**Факултет:** ФАКУЛТЕТ ПО ТЕЛЕКОМУНИКАЦИИ И МЕНИДЖМЪНТ
**Катедра:** ОБЩО ИНЖЕНЕРСТВО
**Професионално направление:** 5.13. ОБЩО ИНЖЕНЕРСТВО
**Специалност:** ДИГИТАЛНО ПРЕДПРИЕМАЧЕСТВО
**Образователно-квалификационна степен:** БАКАЛАВЪР

## ОПИСАНИЕ НА ЛЕКЦИОНЕН КУРС

**1. Наименование на курса:** *ГОЛЕМИ МАСИВИ ОТ ДАННИ*
**2. Код на курса:** *B.48.093.3*
**3. Вид на курса:** *факултативен*
**4. Ниво:** *бакалавър*
**5. Година на изучаване:** *втора*
**6. Семестър:** *трети*
**7. Брой кредити:** 3
**8. Име на лектора:** *Никола Бакалов*
**9. Резултати от обучението за дисциплината:**

*Целта на курса* е да се придобият задълбочени знания за фундаменталните анализиране на големи масиви от данни.

Задачите се свеждат до:

- ✓ Проектиране на системи за управление на „големи масиви от данни“ използвайки Hadoop;
- ✓ Създаване на скриптове за обработка на данни върху Hadoop клъстер използвайки Pig и Spark;
- ✓ Анализиране на нерелационни данни Cassandra и MongoDB;
- ✓ Анализиране на релационни данни използвайки MySQL.

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

*Основни знания* за фундаменталните проблеми и предизвикателства пред анализирането на големи масиви от данни; основните инструменти за анализ и обработка на големи масиви от данни; анализ и обработка на данни от релационни и нерелационни бази от данни.

*Практически умения* за боравене с инструменти за анализ и редакция на големи масиви от данни; публикуване на данни в Hadoop клъстер; Анализ и обработка на данни от релационни и нерелационни бази данни.

*Компетенции:* След приключване на курса *студентите, трябва да са усвоили* основните принципи, модели, методи и подходи за анализ и обработка на големи масиви от данни.

**10. Начин на преподаване:**

Методът на преподаване е „учене на база практика“, като необходимата теория се обяснява в лекциите, по време на които се дават практически примери.

**11. Предварителни изисквания:**

*Студентите следва да имат задълбочени познания* по следните дисциплини: Основи на програмирането – I част, основи на програмирането – II част.

*Курсове, ползващи дисциплината са* Език за управление на бази от данни**.**

**12. Съдържание на курса:**

Дисциплината формира под формата на семинар основни познания за проектиране на системи, които управляват „големи масиви от данни“, използвайки Apache Hadoop и свързани технологии за инженеринг на данни.

Цялото обучение е структурирано под формата на теория и практически задачи с цел учене на база опит.

Задачите по време на обучението се свеждат до:

- ✓ запознаване с основните дейности за анализ на големи масиви от данни;
- ✓ придобиване на знания за обработка и анализ на големи масиви от данни;
- ✓ изграждане на аналитични способности за работа с релационни и нерелационни бази данни;
- ✓ надграждане на уменията чрез управление на клъстера.

**13. Библиография**

***Основна литература:***

**Jeyaraj, R., Pugalendhi, G., Paul, A.** Big Data with Hadoop MapReduce: A Classroom Approach. Apple Academic Press. 2020.

**Kumar, N.** Big data using Hadoop and Hive. Mercury Learning and Information. 2021.

**Prajapati, V.** Big data analytics with R and Hadoop. Packt Publishing. 2013.

***Допълнителна литература:***

**Stokes, M.** Big data Hadoop made simple: A beginner's guide to programming. 2023.

**Bhushan, M.** Big data and Hadoop: Learn by example. BPB Publications. 2018.

**Spivey, B., Echeverria, J.** Hadoop security: Protecting your big data platform. O'Reilly Media. 2015.

**14. Методи и критерии на оценяване:**

Дисциплината завършва с изпит. Постигането на поставената цел на обучението се контролира през семестъра чрез текущ контрол. Оценката се формира чрез точкова скала, като се взема предвид участието в практическите занятия, и индивидуалната работа на студента през семестъра.

*14.1. Изисквания при формиране на оценката по шестобалната система*

Една част от оценка по дисциплината се оформя на базата на точкова система като максималният брой точки, който студентите могат да получат е 100 (таблица 1).

Окончателната оценка се формира от решаване на практически задачи по време на изпитната сесия като под внимание се взема и активността на студентите по време на самия обучителен процес през семестъра.

Тежест на оценяването:

- ✓ Практически изпит от 4 задачи с продължителност 120 минути – носи 100% тежест в крайната оценка. Оценяването на изпита е съобразно следните критерии: 0 верни задачи – Слаб 2; 1 вярна задача – Среден 3; 2 верни задачи – Добър 4; 3 верни задачи – Много добър 5 и 4 верни задачи - Отличен 6).

*14.2. Формиране на крайната оценка*

Формирането на крайната оценка (КО) е на базата на получените резултати от практическия изпит (ПИ), като съответствието между броя точки и оценката по шестобалната система е дадено в *Таблица 1*.

КО = ПИ

Усвоените знания от студентите, изразени в проценти, се оценяват по шестобалната система, както следва:

*Таблица 1*.

| | |
|---|---|
| До 50% | Слаб (2) |
| 51% – 70% | Среден (3) |
| 71% – 80% | Добър (4) |
| 81% – 90% | Мн. добър (5) |
| Над 90% | Отличен (6) |

**15. Език на преподаване:** *Български/Английски*