

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
Львівський національний університет імені Івана Франка
Механіко-математичний факультет
Кафедра математичної статистики і диференціальних рівнянь

Затверджено

На засіданні

кафедри алгебри, топології

та основ математики

механіко-математичного факультету

Львівського національного університету

імені Івана Франка

(протокол № 1 від 28.08.2020 р.)

В.о. завідувача кафедри: Тарас БАНАХ



Силабус з навчальної дисципліни
“ Графові дані та основи NoSQL”,
що викладається в межах ОПП “ Статистичний аналіз даних ”
першого (бакалаврського) рівня вищої освіти для здобувачів з
спеціальності 112 - Статистика

Львів 2020 р.

Назва дисципліни	Графові дані та основи NoSQL
Адреса викладання дисципліни	Головний корпус ЛНУ ім. І. Франка м. Львів, вул. Університетська 1, 79000
Факультет та кафедра, за якою закріплена дисципліна	Механіко-математичний факультет Кафедра математичної статистики і диференціальних рівнянь
Галузь знань, шифр та назва спеціальності	11 - Математика та статистика 112 - Статистика
Викладачі дисципліни	Романів О.М., кандидат фізико-математичних наук, доцент кафедри алгебри, топології та основ математики
Контактна інформація викладачів	oromaniv@lnu.edu.ua , https://new.mmf.lnu.edu.ua/employee/romaniv-o-m
Консультації з питань навчання по дисципліні відбуваються	Консультації в день проведення лекцій/практичних занять (за попередньою домовленістю). Головний корпус ЛНУ ім. І. Франка, каб. 267. м. Львів, вул. Університетська, 1
Інформація про дисципліну	Дисципліна “Графові дані та основи NoSQL” є дисципліною вільного вибору студента з спеціальності 112 – Статистика для освітньої програми “Статистичний аналіз даних”, яка викладається в 7-му семестрі в обсязі 5-ох кредитів (за Європейською Кредитно-Трансферною Системою ECTS).
Коротка анотація дисципліни	Курс розроблено для ознайомлення студентів з основними поняттями та методами сучасних графових та NoSQL баз даних.
Мета та цілі дисципліни	<i>Мета:</i> формування у майбутніх спеціалістів повноцінних теоретичних знань та практичних навичок по застосуванню графові даних та NoSQL баз даних для зберігання і опрацювання інформації <i>Цілі:</i> викласти основні теоретичні положення по роботі з графовими даними та базові методи зберігання і опрацювання даних з використанням NoSQL, сформувати вміння підбирати оптимальну базу даних залежно від типу даних та інсталяція і налаштування необхідної бази.
Література для вивчення дисципліни	1) Robinson I., Webber J., Eifrem E. <i>Graph Databases: New Opportunities for Connected</i> . 2nd ed. O'Reilly Media, Inc., 2015 2) Gosnell D. K., Broecheler M. <i>The Practitioner's Guide to Graph Data</i> . O'Reilly Media, Inc., 2020 3) Vukotic A., Watt N. <i>Neo4j in Action</i> . Manning Publications Co., 2015. 4) Meier A., Kaufmann M. <i>SQL & NoSQL Databases: Models, Languages, Consistency Options and Architectures for Big Data Management</i> . Springer Vieweg; 1st ed., 2019. 5) Bradshaw S., Brazil E., Chodorow K. <i>MongoDB: The Definitive Guide: Powerful and Scalable Data Storage</i> . O'Reilly Media; 3rd edition, 2019.
Обсяг курсу	Загальний обсяг: 150 годин. Аудиторних занять: 64 год., з них 32 год. лекційних та 32 год. практичних занять. Самостійної роботи: 86 год.
Очікувані	У результаті вивчення даного курсу студент буде:

результати навчання	<p>знати: методи очищення, інтеграція та перетворення даних, локальні принципи роботи з даними, аспекти вибору алгоритмів обробки даних і відповідних структур даних, технології розподілення і зберігання даних, технології SQL і NoSQL, графові бази даних;</p> <p>вміти: будувати і обирати моделі даних, виконувати попередню обробку даних (очищення, інтеграція, перетворення даних та інше), уміти обирати ефективний алгоритм / технологію обробки даних, а також структури для роботи з даними (локально і глобально), застосовувати методологію розподілення і зберігання великих даних, уміти використовувати сучасні бази даних (перш за все NoSQL і графові бази даних), виконувати дослідження реальних сховищ даних, презентувати результатів наукових досліджень; здійснювати дослідження з використанням цих ресурсів та технологій.</p>
Ключові слова	SQL, NoSQL, графові бази даних, MongoDB, Neo4j.
Формат курсу	Очний
Теми	Див. Схема курсу
Підсумковий контроль, форма	Залік у кінці семестру
Пререквізити	<p>Для вивчення даного курсу студенти потрібні базові знання з:</p> <ul style="list-style-type: none"> - Дискретна математика. - Дискретні моделі статистики - Інформатика та програмування (Python) - Бази даних та основи SQL - Алгоритми і структури даних
Навчальні методи та техніки, які будуть використовуватися під час викладання курсу	Інформаційні методи (лекція, бесіда, ілюстрація, демонстрація); дедуктивні методи на основі узагальнень; евристичні методи (проблемна лекція); інтерактивні методи (дискусія)
Необхідне обладнання	<p>Для проведення лекційних занять: комп'ютер (мінімальні характеристики: процесор Intel Core i5, 8ГБ оперативної пам'яті), доступ до мережі Internet, засоби мультимедіа (в т.ч. проектор).</p> <p>Для проведення практичних/лабораторних занять: комп'ютер (мінімальні характеристики: процесор Intel Core i5, 8ГБ оперативної пам'яті), доступ до мережі Internet.</p> <p>Необхідне програмне забезпечення включає в себе ОС Windows 10, програмні додатки (MS Teams, MongoDB, MySQL, Neo4j).</p>
Критерії оцінювання (окремо для кожного виду навчальної діяльності)	<p>Оцінювання проводиться за 100-бальною шкалою. Бали нараховуються за наступним співвідношенням:</p> <ul style="list-style-type: none"> • Змістовий модуль 1: 8% семестрової оцінки за активну роботу на заняттях, 30% семестрової оцінки за виконання практичних аудиторних і домашніх завдань, 10% семестрової оцінки за колоквиум, максимальна кількість балів 48. • Змістовий модуль 2: 7% семестрової оцінки за активну роботу на заняттях, 30% семестрової оцінки за виконання практичних аудиторних і домашніх завдань, максимальна кількість балів 37.

• Контрольне тестування: 15% семестрової оцінки, максимальна кількість балів 15.

Підсумкова максимальна кількість балів 100.

Академічна доброчесність: Очікується, що роботи студентів будуть оригінальними дослідженнями чи міркуваннями. Списування та втручання в роботу інших студентів становлять, але не обмежують, приклади можливої академічної недоброчесності. Виявлення ознак академічної недоброчесності в написанні завдань є підставою для її незарахування викладачем, незалежно від масштабів плагіату чи обману.

Жодні форми порушення академічної доброчесності не толеруються.

Відвідання занять є важливою складовою навчання. Очікується, що всі студенти відвідають усі лекції та практичні/лабораторні заняття курсу. Студенти повинні інформувати викладача про неможливість відвідати заняття. У будь-якому випадку студенти зобов'язані дотримуватися термінів визначених для виконання всіх видів робіт, передбачених курсом.

Література. Уся література, яку студенти не зможуть знайти самостійно, буде надана викладачем виключно в освітніх цілях без права її передачі третім особам. Студенти заохочуються до використання також й іншої літератури та джерел, яких немає серед рекомендованих.

Політика виставлення балів. Враховуються бали, набрані при поточному контролі та бали підсумкового тестування. При цьому обов'язково враховуються присутність на заняттях та активність студента під час практичного заняття; недопустимість пропусків та запізнень на заняття; користування мобільним телефоном, планшетом чи іншими мобільними пристроями під час заняття в цілях не пов'язаних з навчанням; списування та плагіат; несвоєчасне виконання поставленого завдання і т. ін.

Оцінювання практичних робіт відбувається шляхом оцінки роботи студента під час проведення практичної роботи в аудиторії та захисту написаної студентом вдома практичної роботи.

Бали оцінювання аудиторного виконання практичних робіт нараховуються за наступним співвідношенням:

100% – студент в повному обсязі володіє навчальним матеріалом, має повне розуміння розглянутої теми, надає правильні відповіді на запитання по темі, код програми функціонує відповідно до завдання;

75% – студент достатньо розуміє розглянутий матеріал та принципи написаного ним коду програми, присутні неточності та незначні помилки у відповідях на запитання по темі, код програми функціонує відповідно до завдання;

50% – студент не досить добре розуміє розглянутий матеріал та написаний ним код програми, вагається та надає неточні/не конкретні відповіді на запитання по темі, код програми функціонує з помірними недоліками;

25% – студент погано розуміє розглянутий матеріал та написаний ним код програми, студент в більшості надає помилкові відповіді на питання по темі, код програми не функціонує належним чином;

0 - студент зовсім не засвоїв розглянутий матеріал, написаний ним код програми не відповідає темі/не функціонує взагалі.

Бали оцінювання домашнього завершення виконання практичних робіт та наданого звіту нараховуються за наступним співвідношенням:

100% – звіт цілком і повністю відображає індивідуальне завдання

	<p>студента, містить правильні висновки, ілюстрований (за потреби) відповідними графіками і таблицями які правильно відображають суть виконаного завдання, студент має повне розуміння розглянутої теми, надає правильні відповіді на запитання по темі, код програми функціонує відповідно до завдання;</p> <p>75% – звіт в достатній мірі відображає індивідуальне завдання студента, містить допустимі висновки, ілюстрований (за потреби) відповідними графіками і таблицями які частково відображають суть виконаного завдання, студент достатньо розуміє принципи написаного ним коду програми, присутні неточності та незначні помилки у відповідях на запитання по темі, код програми функціонує відповідно до завдання;</p> <p>50% – звіт містить загальні формулювання завдання, висновки нечіткі, необхідні ілюстрації чи таблиці відсутні, студент не досить добре розуміє розглянутий матеріал та представлений код програми, надає неточні/не конкретні відповіді на запитання по темі, код програми функціонує з помірними недоліками;</p> <p>25% – звіт не містить формулювання завдання, висновки необґрунтовані чи неповні, необхідні ілюстрації чи таблиці відсутні, студент погано розуміє розглянутий матеріал та представлений код програми, студент в більшості надає помилкові відповіді на питання по темі, код програми не функціонує належним чином;</p> <p>0 – звіт відсутній/не відповідає темі, студент зовсім не засвоїв розглянутий матеріал, написаний ним код програми не відповідає темі/не функціонує взагалі.</p> <p>Оцінювання контрольного тестування відбувається шляхом оцінки письмових відповідей студента на поставлені запитання.</p> <p>Відсотки нарахування балів оцінювання відповіді на кожне запитання нараховуються за наступним співвідношенням:</p> <p>75-100% – тема відтворюється в повному обсязі, правильно, обґрунтовано, логічно;</p> <p>50-75% – відтворюється значна частина розглянутої теми, проте присутні неточності та/або невідповідності;</p> <p>25-50% – виявлено множинні неточності та невідповідності, пояснення відсутні чи частково помилкові;</p> <p>0-25% – тему майже не розкрито, кількість викладеного матеріалу не відповідає загальним нормам обраного виду роботи.</p> <p>Критерії оцінювання результатів неформальної освіти:</p> <p>Нарахування балів відбувається за публікацію студентом тез доповідей на конференціях, наукових статей, за участь студента у діяльності наукових гуртків, семінарів, круглих столів, конкурсів, участь у заходах неформальної освіти, за отримання сертифікатів про проходження навчання на різних освітніх платформах (Coursera, Prometheus тощо), курсах провідних ІТ компаній за тематикою навчальної дисципліни. Кількість балів визначається відсотком покриття результатів відповідної активності до вимог результатів навчання з навчальної дисципліни</p>
<p>Питання для контрольного тестування</p>	<p>Основні поняття концепції SQL і NoSQL, теорема CAP; числові характеристики різних баз даних; поняття масиву (array) їх види та синтаксис опису в JSON форматі.</p>
<p>Опитування</p>	<p>Анкету-оцінку з метою оцінювання якості курсу буде надано по завершенню курсу.</p>

**Схема курсу “Графові дані та основи NoSQL ”
для студентів спеціальності 112 - Статистика**

Тижні	Лекційний курс		Практичні заняття		К-сть год СР	Літе- ратура
	Назва теми	К-сть год	Назва теми	К-сть год		
1	2	3	4	5	6	7
1	Основи концепцій нереляційних баз даних NoSQL	2	Визначення основних концепцій нереляційних баз даних NoSQL в порівнянні з SQL	2	5	[1]-[5]
2	Документно-орієнтовані бази даних NoSQL	2	Установка та налаштування бази даних на прикладі MongoDB	2	6	[1]-[5]
3	Основи створення документно-орієнтованої бази даних	2	Знайомство з форматом даних JSON	2	5	[1]-[5]
4	Моделювання даних в середовищі документно-орієнтованих баз даних	2	Створення бази даних, колекцій та документів в середовищі MongoDB	2	6	[1]-[5]
5	Методи обробки документних даних у документно-орієнтованих базах даних	2	Робота з документами в середовищі MongoDB	2	5	[1]-[5]
6	Методи обробки масивів даних у документно-орієнтованих базах даних	2	Робота із колекціями та документами у середовищі Compass	2	6	[1]-[5]
7	Застосування регулярних виразів у документно-орієнтованих базах даних	2	Робота з масивами та регулярними виразами в середовищі MongoDB	2	5	[1]-[5]
8	Теоретичний колоквіум	2	Проектування БД з контролю виконання навантаження викладачами ВУЗу	2	6	[1]-[5]
9	Графове мислення. Еволюція від реляційного до графового мислення	2	Установка та перше знайомство з Neo4j	2	5	[1]-[5]
10	Розділення сутностей у графових базах даних	2	Побудова основних елементів графових даних (околиць, дерев, шляхів) у Neo4j	2	6	[1]-[5]
11	Графові моделі бази даних NoSQL	2	Об'єднання кількох наборів даних в один граф у Neo4j	2	5	[1]-[5]
12	Основи графової бази даних в середовищі SQL Server 2017	2	Створення та робота з графовою базою даних NoSQL в середовищі Microsoft SQL Server 2017	2	6	[1]-[5]
13	Основи колонко-орієнтованих баз даних	2	Інсталяція та перше знайомство з Redis	2	5	[1]-[5]
14	Методи застосування моделей ключ-значення	2	Побудова моделей ключ-значення з використанням Redis	2	6	[1]-[5]
15	Ітераційні процеси в	2	Робота з даними у форматі JSON	2	5	[1]-[5]

	розвитку баз даних		в SQL Server			
16	Підсумковий колоквіум. Підсумкове заняття	2	Проектування БД туристичного агенства	2	4	[1]-[5]
	Разом	32		32	86	