

Спецификација предмета за књигу предмета

Студијски програм		Вештачка интелигенција и машинско учење		
Изборно подручје (модул)				
Врста и ниво студија		Мастер академске студије		
Назив предмета		Прикупљање и предобрада података за машинско учење		
Наставник (за предавања)		Станимировић С. Александар, Велинов Р. Горан		
Наставник/сарадник (за вежбе)		Станимировић С. Александар, Велинов Р. Горан		
Наставник/сарадник (за ДОН)				
Број ЕСПБ	6	Статус предмета (обавезни/изборни)	Обавезни	
Услов				
Циљ предмета	Циљ предмета је да упозна студенте са основним концептима и принципима прибављања и препреме података за машинско учење. Предмет треба да обезбеди знање неопходно за припрему оригиналних скупова података за креирање различитих модела машинског учења.			
Исход предмета	Теоријска и практична знања о припреми скупова података за процес машинског учења. Предмет оспособљава студенте за коришћење различитих напредних алата за трансформацију оригиналност скупа података у скуп података погодан за примену техника машинског учења. Након одслушаног курса студенти ће поседовати практична знања коришћења техника и алата за прикупљање и предобраду података за машинско учење.			
Садржај предмета				
Теоријска настава	Основни концепти везани за обраду података: објекти и типови атрибута. Основне технике за статистичку анализу скупова података. Мере сличности између података. Дескриптивна анализа података и визуелизација података. Мерење и поређење перформанси различитих модела машинског учења. Основне технике за припрему података. Интеграција података и разрешавање конфликта. Чишћење података и процес чишћења података. Проблем недостајућих података (missing data). Проблем „шума” у подацима (noisy data). Технике за трансформацију података. Нормализација података. Редукција података. Избор атрибута. Избор инстанци података. Дискретизација података.			
Практична настава (вежбе, ДОН, студијски истраживачки рад)	На практичној настави студенти ће се упознати са савременим алатима и програмским језицима који се користе у процесу машинског учења и процесу припреме података за машинско учење. Практични пприпери везани за основне технике статистичке и дескриптивне анализе купова података. Практични примери везни за визуелизацију скупова података. Практични примери који илуструју технике за припрему и чишћење података пре процеса машинског учења. Практичне технике за решавање проблема недостајућих података и проблема „шума” у подацима. Практични примери трансформације и нормализације података. Практична примена техника за редукцију димензионалности скупа података.			
Литература				
1	Jiawei Han, Micheline Kamber and Jian Pei, Data Mining: Concepts and Techniques, Third Edition, 2012, Morgan Kaufmann			
2	Salvador García , Julián Luengo and Francisco Herrera, Data Preprocessing in Data Mining, 2015, Springer			
3	P.-N.Tan, M. Steinbach, A. Karpatne, V. Kumar: Introduction to data mining, Addison Wesley, Second edition, 2019.			
4	Ihab F. Ilyas, Xu Chu, Data Cleaning, 2019, ACM Books			
5	Jason W. Osborne, Best Practices in Data Cleaning: A Complete Guide to Everything You Need to			
Број часова активне наставе недељно током семестра/триместра/године				
Предавања	Вежбе	ДОН	Студијски истраживачки рад	Остали часови
2	2			
Методе извођења наставе	Предавања, вежбе, самостални рад студента на изради домаћих задатака и пројеката			
Оцена знања (максимални број поена 100)				
Предиспитне обавезе	поена	Завршни испит		поена
активност у току предавања		писмени испит		
практична настава	70	усмени испит		30
колоквијуми				
семинари				