

СИЛАБУС НАВЧАЛЬНОЇ ДИСЦИПЛІНИ **ТЕХНОЛОГІЇ ОБРОБКИ ВЕЛИКИХ ДАНИХ**

Рівень вищої освіти другий (магістерський)

Освітня програма Комп'ютерні науки, Мехатроніка та робототехніка, Обладнання легкої промисловості та побутового обслуговування, Комп'ютерно-інтегровані технологічні процеси і виробництва, Якість, стандартизація та сертифікація

1. Анотація курсу

Семестр: 2.

Обсяг модуля: загальна кількість годин – 180 з них лекції – 20 год., практичні – 20 год., самостійна робота – 140 год.; кількість кредитів ЄКТС – 6.

Мета курсу – оволодіння компетентностями здатність забезпечити організацію обчислювальних процесів витягнення та інтерпретації прихованих знань з використанням пакетів прикладних програм різних типів (tensor toolbox, Nadoor, Map reduce та ін.), орієнтованих на середовище МатЛаб; здатність до розробки інтерфейсного програмного забезпечення для поєднання джерела великих даних (ВД) з відповідним програмним середовищем; здатність реалізовувати обчислювальний процес на основі хмарних сервісів і технологій, паралельних і розподілених обчислень. набуття знань, практичних навичок та умінь використання сучасних методів, алгоритмів обробки високо вимірних, велико об'ємних, швидко змінюваних різноформатних даних, в т. ч. спотворених зовнішніми неконтрольованими завадами з урахуванням умов неповної визначеності, для витягнення прихованих знань, необхідних для підтримки прийняття рішень;

Результати навчання дисципліни:

знати: характеристики нового об'єкта інформаційних технологій, відомого як ВД, розуміти його роль та місце в системах автоматизованого проектування в легкій промисловості, основні методи, моделі, програмно-технічні засоби, що дозволяють реалізувати обробку багато вимірних, швидко-змінюваних та високо-об'ємних масивів даних;

вміти: застосувати набуті знання в професійній діяльності під час розробки, налагодження та експлуатації інформаційних систем з використанням штучного інтелекту та за нестандартних умов проведення дослідження;

володіти: навичками застосування математичних моделей представлення ВД, використання сучасних програмних засобів обробки ВД, інтерпретації отриманих результатів аналізу, пошуку прихованих знань, розумінням шляхів розвитку інтелектуальних технологій при дослідженні складних об'єктів, систем, обладнання та технологічних процесів в промисловості;

здатен продемонструвати: методи обробки ВД на прикладах представлення високо-вимірних (3D) об'єктів моделями низької (1D-2D) вимірності, реальну підтримку прийняття рішень при наявності високо об'ємних масивів даних.

Необхідні навчальні компоненти (пререквізити, кореквізити і постреквізити): вища математика, алгоритмізація і програмування, фундаментальні принципи розробки програмного забезпечення, основні системні принципи та інтелектуальні системи .

Зміст дисципліни: **Тема 1.** Аналітика Великих Даних. Вступ в аналітику Великих Даних (ВД). Основні поняття і визначення. Історія питання. Джерела ВД. Технології формування і обробки ВД. Застосування ВД: в економіці, бізнесі, охороні здоров'я, промисловості, приклади використання. Великі дані в наукових областях, Особливості застосування. Big Data, Data Mining. Datafication. МатЛаб як потужний і головний засіб обробки ВД і прийняття рішень за умов невизначеності в середовищі ВД.

Тема 2. Управління даними, життєвий цикл ВД. Фази життєвого циклу даних, умовність циклу, принципи формування. Управління даними, 7 головних міркувань у створенні гібридної стратегії інтеграції.

Тема 3. Проблеми уявлення і моделювання знань і їхній зв'язок з проблемою ВД. Комп'ютерні знання – база даних + база правил. Концептуальні та емпіричні моделі знання. Data Mining - дослідження і виявлення «машиною» в сирих даних прихованих знань.

«Штучний інтелект» як синтез "data mining", "knowledge discovery" і "hidden knowledge".
Методи вилучення нових знань з баз фактів. Матричні декомпозиції та їхня роль в аналізі даних.

Тема 4. Багатовимірні масиви, тензорні моделі. Традиційна DataScience - Big Data - загальні риси та відмінності. Тенденція сучасності – від 2D матриць до 3D тензорів. Доцільність і необхідність машинного навчання при опрацюванні результатів обробки ВД. База головного понятійного апарат для аналізу ВД: методи чисельного аналізу на підставі тензорних представлень даних; новітні методи і моделі обробки великих даних, заснованих на тензорних мережах та тензорних декомпозиціях. Тензор – головний об'єкт аналізу багатовимірних великих даних. Матриці, матричні розкладення.

Тема 5. Тензорна алгебра, застосування тензорів. Коротка історія сучасного тензорного аналізу. Тензори, основа, нотація, мотивація. Волокна та слайси – головні форми представлення багатовимірних тензорів. Узагальнення математичних операцій в тензорній та матричній алгебра. Ранг матриці і ранг тензора. Тензорні декомпозиції, високо порядкові декомпозиції. Оператори векторизації та реорганізації. Перетворення тензорів у вектори. Розгортання тензорів засобами МатЛаб. Суми rank-1 матриць та rank-1 тензорів. Розгортання 3D тензорів по одному з вимірів.

Тема 6. Інструментальні та програмні засоби роботи з великими даними. Основні принципи роботи з даними; інструменти; приклади розв'язання практичних завдань. Машинне навчання. Принципи роботи з великими даними, парадигма MapReduce, Hadoop і корпоративні системи, Hadoop і СУРБД. Сучасні головні засоби МатЛаб, орієнтовані на роботу з ВД: Parallel Computing Toolbox, Database Toolbox, MATLAB Distributed Computing Server. Пакети прикладних програм (Tensor Matlab Toolbox) для роботи з 3D тензорами.

Види навчальних занять: лекція, практичне, консультація.

Форми навчання: денна, заочна, заочна (дистанційна).

Методи навчання: словесний, пояснювально-демонстраційний, метод проблемного викладання, репродуктивний.

Методи контролю: усний, письмовий, тестовий.

Форми підсумкового контролю: екзамен (семестр 2).

Засоби діагностики успішності навчання: індивідуальні завдання, питання для поточного контролю, тести, питання для підсумкового контролю.

Мова навчання: українська, англійська, російська

2. Оцінювання

Розподіл балів, які отримують студенти

Поточне оцінювання та самостійна робота							Екзамен	Сума
T1	T2	T3	T4	T5	T6	МК (тест.)		
10	10	10	10	10	10	30	10	90
Всього з дисципліни								100

Розподіл балів з дисципліни

Види оцінювання	T1	T2	T3	T4	T5	T6	Усього
Виконання практичного завдання	10	10	10	10	10	10	60
Тестовий контроль				15		15	30
Екзамен	10						10
Всього з дисципліни							100

Критерії оцінювання екзамену

Опис відповіді	Питання 1	Питання 2	Задача	Максимум
Відповідь повна без помилок	3			3
Відповідь неповна з незначною кількістю несуттєвих помилок	2			
Відповідь неповна з суттєвими помилками	1			
Відповідь повна без помилок		3		3
Відповідь неповна з незначною кількістю несуттєвих помилок		2		
Відповідь неповна з суттєвими помилками		1		
Задача розв'язана, пояснення повні			4	3
Задача розв'язана, пояснення неповні			3	
Задача розв'язана, пояснення містять несуттєві помилки			2	
Задача розв'язана, пояснення містять суттєві помилки			1	
Разом, максимум				10

Відповідність шкал оцінок якості засвоєння навчального матеріалу

Оцінка за національною шкалою	Оцінка в балах	Оцінка за шкалою ECTS	Пояснення
відмінно	90-100	A	Відмінно (відмінне виконання лише з незначною кількістю помилок)
добре	82-89	B	Дуже добре (вище середнього рівня з кількома помилками)
	74-81	C	Добре (в загальному вірне виконання з певною кількістю суттєвих помилок)
задовільно	64-73	D	Задовільно (непогано, але зі значною кількістю недоліків)
	60-63	E	Достатньо (виконання відповідає мінімальним критеріям)
незадовільно	35-59	FX	Незадовільно (з можливістю повторного складання)
	0-34	F	Незадовільно (з обов'язковим повторним вивченням дисципліни)

3. Політика курсу

Обов'язкове дотримання академічної доброчесності студентами, а саме:

- самостійне виконання всіх видів робіт, завдань, форм контролю, передбачених робочою програмою даної навчальної дисципліни;
- 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.

У випадку порушення норм академічної доброчесності під час виконання завдання, зокрема академічного плагіату, студент отримує 0 (нуль) балів за завдання. При цьому викладач повинен надати докази факту порушення

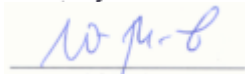
Обговорено та рекомендовано на засіданні кафедри комп'ютерних наук та технологій.
Протокол від 03.02 2020 року № 6

Завідувач кафедри КНТ



В.Ю. Щербань

Викладач



Ю.М. Мінаєв