

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

### Аналіз даних

рівень вищої освіти \_\_\_\_\_перший (бакалаврський) рівень\_\_\_\_\_

галузь знань 11 – Математика та статистика

спеціальність 113 Прикладна математика

освітня програма Прикладна математика

спеціалізація \_\_\_\_\_

вид дисципліни \_\_\_\_\_обов'язкова\_\_\_\_\_

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: *Стспанова Катерина Вадимівна*, канд. фіз.-мат. наук, доцент, доцент 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 «Аналіз даних» складена відповідно до освітньо-професійної програми підготовки бакалавра спеціальності 113 Прикладна математика

### 1. Опис навчальної дисципліни

1.1. Метою викладання навчальної дисципліни є надання майбутнім спеціалістам теоретичних знань та практичних навичок з методикою аналізу даних і основними принципами для розв'язання різних прикладних задач, що виникають при програмуванні, розробці та використанні сучасних інформаційних технологій.

#### 1.2. Основні завдання вивчення дисципліни

- ознайомити студентів із основними поняттями аналізу даних;
- ознайомити з основними типами структур даних та їх використанням;
- викласти основи аналізу даних в задачах моделювання й аналізу даних;
- вивчити використання методів аналізу даних при розв'язанні прикладних задач.

##### 1.2.1. Формування наступних інтегральної та загальних компетентностей:

ІК01. Здатність розв'язувати складні спеціалізовані задачі та практичні проблеми прикладної математики у професійній діяльності або у процесі навчання, що передбачає застосування математичних теорій та методів і характеризується комплексністю та невизначеністю умов.

ЗК02. Здатність застосовувати знання у практичних ситуаціях.

ЗК03. Здатність генерувати нові ідеї.

ЗК08. Знання та розуміння предметної області та розуміння професійної діяльності.

ЗК10. Навички у використанні інформаційних і комунікаційних технологій.

##### 1.2.2. Формування наступних фахових компетентностей:

ФК01. Здатність використовувати й адаптувати математичні теорії, методи та прийоми для доведення математичних тверджень і теорем.

ФК07. Здатність експлуатувати та обслуговувати програмне забезпечення автоматизованих та інформаційних систем різного призначення.

ФК09. Здатність до проведення математичного і комп'ютерного моделювання, аналізу та обробки даних, обчислювального експерименту, розв'язання формалізованих задач за допомогою спеціалізованих програмних засобів.

ФК13. Здатність зрозуміти постановку завдання, сформульовану мовою певної предметної галузі, здійснювати пошук та збір необхідних вихідних даних.

ФК14. Здатність сформулювати математичну постановку задачі, спираючись на постановку мовою предметної галузі, та обирати метод її розв'язання, що забезпечує потрібні точність і надійність результату.

#### 1.3. Кількість кредитів - 4

## 1.4. Загальна кількість годин - 120

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

## 1.6. Заплановані результати навчання. Студент повинен:

- Знати:* основні принципи та методи аналізу даних;  
основи синтаксису і особливості роботи в DataLore
- Вміти:* проводити аналіз даних;  
застосовувати вивчені методи аналізу даних;  
провести аналіз складності алгоритму, довести його коректність;  
застосовувати сучасні програми, які містять алгоритми для аналізу даних.
- Володіти:* методами аналізу даних.

Згідно з вимогами освітньо-професійної програми студенти повинні досягти таких програмних результатів навчання:

РН01. Демонструвати знання й розуміння основних концепцій, принципів, теорій прикладної математики і використовувати їх на практиці.

РН07. Вміти проводити практичні дослідження та знаходити розв'язок некоректних задач.

РН08. Поєднувати методи математичного та комп'ютерного моделювання з неформальними процедурами експертного аналізу для пошуку оптимальних рішень.

РН12. Розв'язувати окремі інженерні задачі та/або задачі, що виникають принаймні в одній предметній галузі: в соціології, економіці, екології та медицині.

РН13. Використовувати в практичній роботі спеціалізовані програмні продукти та програмні системи комп'ютерної математики.

РН19. Збирати та інтерпретувати відповідні дані й аналізувати складності в межах своєї спеціалізації для донесення суджень, які відбивають відповідні соціальні та етичні проблеми.

## 2. Тематичний план навчальної дисципліни

### Тема 1. Вступ до аналізу даних.

Загальна інформація. Різні типи даних. Ресурси пошуку даних. Сучасні методи аналізу даних. Особливості синтаксису при роботі з даними, підключення модулів та бібліотек в DataLore. Приклади програм.

### Тема 2. Візуальний аналіз даних.

Інтернет технологій. Виконання візуального аналізу даних – VisualMining. Характеристики засобів візуалізації даних. Методи візуалізації.

### Тема 3. Інтелектуальний аналіз даних.

Добування даних – DataMining та його практичне застосування. Основні методи DataMining. Проблеми аналізу інформації із Web. Передбачувані моделі. Пошук web-документів за допомогою гіперпосилань. Дослідження web-ресурсів.

### Тема 4. Аналіз текстової інформації та даних в реальному часі.

Задача аналізу текстів. Класифікація текстових документів. Задачі анотації текстів. Засоби аналізу текстової інформації. Пошук інформації в реальному часі. Рекомендаційні системи. Інструменти DataMining в реальному часі.

### Тема 5. Елементи дисперсійного аналізу.

Однофакторний дисперсійний аналіз. Двофакторний дисперсійний аналіз. Особливості і приклади використання.

## 3. Структура навчальної дисципліни

Назви розділів і тем	Кількість годин					
	денна форма					
	усього	у тому числі				
л		п	лаб.	інд.	с. р.	
<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>	<b>6</b>	<b>7</b>
Тема 1. Вступ до аналізу даних	14	4	4			6
Тема 2. Візуальний аналіз даних	23	6	4			13
<i>Виконання індивідуального завдання</i>	2		2			
Тема 3. Інтелектуальний аналіз даних	40	12	10			18
<i>Виконання індивідуального завдання</i>	2		2			
Тема 4. Аналіз текстової інформації та даних в реальному часі	23	6	4			13
<i>Виконання індивідуального завдання</i>	2		2			
Тема 5. Елементи дисперсійного аналізу	14	4	4			6

<i>Усього годин</i>	120	32	32			56
---------------------	-----	----	----	--	--	----

#### 4. Теми практичних занять

№ з/п	Назва теми	Кількість годин
1	Ресурси пошуку даних. Сучасні методи аналізу даних. Особливості синтаксису при роботі з даними, підключення модулів та бібліотек в DataLore. Приклади програм.	4
2	<i>Персональний захист студентами виконаного індивідуального завдання</i>	2
3	Характеристики засобів візуалізації даних. Методи візуалізації VisualMining.	6
4	<i>Персональний захист студентами виконаного індивідуального завдання</i>	2
5	Добування даних – DataMining та його практичне застосування. Основні методи DataMining. Проблеми аналізу інформації із Web. Передбачувані моделі. Пошук web-документів за допомогою гіперпосилань. Дослідження web-ресурсів.	6
6	Задача аналізу текстів. Класифікація текстових документів. Задачі анотації текстів. Засоби аналізу текстової інформації. Пошук інформації в реальному часі. Рекомендаційні системи. Інструменти DataMining в реальному часі.	6
7	<i>Персональний захист студентами виконаного індивідуального завдання</i>	2
8	Елементи дисперсійного аналізу	4
	<b>Разом</b>	<b>32</b>

#### 5. Завдання для самостійної роботи

№ з/п	Види, зміст самостійної роботи	Кількість годин
1	Виконання домашніх завдань з ознайомленням особливостей синтаксису, дослідження web-ресурсів.	6
2	Виконання домашніх завдань і написання найпростіших програм, проведення візуального аналізу даних досліджуваної моделі	13
3	Виконання домашніх завдань з опануванням практичного аспекту застосування DataMining та його основних методів. Усвідомлення проблеми аналізу інформації із Web. Побудова та розгляд прикладів передбачуваних моделей.	18
4	Виконання домашніх завдань по аналізу текстової інформації та даних в реальному часі	13

5	Виконання домашніх завдань з дисперсійного аналізу	6
	<b>Разом</b>	<b>56</b>

### 6. Індивідуальні завдання

Для студентів запланована реалізація та захист трьох індивідуальних завдань за темами: «Попередня обробка реальних даних для подальшого аналізу»; «Різноманітні методи візуалізації реальних даних» та «Презентація побудованих та візуалізованих даних шляхом застосовувати сучасних програм, які містять алгоритми для аналізу даних».

### 7. Методи навчання

Частково-пошуковий, дослідницький, пояснювально-ілюстративний, репродуктивний. Лекційні та практичні заняття.

### 8. Методи контролю

- перевірка виконання домашніх завдань
- поточне опитування за лекційним матеріалом
- перевірка індивідуальних завдань
- перевірка звітів-результатів самостійної роботи студентів
- залік

### 9. Схема нарахування балів

Поточний контроль, самостійна робота, індивідуальні завдання							Залік	Сума
T1	T2	T3	T4	T5	Індивідуальні завдання	Разом		
6	6	6	6	6	3 x 10	60	40	100

T1, ..., T5 – теми.

Для допуску до складання підсумкового контролю (заліку) здобувач вищої освіти повинен набрати не менше 10 балів з навчальної дисципліни протягом семестру.

### Критерії оцінювання навчальних досягнень

#### **Поточний контроль:**

бали нараховуються протягом курсу за виконання домашніх завдань, трьох індивідуальних завдань і активність студентів під час лекцій (через їх поточне опитування) та практичних занять (де в тому числі схвалюється, враховується та винагороджується балами їх самостійна робота й ініціативи).

За реалізоване індивідуальне завдання бали нараховуються таким чином:

- максимальний бал у разі правильної обґрунтованої відповіді без помилок;
- за незначні помилки оцінка зменшується від 10 до 30 відсотків;
- за значні логічні помилки оцінка зменшується до 50 відсотків, якщо хід розв'язання/міркування в цілому правильний;
- у разі частково правильних міркувань за відсутності обґрунтованої відповіді

виставляється до 30 відсотків від максимальної кількості балів;

- якщо відповідь не відповідає жодному з критеріїв, які сформульовані вище, – виставляється 0 балів.

### **Залік:**

Залікова робота складається із двох частин/блоків, що відповідає двом питанням/завданням. Максимальна оцінка за кожне – 20 балів, разом за обидва – 40 балів.

Для отримання максимальної оцінки (20 балів) за одне питання потрібно вміти сформулювати відповідні означення, знати основні поняттями аналізу даних, вміти проводити аналіз даних та/або алгоритми і обов'язково навести приклад(и). Замість прикладів можуть бути наведені міркування щодо змісту відповідних понять і результатів, а також реалізована візуалізація (тобто надані інтерактивні графіки в DataStudio або plotlineorpiechart; plotbarchart; plotbox-plot в DataLore) в результаті проведеного аналізу даних. Студент має продемонструвати вміння виявляти, коментувати та інтерпретувати закономірності в даних; при цьому викладення основних концепцій застосування сучасних програм, які містять алгоритми для аналізу даних та знання синтаксису і особливостей роботи в DataLore, що є обов'язковими.

Формулювання неповне або зі значними помилками: 5-15 балів.

Формулювання без прикладів або міркувань: 10-15 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. Барковський В.В. Теорія ймовірності та математична статистика. Навчальний посібник. – К: Центр учбової літератури, 2017. – 424 с.
2. Ганна Кармелюк. Теорія ймовірності та математична статистика. Посібник з розв'язування задач: Навчальний посібник рекомендовано МОН України. – К: Центр учбової літератури, 2017. – 576 с.
3. Галина Квасниця, Микола Притула, Ольга Прядко. Теорія ймовірностей та математична статистика, Частина 1 (Теорія ймовірностей). – Видавництво ЛНУ ім. Івана Франка, 2019. – 150 с.
4. Жлуктенко В. І., Наконечний С. І., Савіна С. С. Теорія ймовірностей і математична статистика: Навч.-метод. посібник: У 2-х ч. – Ч. II. Математична статистика. – К.: КНЕУ, 2001. – 336 с.
5. К. В. Степанова, Е. Ю. Железнякова, І. Л. Лебедева, Л. О. Норік. Лабораторний практикум з навчальної дисципліни «Теорія ймовірностей та математична статистика»: навчальний посібник. – Харків: ХНЕУ ім. С.Кузнеця, 2016.– 184 с. – Режим доступу: <http://www.repository.hneu.edu.ua/jspui/handle/123456789/14255>



6. Литвин В.В., Нікольський Ю.В., Пасічник В.В. Аналіз даних та знань. Навчальний посібник. – Магнолія 2006, Серія: Комп'ютинг, 2021. – 276 с.
7. Черняк О.І. Інтелектуальний аналіз даних: підручник. – К: Знання, 2014. – 599с.
8. Данильченко О.М., Данильченко А.О. Інтелектуальний аналіз даних: Навч. посібник. – Житомир: ЖДТУ, 2009. – 405 с.
9. Бахрушин В.Є. Методи аналізу даних: навчальний посібник для студентів. – Запоріжжя: КПУ, 2011. – 268 с.
10. Ситник В. Ф., Краснюк М. Т. Інтелектуальний аналіз даних (дейтамайнінг): Навч. посібник. — К.: КНЕУ, 2007. — 376 с.

#### **Допоміжна література**

1. A. Bryman, Handbook of Data Analysis. Pub. by Sage Publications, 2009, 728 p.
2. Daniel T. Larose, Data Mining methods and models. Wiley, 2006, p.322
3. Daniel T. Larose, Discovering knowledge in data. Wiley, 2005, p.222
4. Alexander Paprotny, Michael Thess, Realtime data mining. Springer, 2013, p.313

#### **11. Посилання на інформаційні ресурси в Інтернеті**

1. Recommendation system video lectures Xavier Amatrian  
<http://technocalifornia.blogspot.com/2014/08/introduction-to-recommender-systems-4.html>
2. Data Science  
<https://opendatascience.com>
3. Knowledge is covering  
<http://www.kdnuggets.com>
4. Neo4j DataBase  
<https://neo4j.com>
5. Rapid Miner  
<https://rapidminer.com>