

Студијски програм : Биоинформатика			
Назив предмета: Data mining за биоинформатичаре			
Наставник/наставници: Милош М. Савић			
Статус предмета: Обавезан			
Број ЕСПБ: 6			
Услов: нема			
Циљ предмета Упознавање студента са основама <i>data mining</i> -а: напредних техника за анализу комплексних података применом метода машинског учења. Оспособљавање студената за њихову имплементацију и практичну примену у реалним проблемима из домена биоинформатике.			
Исход предмета На крају курса, очекује се да успешан студент демонстрира суштинско и свеобухватно разумевање принципа <i>data mining</i> -а кроз критичку анализу, избор, имплементацију и примену техника у реалним и истраживачким проблемима.			
Садржај предмета <i>Теоријска настава</i> Основни <i>data mining</i> концепти и њихови односи са вештачком интелигенцијом, машинским учењем, статистиком, анализом и визуелизацијом података. Типови комплексних података: виско-димензионални табуларни подаци, графови и мреже, временске серије. Алгоритми супервизираних, полусупервизираних и несупервизираних машинског учења. Класификација: алгоритми, мерење перформанси, <i>overfitting</i> . Редукција димензионалности, одабир атрибута. Кластеринг: алгоритми, мерење перформанси. Анализа нумеричких података: предвиђање нумеричког параметра, регресија, неуралне мреже. Учење асоцијација. Трансформација и припрема података за примену <i>data mining</i> техника. <i>Практична настава</i> Примена савладаних техника на проблемима из домена биоинформатике: анализа секвенци генома, анализа стрингова аминокиселина-протеина, анализа података из микронизова (Microarrays), анализа биолошких мрежа (мреже протеинских интеракција, метаболичке мреже, итд.)			
Литература I. H. Witten, E. Frank, M. A. Hall. Data Mining: Practical Machine Learning Tools and Techniques. Morgan Kaufmann Publishers, 2011.			
Број часова активне наставе	Теоријска настава: 2	Практична настава: 0+2+0	
Методe извођења наставе Фронтална настава путем мултимедијалних презентација. Вежбе са индивидуалним радом на рачунару. Поред самог решавања практичних проблема, студенти на бази њихових интересовања у биолошком домену такође учествују и у формирању истраживачких задатака, како би се унапредило њихово активно учествовање у процесу наставе и мотивисало за даљи научно-истраживачки рад.			
Оцена знања (максимални број поена 100)			
Предиспитне обавезе	поена	Завршни испит	поена
активност у току предавања			
практична настава		усмени испт	40
колоквијум-и	60		
семинар-и			
Начин провере знања могу бити различити наведено у табели су само неке опције: (писмени испити, усмени испт, презентација пројекта, семинари итд.....			
*максимална дужна 2 странице А4 формата			