

Назив предмета: Технике и модели за дубинску анализу података			
Наставник или наставници: Дејан Милић/Часлав Стефановић			
Статус предмета: Изборни			
Број ЕСПБ: 15			
Услов: /			
Циљ предмета Циљ предмета је стицање дубљих и потпунијих знања о техникама и методама који се примењују у претраживању, класификацији и кластеризацији Web докумената и могућностима њихове примене у практичним реализацијама.			
Исход предмета По завршетку курса студенти ће стећи способност за реализацију пројеката заснованих на примени техника Data mining-а и откривања информација.			
Садржај предмета <i>Теоријска настава</i> Мерења и подаци, Анализа поузданости података; Обрада структурираних података: концепата, база података, графова и стабала; Обрада неструктурираних података; Анализа асоцијације; Кластеризација; Класификација; Регресија; Обрада секвенцијалних података; Скривени Марковљев модел, Адаптација параметара скривеног Марковљевог модела; проширени Калманов филтар и рекурентне неуронске мреже; Предикција нестационарних временских серија; Слепе технике обраде сигнала: слепа сепарација, конволутивна сепарација и деконволуција. Препроцесирање: токенизација, издвајање реченица, класификација токена; Издвајање информација: издвајање ентитета, релација и догађаја; Support Vector Machine: алгоритам инкременталног учења и примена у обради текста; Претраживање информација; Аутоматско креирање апстракта документа; Класификација текста; Семантичко претраживање. Преглед основних проблема у области Web mining-а. Таксонометрија Web-а. Анализа садржаја Web докумената. Прилагођавање техника машинског учења анализи Web докумената. Рангирање и аутоматско оцењивање докумената. Анализа утисака. Анализа структуре Web-а. Откривање образаца понашања корисника. Алати који се користе за анализу Web-а и њихова примена. Примери конкретних пројеката заснованих на примени техника Web mining-а. <i>Практична настава</i> Самосталан истраживачки рад, семинарски радови и пројекти.			
Препоручена литература 1. R. Feldman, J. Sanger, The Text Mining Handbook, Cambridge University Press, 2007. 2. D. Hand, H. Mannila and P. Smyth, Principles of Data Mining, The MIT Press, 2001. 3. B. Liu: Sentiment Analysis: Mining Opinions, Sentiments, and Emotions, Cambridge University Press, 2015. 4. C. D. Manning, P. Raghavan, H. Schütze: An Introduction to Information Retrieval, Cambridge University Press, Cambridge, England, 2009.			
Број часова активне наставе		Предавања: 8	Студијски истраживачки рад: 4
Методe извођења наставе Фронтални, групни, индивидуални и практични.			
Оцена знања (максимални број поена 100)			
Предиспитне обавезе	Поена	Завршни испит	Поена
Семинар-и	40	Усмени испит	60