



Приватний вищий навчальний заклад
«Буковинський університет»

Факультет інформаційних технологій та економіки
Кафедра комп'ютерних систем і технологій

СХВАЛЕНО
на засіданні науково-методичної
ради факультету
протокол № 1 від 29 серпня 2024 р.



СИЛАБУС

обов'язкової навчальної дисципліни

«Технології Big Data»

Освітньо-професійна програма:	<u>Комп'ютерні науки</u>
Спеціальність:	<u>122 «Комп'ютерні науки»</u>
Галузь знань:	<u>12 Інформаційні технології</u>
Рівень вищої освіти:	<u>другий (магістерський)</u>
Факультет:	<u>Інформаційних технологій та економіки</u>
Мова навчання:	<u>українська</u>
Розробник:	<u>Артеменко Ольга Іванівна, кандидат технічних наук, доцент</u>
Профайл викладача	<u>https://bukuniver.edu.ua/university/faculties-and-departments/ite-faculty/department-of-computer-systems-and-technologies/</u>
E-mail:	<u>olgaartemenko@bukuniver.edu.ua</u>
Консультації	<u>четвер з 10.00 до 16.00</u>

1. Анотація дисципліни (призначення навчальної дисципліни).

Дисципліна «Технології Big Data» формує у здобувачів необхідний обсяг теоретичних і практичних знань про великі дані Big Data, методи та засоби їхнього аналізу та обробки, основних принципів побудови систем управління великими даними. Виявлення складу та вивчення можливостей програмного забезпечення сучасних систем, які використовуються для збереження та обробки великих даних в інформаційних системах.

Після вивчення дисципліни студенти повинні:

Знати:

- основні поняття та визначення в теорії Big Data;
- моделі та методи аналізу великих даних;
- сучасні програмні засоби для проектування і розробки систем аналізу великих даних;
- методологію та методи оцінки достовірності отриманих результатів.

Вміти:

- формулювати задачу обробки великих інформаційних масивів відповідно до задачі, що вирішується;
- формувати бази знань (сховища даних) з урахування специфіки предметної галузі досліджень;
- розробляти структуру інформаційної моделі об'єкта дослідження;
- будувати інформаційну модель (модель інформаційних потоків) об'єкту або процесу, який вивчається;
- виконати реструктуризацію моделі або, при необхідності, її декомпозицію;
- оцінити рівень достовірності отриманих результатів;
- модифікувати побудовані інформаційні моделі у відповідності до вимог і цілей замовника досліджень з метою створення спектру моделей;
- використовувати існуючі на ринку інструментальні та програмно-алгоритмічні засоби, що можуть бути використані для вирішення задач обробки великих даних.

2. Мета та завдання вивчення дисципліни:

Мета вивчення дисципліни: формування у слухачів знань, вмінь та навичок аналізу даних, застосування відповідних інформаційних технологій для обробки даних.

Завдання:

- сформувати знання та отримати практичні навички для використання інформаційних технологій аналізу великих даних при вирішенні задач аналізу даних, розпізнавання образів, пошуку знань та ін;
- огляд методів, програмних продуктів та різних інструментальних засобів, що використовуються для побудови інтелектуальних систем аналізу великих даних.

3. Пререквізити.

Знання за програмою єдиного фахового вступного випробування за спеціальністю.

4. Компетентності та результати навчання

Інтегральна компетентність

Здатність розв'язувати задачі дослідницького та/або інноваційного характеру у сфері комп'ютерних наук.

Загальні компетентності (ЗК)

ЗК02. Здатність застосовувати знання у практичних ситуаціях.

ЗК07. Здатність генерувати нові ідеї (креативність).

Спеціальні (фахові) компетентності:

СК03. Здатність використовувати математичні методи для аналізу формалізованих моделей предметної області.

СК04. Здатність збирати і аналізувати дані (включно з великими), для забезпечення якості прийняття проектних рішень.

СК07. Здатність розробляти програмне забезпечення відповідно до сформульованих вимог з урахуванням наявних ресурсів та обмежень.

СК09. Здатність розробляти та адмініструвати бази даних та знань.

СК12. Здатність застосовувати методи, моделі та алгоритми для виявлення прихованих корисних закономірностей у великих наборах даних щодо вирішення задач прогнозування, керування, класифікації та інтелектуального аналізу.

Отримуються наступні програмні результати навчання:

РН8. Розробляти математичні моделі та методи аналізу даних (включно з великим).

РН9. Розробляти алгоритмічне та програмне забезпечення для аналізу даних (включно з великими).

РН12. Проектувати та супроводжувати бази даних та знань.

РН17. Виявляти та усувати проблемні ситуації в процесі експлуатації програмного забезпечення, формулювати завдання для його модифікації або реінжинірингу.

РН20. Виявляти, ідентифікувати, класифікувати закономірності в великих наборах даних та ефективно їх використовувати в міжгалузевих задачах.

5. Зміст навчальної дисципліни

Змістовий модуль 1.

Тема 1. Вступ в аналіз і обробку великих даних.

Тема 2. HADOOP. Архітектура HADOOP. Робота з HDFS.

Тема 3. Парадигма map-reduce. Модель обчислень. Map, Shuffle и Reduce фази. Опис job, task, розбиття операцій на задачі.

Тема 4. Життєвий цикл проєкту з аналізу Big Data.

Змістовий модуль 2.

Тема 5. Основні технології та інструменти для аналізу великих даних.

Тема 6. Аналіз і візуалізація великих даних.

Тема 7. Базові методи аналізу даних з використанням R.

Тема 8. Наукові проблеми аналізу великих даних у складних соціально-економічних системах для досягнення цілей сталого розвитку.

6. Методи навчання. Система контролю та оцінювання результатів навчання

Методи навчання:

- вербальні методи (лекція, диспут, пояснення, розповідь);
- практичні методи (практичні роботи);
- наочні методи (демонстрація, ілюстрація);
- робота з інформаційними ресурсами: з навчально-методичною, науковою, нормативною літературою та інтернет-ресурсами;
- самостійна робота над індивідуальним завданням або за програмою навчальної дисципліни;
- метод мозкового штурму;
- кейс-технології.

Форми та методи оцінювання:

- усне опитування;
- тестування;
- презентація результатів виконаних завдань;
- аналітичні звіти;
- розв'язування практичних ситуацій;

Підсумковий контроль знань проводиться у формі екзамена.

На екзамен виносяться вузлові питання програми та типові задачі, що потребують творчої відповіді та уміння синтезувати отримані знання і застосувати їх при вирішенні практичних задач. Максимально можлива оцінка за екзамен 30 балів.

Критерії оцінювання результатів навчання з навчальної дисципліни

Шкала оцінювання: національна та ECTS

СУМА БАЛІВ	ОЦІНКА ECTS	ОЦІНКА ЗА НАЦІОНАЛЬНОЮ ШКАЛОЮ	
		екзамен	залік
90-100	A	відмінно	відмінно
82-89	B	добре	добре

74-81	C		
64-73	D	задовільно	задовільно
60-63	E		
35-59	FX	незадовільно	незадовільно
1-34	F		

Розподіл балів з навчальної дисципліни

Поточний контроль		Іспит	Загальна кількість балів
Модуль I	Модуль II		
35	35	30	100

7. Література

1. Гороховатський В.О., Творошенко І.С. Методи інтелектуального аналізу та оброблення даних: навч. посібник. – Харків: ХНУРЕ, 2021. – 92 с.
2. Майборода Р. Комп’ютерна статистика. Київ : ВПЦ «Київський університет», 2019. 589 с. URL: <http://probability.univ.kiev.ua/userfiles/mre/cscolor.pdf>
3. Майборода Р. Багатовимірна дескриптивна статистика. Методичні рекомендації по курсу. Київ : КНУ ім. Тараса Шевченка, 2020. 35 с. URL: <https://probability.knu.ua/userfiles/mre/descrtasks.pdf>
4. Zgurovsky M.Z. Big Data: Conceptual Analysis and Applications. / M.Z. Zgurovsky, Y.P. Zaychenko // Springer, 2020, 298 p.
5. Rajkumar Buyya. Big Data. Principles and Paradigms. — Elsevier, 2016. – 496p.

Інформаційні ресурси в інтернеті

1. TextMining with R. Share and Discover Knowledge on SlideShare. URL: <http://www.slideshare.net/whitish/textmining-with-r> (date of access: 1.09.2024).
2. R: The R Project for Statistical Computing. R: The R Project for Statistical Computing. URL: <https://www.r-project.org> (date of access: 1.09.2024).
3. Package "timeSeries". The Comprehensive R Archive Network. URL: <https://cran.r-project.org/web/packages/timeSeries/timeSeries.pdf> (date of access: 1.09.2024).
4. Create Elegant Data Visualisations Using the Grammar of Graphics. Create Elegant Data Visualisations Using the Grammar of Graphics • ggplot2. URL: <https://ggplot2.tidyverse.org/> (date of access: 1.09.2024).