

**ОДЕСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ ІМЕНІ І.І. МЕЧНИКОВА
ФАКУЛЬТЕТ МАТЕМАТИКИ, ФІЗИКИ ТА ІНФОРМАЦІЙНИХ ТЕХНОЛОГІЙ
КАФЕДРА МАТЕМАТИЧНОГО ЗАБЕЗПЕЧЕННЯ КОМП'ЮТЕРНИХ СИСТЕМ**

**Силабус курсу «Системи оперативної аналітичної обробки даних»
(обов'язкова компонента)**

Обсяг	загальна кількість: змістовних модулів – 8
Семестр	осінній (1 курс магістратури або 4 курс бакалаврату)
Дні, Час, Місце	за розкладом занять
Викладач(и)	Малахов Євгеній Валерійович, доктор технічних наук, професор, завідувач кафедри математичного забезпечення комп'ютерних систем Петрушина Т.І., кандидат фізико-математичних наук, доцент, доцент кафедри математичного забезпечення комп'ютерних систем
Контактний телефон	(048)7340723
E-mail	eugene_malakhov@i.ua, tatiana.petrushyna@gmail.com
Робоче місце	кафедра математичного забезпечення комп'ютерних систем
Консультації	очні консультації: понеділок 13.00-14.00 on-line консультації: ZOOM (посилання генерується на початку занять)

КОМУНІКАЦІЯ

Комунікація зі студентами буде здійснюватися електронною поштою, в аудиторії або через ZOOM.

АНОТАЦІЯ КУРСУ

Предметом вивчення курсу є основні типи, структури і моделі даних, які покладено в основу багатовимірних сховищ даних, що використовуються при створенні інформаційних систем оперативної аналітичної обробки даних (OLAP/OLTP), поняття озер даних, ETL/ELT-конвеєрів, а також засоби реалізації таких систем.

Пререквізити курсу

Матеріал курсу ґрунтується на раніше отриманих студентами знаннях, практичних вміннях та навичках з тем та напрямів щодо алгоритмів, структур даних, реляційних баз даних, мови SQL. Тому обирати курс рекомендується не раніше 4 року навчання (4 курс бакалаврату або 1 курс магістратури).

Постреквізити курсу

Цей курс доповнює дисципліну «Аналіз та візуалізація надвеликих масивів даних (Big Data)» в сфері аналізу та обробки даних і є базою для засвоєння наступних дисциплін освітньо-професійної програми підготовки магістрів: «Професійно-дослідницька практика», «Виконання кваліфікаційної роботи магістра».

Метою курсу є формування системних знань щодо основних понять і принципів планування, проектування, побудови, заповнення і супроводу сховищ даних.

Зміст курсу

Розглядаються:

- Поняття системи підтримки прийняття рішень, онлайн обробки транзакцій, сховища даних. Архітектура корпоративної OLAP-системи.
- Поняття OLAP. Способи аналітичної обробки даних. Концепція оперативної аналітичної обробки даних. Вимоги до засобів оперативної аналітичної обробки.
- Сховища даних Категорії даних у сховищах даних. Інформаційні потоки у сховищах даних.
- Структура OLAP-куба. Терміни та поняття структури куба. Операції над гіперкубом. Відмінність обробки транзакцій від обробки аналітики.
- Багатовимірний аналіз даних за допомогою служби SQL Server Analysis Services
- Використання служб Integration Services із службами Analysis Services

- Види даних в організації: Структуровані дані. Текстові дані. Текстовий ETL. Аналогові дані/дані Інтернету речей. Інші типи неструктурованих даних.
- Data Lake (Озеро Даних) та Data Lakehouse. Архітектура Lakehouse. Порівняння Сховища Даних та Озера Даних з Data Lakehouse. Машинне навчання та Data Lakehouse.
- Перспективні дані або дані, орієнтовані на майбутнє (Data future-proofing™) в Data Lakehouse.
- Платформа Databricks Lakehouse. Основи Delta Lake. Архітектура Медальйону. Підтримка потоків архітектурою Медальйону

ОЧІКУВАНІ РЕЗУЛЬТАТИ

У результаті вивчення курсу студент повинен

знати: основні поняття і термінологію технології сховищ даних, технології та сфери застосування інструментів OLAP і методів видобутку даних (Data Mining), ролі сховищ даних у системах автоматизації прийняття рішень.

вміти: проектувати сховища даних, застосовувати сучасні інструменти побудови сховищ даних і аналітичної обробки даних, ETL-пайплайни.

бути здатним: розробляти математичні, інформаційні та комп'ютерні моделі об'єктів і процесів інформатизації, використовувати сучасні технології аналізу даних для оптимізації процесів в інформаційних системах, застосовувати сучасні моделі та методи нечіткого логічного виводу на підставі форм представлення знань і способів організації пошуку рішень; проектувати та розробляти експертні систем, моделювати архітектуру, поведінку та процеси функціонування спеціалізованих, автономних та розподілених інтелектуальних систем автоматизованого пошуку і аналізу інформації.

Компетентності, які отримує студент у результаті вивчення курсу:

- здатність розробляти математичні, інформаційні та комп'ютерні моделі об'єктів і процесів інформатизації.
- здатність використовувати сучасні технології аналізу даних для оптимізації процесів в інформаційних системах.
- здатність застосовувати сучасні моделі та методи нечіткого логічного виводу на підставі форм представлення знань і способів організації пошуку рішень; проектувати та розробляти експертні систем..

Результати навчання: по завершенню курсу студент матиме навички

- розробляти моделі інформаційних процесів та систем різного класу, використовувати методи моделювання, формалізації, алгоритмізації та реалізації моделей з використанням сучасних комп'ютерних засобів.
- розробляти і використовувати сховища даних, здійснювати аналіз даних для підтримки прийняття рішень.
- використовувати сучасні нечіткі моделі, методи та засоби штучного інтелекту в системах прийняття рішень, застосовувати інтелектуальні алгоритми з використанням нечітких моделей для розв'язання задач штучного інтелекту.
- розробляти та забезпечувати підтримку автономних розподілених інтелектуальних системи автоматизованого пошуку та аналізу інформації, створювати оптимізовані пайплайни для підготовки даних до подальшого зберігання та обробки.

ФОРМИ І МЕТОДИ НАВЧАННЯ

Курс буде викладений у формі лекцій та лабораторних занять, а також організації самостійної роботи студентів.

Основна підготовка студентів здійснюється на лекційних та лабораторних заняттях.

Під час викладання курсу використовуються такі **методи навчання:** словесні (лекція, пояснення); наочні (презентація Power Point); практичні (лабораторні роботи); робота з літературними джерелами (самостійна робота студентів).

ВІДПОВІДНІСТЬ ЦІЛЯМ СТАЛОГО РОЗВИТКУ ДО 2030 РОКУ

Якісна освіта (ЦСР 4), а саме: 4.3. Забезпечення рівного доступу до якісної вищої освіти; 4.4. Збільшення кількості молоді з відповідними навичками для гідної праці та підприємництва. Курс формує критично важливі навички роботи з аналітичними системами, що підвищує готовність студентів до роботи в ІТ та аналітичних сферах.

Гідна праця та економічне зростання (ЦСР 8), а саме: 8.2. Підтримка інновацій та підвищення економічної продуктивності через сучасні ІТ-технології. Навички роботи з OLAP-системами підвищують конкурентоспроможність випускників на ринку праці, оскільки аналіз даних є ключовим елементом у багатьох галузях, включаючи фінанси, маркетинг, логістику та управління.

Індустріалізація, інновації та інфраструктура (ЦСР 9), а саме: 9.5. Підтримка наукових досліджень і розвиток інноваційної діяльності. OLAP-системи використовуються для обробки великих обсягів даних, що актуально для наукових досліджень, фінансової аналітики, управління проектами та інновацій.

Відповідальне споживання і виробництво (ЦСР 12), а саме: 12.6. Заохочення підприємств до впровадження сталих практик через аналітичні системи. OLAP-системи дозволяють оптимізувати використання ресурсів за рахунок аналізу даних та виявлення неефективних процесів, що сприяє більш відповідальному використанню ресурсів і зменшенню втрат у виробництві та логістиці.

Партнерство заради сталого розвитку (ЦСР 17), а саме: 17.18. Посилення можливостей для збору, аналізу та використання якісних статистичних даних. Курс сприяє підготовці фахівців, здатних працювати з комплексними даними для підтримки прийняття обґрунтованих рішень на глобальному рівні.