



СИЛАБУС НАВЧАЛЬНОЇ ДИСЦИПЛІНИ

СХОВИЩА ВЕЛИКИХ ДАНИХ

ID 178

Шифр, назва спеціальності та освітній рівень	122 Комп'ютерні науки (бакалавр)	Назва освітньої програми	Комп'ютерні науки (2024)
Тип програми	Освітньо-професійна	Мова викладання	Українська
Факультет	Факультет комп'ютерно-інформаційних систем і програмної інженерії (ФІС)	Кафедра	Каф. комп'ютерних наук (КН)

Викладач/викладачі

Боднарчук Ігор Орестович, канд. техн. наук, доцент, доцент, завідувач кафедри КН, [профіль на порталі "Науковці ТНТУ"](#)

Загальна інформація про дисципліну

Мета курсу	Мета вивчення навчальної дисципліни ознайомлення студентів з основами проектування сховищ даних, доступу до NoSQL баз даних, підготовка до вивчення матеріалу освітньої компоненти "Методи та засоби опрацювання великих даних".
Формат курсу	Заняття проводяться у формі лекцій та лабораторних робіт. Передбачено виконання здобувачами освіти курсової роботи
Компетентності ОП	ЗК1. Здатність до абстрактного мислення, аналізу та синтезу. ЗК2. Здатність застосовувати знання у практичних ситуаціях. ЗК3. Знання та розуміння предметної області та розуміння професійної діяльності. ЗК4. Здатність спілкуватися державною мовою як усно, так і письмово. ЗК7. Здатність до пошуку, оброблення та аналізу інформації з різних джерел. ЗК9. Здатність працювати в команді. СК8. Здатність проектувати та розробляти програмне забезпечення із застосуванням різних парадигм програмування: узагальненого, об'єктно-орієнтованого, функціонального, логічного, з відповідними моделями, методами й алгоритмами обчислень, структурами даних і механізмами управління. СК9. Здатність реалізувати багаторівневу обчислювальну модель на основі архітектури клієнт-сервер, включаючи бази даних, знань і сховища даних, виконувати розподілену обробку великих наборів даних на кластерах стандартних серверів для забезпечення обчислювальних потреб користувачів, у тому числі на хмарних сервісах. СК15. Здатність до аналізу та функціонального моделювання бізнес-процесів, побудови та практичного застосування функціональних моделей організаційно-економічних і виробничо-технічних систем, методів оцінювання ризиків їх проектування. СК16. Здатність реалізовувати високопродуктивні обчислення на основі хмарних сервісів і технологій, паралельних і розподілених обчислень при розробці й експлуатації розподілених систем паралельної обробки інформації. СК17. Здатність проектувати, розробляти та/чи супроводжувати програмно-алгоритмічні засоби для отримання, зберігання, обробки, аналізу та візуалізації різнотипових наборів та колекцій великих даних; впроваджувати рішення, які можуть ефективно масштабуватися при зростанні обсягів даних.
Програмні результати навчання з ОП	ПР10. Використовувати інструментальні засоби розробки клієнт-серверних застосувань, 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. ПР14. Застосовувати знання методології та CASE-засобів проектування складних систем, методів структурного аналізу систем, об'єктно-орієнтованої методології проектування при розробці і дослідженні функціональних моделей організаційно-економічних і виробничо-технічних систем. ПР17. Використовувати знання архітектури великих даних, структурних моделей для забезпечення ефективності та швидкодії

	обробки та візуалізації великих за обсягом даних. ПР18. Застосовувати знання інформаційних технологій обробки, аналізу та візуалізації великих даних на практиці, розробляти стратегії для підтримки та покращення якості великих даних протягом їхнього життєвого циклу. ПР19. Використовувати спеціалізовані інструменти та технології для обробки та візуалізації великих обсягів даних, такі як Apache Hadoop, Spark, або інші фреймворки для роботи з Big Data.
Обсяг курсу	Очна (денна) форма здобуття освіти: Кількість кредитів ECTS — 4; лекції — 32 год.; лабораторні заняття — 32 год.; самостійна робота — 56 год.; Заочна форма здобуття освіти: Кількість кредитів ECTS — 4; лекції — 6 год.; лабораторні заняття — 6 год.; самостійна робота — 108 год.;
Ознаки курсу	Рік навчання — 3; семестр — 5; Обов’язкова (для здобувачів інших ОП може бути вибірковою) дисципліна; кількість модулів — 2;
Форма контролю	Поточний контроль: захист лабораторних робіт; 2 модульні тести в системі Atutor. Підсумковий контроль: екзамен
Компетентності та дисципліни, що є передумовою для вивчення	Об’єктно-орієнтоване програмування Організація баз даних Програмування
Матеріально-технічне та/або інформаційне забезпечення	Персональний комп’ютер Підключення до мережі Інтернет Microsoft SQL Server Developer edition 2022 Visual Studio Community Edition 2022 Mongo DB Cassandra Postgres SQL

СТРУКТУРА КУРСУ

Теми занять, короткий зміст

Лекційний курс	Годин	
	ОФЗО	ЗФЗО
Лекція 1. Вступ до великих даних і сховищ даних.	4	1
Лекція 2. Архітектура та моделі сховищ великих даних.	4	1
Лекція 3. Методи проектування та оптимізації баз даних для великих даних.	4	0,5
Лекція 4. Технології та платформи для роботи з великими даними.	4	0,5
Лекція 5. Реалізація клієнт-серверних застосувань для роботи з великими даними.	2	0,5
Лекція 6. Функціональне моделювання бізнес-процесів у сфері великих даних	2	0,5
Лекція 7. Методи аналізу та обробки великих даних.	4	0,5
Лекція 8. Безпека та управління якістю великих даних.	2	0,5
Лекція 9. Візуалізація великих даних та аналітика.	4	0,5
Лекція 10. Тенденції та перспективи розвитку технологій великих даних.	2	0,5
	РАЗОМ:	32 6
Лабораторний практикум (теми)	Годин	
	ОФЗО	ЗФЗО
1. Створення та управління сховищем даних.	6	1
2. Оптимізація запитів у великих базах даних.	6	1
3. Обробка великих даних за допомогою Apache Hadoop.	6	1
4. Робота з Apache Spark для обробки великих даних.	6	1
5. Візуалізація великих даних.	4	1
6. Робота з хмарними сховищами великих даних	4	1

Курсова робота/проект

Мета виконання курсової роботи	Метою виконання курсової роботи з дисципліни «Сховища великих даних» є систематизація, закріплення та розширення теоретичних знань, їхнє застосування для вирішення конкретного практичного завдання відповідно до вимог формування компетентностей згідно освітньої програми «Сховища великих даних».
Завдання курсової роботи	1. Практичне освоєння реляційних СКБД для побудови сховищ даних. 2. Практичне ознайомлення з NoSQL СКБД 3. Робота з даними з різних типів СКБД. 4. Основи використання фреймворків для Big Data.
Структура курсової роботи	Титульний лист; завдання на курсову роботу; анотація; зміст; перелік умовних позначень; вступ; основна частина; висновки; список використаних джерел; додатки.
Обсяг курсової роботи	Рекомендований обсяг - 30-50 сторінок.
Етапи виконання	Вибір та затвердження теми курсової роботи; критичний аналіз нормативно-правової бази, спеціальної літератури з проблем, що розглядаються, пошук додаткових джерел інформації; складання плану курсової роботи; узагальнення та аналіз накопиченого матеріалу, обробка даних, обґрунтування пропозицій; написання тексту і оформлення курсової роботи; захист курсової роботи згідно з встановленим графіком.
Оцінювання курсової роботи	Зміст курсової роботи – 75 балів, захист курсової роботи – 25 балів.
Форма контролю	Захист курсової роботи передбачає: - стислу доповідь (5 хв.) магістранта, в якій необхідно відокремити мету, об'єкт, предмет дослідження та коротко висвітлити зміст одержаних результатів дослідження. Зробити акцент на висновках та рекомендаціях. Бажано, щоб доповідь магістранта під час захисту супроводжувалась презентацією результатів, підготовленою за допомогою засобів «Microsoft PowerPoint»; - співбесіду і відповіді на запитання наукового керівника та членів комісії. Курсова робота та її захист оцінюється відповідно до вимог кредитно-модульної системи.

Технічне й програмне
забезпечення

Технічні засоби для демонстрування результатів виконання курсової роботи (ноутбук, проектор). Пакет програмних продуктів Microsoft Office.

ІНШІ ВИДИ РОБІТ

Теми, короткий зміст

Інформаційні джерела для вивчення курсу

Базова література

1. Шаховська Н. Б. Сховища та простори даних: [монографія] / Н. Б. Шаховська, В. В. Пасічник; М-во освіти і науки України, Нац. ун-т «Львів. політехніка». – Львів : Вид-во Нац. ун-ту «Львів. політехніка», 2009. – 244 с.: іл.
2. Пасічник В.В., Шаховська Н.Б. Сховища даних: Навчальний посібник. – 2-ге вид., стер. – Львів: “Магнолія 2006”, 2024. – 492 с.
3. Ерл Т. Основи Big Data. Концепції, алгоритми і технології. – Київ: Діалектика, 2021. – 384 с.
4. Грус Дж. Data Science з нуля: перші принципи з Python. – Київ: Наш Формат, 2020. – 352 с.
5. Піацці Л. Data Science для чайників. – Харків: Віват, 2019. – 340 с.
6. Клеппманн М. Розробка додатків з інтенсивним використанням даних. – Київ: Діалектика, 2022. – 456 с.
7. Майєр-Шенбергер В., Кук'є К. Великі дані: революція, яка змінить те, як ми живемо, працюємо та мислимо. – Київ: Наш Формат, 2018. – 312 с.
8. Кнафлік К. Н. Розповідання історій з даними: посібник з візуалізації даних для бізнес-професіоналів. – Київ: Наш Формат, 2019. – 256 с.
9. Kimball, R., & Ross, M. (2018). The data warehouse toolkit: The definitive guide to dimensional modeling. John Wiley & Sons.
10. Rebuilding Reliable Data Pipelines Through Modern Tools by Ted Malaska. Released July 2019. Publisher(s): O'Reilly Media, Inc.
11. Hadoop: The Definitive Guide: Storage and Analysis at Internet Scale 4th Edition by Tom White. 2015. O'Reilly.
12. Fundamentals of Data Engineering: Plan and Build Robust Data Systems. Written by Matt Housley and Joe Reis. 2022. O'Reilly.
13. Kleppmann M. Designing Data-Intensive Applications. – O'Reilly Media, 2017. – 616 p.
14. Provost F., Fawcett T. Data Science for Business: What You Need to Know about Data Mining and Data-Analytic Thinking. – O'Reilly Media, 2013. – 414 p.
15. McKinney W. Python for Data Analysis: Data Wrangling with Pandas, NumPy, and Jupyter. – O'Reilly Media, 2022. – 550 p.
16. Bruce P., Bruce A., Gedeck P. Practical Statistics for Data Scientists: 50+ Essential Concepts Using R and Python. – O'Reilly Media, 2020. – 368 p.
17. Wickham H., Grolemund G. R for Data Science: Import, Tidy, Transform, Visualize, and Model Data. – O'Reilly Media, 2017. – 522 p.
18. Loshin D. Big Data Analytics: From Strategic Planning to Enterprise Integration with Tools, Techniques, NoSQL, and Graph. – Morgan Kaufmann, 2013. – 416 p.
19. Few S. Show Me the Numbers: Designing Tables and Graphs to Enlighten. – Analytics Press, 2012. – 351 p.

Допоміжна література та інформаційні ресурси

1. "Integrating Big Data Analytics into Business Process Modelling". URL: https://thesai.org/Downloads/Volume13No6/Paper_57-Integrating_Big_Data_Analytics_into_Business_Process_Modelling.pdf
2. "Big Data Analytics and Business Process Innovation". URL: <https://www.emerald.com/insight/content/doi/10.1108/bpmj-02-2017-0046/full/html>
3. "A Modeling Framework for Business Process Reengineering Using Big Data Analytics and a Goal-Oriented Approach". URL: <https://ieeexplore.ieee.org/document/7956514>
4. "Business Process Modeling Extension and Scheduling Strategy Compatible with Big Data". URL: <https://ieeexplore.ieee.org/document/10141025>
5. "Modeling Data for Business Processes". URL: <https://ieeexplore.ieee.org/document/6816722>
6. <https://www.tutorialspoint.com/dwh/index.htm> Data Warehousing Tutorial.
7. <https://www.guru99.com/data-warehousing.html> What is Data Warehouse? Types, Definition & Example.

Політики курсу

Політика контролю	Використовуються такі засоби оцінювання та методи демонстрування результатів навчання: поточне опитування; тестування; виконання індивідуальних завдань та презентацій; оцінювання результатів виконаних самостійних робіт; бесіди та обговорення проблемних питань; дискусії; індивідуальні консультації; екзамен. Можливий ректорський контроль.
Політика щодо консультування	Консультації при вивченні дисципліни проводяться згідно затвердженого на кафедрі . Консультування передбачено як очно ,так і з використанням ресурсів електронного навчального курсу у середовищі електронного навчання університету.
Політика щодо перескладання	Студент має право на повторне складання модульного контролю з метою підвищення рейтингу протягом тижня після складання модульного контролю за графіком. Перескладання екзамену відбувається в терміни, визначені графіком освітнього процесу. Здобувач ВО має право на зарахування результатів навчання здобутих у неформальній чи інформальній освіті.
Політика щодо академічної доброчесності	При складанні усіх видів контролю у середовищі електронного навчання завжди активується система розпізнавання особи, що складає контроль. Усі практичні роботи у ЕНК перевіряються вбудованою системою Антиплагіат. При складанні усіх форм контролю забороняється списування, у тому числі з використанням сучасних інформаційних технологій.
Політика щодо відвідування	Відвідування занять є обов'язковим компонентом освітнього процесу. За наявності поважних причин (наприклад, хвороба, особливі потреби, відрядження, сімейні обставини, участь у програмах академічної мобільності тощо) навчання може здійснюватися за індивідуальним графіком, погодженим з деканом факультету.

СИСТЕМА ОЦІНЮВАННЯ

Розподіл балів, які отримують студенти за курс

Модуль 1			Модуль 2			Підсумковий контроль		Разом з дисципліни
Аудиторна та самостійна робота			Аудиторна та самостійна робота			Теоретичний курс	Практичне завдання	100
Теоретичний курс (тестування)	Лабораторна робота		Теоретичний курс (тестування)	Лабораторна робота				
15	25		10	30		15	10	
№ лекції	Види робіт	К-ть балів	№ лекції	Види робіт	К-ть балів			
Тема 1	Лабораторна робота №1	8	Тема 6	Лабораторна робота №4	10			
Тема 2	Лабораторна робота №2	8	Тема 7	Лабораторна робота №5	10			
Тема 3	Лабораторна робота №3	9	Тема 8	Лабораторна робота №6	10			
Тема 4			Тема 9					
Тема 5			Тема 10					

Розподіл балів, які отримують студенти за виконання та захист КР

Модуль 1		Модуль 2		Модуль 3		Підсумковий контроль	Разом за КР
Виконання розділу 1		Виконання розділу 2		Виконання розділу 2		Захист КР	100
10		45		20		25	
Види робіт	К-ть балів	Види робіт	К-ть балів	Види робіт	К-ть балів		
Розділ 1	10	Розділ 2	45	Розділ 3	20		

Розподіл оцінок

Сума балів за навчальну діяльність	Шкала ECTS	Оцінка за національною шкалою
90-100	A	Відмінно
82-89	B	Добре
75-81	C	Добре
67-74	D	Задовільно
60-66	E	Задовільно
35-59	FX	Незадовільно з можливістю повторного складання
1-34	F	Незадовільно з обов’язковим повторним вивченням дисципліни

Затверджено рішенням кафедри КН, протокол №1 від «26» серпня 2024 року.
ПОГОДЖЕНО
Гарант освітньої програми канд. техн. наук, доцент кафедри КН
Леся ДМИТРОЦА