



СИЛАБУС ДИСЦИПЛІНИ «ТЕХНОЛОГІЇ BIGDATA»

Ступінь вищої освіти - Магістр
Спеціальність 122«Комп'ютерні науки»
Освітня програма «Інформаційні управляючі системи і технології»
Рік навчання 1, семестр 1
Форма навчання денна
Кількість кредитів ЄКТС 4
Мова викладання українська

Лектор курсу

Хиленко Володимир Васильович, д.т.н., професор

Контактна інформація
лектора (e-mail)

Кафедра комп'ютерних наук,
навчальний корпус 15, к. 236, 237,
тел.: (044) 527-87-23
E-mail: y-khilenko@it.nubip.edu.ua

Сторінка курсу veLearn

ОПИС ДИСЦИПЛІНИ

Мета: Навчити студента використовувати відомі методи роботи з великими об'ємами даних при вирішенні професійних (виробничих) задач. Ознайомити студента з відповідними існуючими програмними продуктами. Надати базові знання щодо математичного апарату (математичних методів та алгоритмів) які використовуються для обробки великих об'ємів даних, методів аналітичної обробки великих даних та імітаційного моделювання. Навчити студента формулювати задачі створення баз даних, баз знань, сховищ даних для розв'язання аналітичних задач (системного аналізу / Data Analysis, імітаційного моделювання) з урахуванням особливостей предметної області.

У результаті вивчення навчальної дисципліни студент повинен

Знати:

- принципи та методи побудови інформаційних моделей;
- принципи та методи побудови сховищ даних та баз знань;
- принципи і підходи до організації процедур імітаційного моделювання,
- особливості математичних методів, математичного і алгоритмічного забезпечення яке може використовуватись для роботи з великими об'ємами даних;
- методології аналізу результатів імітаційного моделювання;
- методологію та методи оцінки достовірності отриманих результатів;
- методи побудови та дослідження математичних моделей складних динамічних систем.

Уміти:

- формулювати задачу обробки великих інформаційних масивів відповідно до задачі, що вирішується;
- формувати бази знань (сховища даних) з урахування специфіки предметної галузі досліджень;
- розробляти структуру інформаційної моделі об'єкта дослідження;
- будувати інформаційну модель (модель інформаційних потоків) об'єкту або процесу, який вивчається;
- виконати реструктуризацію моделі або, при необхідності, її декомпозицію;
- оцінити рівень достовірності отриманих результатів;
- модифікувати побудовані інформаційні моделі у відповідності до вимог і цілей замовника досліджень з метою створення спектру моделей;
- використовувати існуючі на ринку інструментальні та програмно-алгоритмічні засоби, що можуть бути використані для вирішення задач обробки великих інформаційних масивів.

Інтегральна компетентність (ІК):

Здатність розв'язувати задачі в галузі професійної та/або дослідницько-інноваційної діяльності у сфері комп'ютерних наук, що передбачає як вільне володіння наявними знаннями, так і спроможність їх застосування у професійній практиці.

Загальні компетентності (ЗК):

ЗК3. Здатність спілкуватися державною мовою як усно, так і письмово.

Фахові (спеціальні) компетентності (СК):

СК1. Розуміння теоретичних засад комп'ютерних наук для об'єктивного оцінювання можливостей використання обчислювальної техніки в певних процесах людської діяльності і визначення перспективних інформаційних технологій, у тому числі, у природоохоронній галуззі.

СК6. Здатність збирати і аналізувати дані (включно з великими), насамперед, пов'язаних з природоохоронною галуззю, для забезпечення якості прийняття рішень.

СК8. Здатність застосовувати існуючі і розробляти нові алгоритми

СК11. Здатність розробляти та адмініструвати бази даних та знань, володіти сучасними теоріями та моделями даних та знань, методами їх інтерактивної та автоматизованої розробки, технологіями обробки та візуалізації.

Програмні результати навчання (РН):

РН8. Розробляти та забезпечувати заходи з моніторингу, оптимізації, технічного обслуговування, виявлення відмов тощо.

СТРУКТУРА КУРСУ

Тема	Години (лекції/лабораторні)	Результати навчання	Завдання	Оцінювання
Модуль 1 - Формулювання і формалізація задач імітаційного моделювання				
Тема 1. Вступ до дисципліни. Сучасні тенденції в розвитку систем обробки великих обсягів даних	2/-	Знати цілі та задачі BigData. Розуміти зв'язок з іншими навчальними дисциплінами та які знання необхідні для вирішення задач, які передбачають використання великих обсягів даних	Вивчення теоретичного матеріалу.	-
Тема 2. Використання сучасних інформаційно-комунікаційних технологій для формування інформаційної бази (спектру) моделей. Масштабування БД	4/2	Знати апаратні засоби та технології для обробки великих обсягів даних. Уміти формувати і використовувати вузькопрофільну базу знань, сховище даних. Уміти обирати спеціальне програмне забезпечення для моделювання предметної області.	Здача лабораторної роботи. Вивчення теоретичного матеріалу.	10
Тема 3. Парсинг. Імітаційний експеримент і побудова імітаційних систем	4/2	Знати класифікацію вимог, стратегії вибору рішення. Уміти формувати вимоги до системи та результатів імітаційних експериментів, розробляти структуру системи. Аналізувати очікувані результати, цілі. Здійснювати оптимізацію процесу вибору рішення.	Здача лабораторної роботи. Вивчення теоретичного матеріалу.	15
Тема 4. Стиснення даних. Математична формалізація задач в імітаційному моделюванні	2/2	Знати алгоритми стиснення даних, поняття «повноти моделі». Уміти використовувати алгоритми стиснення для вирішення практичних задач. Аналізувати економічну доцільність обраного типу моделі відповідно до вимог замовника.	Здача лабораторної роботи. Вивчення теоретичного матеріалу.	15
Тема 5. Аналіз достовірності отриманих даних	2/2	Проводити аналіз БД щодо об'єкту дослідження. Обґрунтовувати доцільність використання обраного класу математичних моделей з точки зору забезпечення ефективності імітаційного моделювання.	Здача лабораторної роботи. Вивчення теоретичного матеріалу.	15
Тема 6. Реструктуризація моделей. Декомпозиція	4/4	Знати та використовувати у практичних задачах методи корегування моделей у відповідності до цілі дослідження. Аналізувати предметну область. Проводити декомпозицію, реструктуризацію моделей відповідно до вимог та цілей замовника. Обґрунтовувати вибір ступеня деталізації математичної моделі.	Здача лабораторної роботи. Вивчення теоретичного матеріалу.	15
<i>Модульний контроль №1</i>			<i>Тестування</i>	30

Модуль 2 - Програмно-алгоритмічне забезпечення для вирішення задач обробки великих обсягів даних				
Тема 7. Математичний апарат для вирішення задач обробки великих обсягів даних. Обґрунтування доцільності використання алгоритмів стиснення даних	2/4	Знати типові задачі системного аналізу. Знати та вміти використовувати методи системного аналізу, математичні методи дослідження операцій.	Здача лабораторної роботи. Вивчення теоретичного матеріалу.	15
Тема 8. Побудова системи імітаційного моделювання. Вибір структури	2/2	Уміти оптимізувати вибір структури системи відповідно до цілей модельного експерименту. Визначати критерії відповідності та прийнятності, суттєві параметри. Будувати діаграми потоків даних, станів. Аналізувати вплив факторів невизначеності.	Здача лабораторної роботи. Вивчення теоретичного матеріалу.	10
Тема 9. Сучасні програмно-алгоритмічні засоби для роботи з великими даними	2/4	Знати можливості та функції ApacheHadoop, MapReduce. Уміти використовувати бібліотеки та стандартні (ринкові) програмні продукти для роботи з великими даними.	Здача лабораторної роботи. Вивчення теоретичного матеріалу.	15
Тема 10. Особливості екологічних систем як об'єктів імітаційного моделювання з використанням технологій BigData	2/4	Знати особливості типових математичних моделей екологічних систем. Використовувати готові рішення для типових проблемних ситуацій процесу імітаційного моделювання екологічних систем. Обґрунтовувати тип обраної моделі відповідно до цілей імітаційного експерименту.	Здача лабораторної роботи. Вивчення теоретичного матеріалу.	15
Тема 11. Модифікація інформаційних та математичних моделей	4/4	Здійснювати аналіз чутливості моделей до зміни окремих параметрів та коефіцієнтів. Оцінювати об'єм досліджень різних моделей відповідно до градації частоти вирішення практичних задач.	Здача лабораторної роботи. Вивчення теоретичного матеріалу.	15
Модульний контроль №2			Тестування	30
Всього за семестр				70
Іспит			Підсумковий тест, практичне і теоретичне завдання	30
Всього за курс				100

ПОЛІТИКА ОЦІНЮВАННЯ

<i>Політика щодо дедлайнів та перескладання:</i>	Роботи, які здаються із порушенням термінів без поважних причин, оцінюються на нижчу оцінку. Перескладання модулів відбувається із дозволу лектора за наявності поважних причин (наприклад, лікарняний).
<i>Політика щодо академічної доброчесності:</i>	Списування під час контрольних робіт та екзаменів заборонені (в т.ч. із використанням мобільних девайсів). Курсові роботи, реферати повинні мати коректні текстові посилання на використану літературу
<i>Політика щодо відвідування:</i>	Відвідування занять є обов'язковим. За об'єктивних причин (наприклад, хвороба, міжнародне стажування) навчання може відбуватись індивідуально (в он-лайн формі за погодженням із деканом факультету)

ІШКАЛА ОЦІНЮВАННЯ СТУДЕНТІВ

Рейтинг здобувача вищої освіти, бали	Оцінка національна за результати складання екзаменів заліків	
	екзаменів	заліків
90-100	відмінно	зараховано
74-89	добре	
60-73	задовільно	
0-59	незадовільно	незараховано