

ІНТЕЛЕКТУАЛЬНИЙ АНАЛІЗ ДАНИХ

Викладач: кандидат технічних наук, доцент Добровольський Геннадій Анатолійович

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

Е-mail: gen.dobr@gmail.com

Телефон: (061) 289-12-57

Інші засоби зв'язку: Moodle (форум курсу, приватні повідомлення), Telegram:

https://t.me/gen_dobr

Освітня програма, рівень вищої освіти:		Інформаційні системи та технології Бакалавр					
Статус дисципліни:		Вибіркова					
Кредити ECTS	5	Навч. рік:	2022-23	Рік навчання	3	Тижні	14
Кількість годин	150	Кількість змістових модулів¹	8	Лекційні заняття – 28 Лабораторні заняття – 28 Самостійна робота – 94			
Вид контролю:		Залік					
Посилання на курс в Moodle			https://moodle.znu.edu.ua/course/view.php?id=13793				
Консультації:							

ОПИС КУРСУ

В межах дисципліни «Інтелектуальний аналіз 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 «Інтелектуальний аналіз 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 «Інтелектуальний аналіз даних» є вивчення методів, виявлення груп, аналізу зв'язків, аналізу потоків даних:

- 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) та компетентностей:

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



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;
- проектувати рекомендаційні системи.

Згідно з вимогами освітньо-професійної програми студенти повинні досягти таких **компетенцій**:

- ІК Здатність розв'язувати складні спеціалізовані задачі та практичні проблеми в області інформаційних систем та технологій, або в процесі навчання, що характеризуються комплексністю та невизначеністю умов, які потребують застосування теорій та методів інформаційних технологій.
- КЗ 1. Здатність до абстрактного мислення, аналізу та синтезу;
- КЗ 2. Здатність застосовувати знання у практичних ситуаціях;
- КЗ 3. Здатність до розуміння предметної області та професійної діяльності;
- КЗ 5. Здатність вчитися і оволодівати сучасними знаннями;
- КЗ 6. Здатність до пошуку, оброблення та узагальнення інформації з різних джерел;
- КЗ 8. Здатність оцінювати та забезпечувати якість виконуваних робіт;
- КС 1. Здатність аналізувати об'єкт проектування або функціонування та його предметну область;
- КС 13. Здатність проводити обчислювальні експерименти, порівнювати результати експериментальних даних і отриманих рішень;
- КС 17. Здатність розробляти та використовувати методи машинного навчання для аналізу даних і прогнозування в інформаційних системах.

та наступних **програмних результатів навчання**:

- ПР 1. Знати лінійну та векторну алгебру, диференціальне та інтегральне числення, теорію функцій багатьох змінних, теорію рядів, диференціальні рівняння для функції однієї та багатьох змінних, операційне числення, теорію ймовірностей та математичну статистику в обсязі, необхідному для розробки та використання інформаційних систем, технологій та інфокомунікацій, сервісів та інфраструктури організації.
- ПР 2. Застосовувати знання фундаментальних і природничих наук, системного аналізу та технологій моделювання, стандартних алгоритмів та дискретного аналізу при розв'язанні задач проектування і використання інформаційних систем та технологій.
- ПР 15. Розробляти та використовувати моделі машинного навчання для обробки даних і прогнозування в інформаційних системах.

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 – розробка рекомендаційної системи.

Підсумковий тест – підсумкове тестування з курсу за обмежений час у системі Moodle.

Контрольний захід		Термін виконання	% від загальної оцінки
Поточний контроль (max 60%)			
Змістовий модуль 1 (розділ 1)	Опитування з теоретичного матеріалу	Тиждень 1	2
	Лабораторна робота 1	Тиждень 1	4
Змістовий модуль 2 (розділ 1)	Опитування з теоретичного матеріалу	Тиждень 2, 3	2
	Лабораторна робота 2	Тиждень 2, 3	4
Змістовий модуль 3 (розділ 1)	Опитування з теоретичного матеріалу	Тиждень 4, 5	2
	Лабораторна робота 3	Тиждень 4, 5	4
Змістовий модуль 4 (розділ 1)	Опитування з теоретичного матеріалу	Тиждень 6, 7	2
	Лабораторна робота 4	Тиждень 6, 7	4
	Поточний тест	Тиждень 7	6
Змістовий модуль 5 (розділ 2)	Опитування з теоретичного матеріалу	Тиждень 8, 9	2
	Лабораторна робота 5	Тиждень 8, 9	4
Змістовий модуль 6 (розділ 2)	Опитування з теоретичного матеріалу	Тиждень 10,11	2
	Лабораторна робота 6	Тиждень 10,11	4

ЗАПОРІЗЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ
МАТЕМАТИЧНИЙ ФАКУЛЬТЕТ
Силабус навчальної дисципліни



Змістовий модуль 7 (розділ 2)	Опитування теоретичного матеріалу з	Тиждень 12, 13	2
	Лабораторна робота 7	Тиждень 12, 13	4
Змістовий модуль 8 (розділ 2)	Опитування теоретичного матеріалу з	Тиждень 14	2
	Лабораторна робота 8	Тиждень 14	4
	Поточний тест	Тиждень 14	6
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. MapReduce.		
Тиждень 1 Лабораторна робота 1	MapReduce	Усне опитування з теми. Виконання завдань лабораторної роботи	6
Змістовий модуль 2.			
Тиждень 2 Лекція 2	Apache Spark. ч1		
Тиждень 3 Лекція 3	Apache Spark. ч2		
Тиждень 2, 3 Лабораторна робота 2	Apache Spark	Усне опитування з теми. Виконання завдань лабораторної роботи	6
Змістовий модуль 3.			

ЗАПОРІЗЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ
МАТЕМАТИЧНИЙ ФАКУЛЬТЕТ
Силабус навчальної дисципліни



Тиждень 4 Лекція 4	Зменшення розмірності даних		
Тиждень 5 Лекція 5	Пошук схожих об'єктів		
Тиждень 4, 5 Лабораторна робота 3	Очищення даних	Усне опитування з теми. Виконання завдань лабораторної роботи	6
Змістовий модуль 4.			
Тиждень 6 Лекція 6	Кластеризація даних		
Тиждень 7 Лекція 7	Пошук груп об'єктів		
Тиждень 6, 7 Лабораторна робота 4	Пошук асоціативних правил	Усне опитування з теми. Виконання завдань лабораторної роботи	6
Тиждень 7 Поточний тест		Тестові завдання в системі Moodle	6
Змістовий модуль 5.			
Тиждень 8 Лекція 8	Рекомендаційні системи, ч1		
Тиждень 9 Лекція 9	Рекомендаційні системи, ч2		
Тиждень 8, 9 Лабораторна робота 5	Рекомендаційна система	Усне опитування з теми. Виконання завдань лабораторної роботи	6
Змістовий модуль 6.			
Тиждень 10 Лекція 10	Аналіз зв'язків між об'єктами, ч1		
Тиждень 10 Лекція 10	Аналіз зв'язків між об'єктами, ч2		
Тиждень 10, 11 Лабораторна робота 6	Дослідження випадкового графа	Усне опитування з теми. Виконання завдань лабораторної роботи	6
Змістовий модуль 7.			
Тиждень 12 Лекція 12	Аналіз соціальних мереж. ч1		
Тиждень 13 Лекція 13	Аналіз соціальних мереж. ч2		
Тиждень 12, 13 Лабораторна робота 7	Пошук спільнот	Усне опитування з теми. Виконання завдань лабораторної роботи	6
Змістовий модуль 8.			
Тиждень 14	Дослідження потоків даних		



Лекція 14			
Тиждень 14 Лабораторна робота 8	Фільтр Блума	Усне опитування з теми. Виконання завдань лабораторної роботи	6
Тиждень 14 Поточний тест		Тестові завдання в системі Moodle	6

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.mmids.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. Гладун А. Я., Рогущина Ю. В. Data Mining: пошук знань в даних. Київ: АДЕФ-Україна, 2016. 452 с.
4. Акіменко В.В. Прикладні задачі інтелектуального аналізу даних (DATA MINING). Київ: КНУ ім. Тараса Шевченка, 2018. 152 с.

ДОДАТКОВІ ДЖЕРЕЛА

1. Кононова К. Ю. Машинне навчання: методи та моделі. Харків : ХНУ імені В. Н. Каразіна, 2020. 301 с.
2. Могильний С. Б. Машинне навчання з використанням мікрокомп'ютерів. Київ : Національний центр «Мала академія наук України». 2019. 226 с.
3. Sarkar D., Bali R., Sharma T. Practical Machine Learning with Python. New York: Apress, 2018. 545 p.
4. Хенрик Бринк & Джозеф Ричардс & Марк Феверолф, Машинное обучение: Питер, 2017, 338 с.

ІНФОРМАЦІЙНІ РЕСУРСИ

1. CS246: Mining Massive Data Sets. Stanford. URL: <http://web.stanford.edu/class/cs246/>.
2. Видобуток даних: процес, методи та основні проблеми аналізу даних. URL: <https://uk.myservername.com/data-mining-process>

РЕГУЛЯЦІЯ І ПОЛІТИКИ КУРСУ²

Відвідування занять. Регуляція пропусків.

Відвідування лекційних і лабораторних занять є обов'язковим. Студенти, які за певних обставин не можуть відвідувати лабораторні заняття регулярно, мусять впродовж тижня узгодити із викладачем графік індивідуального відпрацювання пропущених занять.

Політика академічної доброчесності

Усі письмові роботи, що виконуються слухачами під час проходження курсу, перевіряються на наявність плагіату за допомогою спеціалізованого програмного забезпечення UniCheck. Відповідно до чинних правових норм, плагіатом вважатиметься: копіювання чужої наукової роботи чи декількох робіт та оприлюднення результату під своїм іменем; створення суміші власного та запозиченого тексту без належного цитування джерел; рерайт (перефразування чужої праці без згадування оригінального автора). Будь-яка ідея, думка чи речення, ілюстрація чи фото, яке ви запозичуєте, має супроводжуватися посиланням на першоджерело.

² Тут зазначається все, що важливо для курсу: наприклад, умови допуску до лабораторій, реактивів і т.д. Викладач сам вирішує, що треба знати студенту для успішного проходження курсу!



Роботи, у яких виявлено ознаки плагіату, до розгляду не приймаються і відхиляються без права перескладання. Якщо ви не впевнені, чи підпадають зроблені вами запозичення під визначення плагіату, будь ласка, проконсультуйтеся з викладачем.

Висока академічна культура та європейські стандарти якості освіти, яких дотримуються у ЗНУ, вимагають від дослідників відповідального ставлення до вибору джерел. Посилання на такі ресурси, як Wikipedia, бази даних рефератів та письмових робіт (Studopedia.org та подібні) є неприпустимим. Рекомендовані бази даних для пошуку джерел:

Електронні ресурси Національної бібліотеки ім. Вернадського: <http://www.nbuv.gov.ua>

Цифрова повнотекстова база даних англomовної наукової періодики JSTOR: <https://www.jstor.org/>

Використання комп'ютерів/телефонів на занятті

Використання мобільних телефонів, планшетів та інших гаджетів під час лекційних та лабораторних занять дозволяється виключно у навчальних цілях. Будь ласка, не забувайте активувати режим «без звуку» до початку заняття.

Під час виконання заходів контролю (поточних та підсумкового тестів) використання гаджетів заборонено. У разі порушення цієї заборони роботу буде анульовано без права перескладання.

Комунікація

Базовою платформою для комунікації викладача зі студентами є Moodle.

Важливі повідомлення загального характеру розміщуються викладачем на форумі курсу. Для персональних запитів використовується сервіс приватних повідомлень. Відповіді на запити студентів подаються викладачем впродовж трьох робочих днів. Для оперативного отримання повідомлень про оцінки та нову інформацію, розміщену на сторінці курсу у Moodle, будь ласка, переконайтеся, що адреса електронної пошти, зазначена у вашому профайлі на Moodle, є актуальною, та регулярно перевіряйте папку «Спам».

Якщо за технічних причин доступ до Moodle є неможливим, або ваше питання потребує термінового розгляду, направте електронного листа з позначкою «Важливо» на адресу gen.dobr@gmail.com. У листі обов'язково вкажіть ваше прізвище та ім'я, курс та шифр академічної групи.

ДОДАТОК ДО СИЛАБУСУ ЗНУ – 2021-2022

ГРАФІК НАВЧАЛЬНОГО ПРОЦЕСУ. http://sites.znu.edu.ua/navchalnyj_viddil/1635.ukr.html

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>.

ОСВІТНІЙ 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 методику 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/y9pkmmmp5>. 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/ycds57la>.

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/y8gbt4xs>.

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/ycyfws9v>. 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>.

ЗАПОБІГАННЯ КОРУПЦІЇ. Уповноважена особа з питань запобігання та виявлення корупції (Воронков В. В., 1 корп., 29 каб., тел. +38 (061) 289-14-18).

ПСИХОЛОГІЧНА ДОПОМОГА. Телефон довіри практичного психолога (061)228-15-84 (щоденно з 9 до 21).

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

**ЗАПОРІЗЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ
МАТЕМАТИЧНИЙ ФАКУЛЬТЕТ
Силабус навчальної дисципліни**



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 17.00; субота з 09.00 до 15.00.

ЕЛЕКТРОННЕ ЗАБЕЗПЕЧЕННЯ НАВЧАННЯ (MOODLE): [HTTPS://MOODLE.ZNU.EDU.UA](https://moodle.znu.edu.ua)

Якщо забули пароль/логін, направте листа з темою «Забув пароль/логін» за адресами:

- для студентів ЗНУ - moodle.znu@gmail.com, Савченко Тетяна Володимирівна
- для студентів Інженерного інституту ЗНУ - alexvask54@gmail.com, Василенко Олексій Володимирович

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>.