

Силабус курсу Аналіз даних

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

Рік навчання: I, Семестр: 2

Кількість кредитів: 5 Мова викладання: українська

Керівник курсу

ПП

к.е.н., доцент Пасічник Роман Мирославович

Контактна інформація

roman.pasi@gmail.com

Опис дисципліни

Даний курс знайомить студентів з предметною областю великих даних (big data), а також показує взаємозв'язки з наукою про дані (data science) та аналізом даних (data analytics) та забезпечує: оволодіння навичками оперування великими даними шляхом використання спеціалізованих програмно-апаратних засобів, зокрема хмарних сервісів, спеціалізованих систем зберігання даних, розподілених файлових систем, тощо; розвиток уміння працювати з великим обсягом даних з урахуванням їх ключових характеристик: обсягу, різноманітності, мінливості й забезпечення відповідного рівня швидкості опрацювання; розвиток навичок ефективного використання математичного, алгоритмічного та програмного забезпечення для розв'язки основних задач предметної області великих даних.

Студенти зможуть ефективно використовувати парадигми паралельного опрацювання даних, зокрема MapReduce та системи Apache Spark, відповідні хмарні служби Amazon Web Services та IBM Bluemix; розгортати надійні та швидкі сховища для надвеликих обсягів даних; використовувати програмні бібліотеки та фреймворки з ефективними алгоритмами опрацювання надвеликих обсягів даних.

Мета та цілі курсу.

Метою курсу «Аналіз даних» є отримання теоретичних знань та практичних вмінь і навиків, достатніх для успішного виконання професійних обов'язків фахівців в області інформаційних технологій та системного аналізу, для подальшого працевлаштування.

Курс направлений на вивчення і набуття умінь в сфері інженерії даних та знань в технологічних процесах розроблення і підтримки комп'ютерних систем, які базуються на інтелектуальному аналізі великих даних.

В процесі навчання студенти набувають системних практичних навиків, як в технологічній області, так і в області виробничо-дослідницької діяльності.

В результаті навчання студенти отримують сучасні навички, які необхідні для провадження інженерної та наукової діяльності для створення прикладних інформаційних продуктів.

Результати навчання:

1. Використовувати інструментальні засоби інтеграції різнотипових даних у наборах різної розмірності;

2. Здатність до професійної роботи з даними: видобуток (у тому числі повторний), фільтрації, інтеграція, зберігання, перевірки актуальності, валідація і репрезентативність;

3. Видобувати знання шляхом інтеграції та аналізу великих даних, отриманих з різноманітних та різнорідних джерел інформації;

4. Проектувати сховища великих даних, для видобутку даних і знань;

5. Проектувати та розробляти прикладні інформаційні продукти для видобутку корисних знань і вироблення оптимальних стратегій прийняття рішень;

6. Аналізувати науково-технічну інформацію при проведенні науко-дослідних робіт.

4. Перелік тем

1. Вступ в аналіз і обробку великих даних
2. Життєвий цикл проекту по аналітиці великих даних
3. Основні техніки, щодо роботи з великими даними
4. Конгнітивний аналіз даних
5. Аналіз візуалізації великих даних
6. Основні технології та інструменти роботи з великими даними
7. Наукові проблеми у сфері великих даних

Літературні джерела

1. <https://rb.ru/howto/что-такое-big-data/>
2. <https://postnauka.ru/faq/46974>
3. <https://www.datacenterknowledge.com/archives/2015/03/30/big-data-bubble-set-burst>
4. [http://www.tadviser.ru/index.php/Статья:Большие_данные_\(Big_Data\)](http://www.tadviser.ru/index.php/Статья:Большие_данные_(Big_Data))
5. https://ru.wikipedia.org/wiki/Большие_данные
6. <https://intellect.ml/big-data-6821>
7. [http://www.mckinsey.com/insights/business_technology/big_data_the_next_frontier_f or_innovatin](http://www.mckinsey.com/insights/business_technology/big_data_the_next_frontier_f_or_innovatin)
9. <http://www.ogcs.com.ua/index.php/articles/121-big-data-v-promyshlennosti-innovatsii-k-kotorym-pridetsya-priykat>

Політика оцінювання

- **Політика щодо дедлайнів та перескладання:** Роботи, які здаються із порушенням термінів без поважних причин, оцінюються на нижчу оцінку (-20 балів). Перескладання модулів відбувається із дозволу деканату за наявності поважних причин (наприклад, лікарняний).

- **Політика щодо академічної доброчесності:** Усі письмові роботи перевіряються на наявність плагіату і допускаються до захисту із коректними текстовими запозиченнями не більше 20%. Списування під час контрольних робіт та екзаменів заборонені (в т.ч. із використанням мобільних девайсів). Мобільні пристрої дозволяється використовувати лише під час он-лайн тестування (наприклад, програма Kahoot).

- **Політика щодо відвідування:** Відвідування занять є обов'язковим компонентом оцінювання, за яке нараховуються бали. За об'єктивних причин (наприклад, хвороба, міжнародне стажування) навчання може відбуватись в он-лайн формі за погодженням із керівником курсу.

Оцінювання Система оцінювання та вимоги

Заліковий модуль 1	Заліковий модуль 2	Заліковий модуль 3
30%	40%	30%
Усне опитування під час занять (4 теми по 10 балів =40 балів) Модульна контрольна робота (теми 1-4) = 60 балів	Усне опитування під час занять (3 теми по 10 балів = 30 балів) Ректорська контрольна робота (1-9) = 70 балів	Написання КППЗ =50 балів Захист КППЗ=30 балів Участь в тренінгах = 20 балів

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)

9. 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, використанні 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;

списування - 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.

Політика щодо відвідування. Відвідування занять є обов'язковим. За об'єктивних причин (наприклад, хвороба, міжнародне стажування) навчання може відбуватися в он-лайн формі за погодженням із керівником курсу.