

### Загальні відомості про дисципліну

Назва дисципліни	Аналітика Big Data
Освітній ступінь	магістр
Галузь знань	12 Інформаційні технології
Спеціальність	124 Системний аналіз
Освітня програма	Інформаційні технології та бізнес-аналітика (Data Science)
Навчальний рік	2021-2022
Семестр	3
Факультет	ФІТ
Курс	2
Підсумковий контроль	екзамен

### Місце дисципліни в освітній програмі

Фахові компетентності	<p>СК 6. Здатність застосовувати теорію і методи Data Science для здійснення інтелектуального аналізу даних з метою виявлення нових властивостей та генерації нових знань про складні системи.</p> <p>СК 11. Здатність ефективно використовувати теорію і методи Data Science.</p> <p>СК 12. Здатність до здійснення процедур дослідження, аналізу, систематизації та обробки великих даних.</p> <p>СК 13. Здатність розробляти і впроваджувати моделі задач інтелектуального аналізу даних засобами комп'ютерного моделювання.</p>
Програмні результати навчання	<p>РН 2. Будувати та досліджувати моделі складних систем і процесів застосовуючи методи системного аналізу, математичного, комп'ютерного та інформаційного моделювання.</p> <p>РН 6. Застосовувати методи машинного навчання та інтелектуального аналізу даних, математичний апарат нечіткої логіки, теорії ігор та розподіленого штучного інтелекту для розв'язання складних задач системного аналізу.</p> <p>РН 8. Здійснювати ідентифікацію та оцінювання параметрів математичних моделей об'єктів керування.</p> <p>РН 12. Розробляти моделі управління даними та знаннями в складних системах.</p> <p>РН 13. Здійснювати інтелектуальний аналіз та обробку великих даних засобами комп'ютерного моделювання.</p>
Передумови вивчення дисципліни	<p><b>знання</b></p> <ul style="list-style-type: none"> <li>• основ інформаційних технологій (операційна система Windows, бази даних, доступ до веб-ресурсів);</li> <li>• основ вищої математики (вектори, матриці, випадкові величини та їх числові характеристики);</li> <li>• основ роботи у середовищі R.</li> </ul> <p><b>вміння</b></p> <ul style="list-style-type: none"> <li>• вільно працювати з офісними додатками Microsoft Word, Microsoft Excel, Microsoft PowerPoint;</li> <li>• проводити завантаження даних з локальних джерел, здійснювати первинну обробку даних в R, графічний аналіз даних з використанням пакетів base та ggplot2.</li> </ul>

### Забезпечення дисципліни

Основні джерела	Data Science & Big Data Analytics: Discovering, Analyzing, Visualizing and Presenting Data /EMC Education Services. Indianapolis : John Wiley & Sons, Inc, 2015. 432 p.
	Prajapati V. Big Data Analytics with R and Hadoop: Packt Publishing, 2013. 238 p.
	Wade R. Advanced Analytics in Power BI with R and Python: Ingesting, Transforming, Visualizing. Apress, 2020. – 440 p.
	Walkowiak S. Big Data Analytics with R: Packt Publishing, 2016. 491 p.
	5. Wickham H. Mastering Shiny: O'Reilly Media, 2021 352 p.

### Аудиторні заняття

Лектор - д.е.н., проф. Роскладка А. А.

№	Тема лекції
1	Концепція Big Data
2	Концепція Open Data

3	Первинна обробка великих даних у середовищі R
4	Графічний аналіз великих даних
5	Створення інтерактивних веб-додатків для аналізу великих даних у R Shiny
6	Реактивне програмування процедур обробки великих даних у Shiny
7	Аналіз великих даних у Power BI
8	Веб-скрапінг та парсинг великих даних
9	Технології MapReduce, Hadoop і Spark в аналітиці великих даних
10	Методи машинного навчання для надвеликих даних
<b>Загальний обсяг лекцій: 20 год</b>	

Викладач - д.е.н., проф. Роскладка А. А.

№	Тема практичного заняття	Бали
1	Аналіз інформаційного забезпечення галузевої аналітики	4
2	Аналіз інформаційного забезпечення галузевої аналітики	4
3	Імпорт та попередній аналіз великих даних	4
4	Методи аналізу різних типів великих даних у R	5
5	Методи роботи з пропущеними даними	5
6	Базові технології візуалізації даних	5
7	Візуальний аналіз взаємозв'язку великих даних	5
8	Передові методи візуалізації великих даних	5
9	Створення односторінкових макетів обробки великих даних	4
10	Створення односторінкових макетів обробки великих даних	4
11	Створення багатопанельних макетів обробки великих даних	5
12	Створення багатопанельних макетів обробки великих даних	5
13	Реактивні функції в аналітиці великих даних у Shiny	5
14	Реактивні функції в аналітиці великих даних у Shiny	5
15	Обробка геоданих з використанням R Shiny	5
16	Обробка геоданих з використанням R Shiny	5
17	Візуалізація великих даних у Power BI за допомогою R	5
18	Створення інформаційної моделі даних Power BI за допомогою R	5
19	Перетворення великих даних у Power BI за допомогою методів R	5
20	Скрапінг та парсинг веб-сторінок з використанням пакету rvest	5
21	Скрапінг та парсинг веб-сторінок з використанням пакету rvest	5
<b>Загальний обсяг практичних занять 42 год</b>		<b>100</b>

#### Політика дисципліни

Відвідування занять	Відвідування лекційних та практичних занять з дисципліни є обов'язковим для всіх студентів
---------------------	--

Відпрацювання пропущених занять	Студент, який пропустив практичне заняття, самостійно вивчає матеріал за наведеними в силабусі джерелами, виконує завдання і здає його викладачу. За умови неповажної причини пропуску заняття, оцінка за практичне завдання буде знижена.
---------------------------------	--

Допуск до екзамену	Згідно з Положенням про організацію освітнього процесу всі студенти допускаються до іспиту
Підсумкова модульна оцінка	Підсумкова модульна оцінка за семестр є сумою оцінок, отриманих студентом за виконання практичних завдань. Максимальна модульна оцінка становить 100 балів. Студент, який отримав підсумкову модульну оцінку менше за 20 балів, при будь-якій екзаменаційній оцінці не може отримати задовільну підсумкову оцінку з дисципліни і буде ліквідувати академічну заборгованість під час додаткової сесії.
Екзаменаційна оцінка	Максимальна екзаменаційна оцінка становить 100 балів
Підсумкова оцінка з дисципліни	Підсумкова оцінка з дисципліни обчислюється як середнє арифметичне підсумкової модульної та екзаменаційної оцінки.