



ДОБУВАННЯ ДАНИХ

Викладач: к. т. н., ст. викл. кафедри комп'ютерних наук, Добровольський Геннадій Анатолійович

Кафедра: кафедра комп'ютерних наук, I корпус, ауд. 39

E-mail: gen.dobr@gmail.com

Телефон: (061) 289-12-57

Інші засоби зв'язку: Telegram: https://t.me/gen_dobr

Освітня програма, рівень вищої освіти:		Інформаційні системи та технології Магістр					
Статус дисципліни:		Вибіркова					
Кредити ECTS	3	Навч. рік:	2023-24	Рік навчання	2	Тижні	12
Кількість годин	90	Кількість змістових модулів¹	4	Лекційні заняття – 24 Лабораторні заняття – 24 Самостійна робота – 42			
Вид контролю:		Залік					
Посилання на курс в Moodle			https://moodle.znu.edu.ua/course/view.php?id=15093				
Консультації:							

ОПИС КУРСУ

В межах дисципліни «Добування даних» вивчаються методи виявлення у даних раніше невідомих, нетривіальних, практично корисних та доступних для інтерпретації знань, необхідних для прийняття рішень. Акцент зроблено на MapReduce і Spark - інструментах для створення паралельних алгоритмів, які можуть обробляти дуже великі обсяги даних.

Метою вивчення дисципліни «Добування даних» є оволодіння знаннями методів виявлення структури та закономірностей у даних, набуття вмінь та навичок ефективного застосування методів виявлення структури та закономірностей для виявлення у даних раніше невідомих, нетривіальних, практично корисних та доступних для інтерпретації знань, необхідних для прийняття рішень.

Основними завданнями вивчення дисципліни «Добування даних» є:

- оволодіння базовими теоретичними знаннями щодо методів виявлення структури та закономірностей у даних, аналізу графів, аналізу потоків даних;
- набуття вмінь та навичок розвідувального аналізу даних, зниження розмірності даних;
- набуття вмінь та навичок зниження розмірності даних;
- набуття вмінь та навичок аналізу соціальних графів;
- набуття вмінь та навичок аналізу потоків даних;
- набуття вмінь та навичок виявлення груп у даних;
- набуття навичок побудови рекомендаційних систем.

У результаті вивчення навчальної дисципліни студент повинен набути таких результатів навчання (знання, уміння тощо)

- РН 8 Розробляти моделі інформаційних процесів та систем різного класу, використовувати методи моделювання, формалізації, алгоритмізації та реалізації моделей з використанням сучасних комп'ютерних засобів.

¹ 1 змістовий модуль = 15 годин (0,5 кредита ECTS)



- РН 9 Розробляти і використовувати сховища даних, здійснювати аналіз даних для підтримки прийняття рішень.
- РН 12 Розробляти та використовувати методи штучного інтелекту.

та компетентностей:

- ЗК 1 Здатність до абстрактного мислення, аналізу та синтезу.
- ЗК 5 Здатність оцінювати та забезпечувати якість виконуваних робіт.
- СК 4 Здатність розробляти математичні, інформаційні та комп'ютерні моделі об'єктів і процесів інформатизації.
- СК 5 Здатність використовувати сучасні технології аналізу даних для оптимізації процесів в інформаційних системах.
- СК 8 Здатність розробляти та використовувати методи штучного інтелекту для аналізу даних і прогнозування в інформаційних системах.
- СК 9 Здатність використовувати сучасні технології візуалізації даних.

ОЧІКУВАНІ РЕЗУЛЬТАТИ НАВЧАННЯ

У результаті вивчення дисципліни студенти повинні:

знати:

метод випадкових проекцій;
методи зниження розмірності даних;
методи пошуку асоціативних правил;
методи кластеризації;
методи аналізу зв'язків;
методи аналізу потоків даних;
побудови рекомендаційних систем.

вміти:

використовувати MapReduce для обробки даних;
використовувати Apache Spark для обробки даних;
виявляти прості структури та закономірності у даних;
проекувати рекомендаційні системи.

ОСНОВНІ НАВЧАЛЬНІ РЕСУРСИ

Презентації лекцій, методичні рекомендації до виконання лабораторних робіт та індивідуального завдання, тести у системі Moodle

КОНТРОЛЬНІ ЗАХОДИ

Поточні контрольні заходи

Поточний контроль передбачає такі **теоретичні** завдання:

- усне опитування з теоретичного матеріалу за темою на початку кожного лабораторного заняття. Перелік питань з кожної лабораторної роботи розміщено у файлі з завданням до лабораторної роботи у системі Moodle.
- поточний тест за пройденим матеріалом.

Поточний контроль передбачає таке **практичне** завдання:

- виконання лабораторних робіт.

Підсумкові контрольні заходи:

Індивідуальне завдання – Чат-бот.

Підсумковий тест – підсумкове тестування з курсу у системі Moodle.

ЗАПОРІЗЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ
МАТЕМАТИЧНИЙ ФАКУЛЬТЕТ
Силабус навчальної дисципліни



Контрольний захід		Термін виконання	% від загальної оцінки
Поточний контроль (max 60%)			
Змістовий модуль 1	Лабораторна робота 1	Тиждень 1, 2, 3	10
Змістовий модуль 2	Лабораторна робота 2	Тиждень 4, 5, 6	10
	Поточний тест №1	Тиждень 6	10
Змістовий модуль 3	Лабораторна робота 3	Тиждень 7, 8, 9	10
Змістовий модуль 4	Лабораторна робота 4	Тиждень 10, 11, 12	10
	Поточний тест №2	Тиждень 12	10
Підсумковий контроль (max 40%)			
Підсумковий тест			20
Індивідуальне завдання			20
Разом			100%

Шкала оцінювання: національна та ECTS

За шкалою ECTS	За шкалою університету	За національною шкалою	
		Екзамен	Залік
A	90 – 100 (відмінно)	5 (відмінно)	Зараховано
B	85 – 89 (дуже добре)	4 (добре)	
C	75 – 84 (добре)		
D	70 – 74 (задовільно)	3 (задовільно)	
E	60 – 69 (достатньо)		
FX	35 – 59 (незадовільно – з можливістю повторного складання)	2 (незадовільно)	Не зараховано
F	1 – 34 (незадовільно – з обов'язковим повторним курсом)		

РОЗКЛАД КУРСУ ЗА ТЕМАМИ І КОНТРОЛЬНІ ЗАВДАННЯ

Тиждень і вид заняття	Тема заняття	Контрольне завдання	Кількість балів
Змістовий модуль 1.			
Тиждень 1 Лекція 1	Визначення, основні задачі Data Mining, приклади застосування.		
Тиждень 2 Лекція 2	Apache SparkPython pandas, очищення даних		
Тиждень 3 Лекція 3	Розвідувальний аналіз даних		
Тиждень 1, 2, 3 Лабораторна робота 1	Розвідувальний аналіз даних	Усне опитування з теми. Виконання завдань лабораторної роботи	10
Змістовий модуль 2.			
Тиждень 4 Лекція 4	Зменшення розмірності даних		
Тиждень 5	Пошук схожих об'єктів		

ЗАПОРІЗЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ
МАТЕМАТИЧНИЙ ФАКУЛЬТЕТ
Силабус навчальної дисципліни



Лекція 5			
Тиждень 6 Лекція 6	Кластеризація даних		
Тиждень 4, 5, 6 Лабораторна робота 2	Квантування даних	Усне опитування з теми. Виконання завдань лабораторної роботи	10
Тиждень 6 Поточний тест			10
Змістовий модуль 3.			
Тиждень 7, 8 Лекція 7	Аналіз зв'язків між об'єктами		
Тиждень 9 Лекція 8	Аналіз соціальних мереж		
Тиждень 7, 8, 9 Лабораторна робота 4	Пошук спільнот	Усне опитування з теми. Виконання завдань лабораторної роботи	10
Змістовий модуль 4.			
Тиждень 10, 11, 12 Лекція 9	Аналіз послідовностей		
Тиждень 10, 11, 12 Лабораторна робота 5	Фільтр Блума	Усне опитування з теми. Виконання завдань лабораторної роботи	10
Тиждень 12 Поточний тест			10

ОСНОВНІ ДЖЕРЕЛА

1. Leskovec J., Rajaraman A., Ullman J. Mining of Massive Datasets. 3rd Edition. Cambridge: Cambridge University Press, 2020. 603 p. URL: <http://www.mmms.org/>
2. Марченко О.О., Россада Т.В. Актуальні проблеми Data Mining: Навчальний посібник для студентів факультету комп'ютерних наук та кібернетики. Київ. 2017. 150 с.
3. Акіменко В.В. Прикладні задачі інтелектуального аналізу даних (DATA MINING). Київ: КНУ ім. Тараса Шевченка, 2018. 152 с.

ДОДАТКОВІ ДЖЕРЕЛА

4. Кононова К. Ю. Машинне навчання: методи та моделі. Харків : ХНУ імені В. Н. Каразіна, 2020. 301 с.
5. Гладун А. Я., Рогушина Ю. В. Data Mining: пошук знань в даних. Київ: АДЕФ-Україна, 2016. 452 с.
6. Могильний С. Б. Машинне навчання з використанням мікрокомп'ютерів. Київ : Національний центр «Мала академія наук України». 2019. 226 с.
7. Sarkar D., Bali R., Sharma T. Practical Machine Learning with Python. New York: Apress, 2018. 545 p.
8. Хенрик Бринк & Джозеф Ричардс & Марк Феверолф, Машинное обучение: Питер, 2017, 338 с.

ІНФОРМАЦІЙНІ РЕСУРСИ

9. CS246: Mining Massive Data Sets. Stanford. URL: <http://web.stanford.edu/class/cs246/>



РЕГУЛЯЦІЯ І ПОЛІТИКИ КУРСУ

Відвідування занять. Регуляція пропусків.

Відвідування усіх занять є обов'язковим. Студенти зобов'язані дотримуватися усіх строків, визначених для виконання усіх видів робіт, передбачених даною дисципліною. Пропуски та запізнення на заняття є недопустимими.

Політика академічної доброчесності

Кожний студент зобов'язаний дотримуватися принципів академічної доброчесності. Письмові завдання з використанням часткових або повнотекстових запозичень з інших робіт без зазначення авторства – це *плагіат*. Використання будь-якої інформації (текст, фото, ілюстрації тощо) мають бути правильно процитовані з посиланням на автора! Якщо ви не впевнені, що таке плагіат, фабрикація, фальсифікація, порадьтеся з викладачем. До студентів, у роботах яких буде виявлено списування, плагіат чи інші прояви недоброчесної поведінки можуть бути застосовані різні дисциплінарні заходи (див. посилання на Кодекс академічної доброчесності ЗНУ в додатку до силабусу). Неприпустиме складання роботи, виконаної іншою особою.

Використання комп'ютерів/телефонів на занятті

Використання мобільних телефонів, ноутбуків та інших гаджетів під час лекційних та лабораторних занять дозволяється виключно у навчальних цілях (з активованим режимом «без звуку»).

Комунікація

Комунікація викладача зі студентами здійснюється безпосередньо на заняттях та додатково за допомогою месенджерів (наприклад, Telegram), електронної пошти і в СЕЗН Moodle (форум курсу, приватні повідомлення)



ДОДАТОК ДО СИЛАБУСУ ЗНУ – 2023-2024 рр.

ГРАФІК ОСВІТНЬОГО ПРОЦЕСУ 2023-2024 н. р. доступний за адресою: <https://tinyurl.com/yckze4jd>.

АКАДЕМІЧНА ДОБРОЧЕСНІСТЬ. Студенти і викладачі Запорізького національного університету несуть персональну відповідальність за дотримання принципів академічної доброчесності, затверджених **Кодексом академічної доброчесності ЗНУ**: <https://tinyurl.com/ya6yk4ad>. Декларація академічної доброчесності здобувача вищої освіти (додається в обов'язковому порядку до письмових кваліфікаційних робіт, виконаних здобувачем, та засвідчується особистим підписом): <https://tinyurl.com/y6wzzlu3>.

НАВЧАЛЬНИЙ ПРОЦЕС ТА ЗАБЕЗПЕЧЕННЯ ЯКОСТІ ОСВІТИ. Перевірка набутих студентами знань, навичок та вмінь (атестації, заліки, іспити та інші форми контролю) є невід'ємною складовою системи забезпечення якості освіти і проводиться відповідно до Положення про організацію та методіку проведення поточного та підсумкового семестрового контролю навчання студентів ЗНУ: <https://tinyurl.com/y9tve4lk>.

ПОВТОРНЕ ВИВЧЕННЯ ДИСЦИПЛІН, ВІДРАХУВАННЯ. Наявність академічної заборгованості до 6 навчальних дисциплін (в тому числі проходження практики чи виконання курсової роботи) за результатами однієї екзаменаційної сесії є підставою для надання студенту права на повторне вивчення зазначених навчальних дисциплін. Порядок повторного вивчення визначається Положенням про порядок повторного вивчення навчальних дисциплін та повторного навчання у ЗНУ: <https://tinyurl.com/y9pkmmp5>. Підстави та процедури відрахування студентів, у тому числі за невиконання навчального плану, регламентуються Положенням про порядок переведення, відрахування та поновлення студентів у ЗНУ: <https://tinyurl.com/y cds57la>.

НЕФОРМАЛЬНА ОСВІТА. Порядок зарахування результатів навчання, підтверджених сертифікатами, свідоцтвами, іншими документами, здобутими поза основним місцем навчання, регулюється Положенням про порядок визнання результатів навчання, отриманих у неформальній освіті: <https://tinyurl.com/y8ggt4xs>.

ВИРІШЕННЯ КОНФЛІКТІВ. Порядок і процедури врегулювання конфліктів, пов'язаних із корупційними діями, зіткненням інтересів, різними формами дискримінації, сексуальними домаганнями, міжособистісними стосунками та іншими ситуаціями, що можуть виникнути під час навчання, регламентуються Положенням про порядок і процедури вирішення конфліктних ситуацій у ЗНУ: <https://tinyurl.com/57wha734>. Конфліктні ситуації, що виникають у сфері стипендіального забезпечення здобувачів вищої освіти, вирішуються стипендіальними комісіями факультетів, коледжів та університету в межах їх повноважень, відповідно до: Положення про порядок призначення і виплати академічних стипендій у ЗНУ: <https://tinyurl.com/yd6bq6p9>; Положення про призначення та виплату соціальних стипендій у ЗНУ: <https://tinyurl.com/y9r5dpwh>.

ПСИХОЛОГІЧНА ДОПОМОГА. Телефон довіри практичного психолога Марті Ірини Вадимівни (061)228-15-84, (099)253-78-73 (щоденно з 9 до 21).



УПОВНОВАЖЕНА ОСОБА З ПИТАНЬ ЗАПОБІГАННЯ ТА ВИЯВЛЕННЯ КОРУПЦІЇ Запорізького національного університету: **Борисов Костянтин Борисович**
Електронна адреса: uv@znu.edu.ua Гаряча лінія: Тел. [\(061\) 228-75-50](tel:0612287550)

РІВНІ МОЖЛИВОСТІ ТА ІНКЛЮЗИВНЕ ОСВІТНЄ СЕРЕДОВИЩЕ. Центральні входи усіх навчальних корпусів ЗНУ обладнані пандусами для забезпечення доступу осіб з інвалідністю та інших маломобільних груп населення. Допомога для здійснення входу у разі потреби надається черговими охоронцями навчальних корпусів. Якщо вам потрібна спеціалізована допомога, будь-ласка, зателефонуйте (061) 228-75-11 (начальник охорони). Порядок супроводу (надання допомоги) осіб з інвалідністю та інших маломобільних груп населення у ЗНУ: <https://tinyurl.com/ydhcsagx>.

РЕСУРСИ ДЛЯ НАВЧАННЯ. Наукова бібліотека: <http://library.znu.edu.ua>. Графік роботи абонементів: понеділок – п`ятниця з 08.00 до 16.00; вихідні дні: субота і неділя.

ЕЛЕКТРОННЕ ЗАБЕЗПЕЧЕННЯ НАВЧАННЯ (MOODLE):

<https://moodle.znu.edu.ua>

Якщо забули пароль/логін, направте листа з темою «Забув пароль/логін» за адресою: moodle.znu@znu.edu.ua.

У листі вкажіть: прізвище, ім'я, по-батькові українською мовою; шифр групи; електронну адресу.

Якщо ви вказували електронну адресу в профілі системи Moodle ЗНУ, то використовуйте посилання для відновлення паролю <https://moodle.znu.edu.ua/mod/page/view.php?id=133015>.

Центр інтенсивного вивчення іноземних мов: <http://sites.znu.edu.ua/child-advance/>

Центр німецької мови, партнер Гете-інституту:

<https://www.znu.edu.ua/ukr/edu/ocznu/nim>

Школа Конфуція (вивчення китайської мови): <http://sites.znu.edu.ua/confucius>