

ОБРОБКА ВЕЛИКИХ ДАНИХ (BIG DATA)

АНОТАЦІЯ

В теперішній час дані стали таким же важливим фактором виробництва, як трудові ресурси чи виробничі активи. За рахунок використання великих даних, компанії можуть отримувати відчутні конкурентні переваги. Технології Big Data є корисними при вирішенні наступних задач: – прогнозування ринкової ситуації; – маркетинг і оптимізація продажів; – вдосконалення продукції; – ухвалення управлінських рішень; – підвищення продуктивності праці; – ефективна логістика; – моніторинг стану основних фондів. На виробничих підприємствах великі дані генеруються також внаслідок впровадження технологій Промислового інтернету речей. При цьому основні вузли і деталі станків і машин оснащуються датчиками, виконавчими пристроями, контролерами та процесорами, здатними виробляти обмежені обчислення. Під час виконання технологічних процесів здійснюється постійний збір даних, їх попередня обробка (фільтрація, перетворення та ін.). Аналітичні платформи обробляють результати у найбільш зручному для сприйняття вигляді і зберігають для подальшого використання. На основі аналізу отриманих даних робляться висновки про стан обладнання, ефективність внесених змін у технологічні процеси і т.д.

Основними особливостями обробки великих даних є застосування методів аналізу, придатних до обробки різнотипних даних (числові дані, геодані, слабоструктуровані звіти тощо), потреба у значних людських ресурсах для підтримки процесу аналізу даних, висока обчислювальна складність наявних алгоритмів аналізу та стрімке зростання обсягу зібраних даних. Відбувається активний розвиток засобів для оперативного збору різнотипних даних, завантаження їх у сховище даних, аналізу та прогнозування. Опрацювання та аналіз різнотипних даних використовується для моделювання розвитку подій та ситуацій, а також в системах підтримки прийняття рішень.

Метою вивчення дисципліни «Обробка великих даних (Big Data)» є освоєння студентами інформаційних технологій, які дозволяють проектувати, розгортати та використовувати інформаційні системи для накопичення та обробки Big Data з подальшою можливістю здобування знань з даних і підтримки прийняття рішень.

Під час вивчення дисципліни студенти отримують знання, вміння та навички: архітектури сучасних сховищ великих даних; технологій їх проектування та побудови, методів та інструментів обробки, в тому числі розподіленої, великих даних; методів та інструментів використання сховищ великих даних, в тому числі відкритих, для видобутку закономірностей функціонування предметних областей.



Освітній рівень

МАГІСТР

Кількість кредитів

5,0

(як обов'язкова)

5,0

(як вибіркова)

Мова навчання

УКРАЇНСЬКА
(ОКРЕМІ ДЖЕРЕЛА
ІНФОРМАЦІЇ –
ЧАСТКОВО
АНГЛІЙСЬКОЮ
МОВОЮ)

Назва кафедри,
яка пропонує
дисципліну

Кафедра
цифрових
технологій та
проектно-
аналітичних
рішень

САГАЙДА Павло

Доктор технічних наук, доцент,
фахівець у сфері організації баз даних і знань,
інтелектуальної обробки даних та Data Science

pavlo.sahaida@mipolytech.education



ВИМОГИ ДО ПОПЕРЕДНЬОГО РІВНЯ ЗНАНЬ

- Базові знання із вищої математики та математичного аналізу, теорії ймовірності та математичної статистики, основ теорії реляційних баз даних.
- Знання щодо вибору архітектури програмного рішення та навички проектування програмного забезпечення, в тому числі реляційних баз даних, побудови запитів на мові SQL
- Знання основ програмування, в тому числі об'єктно-орієнтованого програмування

РЕЗУЛЬТАТИ НАВЧАННЯ

- Мати спеціалізовані концептуальні знання, що включають сучасні наукові здобутки у сфері комп'ютерних наук і є основою для оригінального мислення та проведення досліджень, критичне осмислення проблем у сфері комп'ютерних наук та на межі галузей знань.
- Здатність розробляти концептуальну модель інформаційної або комп'ютерної системи, в тому числі з використанням знань щодо етапів життєвого циклу систем цифрового інтелекту
- Вміння аналізувати існуючі цифрові технології, проектувати, розробляти та впроваджувати на підприємствах різних галузей економіки системи обробки великих даних, використовуючи сучасні знання бізнес-аналізу, методів інтелектуальної обробки даних, технологій видобування знань предметної області, підтримки інженерної та організаційної діяльності
- Здатність використовувати сучасні засоби організації сховищ великих даних, інструменти аналізу великих даних та відповідні системи розробки програмних комплексів для проектування, розгортання та застосування інформаційних систем для накопичення та обробки Big Data з подальшою можливістю здобування знань з даних і підтримки прийняття рішень.

МЕТОДИ І ФОРМИ НАВЧАННЯ

Освітній процес будується як комбінація лекцій та самостійного вивчення навчального матеріалу на платформі Moodle — з одного боку, та практичних занять з відпрацювання навичок проектування та програмування — з іншого. Лабораторні та практичні заняття передбачають виконання завдань з розробки даталогічних моделей сховищ великих даних, побудови сховищ даних з використанням сучасних СУБД та інтегрованих служб аналізу даних, рішення типових завдань аналізу даних, в тому числі відкритих великих даних. Окрім роботи на цих заняттях студент повинен виконати індивідуальні завдання, результати яких враховуються у модульних контрольних роботах, а також дозволяють перевірити отримані студентом компетентності в рамках даної дисципліни. Опціонально доступні індивідуальні та групові консультації.

ПІДХОДИ ДО ОЦІНЮВАННЯ

*Складові оцінювання успішності для здобувачів освіти за програмою
«Комп'ютерні науки та цифровий інтелект»*

Назва і стислий зміст контрольного заходу	Кількість балів денна форма
Робота на практичних заняттях	60
Виконання індивідуальних аналітично-розрахункових завдань	20
Модульні контрольні роботи	20
Всього (О)	100

*Складові оцінювання успішності для здобувачів освіти, які обрали дисципліну
«Обробка великих даних (Big data)» в рамках індивідуальної освітньої траєкторії*

Назва і стислий зміст контрольного заходу	Кількість балів денна форма
Робота на практичних заняттях	60
Виконання індивідуальних аналітично-розрахункових завдань	20
Модульні контрольні роботи	20
Всього (О)	100

- Модульні контрольні роботи складаються на практичних заняттях за розкладом, графік складання контрольних точок повідомляється викладачем на початку викладання освітнього компоненту, однак вони мають бути захищені не пізніше, як за один тиждень до закінчення семестру (теоретичного навчання) для виставлення оцінки за поточну успішність (О).
- Підсумкова оцінка (ПО) за освітній компонент здобувачам освіти визначається на момент закінчення сесійного контролю за результатами остаточної оцінки всіх контрольних заходів, в т.ч. тих, які були складені після завершення теоретичного навчання, а в разі невиконання вимог даної робочої програми – у встановлені терміни ліквідації академічної заборгованості. Переведення кількості балів у шкалу ECTS (A, B, C, D, E, F, FX) та інші шкали здійснюється відповідно до регламентів Університету.
- В рамках процедур визнання та перезарахування кредитів, отриманих в рамках формальної освіти, враховуються кредити та оцінка результатів навчання з дисциплін, споріднених за змістом, отримані на бакалавському рівні освіти при відновленні/переведенні на навчання до Університету.

ПІДХОДИ ДО ОЦІНЮВАННЯ

- Результати неформальної або інформальної освіти можуть бути визнані в рамках оцінювання окремих індивідуальних завдань за узгодженням з викладачем.
- Результати участі у науковій роботі (статті, тези виступів, конкурсні наукові роботи тощо) можуть бути визнані в рамках оцінювання окремих індивідуальних завдань і модульних контрольних робіт за узгодженням з викладачем.

ЛІТЕРАТУРА

1. Пасічник, В.В. Організація баз даних та знань / В. В. Пасічник, В. А. Резниченко. – К.: Видавнича група BHV, 2006. – 384 с.
2. Date, C.J. An Introduction to Database Systems (8th. ed.). Addison-Wesley Longman Publishing Co., Inc., USA, 2003.
3. Ben-Gan, I. T-SQL Fundamentals. Microsoft Press, 2023. ISBN: 978-0138102104
4. Petkovic, D. Microsoft SQL Server 2019: A Beginner's Guide. McGraw-Hill Education, 2020. ISBN: 978-1260458879.
5. Byrne, G. Target C#: Simple Hands-On Programming with Visual Studio 2022. Apress, 2022. ISBN: 978-1484286180.
6. IoT Fundamentals: Big Data & Analytics // Електронний ресурс. Режим доступу: <https://www.netacad.com/courses/iot/big-data-analytics>
7. Containerization // Електронний ресурс. Режим доступу: <https://www.ibm.com/cloud/learn/containerization>
8. Apache Spark // Електронний ресурс. Режим доступу: <https://spark.apache.org/>
9. MapReduce Tutorial // Електронний ресурс. Режим доступу: https://hadoop.apache.org/docs/r1.2.1/mapred_tutorial.html
10. Harinath, S., Pihlgren, R., Lee, D. Professional Microsoft SQL Server 2012 Analysis Services with MDX and DAX. Wrox, 2012. ISBN: 978-1118101100

АКАДЕМІЧНІ ПОЛІТИКИ

Як член студентської спільноти Технічного університету «МЕТІНВЕСТ ПОЛІТЕХНІКА» Ви маєте дотримуватися певних стандартів та академічної політики

[Академічні політики - Polytechnic \(metinvest.university\)](https://www.metinvestpolytechnic.edu.ua/academy)

- Шахрайство та плагіат заборонені.
- Матеріали в рамках курсу, захищені авторським правом, можуть бути використані лише тільки здобувачами освіти, яким призначено даний курс. зарахованих на курс для цілей, пов'язаних з цим курсом і не можуть поширюватися.
- Спілкування з однокурсниками та викладачем має бути професійним та ввічливим.
- Очікується, що Ви перевірятимете всі Ваші письмові повідомлення, включаючи поштові повідомлення, на коректність змісту та мови.
- Університет прагне підтримувати середовище, вільне від дискримінації або дискримінаційних домагань, спрямованих на будь-яку людину або групу в межах своєї спільноти - здобувачів освіти, співробітників або відвідувачів.