



ВЕЛИКІ ДАНІ. НАУКА ПРО ДАНІ

Викладач: к.т.н., доцент Кеберле Наталія Геннадіївна

Кафедра: комп'ютерних технологій, 1 корпус ЗНУ, ауд. 39

E-mail: nkeberle@gmail.com

Телефон: +38(093) 695-53-32

Інші засоби зв'язку: Viber, Telegram, WhatsApp

Консультації: особисті – п'ятниця, з 17:40 до 19:00, 1 корпус, ауд. 39

Освітня програма, рівень вищої освіти	Комп'ютерні науки. Магістр						
Статус дисципліни	Вибіркова дисципліна блоку дисциплін вільного вибору студентів						
Кредити ECTS	4	Навч. рік	2022-2023	Рік навчання	2	Тижні	12
Вид контролю	Залік						
Посилання на курс в Moodle	https://moodle.znu.edu.ua/course/view.php?id=4052						

ОПИС КУРСУ

Метою викладання навчальної дисципліни «Великі дані. Наука про дані» є ознайомлення з галуззю великих даних (Big Data), пов'язаних даних (Linked Data) та з інструментарієм, що використовується для створення, зберігання, публікації, пошуку, аналізу великих даних..

Основними **завданнями** вивчення дисципліни «Великі дані. Наука про дані» є: оволодіння основними поняттями в галузі великих даних, ознайомлення з принципами організації сховищ великих даних, із базовими алгоритмами збереження та пошуку у сховищі великих даних, ознайомлення із переліком задач, що відносяться до організації роботи з великими даними, та способами рішення таких задач, формування навичок з реалізації сховища великих даних, організації процесів роботи зі сховищем, типового аналізу великих даних.

ОЧІКУВАНІ РЕЗУЛЬТАТИ НАВЧАННЯ

ПР4 Використовувати методи обчислювального інтелекту, машинного навчання, нейромережевої та нечіткої обробки даних, генетичного та еволюційного програмування для розв'язання задач розпізнавання, прогнозування, класифікації, ідентифікації об'єктів керування тощо

ПР10 Використовувати інструментальні засоби розробки клієнт-серверних застосунків, проектувати концептуальні, логічні та фізичні моделі баз даних, розробляти та оптимізувати запити до них, створювати розподілені бази даних, сховища та вітрини даних, бази знань, у тому числі на хмарних сервісах, із застосуванням мов веб-програмування

ПР12 Застосовувати методи та алгоритми обчислювального інтелекту та інтелектуального аналізу даних в задачах класифікації, прогнозування, кластерного аналізу, пошуку асоціативних правил з використанням програмних інструментів підтримки багатовимірного аналізу даних на основі технологій DataMining, TextMining, WebMining



ПР19 Організувати свою працю для досягнення результату у області комп'ютерних наук, виконання розумових і практичних дій, прийомів та операцій, усвідомлення відповідальності за результати своєї діяльності, застосування самоконтролю й самооцінки

ОСНОВНІ НАВЧАЛЬНІ РЕСУРСИ

1. Майер-Шенбергер В., Кукьер К. Большие данные. Революция, которая изменит то, как мы живем, работаем и мыслим. Москва : Издательство «Манн, Иванов и Фербер», 2014. 240 с.
2. Силен Д., Мейсман А., Али М. Основы Data Science и Big Data: Python и наука о данных. Санкт-Петербург : ПИТЕР, 2017. 336 с.

Презентації лекцій, плани практичних занять, методичні рекомендації до практичних занять розміщені на платформі Moodle: <https://moodle.znu.edu.ua/course/view.php?id=4052>

РОЗКЛАД КУРСУ ЗА ТЕМАМИ І КОНТРОЛЬНІ ЗАВДАННЯ

Поточні контрольні заходи:

Обов'язковими поточними контрольними заходами кожного змістового модуля є: практичні завдання та самостійні роботи, які виконуються після опанування відповідної теми курсу. Перелік практичних завдань, відповідні методичні рекомендації розміщено в СЕЗН Moodle: <https://moodle.znu.edu.ua/course/view.php?id=4052>

Підсумкові контрольні заходи: тест (до 40 балів), питання: <https://moodle.znu.edu.ua/course/view.php?id=4052>

Тиждень і вид заняття	Тема змістового модулю	Контрольний захід	Кількість балів
Змістовий модуль 1.			
Тиждень 1 Лекція 1	Лекція 1. Linked Data. Причини появи. Методи обробки. SPARQL сховища триплів (триплстори). Federated SPARQL.	Виконання лабораторної роботи 1.	7
Тиждень 2 Практичне заняття 1	Лабораторна робота 1. Виконання SPARQL-запитів з візуалізацією у HTML. Знайомство з бібліотекою Jena API. Програмування простого додатку виконання SPARQL-запитів за допомогою Jena API.		
Змістовий модуль 2.			
Тиждень 3 Лекція 2	Лекція 2. Створення пов'язаних даних та аналіз пов'язаних даних у аналітичних платформах.	Самостійна робота 1. Встановлення та налагодження	6
Тиждень 4 Практичне заняття 2	Лабораторна робота 1. Виконання SPARQL-запитів з візуалізацією у HTML. Знайомство з бібліотекою Jena API. Програмування простого додатку виконання SPARQL-запитів за допомогою Jena API.	графової семантичної бази даних (сховища триплів) Ontotext GraphDB	



Змістовий модуль 3.			
Тиждень 5 Лекція 3	Лекція 3. Визначення "наука про дані". Властивості великих даних (Big Data). Джерела великих даних та основні шляхи використання.	Виконання лабораторної роботи 2.	5
Тиждень 6 Практичне заняття 3	Лабораторна робота 2. Аналіз даних в RapidMiner. Створення процесу завантаження та пошуку моди/медіани у наборі даних.	Самостійна робота 2. Опрацювання набору даних у форматі CSV для перетворення у RDF за допомогою Ontotext GraphDB OpenRefine.	6
Змістовий модуль 4.			
Тиждень 7 Лекція 4	Лекція 3. Визначення "наука про дані". Властивості великих даних (Big Data). Джерела великих даних та основні шляхи використання.	Самостійна робота 3. Встановлення RapidMiner та його налагодження для опрацювання RDF даних.	6
Тиждень 8 Практичне заняття 4	Лабораторна робота 2. Аналіз даних в RapidMiner. Створення процесу завантаження та пошуку моди/медіани у наборі даних.		
Змістовий модуль 5.			
Тиждень 9 Лекція 5	Лекція 5. Інструменти пошуку та аналітичної обробки даних у сховищі великих даних.	Лабораторна робота 3. Знайомство зі стеком технологій Apache Spark.	10
Тиждень 10 Практичне заняття 5	Лабораторна робота 3. Знайомство зі стеком технологій Apache Spark.	Самостійна робота 4. Реєстрація у Google Cloud Platform. Знайомство з можливостями платформи GCP.	5
Змістовий модуль 6.			
Тиждень 11 Лекція 6	Лекція 5. Інструменти пошуку та аналітичної обробки даних у сховищі великих даних.	Самостійна робота 5. Застосування GCP BigQueryML для	5

ЗАПОРІЗЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ
МАТЕМАТИЧНИЙ ФАКУЛЬТЕТ
Силабус навчальної дисципліни



		аналітичних задач	
Тиждень 12 Практичне заняття 6	Лабораторна робота 4. Інструменти аналітичної обробки даних у Google Cloud Platform.	Лабораторна робота 4. Інструменти аналітичної обробки даних у Google Cloud Platform.	10

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)		



ОСНОВНІ ДЖЕРЕЛА
Змістові модулі 1-6

Навчальні підручники, посібники:

1. Майер-Шенбергер В., Кукьер К. Большие данные. Революция, которая изменит то, как мы живем, работаем и мыслим. Москва : Издательство «Манн, Иванов и Фербер», 2014. 240 с.
2. Силен Д., Мейсман А., Али М. Основы Data Science и Big Data: Python и наука о данных. Санкт-Петербург : ПИТЕР, 2017. 336 с.
3. RapidMiner 4.2 User Guide Operator Reference Developer Tutorial. URL : http://moodle.znu.edu.ua/pluginfile.php?file=/89874/mod_resource/content/1/15967723-Rapidminer-4-2-Tutorial.pdf (дата звернення 20.08.2018)
4. Дьяконов А.Г. Анализ данных, обучение по прецедентам, логические игры, системы WEKA, RapidMiner и MatLab (Практикум на ЭВМ кафедры математических методов прогнозирования): учебное пособие. Москва : Издательский отдел факультета ВМК МГУ имени М.В. Ломоносова, 2010. 278 с.
5. Zikopoulos P., deRoos D., Bienko C., Buglio R., Andrews M. Big Data Beyond the Hype. A Guide to Conversations for Today's Data Center. McGrawHill Education. 2015. 393 p.
6. Sathi A. Big Data Analytics. Disruptive Technologies for Changing the Game. MC Press Online LLC. 2013. 93 p.

Интернет-посилання:

1. GraphDB OntoText. URL: <http://graphdb.ontotext.com/documentation/free/introduction-to-semantic-web.html>
2. KDNuggets: Data Mining Community Top Resource for Analytics, Data Mining, and Data Science Software, Companies, Data, Jobs, Education, News, and more. URL: <http://www.kdnuggets.com>
3. The Data Mine. URL: <http://www.the-data-mine.com>
4. Google Cloud. <https://cloud.google.com/>

РЕГУЛЯЦІЯ І ПОЛІТИКИ КУРСУ¹

Відвідування занять. Регуляція пропусків.

Інтерактивний характер курсу передбачає обов'язкове відвідування практичних занять. Аспіранти, які за певних обставин не можуть відвідувати практичні заняття регулярно, мусять впродовж тижня узгодити із викладачем графік індивідуального відпрацювання пропущених занять. Окремі пропущені завдання мають бути відпрацьовані на найближчій консультації впродовж тижня після пропуску. Відпрацювання занять здійснюється шляхом виконання індивідуального письмового завдання. Аспіранти, які станом на початок екзаменаційної сесії мають понад 70% невідпрацьованих пропущених занять, до сесії не допускаються.

Політика академічної доброчесності

Усі письмові роботи, що виконуються аспірантами 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 (перепарафразування чужої праці без згадування автора оригінального тексту). Будь-яка ідея, думка чи речення, ілюстрація чи фото, яке ви запозичуєте, має супроводжуватися посиланням на першоджерело. Приклади оформлення цитувань див. на платформі Moodle: <https://moodle.znu.edu.ua/mod/resource/view.php?id=103857>

Виконавці індивідуальних завдань обов'язково додають до текстів своїх робіт власноруч підписану Декларацію академічної доброчесності (див. посилання у Додатку до силабусу).

Роботи, в 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 в 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>; наукометрична база Scopus: <https://www.scopus.com>; наукометрична база Web of Science: <https://apps.webofknowledge.com>

Використання комп'ютерів/телефонів на занятті

Використання мобільних телефонів, планшетів та інших гаджетів під час лекційних та практичних занять дозволяється виключно у навчальних цілях (для уточнення певних даних, перевірки правопису, отримання довідкової інформації тощо). Будь ласка, не забувайте активувати режим «без звуку» до початку заняття.

Комунікація

Базовою платформою для комунікації викладача з аспірантами є платформа Moodle. Важливі повідомлення загального характеру регулярно розміщуються викладачем на форумі курсу. Для персональних запитів використовується сервіс приватних повідомлень. Відповіді на запити аспірантів подаються викладачем впродовж трьох робочих днів. Для

¹ Тут зазначається все, що важливо для курсу: наприклад, умови допуску до лабораторій, реактивів тощо. Викладач сам вирішує, що треба знати аспіранту для успішного проходження курсу!

ЗАПОРІЗЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ
МАТЕМАТИЧНИЙ ФАКУЛЬТЕТ
Силабус навчальної дисципліни



оперативного отримання повідомлень про оцінки та нову інформацію, розміщену на платформі 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 на пошту або у зазначені меседжери викладача. У листі обов'язково вкажіть ваше прізвище, ім'я та курс.



ДОДАТОК ДО СИЛАБУСУ

АКАДЕМІЧНА ДОБРОЧЕСНІСТЬ. Аспіранти і викладачі Запорізького національного університету несуть персональну відповідальність за дотримання принципів академічної доброчесності, затверджених **Кодексом академічної доброчесності ЗНУ:** <https://tinyurl.com/ya6yк4ad>. *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 аспірантами 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 аспіранту 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 аспірантів, 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/yсyfws9v>. 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 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 адресою:

· для аспірантів ЗНУ - moodle.znu@gmail.com, Савченко Тетяна Володимирівна

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 використуйте посилання для відновлення паролю <https://moodle.znu.edu.ua/mod/page/view.php?id=133015>.

ЗАПОРІЗЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ
МАТЕМАТИЧНИЙ ФАКУЛЬТЕТ
Силабус навчальної дисципліни



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>.