

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

2. A tantárgy adatai

2.1 A tantárgy neve		Információ-visszakeresés / Regăsirea informației / Information retrieval					
2.2 Az előadásért felelős tanár neve			Bodó Zalán				
2.3 A szemináriumért felelős tanár neve			Bodó Zalán				
2.4 Tanul- mányi év	2	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	1+2
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					40
Könyvtárban, elektronikus adatbázisokban vagy terepen való további tájékozódás					30
Szemináriumok / laborok, házi feladatok, portfóliók, referátumok, esszék kidolgozása					40
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	Algoritmika, programozási készségek, matematikai alapismeretek (algebra, valószínűségszámítás).

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 illusztráció számára tábla szükséges.
5.2 A szeminárium / labor lebonyolításának feltételei	<ul style="list-style-type: none"> • A laboratóriumi órák alatt a diákok a számítógépet, az oktató a táblát használja.

6. Elsajátítandó jellemző kompetenciák

Szakmai kompetenciák	<ul style="list-style-type: none"> • Szöveges gyűjteményekből történő információkeresés alapjai • Különböző tárolási formátumok ismerete, azok kezelése, és konverziója • Algoritmusok elemzése és fejlesztése, melyek szöveges adatbázisokból nyernek ki információkat
Transzverzális kompetenciák	<ul style="list-style-type: none"> • Feladatmegoldó készségek fejlesztése

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

7.1 A tantárgy általános célkitűzése	<ul style="list-style-type: none"> • A tantárgy célja az információkeresés általános bemutatása és az információkeresési algoritmusok bemutatása és elemzése.
7.2 A tantárgy sajátos célkitűzései	<ul style="list-style-type: none"> • Az információkeresés fogalmai: <ul style="list-style-type: none"> ○ Alapfogalmak: dokumentum- és term-listák, indexek ○ Indexek építése ○ Bináris visszakeresés ○ Valószínűségi módszerek használata ○ A vektortér modell ○ Felügyelt és felügyelet nélküli módszerek ○ Weben történő keresés, link-ek elemzése • Keresőmotorok tervezése és használata

8. A tantárgy tartalma

8.1 Előadás	Didaktikai módszerek	Megjegyzések
1. Bevezető fogalmak, definíciók, az információ-visszakereső rendszerek bemutatása.	tanári magyarázat, rávezetés,	
2. Indexelési technikák.	tanári magyarázat, rávezetés, munkáltatás	
3-4. A vektortér modell.	tanári magyarázat, rávezetés, munkáltatás	
5. Információ-visszakereső rendszerek kiértékelése.	tanári magyarázat, rávezetés, munkáltatás	
6. Valószínűségi módszerek az információ-visszakeresésben.	tanári magyarázat, rávezetés	
7. Nyelv-modellek az információ-visszakeresésben.	tanári magyarázat, rávezetés	
8-9. Osztályozási módszerek az információ-	tanári magyarázat, rávezetés,	

visszakeresésben: naiv Bayes, Rocchio, regularizált legkisebb négyzetek (RLS), szupport vektor gépek (SVM).	munkáltatás	
10-11. Felügyelet nélküli módszerek, klaszterezési technikák.	tanári magyarázat, rávezetés, munkáltatás	
12-13. Dimenzióredukciós módszerek, mátrixfaktorizálások.	tanári magyarázat, rávezetés	
14. Webes keresési módszerek, link-analízis.	tanári magyarázat, rávezetés	
Könyvészet		
[1] MANNING C.D., RAGHAVAN P., SCHÜTZE H. <i>Introduction to Information Retrieval</i> . Cambridge University Press, 2009.		
[2] BAEZA-YATES R., RIBEIRO-NETO B. <i>Modern Information Retrieval</i> . Addison-Wesley, 1999.		
[3] VAN RIJSBERGEN C. J. <i>Information Retrieval</i> (2nd ed.). Butterworths, 1979.		
[4] DOMINICH S. <i>The Modern Algebra of Information Retrieval</i> . Springer, 2008.		
[5] BODON F. <i>Adatbányászati algoritmusok</i> . GNU Free Documentation License, 2010 (http://www.cs.bme.hu/~bodon/magyar/adatbanyaszat/tanulmany/adatbanyaszat.pdf).		
8.2 Szeminárium / Labor	Didaktikai módszerek	Megjegyzések
1. A Perl és Python programozási nyelvek használata.	munkáltatás, demonstráció, példák	
2. Adatok indexelése.	munkáltatás, demonstráció, példák	
3. Osztályozási módszerek: naiv Bayes, Rocchio, SVM.	munkáltatás, demonstráció, példák	
4-5. Szövegek kategorizálása: alkalmazások a spamszűrésben.	munkáltatás, demonstráció, példák	
6. A Solr, ElasticSearch, Sphinx és Xpian keresőmotorok.	munkáltatás, demonstráció, példák	
7. Összefoglalás, feladatok bemutatása.		Szemináriumi bemutatók, melyeket a diákok tartanak a témáról és a tartalomról történt egyeztetés után.
Könyvészet		
[1]–[5] +		
[6] MANNING C. D., SCHÜTZE H. <i>Foundations of statistical language processing</i> . MIT Press, Cambridge, 1999.		
[7] SEBASTIANI F. Machine Learning in Automated Text Categorization. ACM Computing Surveys, 2002, vol. 34, pp. 1–47.		
[8] http://nlp.stanford.edu/IR-book/		
[9] http://www.stanford.edu/class/cs276/		

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

- Az előadás váza a Stanfordi egyetem honlapján található „Information Retrieval” jegyzet, mely elérhető on-line.

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

Előadás felelőse

Labor / praktika felelőse

2015.04.28

Bodó Zalán

Bodó Zalán

Az intézeti jóváhagyás dátuma

Intézetigazgató

.....

Szenkovits Ferenc