

Міністерство освіти і науки України

Харківський національний університет імені В.Н. Каразіна

Кафедра прикладної математики

“ЗАТВЕРДЖУЮ”

Декан факультету
математики і інформатики

Григорій ЖОЛТКЕВИЧ

“29” серпня 2024 р.



РОБОЧА ПРОГРАМА НАВЧАЛЬНОЇ ДИСЦИПЛІНИ

Прикладні задачі аналізу «великих даних»

рівень вищої освіти _____ перший (бакалаврський) рівень _____

галузь знань 11 – Математика та статистика _____

спеціальність 113 – Прикладна математика _____

освітня програма «Прикладна математика» _____

спеціалізація _____

вид дисципліни _____ за вибором _____

факультет _____ математики і інформатики _____

2024 / 2025 навчальний рік

Програму рекомендовано до затвердження Вченою радою факультету математики і інформатики

“27” серпня 2024 року, протокол № 8

РОЗРОБНИКИ ПРОГРАМИ: **Кізілова Наталія Миколаївна**, доктор фіз.-мат. наук, професор, професор закладу вищої освіти кафедри прикладної математики.

Програму схвалено на засіданні кафедри прикладної математики
Протокол від “26” серпня 2024 року № 8

Завідувач кафедри прикладної математики



Валерій КОРОБОВ

Програму погоджено з гарантом
освітньо-професійної програми «Прикладна математика»

Гарант освітньо-професійної програми «Прикладна математика»

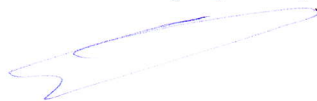


Сергій ПОСЛАВСЬКИЙ

Програму погоджено науково-методичною комісією
факультету математики і інформатики

Протокол від “27” серпня 2024 року № 1

Голова науково-методичної комісії факультету математики і інформатики



Євген МЕНЯЙЛОВ

ВСТУП

Програма навчальної дисципліни “**Прикладні задачі аналізу великих даних**” складена відповідно до освітньо-професійної програми підготовки бакалавра спеціальності 113 Прикладна математика

1. Опис навчальної дисципліни

1.1. Мета викладання навчальної дисципліни

Метою викладання навчальної дисципліни “Прикладні задачі аналізу великих даних” є надання знань в галузі сучасного статистичного аналізу «великих даних», відповідних математичних і аналітичних моделей, їх програмної реалізації та практичних застосувань у розв’язанні конкретних прикладних задач.

1.2. Основні завдання вивчення дисципліни.

Основними завданнями вивчення дисципліни “Прикладні задачі аналізу великих даних” є вивчення студентами основних статистичних методів, математичних моделей та підходів у аналізі різних типів «великих даних», знання прикладних пакетів програм для розв’язання актуальних прикладних проблем.

1.3. Кількість кредитів – 4

1.4. Загальна кількість годин – 120

1.5. Характеристика навчальної дисципліни	
За вибором	
Денна форма навчання	Заочна (дистанційна) форма навчання
Рік підготовки	
4-й	
Семестр	
8-й	
Лекції	
32 год	
Практичні, семінарські заняття	
32 год	
Лабораторні заняття	
Самостійна робота	
56 год	
у тому числі індивідуальні завдання	

1.6. Заплановані результати навчання

студенти повинні досягти таких результатів навчання:

Знати:

1. основні методи та підходи до аналізу «великих даних»;
2. статистичні та аналітичні методи досліджень «великих даних»;
3. сучасні пакети прикладних програм для статистичного аналізу «великих даних»;

4. структуру та особливості роботи з базами «великих даних» в мережі Інтернет;

Вміти:

1. використовувати сучасні статистичні методи для аналізу «великих даних» в різних прикладних задачах;
2. вміти користуватися існуючими програмними пакетами та розробляти власний програмний код для чисельної реалізації аналізу «великих даних»;
3. розробляти системи штучного інтелекту для аналізу «великих даних»;
4. проводити пошук інформації у вигляді «великих даних» в мережі Інтернет.

2. Тематичний план навчальної дисципліни.

Розділ 1. Статистичні методи та аналітичні моделі аналізу «великих даних».

Тема 1. Види і структура різних типів «великих даних».

Тема 2. Основні статистичні методи аналізу «великих даних».

Тема 3. Головні аналітичні моделі науки о даних.

Тема 4. Методи та алгоритми штучного інтелекту в аналізі «великих даних».

Розділ 2. Конкретні прикладні задачі обробки та аналізу «великих даних».

Тема 5. Методи «великих даних» для аналізу медичної інформації.

Тема 6. Реалізація аналізу «великих даних» в біоінформації.

Тема 7. Реалізація аналізу «великих даних» в геоінформації.

Тема 8. Реалізація методів «великих даних» при аналізі економічної інформації.

3. Структура навчальної дисципліни

Назви розділів і тем	Кількість годин											
	Денна форма						Заочна форма					
	Усього	у тому числі					Усього	у тому числі				
		л	п	лаб	інд	ср		л	п	лаб	інд	ср
1	2	3	4	5	6	7	8	9	10	11	12	13
Розділ 1. Статистичні методи та аналітичні моделі аналізу «великих даних»												
Тема 1. Види і структура різних типів «великих даних»	15	4	4			7						
Тема 2. Основні статистичні методи аналізу «великих даних»	15	4	4			7						
Тема 3. Головні аналітичні моделі	15	4	4			7						

науки о даних											
Тема 4. Методи та алгоритми штучного інтелекту в аналізі «великих даних»	15	4	4			7					
Разом за розділом 1	60	16	16			28					
Розділ 2. Конкретні прикладні задачі обробки та аналізу «великих даних».											
Тема 5. Методи «великих даних» для аналізу медичної інформації	15	4	4			7					
Тема 6. Реалізація аналізу «великих даних» в біоінформатиці	15	4	4			7					
Тема 7. Реалізація аналізу «великих даних» в геоінформатиці	15	4	4			7					
Тема 8. Реалізація методів «великих даних» при аналізі екологічної інформації	15	4	4			7					
Разом за розділом 2	60	16	16			28					
Всього годин	120	32	32			56					

4. Теми семінарських (практичних, лабораторних) занять

№ з/п	Назва теми	Кількість годин
1	Структура масивів «великих даних» на прикладах	4
2	Регресивний, кореляційний, факторний і кластерний аналіз даних.	4
3	Методи аналізу та розпізнавання зображень	4
4	Аналітичні моделі для дослідження «великих даних»	4
5	Алгоритми штучного інтелекту в аналізі «великих даних»	4
6	Типи медичної інформації та засоби її обробки	4
7	Типи біологічної інформації та засоби її обробки	4
8	Типи геофізичної інформації та засоби її обробки	4
	Разом	32

5. Завдання для самостійної роботи

№ з/п	Види, зміст самостійної роботи	Кількість годин
1	Виконання домашніх завдань за темою «Інтернет речей (Internet of Things) та методи аналізу «великих даних»»	20
2	Написання та валідація програмного коду для аналізу конкретної медичної інформації	18
3	Написання та валідація програмного коду для аналізу медичних зображень	18
	Разом	56

6. Індивідуальні завдання

Не передбачено навчальним планом

7. Методи навчання

Частково-пошукові і проблемні лекції, пояснення, розрахунково-ілюстративні і пояснювально-ілюстративні методи при проведенні практичних занять, консультації.

8. Методи контролю

- Перевірка виконання домашніх завдань
- Проведення і перевірка контрольної роботи
- Проведення заліку.

9. Схема нарахування балів

Поточний контроль, самостійна робота, індивідуальні завдання					Залік	Сума
Розділ 1	Розділ 2	Контрольна робота, передбачена навчальним планом	Індивідуальне завдання	Разом		
T1–T4	T5–T8					
25	25	10	-	60	40	100

T1, T2 ... – теми розділів.

Для допуску до складання підсумкового контролю (заліку) здобувач вищої освіти повинен набрати не менше 10 балів з навчальної дисципліни під час поточного контролю, самостійної роботи, індивідуального завдання.

Критерії оцінювання навчальних досягнень

Поточний контроль: бали нараховуються за виконання домашніх завдань і активність під час практичних занять.

Контрольна робота оцінюється в 10 балів.

Залікова робота складається з двох теоретичних і одного практичного залікового завдання. Кожне теоретичне завдання оцінюється у 10 балів, а практичне завдання – у 20 балів максимально.

По кожному завданню нараховується:

- максимальний бал у разі правильної обґрунтованої відповіді;
- за незначні та за арифметичні помилки оцінка зменшується від 10 до 30 відсотків;
- за значні логічні помилки оцінка зменшується до 50 відсотків, якщо хід розв'язання в цілому правильний,
- у разі частково правильних міркувань за відсутності обґрунтованої відповіді виставляється до 30 відсотків від максимальної кількості балів
- розв'язання не відповідає жодному з критеріїв, які сформульовані вище, – виставляється 0 балів.

Шкала оцінювання (дворівнева)

Сума балів за всі види навчальної діяльності протягом семестру	Оцінка
50-100	зараховано
1-49	не зараховано

10. Рекомендована література

Основна література

1. Кізілова Н.М. Методи аналізу «великих даних». Методичні рекомендації з курсу «Прикладні задачі аналізу великих даних» для студентів спеціальності «Прикладна математика». Харків: ХНУ ім. В.Н. Каразіна. 2021. – 96 с.
2. Шаховська Н.Б., Болюбаш Ю.Я., Верес О.М. Організація великих даних у розподіленому середовищі. // Обчислювальна техніка та автоматизація. Вісник ДонНТУ. 2014. – № 2(27). - С.147–155.

Допоміжна література

1. Lowman M. A Practical Guide to Analytics for Governments: Using Big Data for Good. John Wiley & Sons. 2017. 224p.
2. Prasad Y.L. Big Data Analytics. Made Easy. Notial Press. 2016. 228p.

11. Посилання на інформаційні ресурси в Інтернеті, відео-лекції, інше методичне забезпечення

1. Omar H., Jumaa A.K. Big Data Analysis Using Apache Spark MLlib and Hadoop HDFS with Scala and Java
https://www.researchgate.net/publication/332936006_Big_Data_Analysis_Using_Apache_Spark_MLlib_and_Hadoop_HDFS_with_Scala_and_Java/link/5ce5c3c7299bf14d95b1cfc0/download
2. Marr B. Big Data in Practice
https://books.google.pl/books?id=UIU7jgEACAAJ&printsec=frontcover&hl=en&source=gbs_ge_summary_r&cad=0#v=onepage&q&f=false