

Студијски програм: Мастер академске студије ФИЗИКА			
Назив предмета: Обрада и анализа великих база података у физици			
Наставник: Душан Јаковетић, Данијела Боберић Крстићев			
Статус предмета: изборни			
Број ЕСПБ: 8			
Услов:			
Циљ предмета			
Упознавање са информационим технологијама које се користе за обраду, складиштење и анализу великих количина података.			
Исход предмета			
Након одслушаног и научног садржаја предмета студент треба да има развијене:			
- Опште способности: Вештине неопходне за решавање практичних и истраживачких „data-driven” проблема ; Коришћење модерних софтверских алата - Предметно-специфичне способности: Способност примена аналитичких метода и софтверских алата <i>Big Data</i> на практичне и истраживачке проблеме у физици; Познавање, разумевање, и ефективна примена практичних технологија и алата за складиштење, обраду, анализу, и визуализацију <i>Big Data</i>; Способност селекције одговарајућег метода машинског учења или оптимизације на изабрани проблем из физике, као и примена и анализа изабраног метода			
Садржај предмета			
<i>Теоријска настава</i>			
Увод у <i>Big Data</i> и примена у физици. Карактеристике <i>Big Data</i> и димензије скалабилности. Технологије за складиштење података.Упознавање са дистрибуираним фајл системом.Технике паралелне обраде података. Map/Reduce алгоритам. Алати за приступ неструктурираним подацима. Основни алати за статистичку анализу. Изабране методе машинског учења за обраду података – регресија, класификација, кластеровање, минимизација функције емпиријског ризика. Изабране методе оптимизације за обраду података – проксимални методи, дистрибуирани градијентни методи. Визуализација података.			
<i>Практична настава</i>			
Упознавање са Hadoop системом. Имплементација Map/Reduce алгоритма за илустративне проблеме. Имплементација изабраних метода машинског учења и оптимизације за изабране проблеме. Програмирање употребом <i>Apache Pig</i> библиотеке.			
Литература			
- Erl, T., Khattak, W., & Buhler, P. (2016). <i>Big Data Fundamentals: Concepts, Drivers & Techniques</i>. Prentice Hall Press. - Lam, C. (2010). <i>Hadoop in action</i>. Manning Publications Co. - Bishop, C. (2006). <i>Pattern recognition and machine learning</i>, Springer.			
Број часова активне наставе		Теоријска настава: 3	Практична настава:2
Методе извођења наставе			
Предавања (3 часа недељно); вежбе (1 час недељно); израда и презентација семинарског рада (1 час недељно). На предавањима се користе класичне методе наставе уз коришћење пројектора. Објашњавају се карактеристике <i>Big Data</i> и илуструју примери различитих информационих технологија за складиштење, обраду, анализу и визуализацију података. На вежбама се у изабраним технологијама решавају конкретни проблеми који илуструју теоријске концепте презентоване на предавањима.			
Оцена знања (максимални број поена 100)			
Предиспитне обавезе	поена	Завршни испит	поена
први колоквијум	15	писмени испит	30
други колоквијум	15	усмени испит	40