ВИСШЕ УЧИЛИЩЕ ПО ТЕЛЕКОМУНИКАЦИИ И ПОЩИ

Факултет: ФАКУЛТЕТ ПО ТЕЛЕКОМУНИКАЦИИ И МЕНИДЖМЪНТ
Катедра: ОБЩО ИНЖЕНЕРСТВО
Професионално направление: 5.13. ОБЩО ИНЖЕНЕРСТВО
Специалност: ДИГИТАЛНА ПУБЛИЧНА АДМИНИСТРАЦИЯ
Образователно-квалификационна степен: БАКАЛАВЪР

ОПИСАНИЕ НА ЛЕКЦИОНЕН КУРС

1. Наименование на курса: *ГОЛЕМИ МАСИВИ ОТ ДАННИ*
2. Код на курса: B.48.047.3
3. Вид на курса: *факултативна*
4. Ниво: *бакалавър*
5. Година на изучаване: *трета*
6. Семестър: *пети*
7. Брой кредити: *три*
8. Име на лектора: *доц. д-р Красимира Иванова*
9. Резултати от обучението за дисциплината:

Целта на курса е студентите да изградят знания и умения за работа с големи масиви от данни, да бъдат анализирани и извлечена информация от тях.

Студентите, приключили обучението си по дисциплината, *трябва да придобият:*

Основни знания в областта на големите масиви от данни.

Практически умения за опериране с големи масиви от данни.

Компетенции: След приключване на курса студентите, успешно завършили курса, трябва да са запознати с проблемите, свързани с оперирането с големи масиви от данни, както и да владеят някои методи за анализ и графично представяне на резултатите от анализи на "големи данни" във вид на диаграми или анимирани изображения с цел опростяване на интерпретацията и по-лесно разбиране на получените резултати.

10. Начин на преподаване:

Начинът на преподаване: обучението по дисциплината „*Големи масиви от данни*" се извършва през пети семестър съгласно действащия учебен план чрез лекции, групирани по един учебен час седмично в продължение на 15 седмици и практически упражнения, групирани по един учебен час седмично в продължение на 15 седмици.

Проектно – базирано обучение, при което студентите развиват компетенциите си за мултимедийна комуникация. Учебният процес е интерактивен и иновативен, при който студентите участват активно.

11. Предварителни изисквания:

Студентите следва да имат предварителни знания по дисциплините: Основи на информационните технологии, Основи на дигиталното предприемачество, Интегрирани системи за управление на бизнес процеси, Системи за управление на иновациите във високоавтоматизирани производства и др. включени в учебния план на специалността.

Курсове, ползващи дисциплината са: Дигитална среда и бизнес, Иновационен мениджмънт на високите технологии.

12. Съдържание на курса:

В дисциплината се изучават: Големи масиви данни (Big Data) - същност и свързани с тях проблеми; Приложения на големите масиви данни в бизнеса; Подходи за съхраняване на големи масиви данни; Разпределени файлови системи. Въведение в Apache™ Hadoop®; Същност и възможности на парадигмата MapReduce; Основни възможности на Apache™ HBase®; Проблеми при обработката на големи масиви от данни; Методи за разпределена обработка на големи масиви данни; Платформи за анализ на големи масиви данни; Основни възможности на Apache™ Spark®; Големи масиви от свързани данни. Графови бази от данни; Основни характеристики и възможности за съхранение на RDF данни; Извличане на закономерности и анализ на големи масиви от данни; Основни алгоритми за анализ на големи масиви от данни; Визуализация на резултатите от анализа на големи масиви данни.

Практическите упражнения включват: Основни задачи, свързани с методите за извличане на знания от големи масиви от данни; Предварителна обработка на данните; Редуциране на размерността в множество от данни чрез избор на характеристики; Дискретизация на непрекъснати данни; Клъстеризация на големи масиви от данни; Построяване на класификационна дървовидна структура; Анализ на данни с често срещащи се набори от елементи; Анализ на пазарна кошница; Построяване на асоциативни правила; Избор на обучаваща извадка и обучение въз основа на нея; Работа с асоциативни правила; Алгоритъм A-Priori; Откриване на подобни елементи;

Филтрация потокови данни и преброяване на елементи в потокови данни; Откриване на закономерности в потокови данни.

13. Библиография

Основна литература:

1. Huadong Guo, Michael F. Goodchild, Alessandro Annoni (Editors) Manual of Digital Earth. Springer Nature Singapore Pte Ltd. Singapore (2020) ISBN 978-981-32-9914-6 ISBN 978-981-32-9915-3 (eBook). https://doi.org/10.1007/978-981-32-9915-3
2. James D. Miller. Big Data Visualization. Packt Publishing Ltd., 2017. Birmingham, UK. ISBN 978-1-78528-194-5. www.packtpub.com; https://futa.edu.ng/csclibrary/books/big-data-visualization.9781785281945.78390.pdf
3. Holden Karau, Andy Konwinski, Patrick Wendell, and Matei Zaharia. Learning Spark. Databricks. O'Reilly Media, Inc., 2015. ISBN 978-1-449-35862-4 http://index-of.co.uk/Big-Data-Technologies/Learning%20Spark%20%20Lightning-Fast%20Big%20Data%20Analysis%20.pdf
4. Иванов, Мартин П. (2016) Съвременни методи за интелигентен анализ на данни. Working Paper. Научен електронен архив на НБУ, София. http://eprints.nbu.bg/3082/1/%D0%98%D0%BD%D1%82%D0%B5%D0%BB%D0%B8%D0%B3%D0%B5%D0%BD%D1%82%D0%B5%D0%BD%20%D0%B0%D0%BD%D0%B0%D0%BB%D0%B8%D0%B7%20%D0%BD%D0%B0%20%D0%B4%D0%B0%D0%BD%D0%BD%D0%B8%20-%20%D0%BF%D1%83%D0%B1%D0%BB%D0%B8%D0%BA%D0%B0%D1%86%D0%B8%D1%8F.pdf

Допълнителна литература:

1. Oracle. Introduction to Very Large Databases. In: Oracle Database Online Documentation 11g Release 1 (11.1) / Data Warehousing and Business Intelligence. ORACLE Help Center (2020). https://docs.oracle.com/cd/B28359_01/server.111/b32024/intro.htm
2. Pete York and Michael Bamberger. Measuring results and impact in the age of big data: The nexus of evaluation, analytics, and digital technology., The Rockefeller Foundation (2020). https://www.rockefellerfoundation.org/wp-content/uploads/Measuring-results-and-impact-in-the-age-of-big-data-by-York-and-Bamberger-March-2020.pdf
3. Bart van der Sloot, Dennis Broeders, Erik Schrijvers (eds.) Exploring the Boundaries of Big Data. WRR, Amsterdam University Press, The Hague/Amsterdam (2016) ISBN 978 94 6298 358 8, e-ISBN 978 90 4853 393 0 (pdf), e-ISBN 978 90 4853 403 6 (ePub)
4. Martin Mois. Apache Hadoop Tutorial. Exelixis Media P.C., (2016) https://www.wrr.nl/binaries/wrr/documenten/verkenningen/2016/04/28/exploring-the-boundaries-of-big-data-32/V032-Exploring-Boundaries-Big-Data.pdf

14. Методи и критерии на оценяване:

Дисциплината завършва с *текуща оценка.* Проверяват се усвоените от студента знания по дисциплината и едновременно с това се оценяват неговите умения да мисли в реално време за решаване на поставени проблеми. Оценката се формира по точкова система. Контролната работа по време на семестъра проверява продуктивните знания на студента, възможностите му да осмисля наученото. Формирането на оценка в курса на обучение на всеки от посочените етапи, съобразно получени точки и подредбата им по шестобалната система.

	Компонент	**Макс. точки**	**Пример 1**	**Пример 2**
1	Първа контролна работа	30	20	16
2	Втора контролна работа	30	12	25
	Крайна оценка:	0,1*60 = 6	Σ =0,1*32 $\approx$ 3	Σ = 0,1*41 $\approx$ 4

15. Език на преподаване: *Български*