

1.	Наслов на наставниот предмет	<b>Рударење на масивни податоци</b> Mining Massive Data Sets		
2.	Код	F18L3W154		
3.	Студиска програма	Софтверско инженерство и информациски системи, Компјутерски науки		
4.	Организатор на студиската програма (единица, односно институт, катедра, оддел)	Факултет за информатички науки и компјутерско инженерство		
5.	Степен (прв, втор, трет циклус)	прв циклус		
6.	Академска година / семестар 4 / зимски / изборен	7. Број на ЕКТС кредити 6		
8.	Наставник	вонр. проф. д-р Ѓорѓи Маџаров, вонр. проф. д-р Слободан Калајдиски, доц. д-р Ефтим Здравевски		
9.	Предуслови за запишување на предметот	Паралелно и дистрибуирано процесирање		
10.	Цели на предметната програма (компетенции): Студентите да се запознаат со алгоритмите и техниките на податочно рударење и машинско учење за анализа на многу големи податочни множества. Фокус ќе биде даден на дистрибуираните платформи како и на начинот на дефинирање и креирање на алгоритми за процесирање и анализа на многу големи податочни множества.			
11.	Содржина на предметната програма: Вовед во MapReduce, фреквентни множества и асоцијативни правила, барање на најблиски соседи во многу-димензионални податоци, локациски сензитивни хеширања, намалување на димензионалноста (SVD and CUR), системи за препорака, кластерирање, Random Walks with Restarts, надгледувано учење кај масивни податочни множества (K-најблиски соседи, перцептрон, класификациски и регресиски дрва, рударење на податочни потоци, веб рекламирање			
12.	Методи на учење: Предавања со користење на презентации, интерактивни предавања, вежби (користење на опрема и софтверски пакети), тимска работа, пример случаи, поканети гости предавачи, самостојна изработка и одбрана на проектна задача и семинарска работа.			
13.	Вкупен расположив фонд на време	180		
14.	Распределба на расположивото време	30 + 30 + 19 + 15 + 75 = 169 часа		
15.	Форми на наставните активности	15.1.	Предавања- теоретска настава	30 часови
		15.2.	Вежби (лабораториски, аудиториски), семинари, тимска работа	30 часови

16.	Други форми на активности		16.1.	Проектни задачи	19 часови	
			16.2.	Самостојни задачи	15 часови	
			16.3.	Домашно учење	75 часови	
17.	Начин на оценување					
	17.1.	Тестови			0 бодови	
	17.2.	Семинарска работа/ проект ( презентација: писмена и усна)			40 бодови	
	17.3.	Активности и учење			0 бодови	
	17.4.	Завршен испит			60 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			
20.	Јазик на кој се изведува наставата		македонски и англиски			
21.	Метод на следење на квалитетот на наставата		$30+65+30+20+35 = 180$ часа			
22.	Литература					
	22.1.	Задолжителна литература				
		Ред.бр.	Автор	Наслов	Издавач	Година
		1	Anand Rajaraman and Jeffrey Ullman	Mining of Massive Datasets	Cambridge University Press	2014
		2	Avrim Blum, John Hopcroft, and Ravindran Kannan	Foundations of Data Science	Draft version	2017
		3	Jiawei Han, Micheline Kamber, and Jian Pei	Data Mining: Concepts and Techniques, Third Edition	Morgan Kaufmann	2011
	22.2.	Дополнителна литература				
		Ред. број	Автор	Наслов	Издавач	Година

