

Prirodno-matematički fakultet

Vijeću Prirodno-matematičkog fakulteta

Crna Gora
UNIVERZITET CRNE GORE
PRIRODNO-MATEMATIČKI FAKULTET
Broj 33
Podgorica, 15-01-2021 20__ god.

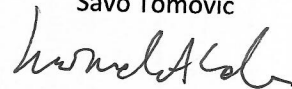
Molim Vijeće Prirodno-matematičkog fakulteta da odobri uvođenje izbornog predmeta *Ekstrakcija informacija iz teksta* na magistarskim studijama računarskih nauka.

Ekstrakcija informacija iz teksta – Text Mining je oblast data mining-a koja se bavi proučavanjem zadataka i metoda za efikasno predstavljanje, analizu i razumijevanje velikih količina podataka predstavljenih u tekstualnoj formi.

Cilj kursa je pregled osnovnih problema i tehnika iz text mining-a i njihova ilustracija kroz izradu praktičkih zadataka.

U Podgorici, 13. 1. 2021. godine

Savo Tomović



Naziv predmeta: Ekstrakcija informacija iz teksta				
Šifra predmeta	Status predmeta	Semestar	Broj ECTS kredita	Fond časova
	Izborni	I	6	4P

Studijski programi za koje se organizuje :

magistarske studije RAČUNARSKIH NAUKA NA PRIRODNO-MATEMATIČKOM FAKULTETU

Uslovljenost drugim predmetima: Nema uslova za prijavljivanje i slušanje predmeta

Ciljevi izučavanja predmeta: Kroz ovaj predmet studenti se upoznaju sa aktuelnim oblastima vezanim za zadatke, probleme i metode ekstrakcije informacija iz teksta (text mining).

Ime i prezime nastavnika i saradnika: Savo Tomović

Metod nastave i savladanja gradiva: Predavanja, konsultacije isamostalna izrada praktičnih zadataka.

SADRŽAJ PREDMETA

Pripremna nedjelja	Priprema i upis semestra
I nedjelja	RapidMiner – analitika teksta
II nedjelja	Rank-frequency distribucije i sekvencijalni windowing
III nedjelja	KNIME – analitika teksta
IV nedjelja	Analiza društvenih mreža
V nedjelja	Analiza nestrukturiranog teksta
VI nedjelja	Analiza sentimenta
VII nedjelja	Seminarski rad
VIII nedjelja	Rudarenje log fajlova i prepoznavanje šablona upotrebe
IX nedjelja	Analiza objava sa Twitter-a u alatu R
X nedjelja	Klasifikacija teksta. Primjer implementacije u Python-u.
XI nedjelja	Kategorizacija teksta. Python implementacija.
XII nedjelja	Ekstrakcija informacija sa veb sajtova za agregiranje tekstualnih sadržaja. Implementacija u Python-u.
XIII nedjelja	Analiza Stack Overflow tagova
XIV nedjelja	Alati za pretraživanje teksta – search engines
XV nedjelja	Seminarski rad
XVI nedjelja	Završni ispit
XVII nedjelja	Ovjera semestra i upis ocjena
XVIII-XXI nedjelja	Dopunska nastava i popravni ispitni rok

Obaveze studenta u toku nastave: Studenti su obavezni da pohađaju nastavu, rade i predaju sve testove, eseje i prezentacije, odrade vježbe u računarskoj sali i rade kolokvijum i završni ispit.

Opterećenje studenta u časovima:

Nedjeljno	u semestru
6 x 40/30 = 6 sati i 40 minuta Predavanja: 4 sata Vježbe: Ostale nastavne aktivnosti: 0	Nastava i završni ispit: (6 sati i 40 minuta) x 16 = 106 sati i 40 minuta Neophodne pripreme (administracija, upis, ovjera prije početka semestra): 2 x (6 sati i 40 minuta) = 13 sati i 20 minuta
Individualni rad studenata: 2 sati i 40 minuta	Ukupno opterećenje za predmet : 5x30 = 150 sati Struktura opterećenja: 106 sati i 40 minuta (Nastava) + 13 sati i 40 minuta (Priprema) + 30 sati (Dopunski rad)

Literatura:

1. Markus Hofmann, Andrew Chisholm: *Text Mining and Visualization Case Studies Using Open-Source Tools*, CRC Press, 2015
2. Cheng Xiang Zhai, Sean Massung: *Text Data Management and Analysis*, Morgan & Claypool Publishers, 2016
3. Christopher D. Manning, Prabhakar Raghavan, Hinrich Schütze: *An Introduction to Information Retrieval*, Cambridge University Press, 2008

Oblici provjere znanja i ocjenjivanje:

- Dva seminarka rada 2 * 25 = 50 poena
- Domaći zadatak 25 poena
- Završni ispit 25 poena.

Prelazna ocjena se dobija ako se kumulativno sakupi najmanje 50 poena.

Ime i prezime nastavnika koji je pripremio podatke: Savo Tomovic