

1.	Наслов на наставниот предмет	Обработка на текстуални податоци	
2.	Код	КН-И-11	
3.	Студиска програма	магистерски студии по Компјутерски науки	
4.	Организатор на студиската програма (единица, односно институт, катедра, оддел)	Факултет за информатички науки и компјутерско инженерство – ФИНКИ	
5.	Степен (прв, втор, трет циклус)	Втор циклус	
6.	Академска година / семестар 5/10	7. Број на ЕКТС кредити 6	
8.	Наставник	проф. д-р Катерина Здравкова, доц. д-р Игор Трајковски	
9.	Предуслови за запишување на предметот	Нема	
10.	Цели на предметната програма (компетенции): По успешното завршување на овој предмет, студентот ќе биде во состојба да ги препознае клучните проблеми што произлегуваат од користењето на текстуалните податоци, да ги запознае најсоодветните алгоритми за нивното разрешување, да ги истражи и анализира текстуалните податоци и од нив да открие ново знаење, да ги запознае и примени постоечките алатки за обработка на текстуалните податоци и да создаде сопствени решенија за реализација на одделни задачи од областа на обработката и рударството на текстуалните документи и податоци.		
11.	Содржина на предметната програма: Вовед во обработката на текстуалните податоци и текстуалното рударство;; Извлекување, класификација, категоризација и групирање на податоците;; Основни техники на текстуалното рударство;; Текстуалното рударство и мапинското учење;; Примери на примена на техниките на текстуалното рударство во различни апликации;; Идните трендови на обработката на текстуалните податоци.		
12.	Методи на учење: предавања, проекти, дискусии, работилници		
13.	Вкупен расположив фонд на време	6 ECTS x 30 часа = 180 часа	
14.	Распределба на расположивото време	30+45+60+15+30 = 180	
15.	Форми на наставните активности	15.1. Предавања- теоретска настава	30 часови
		15.2. Вежби (лабораториски, аудиториски), семинари, тимска работа	45 часови
16.	Други форми на активности	16.1. Проектни задачи	60 часови
		16.2. Самостојни задачи	15 часови
		16.3. Домашно учење	30 часови
17.	Начин на оценување		
	17.1	Тестови	10 бодови
	17.2.	Семинарска работа/ проект ( презентација: писмена и усна)	50 бодови
	17.3.	Активност и учество	40 бодови

18.	Критериуми за оценување (бодови/ оценка)	до 50 бода		5 (пет) (F)		
		од 51 до 60 бода		6 (шест) (E)		
		од 61 до 70 бода		7 (седум) (D)		
		од 71 до 80 бода		8 (осум) (C)		
		од 81 до 90 бода		9 (девет) (B)		
		од 91 до 100 бода		10 (десет) (A)		
19.	Услов за потпис и полагање на завршен испит	Реализирани активности 15, 16				
20.	Јазик на кој се изведува наставата	македонски или англиски				
21.	Метод на следење на квалитетот на наставата					
22.	Литература					
	22.1.	Задолжителна литература				
		Ред. број	Автор	Наслов	Издавач	Година
		1	Miner, G., Elder, J., Hill, T., Nisbet, R., Delen, D., Fast, D.	Practical Text Mining and Statistical Analysis for Non-structured Text Data Applications	Academic Press	2012
		2	Berry, W. M., Kogan, J. (editors)	Text Mining: Application and Theory	Wiley	2010
	3	Bird, S., Klein, E., Loper, E.	Natural Language Processing with Python	O'Reilley Media	2009	
	22.2.	Дополнителна литература				
		Ред. број	Автор	Наслов	Издавач	Година
		1	Jurafsky, D., Martin, J. H.	Speech and Language Processing (2 <sup>nd</sup> edition)	Pearson Prentice Hall	2008
2		University of Manchester	NaCTeM, The National Center for Text Mining	<a href="http://www.nactem.ac.uk/">http://www.nactem.ac.uk/</a>		