

Prirodno-matematički fakultet

Vijeću Prirodno-matematičkog fakulteta

Molim Vijeće Prirodno-matematičkog fakulteta da odobri uvođenje izbornog predmeta *Obrada prirodnih jezika* na doktorskim studijama računarskih nauka.

Obrada prirodnih jezika – Natural language processing je oblast vještačke inteligencije i lingvistike koja se bavi proučavanjem problema interakcije između računara i prirodnih jezika, odnosno dizajnom i implementacijom metoda za efikasno predstavljanje, analizu i razumijevanje velikih količina podataka predstavljenih u formi prirodnih jezika.

Cilj kursa je pregled osnovnih problema i metoda u obradi prirodnih jezika i njihova ilustracija kroz rješavanje praktičkih zadataka.

U Podgorici, 5. 10. 2020. godine

Savo Tomović



Naziv predmeta: Obrada prirodnih jezika				
Šifra predmeta	Status predmeta	Semestar	Broj ECTS kredita	Fond časova
	Izborni	I	5	4P

Studijski programi za koje se organizuje :

doktorske 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 obrade prirodnih jezika (Natural language processing).

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	Uvod. Odnos <i>obrade prirodnih jezika</i> prema drugim naukama, oblastima i disciplinama.
II nedjelja	Regularni izrazi. Normalizacija teksta. Edit distance.
III nedjelja	Modelovanje jezika sa N-gramima.
IV nedjelja	Naive Bayes i sentiment analiza.
V nedjelja	Logistička regresija i sentiment analiza.
VI nedjelja	Part-of-Speech Tagging i Chunking.
VII nedjelja	Seminarski rad
VIII nedjelja	Semantika vektora. Word2Vec.
IX nedjelja	Procesiranje rečenica sa rekurentnim mrežama.
X nedjelja	Značenje riječi. WordNet. Prepoznavanje semantičkih uloga.
XI nedjelja	Coreference resolution.
XII nedjelja	Sistemi za provjeru pravopisa i rječnika (spelling correction).
XIII nedjelja	NLP za socijalne mreže
XIV nedjelja	Uvod u chatbot platforme
XV nedjelja	Sistemi za davanje preporuka (Recommender Systems)
XVI nedjelja	<i>Završni ispit</i>
XVII nedjelja	<i>Ovjera semestra i upis ocjena</i>
XVIII-XXI nedjelja	<i>Dopunska nastava i popravni ispitni rok</i>

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
5 x 40/30 = 6 sati i 40 minuta	Nastava i završni ispit: (6 sati i 40 minuta) x 16 = 106 sati i 40 minuta
Predavanja: 4 sata	Neophodne pripreme (administracija, upis, ovjera prije početka semestra): 2 x (6 sati i 40 minuta) = 13 sati i 20 minuta
Vježbe:	
Ostale nastavne aktivnosti: 0	
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. Daniel Jurafsky, James H. Martin: *Speech and Language Processing*, 2nd Edition, Prentice Hall, 2008
2. Daniel Jurafsky, James H. Martin: *Speech and Language Processing*, 3rd Edition draft, 2020
3. Christopher D. Manning, Prabhakar Raghavan, Hinrich Schütze: *An Introduction to Information Retrieval*, Cambridge University Press, 2008
4. Jacob Perkins: *Python 3 Text Processing with NLTK 3 Cookbook*, PACKT, 2014

Oblici provjere znanja i ocjenjivanje:

- Seminarki rad 25 poena
- Praktični zadatak 35 poena
- Završni ispit 40 poena.

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

Ime i prezime nastavnika koji je pripremio podatke: Savo Tomović

[Signature]