM algoritmus: alsó korlát az adatok illeszkedésére, az alsó korlát maximálása. Alkalmazások.		
14. Ismétlések. Feladatok oldása. Kitekintések.	Kérdés, nyitott vita.	
Könyvészet [1]. Bishop C.M (2006) Pattern Recognition and Machine Learning, Springer Verlag. [2]. Mitchell T (1997) Machine Learning, McGraw Hill. [3]. Murphy K (2012) Machine Learning, a probabilistic perspective, The MIT Press. [4]. Bernardo J.M, Smith A.F.M (2000) Bayesian Theory, John Wiley & Sons. [5]. T. Hastie, J. Friedman, R. Tibshirani: The Elements of Statistical Learning: Data mining, Inference, and Prediction, Springer 2003.		
8.2 Szeminárium / Labor	Didaktikai módszerek	Megjegyzések
1. A Python és Matlab programozási nyelvek használata, csomagok bemutatása, adatok beolvasása és elemzése.	munkáltatás, demonstráció, példák	
2. Adatok előfeldolgozása. A pre-processzálás fontossága és tipikus formái.	munkáltatás, demonstráció, példák	
3. Osztályozási és regressziós modellek tanulmányozása, regressziós modell kiválasztása későbbi bemutatás céljából.	munkáltatás, demonstráció, példák	
4-5. A Bayes-módszer alkalmazásai gépi tanulásnál, alkalmazása feladatoknál (KAGGLE.COM vagy más versenyfeladat tanulmányozása).	munkáltatás, demonstráció, példák	

6. A PCA és ICA modellek: implementáció és alkalmazás.	munkáltatás, demonstráció, példák	
7. Szemináriumi feladatok bemutatása.		Szemináriumi bemutatók, melyeket a diákok tartanak a témáról és a tartalomról történt egyeztetés után.
Könyvészet [1].–[5]. + [6]. MacKay D.J.C (2003) Information Theory, Inference and Learning Algorithms, Cambridge University Press, HTTP: http://wol.ra.phy.cam.ac.uk/mackay/itila/book.html . [7]. A. Webb: Statistical Pattern Recognition, Wiley, 2002 [8].E. Alpaydin: Introduction to Machine Learning, The MIT Press, 2004		

9. Az episztemikus közösségek képviselői, a szakmai egyesületek és a szakterület reprezentatív munkáltatói elvárásainak összhangba hozása a tantárgy tartalmával.

- Az előadás váza a Stanfordi egyetem, az UCL egyetem honlapjain található „Machine learning” előadással megegyezik.

10. Értékelés

Tevékenység típusa	10.1 Értékelési kritériumok	10.2 Értékelési módszerek	10.3 Aránya a végső jegyben
10.4 Előadás	Írásbeli vizsga a félév végén	Írásbeli vizsga	60%
10.5 Labor	Programozási feladatok bemutatása és bemutatók	A megoldások pontozása	40%
10.6 A teljesítmény minimumkövetelményei			
Kötelező a pontok felének összeszedése minden kiértékeléskor (évközi kiértékelés (laborgyakorlatok, szemináriumi bemutatók), végső vizsga).			

Kitöltés dátuma

2018.04.11

Előadás felelőse

Prof. dr. CSATÓ Lehel

Labor / praktika felelőse

Prof. dr. CSATÓ Lehel

Az intézeti jóváhagyás dátuma

.....

Intézetigazgató

dr. András Szilárd, egy. docens