



Силабус курсу

Аналітика великих даних

Освітньо-професійна програма «Системний аналіз»
Ступінь вищої освіти – бакалавр
Спеціальність: 124 «Системний аналіз»

Рік навчання: 4 Семестр: 7

Кількість кредитів: 5 Мова викладання: українська

Керівник курсу

ПШ

к.т.н., старший викладач Олег Адамів

Контактна інформація

oleh.adamiv@wunu.edu.ua

Опис

Предметом дисципліни є вивчення фундаментальних принципів, методів та технологій роботи з великими даних: від збору та зберігання (у розподілених чи хмарних сховищах) до аналітичної обробки, візуалізації та застосування сучасних інструментів (Hadoop, Spark) з урахуванням етичних і правових аспектів роботи з даними.

Мета та цілі курсу

Передбачено вивчення принципів, методів та технологій обробки великих обсягів даних, включно з побудовою сховищ, застосуванням алгоритмів аналітики та інструментів (Hadoop, Spark) для ефективного отримання, аналізу та візуалізації даних у різних галузях.

Завдання дисципліни. В результаті вивчення курсу "Аналітика великих даних" студенти повинні:

- знати концепти та джерела Big Data: ознайомитися з характеристиками великих даних (5V), особливостями роботи з структурованими й неструктурованими даними, а також з поняттям IoT (Інтернет Речей) як головним рушієм зростання обсягів даних;
- опанувати технології та інструменти обробки великих даних: знати принципи та інфраструктуру Hadoop (HDFS, YARN, MapReduce), а також NoSQL-системи для зберігання розподілених даних; навчитися застосовувати Apache Spark (Spark Core, Spark Streaming, Spark SQL, MLlib) для обчислень у пам'яті, потокової обробки та машинного навчання; оволодіти базовими підходами до візуалізації (Big Data Visualization, Gephi) та технологіями хмарних обчислень у контексті сховищ великих даних;
- вміти використовувати результати аналітики для прийняття рішень, дотримуючись етичних і правових норм: застосовувати методи візуалізації, класифікації й аналізу великих даних для формування обґрунтованих висновків та стратегічних рішень у бізнесі та науці; урахувати етичні вимоги, законодавчі обмеження й принципи збереження приватності під час роботи з конфіденційною інформацією.

Перелік тем

1. Введення в аналітику великих даних
2. Джерела великих даних
3. Організація сховищ великих даних
4. Основні технології та інструменти роботи з великими даними
5. Технології Hadoop Big Data. Розподілена обробка MapReduce
6. Apache Spark

Рекомендовані джерела інформації

Основна література

1. Ланде Д. В., Субач І. Ю., Гладун А. Я. Оброблення надвеликих масивів даних (BIG DATA) : навчальний посібник. Київ 2021. 168 с.
2. Литвин В. В., Пасічник В. В., Нікольський Ю. В. Аналіз даних та знань: навчальний посібник. Львів: Магнолія, 2021. 276 с.
3. Тарн В.І., Гордієнко Ю. Г., Стіренко С.Г. / Технології BIG DATA Практикум // Київ: КПІ ім. Ігоря Сікорського, 2022. 56 с.

Додаткова література:

1. Олещенко, Л. М. Технології оброблення великих даних. Конспект лекцій [Електронний ресурс] : навчальний посібник для студентів спеціальності 121 «Інженерія програмного забезпечення» (освітня програма «Інженерія програмного забезпечення мультимедійних та інформаційно-пошукових систем») / Л. М. Олещенко ; КПІ ім. Ігоря Сікорського. – Електронні текстові дані (1 файл: 5,55 Мбайт). – Київ : КПІ ім. Ігоря Сікорського, 2021. – 227 с. // <https://ela.kpi.ua/items/4ad484d3-aa21-4f7e-be09-a6d8e1d1e8bc>
2. Чаплінський В. Р., Кушнір О. К., Свідер О. П. Аналіз великих даних та їх візуалізація для потреб бізнесу. *Ефективна економіка*. 2021. № 6. – URL: <http://www.economy.nayka.com.ua/?op=1&z=8979>. DOI: [10.32702/2307-2105-2021.6.89](https://doi.org/10.32702/2307-2105-2021.6.89)
3. Benjamin S. Baumer, Daniel T. Kaplan, Nicholas J. Horton. Modern Data Science with R. 2nd Edition. Chapman and Hall/CRC, 2021. 632 p.
4. Hadley Wickham, Mine Çetinkaya-Rundel, Garrett Grolemund. R for Data Science: Import, Tidy, Transform, Visualize, and Model. O'Reilly Media, Inc., 2023. 576 p.

Оцінювання

Модуль 1		Модуль 2	Модуль 3
40%	40%	5%	15%
Поточне оцінювання	Модульний контроль	Тренінг	Самостійна робота
Оцінка за даний модуль визначається як середнє арифметичне за роботу на лабораторних заняттях	Підсумкова письмова робота	Визначається як середнє арифметичне за виконання завдань тренінгу	Оцінка за виконання самостійної роботи

Шкала оцінювання:

За шкалою ЗУНУ	За національною шкалою	За шкалою ECTS
90–100	Відмінно	A (відмінно)
85–89	Добре	B (дуже добре)
75–84		C (добре)
65–74	Задовільно	D (задовільно)
60–64		E (достатньо)
35–59	Незадовільно	FX (незадовільно з можливістю повторного складання)
1–34		F (незадовільно з обов'язковим повторним курсом)

Політики курсу

Академічна доброчесність. Дотримання академічної доброчесності студентами передбачає:

- самостійне виконання навчальних завдань, завдань поточного та підсумкового контролю результатів навчання (для осіб з особливими освітніми потребами ця вимога застосовується з урахуванням їхніх індивідуальних потреб і можливостей);
- посилення на джерела інформації у разі використання ідей, розробок, тверджень, відомостей;
- дотримання норм законодавства про авторське право і суміжні права;
- надання достовірної інформації про результати власної навчальної (наукової, творчої) діяльності, використані методики досліджень і джерела інформації.

Порушенням академічної доброчесності вважається:

академічний плагіат - оприлюднення (частково або повністю) наукових (творчих) результатів, отриманих іншими особами, як результатів власного дослідження (творчості) та/або відтворення опублікованих текстів (оприлюднених творів мистецтва) інших авторів без зазначення авторства; самоплагіат - оприлюднення (частково або повністю) власних раніше опублікованих наукових результатів як нових наукових результатів;

фабрикація - вигадкування даних чи фактів, що використовуються в освітньому процесі або наукових дослідженнях;

фальсифікація - свідомо зміна чи модифікація вже наявних даних, що стосуються освітнього процесу чи наукових досліджень;

спісування - виконання письмових робіт із залученням зовнішніх джерел інформації, крім дозволених для використання, зокрема під час оцінювання результатів навчання.

За порушення академічної доброчесності здобувачі освіти можуть бути притягнені до такої академічної відповідальності:

- повторне проходження оцінювання (контрольна робота, іспит, залік тощо);
- повторне проходження відповідного освітнього компонента освітньої програми.

Політика запізнення. За несвоєчасно виконані завдання буде накладено штраф 10 відсотків від загальної кількості балів за це завдання. Примітка. Виключення можуть бути зроблені до невчасно зданих завдань з поважних причин.