

**МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
ЗАХІДНОУКРАЇНСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ
ФАКУЛЬТЕТ КОМП'ЮТЕРНИХ ІНФОРМАЦІЙНИХ ТЕХНОЛОГІЙ**

ЗАТВЕРДЖУЮ
Проректор з наукової роботи

Микола ДИВАК
„30” вересня 2024 р.



РОБОЧА ПРОГРАМА

з дисципліни **«Математичне моделювання та обчислювальні методи»**
Рівень вищої освіти – третій (освітньо-науковий)
Галузь знань 12 «Інформаційні технології»
Спеціальність 125 «Кібербезпека та захист інформації»
Освітньо-наукова програма «Кібербезпека та захист інформації»

Кафедра комп'ютерних наук

Форма навчання	Курс	Семест р	Лекції	Практ	СРС	Разом	Екзамен, залік
Денна	1	1	30	15	105	150	Залік
Денна	1	1	20	10	120	150	Залік

**Тернопіль –
2024**

СТРУКТУРА РОБОЧОЇ ПРОГРАМИ НАВЧАЛЬНОЇ ДИСЦИПЛІНИ

«Математичне моделювання та обчислювальні методи»

Дисципліна «Математичне моделювання та обчислювальні методи»	Галузь знань, спеціальність, СВО	Характеристика навчальної дисципліни
Кількість кредитів – 5	галузь знань 12 Інформаційні технології	Статус дисципліни нормативна Мова навчання українська
Кількість залікових модулів – 1	спеціальність 125 «Кібербезпека та захист інформації»	Рік підготовки: <i>Денна – 1</i> <i>Семестр</i> <i>Денна – 1</i>
Кількість змістових модулів – 2	рівень вищої освіти – третій (освітньо-науковий)	Аудиторні години - <i>Денна форма:</i> <i>Лекції – 30 год</i> <i>Практичні – 15 год</i> <i>Заочна форма:</i> <i>Лекції – 20 год</i> <i>Практичні – 10 год</i>
Загальна кількість годин – 150		Самостійна робота: <i>Денна - 105 год</i> <i>Заочна - 120 год</i>
Тижневих годин – 10, з них аудиторних – 3		Вид підсумкового контролю – залік

2. Мета і завдання вивчення дисципліни «Математичне моделювання та обчислювальні методи»

2.1. Мета вивчення дисципліни

Основною метою викладання дисципліни «Математичне моделювання та обчислювальні методи» є формування у майбутніх фахівців теоретичних основ побудови методів попередньої обробки інформації, яка фіксується у тематичних базах даних, а також набуття навичок щодо використання згаданих методів для видобування корисних знань.

2.2. Завдання вивчення дисципліни

В результаті вивчення курсу «Математичне моделювання та обчислювальні методи» студенти повинні:

- знати основні поняття попереднього аналізу вибірок, метод ієрархічного кластерного аналізу, метод к-середніх, методи побудови дерев рішень, лінійних класифікаторів, випадкових лісів, відбору ознак, основи прогнозування часових рядів;
- вміти здійснювати попередній аналіз вибірок, реалізовувати кластеризацію вибірок, відбір ознак, побудову дерев класифікації, лінійних класифікаторів, випадкових лісів, будувати регресійні залежності прогнозувати часові ряди.

2.3. Найменування та опис компетентностей, формування котрих забезпечує вивчення дисципліни:

Здатність застосовувати сучасні інформаційні технології, бази даних та інші електронні ресурси, спеціалізоване програмне забезпечення у науковій та навчальній діяльності

2.4. Передумови для вивчення дисципліни

Математичний аналіз, теорія імовірностей та математична статистика.

2.5. Найменування та опис компетентностей, формування котрих забезпечує вивчення дисципліни.

СК03. Здатність отримувати нові наукові результати, які створюють нові знання та становлять оригінальний внесок у розвиток інженерії програмного забезпечення та дотичних до неї міждисциплінарних напрямів.

2.6. Результати навчання

РН06. Формулювати і перевіряти гіпотези; використовувати для обґрунтування висновків належні докази, зокрема, результати теоретичного аналізу, експериментальних досліджень і математичного та/або комп'ютерного моделювання, наявні літературні дані.

РН07. Розробляти та досліджувати концептуальні, математичні і комп'ютерні моделі процесів і систем для отримання нових знань та/або створення інноваційних продуктів у інженерії програмного забезпечення та дотичних міждисциплінарних напрямках.

3. ПРОГРАМА НАВЧАЛЬНОЇ ДИСЦИПЛІНИ «МАТЕМАТИЧНЕ МОДЕЛЮВАННЯ ТА ОБЧИСЛЮВАЛЬНІ МЕТОДИ»

Змістовий модуль 1 – Параметризовані моделі систем

Тема 1. Статистичні методи дослідження зв'язків між факторами

Моделі розподіли. Варіаційний ряд. Лінійні відхилення від середнього. Дисперсія. Середньо квадратичне відхилення. Таблиці частот розподілу. Квантілі, проценти́лі, квартилі.

Асиметрія та ексцес. Нормальний розподіл. Рівномірний розподіл. Біноміальний розподіл. Розподіл Пуассона. Перевірка статистичних гіпотез. Нульова гіпотеза. Альтернативна гіпотеза. Однобічна та двобічна гіпотези. Рівень статистичної значущості. Ступені свободи.

Тема 2. Дисперсійний аналіз

Критерій Фішера. Критерій Бартлета для послідовності вибірок. Параметричні та непараметричні критерії. Рангові критерії. Критерій Левена. Критерій Стюдента. Критерій Манна-Уїтні. Критерій Краскела-Валліса.

Тема 3. Перевірка гіпотез про закон розподілу

Критерій χ^2 квадрат. Критерій Колмогорова-Смирнова. Критерій Шапіро-Вілکا.

Тема 4. Багатофакторний дисперсійний аналіз.

Односторонній дисперсійний аналіз. Двосторонній дисперсійний аналіз. Двофакторний дисперсійний аналіз без взаємодії. Двофакторний дисперсійний аналіз із взаємодією. Багатофакторний ANOVA з повторними вимірами.

Тема 5. Аналіз лінійних моделей даних

Лінійна модель даних із розбиттям на підвибірки. Реалізація мультилінійної регресії за допомогою пакету Scikit-learn. Відбір пояснюючих змінних регресійної моделі. Фільтрація змінних за кореляціями.

Тема 6. Гребнева регресія

Загальна гребнева мультилінійна регресійна модель. Мультиколінеарність. Регуляризація співвідношення ідентифікації моделі. Вибір структури гребневої регресійної моделі.

Змістовий модуль 2 – Класифікатори

Тема 7. Реалізація лінійних моделей у статистичних пакетах

Середовище R. Операції в R. Побудова лінійних моделей в R. Порівняння точності лінійних моделей в R. Побудова багатовимірної лінійної моделі методом виключень. Вибір головних лінійно-незалежних компонент.

Тема 8. Задачі класифікації

Простий класифікатор. Логістична класифікація. Логістична класифікація на реальних даних. Лінійна дискримінантна класифікація. Квадратичний дискримінантний аналіз. Класифікація методом головних компонент.

Тема 9. Класифікація методом найближчих сусідів

Аналіз методу найближчих сусідів. Класифікація методом найближчих сусідів. Коефіцієнти кореляції для класифікації. Регресія методом найближчих сусідів.

Тема 10. Машина опорних векторів

Теорія методу машини опорних векторів. Площина максимального розділення. Нелінійна SVM класифікація. Підтримка візуалізації SVM моделі.

Література: 21, 22, 23, 24

4. Структура залікового кредиту дисципліни

"МАТЕМАТИЧНЕ МОДЕЛЮВАННЯ ТА ОБЧИСЛЮВАЛЬНІ МЕТОДИ"

денна форма навчання

	Кількість годин	
	Аудиторні години	Самостійна робота
Змістовий модуль 1. Аналіз часових рядів та інформаційне моделювання		

Тема 1. Статистичні методи дослідження зв'язків між факторами.	4	15
Тема 2 Дисперсійний аналіз	4	10
Тема 3. Перевірка гіпотез про закон розподілу	4	10
Тема 4. Багатофакторний дисперсійний аналіз	4	10
Тема 5. Аналіз лінійних моделей даних	4	10
Тема 6. Гребенева регресія	5	10
Змістовий модуль 2. Методи аналізу даних		
Тема 7. Реалізація лінійних моделей у статистичних пакетах	5	10
Тема 8. Задачі класифікації.	5	10
Тема 9. Класифікація методом найближчих сусідів	5	10
Тема 10. Машина опорних векторів	5	10
Разом	45	105

заочна форма навчання

	Кількість годин	
	Аудиторні години	Самостійна робота
Змістовий модуль 1. Аналіз часових рядів та інформаційне моделювання		
Тема 1. Статистичні методи дослідження зв'язків між факторами.	3	12
Тема 2 Дисперсійний аналіз	3	12
Тема 3. Перевірка гіпотез про закон розподілу	3	12
Тема 4. Багатофакторний дисперсійний аналіз	3	12
Тема 5. Аналіз лінійних моделей даних	3	12
Тема 6. Гребенева регресія	3	12
Змістовий модуль 2. Методи аналізу даних		
Тема 7. Реалізація лінійних моделей у статистичних пакетах	3	12
Тема 8. Задачі класифікації.	3	12
Тема 9. Класифікація методом найближчих сусідів	3	12
Тема 10. Машина опорних векторів	3	12
Разом	30	120

5. Самостійна робота

Самостійна робота "Розробка регресійних та класифікаційних моделей бізнес процесів"

Мета роботи: Розробити моделі для формування рекомендацій на основі статистичних масивів

Завдання:

1. Кожен студент отримує реальні статистичні дані.
2. На основі специфіки обраного статистичного матеріалу, студент проводить дослідження, яке включає:
 - Аналіз природи та ролі атрибутів статистичного масиву
 - Вибір регресійних та класифікаційних моделей
 - Навчання та тестування моделей
 - Формування висновків та рекомендацій для користувачів
3. Студент повинен підготувати звіт, який містить:
 - Опис особливостей отриманих статистичних масивів
 - Детальний опис структури регресійних та класифікаційних моделей
 - Опис результатів навчання моделей
 - Опис результатів тестування моделей
 - Опис висновків і рекомендацій отриманих в результаті моделювання

Роботу необхідно здати викладачу у вигляді письмового звіту та презентації розробленого додатку. Захист роботи відбувається у формі усної презентації з демонстрацією функціоналу додатку.

6. Засоби оцінювання та методи демонстрування результатів навчання

У процесі вивчення дисципліни «Математичне моделювання та обчислювальні методи» використовуються такі засоби оцінювання та методи демонстрування результатів навчання:

- поточне опитування;
- підсумкове тестування по кожному змістовому модулю;
- розрахункові та розрахунково-графічні роботи;
- залік

7. Критерії, форми поточного та підсумкового контролю

Підсумковий бал (за 100-бальною шкалою) з дисципліни «Математичне моделювання та обчислювальні методи» визначається за шкалою оцінювання:

За шкалою ТНЕУ	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)

8. Інструменти, обладнання та програмне забезпечення, використання яких передбачає навчальна дисципліна

№	Найменування	Номер теми
1.	Пакет R, програмна оболонка R Studio, прикладні бібліотеки Python	1-10

РЕКОМЕНДОВАНІ ДЖЕРЕЛА ІНФОРМАЦІЇ

1. О. Дзендзелюк, Л. Костів, В. Рабик. Побудова Arima моделей часових рядів для прогнозування метеоданих на мові програмування R. Електроніка та інформаційні технології. 2013. Випуск 3. С. 211–219.
2. Daniel Lawson and Glenn Marion. An Introduction to Mathematical Modelling. 2008. Режим доступу: https://people.maths.bris.ac.uk/~madjl/course_text.pdf
3. Roderick Melnik. Mathematical and Computational Modeling: With Applications in Natural and Social Sciences, Engineering, and the Arts. Режим доступу: <https://www.wiley.com/en-ie/Mathematical+and+Computational+Modeling%3A+With+Applications+in+Natural+and+Social+Sciences%2C+Engineering%2C+and+the+Arts-p-9781118853986>
4. John Verzani. Using R for Introductory Statistics. Access mode: <https://cran.r-project.org/doc/contrib/Verzani-SimpleR.pdf>
5. Гудар, В., Рудинський, В. (2018). "Статистичний аналіз даних: методи та застосування." Наукова думка.
6. Беренс, Д., Лемер, С. (2019). "Основи статистичного програмування з R." CRC Press.
7. Decision Trees: Theory and Algorithms. <https://www.odbsms.org/wp-content/uploads/2014/07/DecisionTrees.pdf>
8. Emir Demirović. MurTree: Optimal Decision Trees via Dynamic Programming and Search. <https://laas.hal.science/hal-03597741/file/20-520.pdf>
9. Himchan Jeong. Bin ZouLinear. Classifier Models for Binary Classification. https://www.researchgate.net/publication/369033845_Linear_Classifier_Models_for_Binary_Classification
10. Sergey Subbotin. A Random Forest Model Building Using A priori Information for Diagnosis. Access mode: <http://ceur-ws.org/Vol-2353/paper76.pdf>
11. C'edric Beaulac Jeffrey S. Rosenthal. Predicting University Students' Academic Success and Major using Random Forests. . January 15, 2019. <https://arxiv.org/pdf/1802.03418.pdf>
12. Eric Ariel L. Salas ,Sakthi Kumaran Subburayalu. Modified shape index for object-based random forest image classification of agricultural systems using airborne hyperspectral datasets. Published: March 7, 2019. <https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0213356>
13. Diran Basmadjian, Ramin Farnood. The Art of Modeling in Science and Engineering with Mathematica. Режим доступу: <https://www.routledge.com/The-Art-of-Modeling-in-Science-and-Engineering-with-Mathematica/Basmadjian-Farnood/p/book/9780367390464>
14. Gerhard Dangelmayr and Michael Kirby. Mathematical Modeling a Comprehensive Introduction. Режим доступу: <https://www.math.colostate.edu/~gerhard/331book.pdf>
15. Decision Trees: Theory and Algorithms. <https://www.odbsms.org/wp-content/uploads/2014/07/DecisionTrees.pdf>
16. Andy J Ma, Pong C Yuen. Linear Dependency Modeling for Classifier Fusion and Feature Combination. https://www.isee-ai.cn/~majh/publications/LDM_draft.pdf
17. Nataliya Boyko, Roman Omeliukh and Nataliia Duliaba. The Random Forest Algorithm as an Element of Statistical Learning for Disease Prediction. <https://ceur-ws.org/Vol-3422/Paper6.pdf>
18. Gilles Louppe. Understanding random forests. <https://arxiv.org/pdf/1407.7502.pdf>

19. Edmond Zhang. Michael Mayo. Improving Bag-of-Words model with spatial information. https://www.researchgate.net/publication/254051049_Improving_Bag-of-Words_model_with_spatial_information
20. Luc Schoot Uiterkamp. Improving text representations for NLP from bags of words to strings of words. https://essay.utwente.nl/79245/1/Schoot%20Uiterkamp_MA_BMS.pdf
21. C.L. Wilson. Mathematical modeling, clustering algorithms and applications. https://www.researchgate.net/publication/293216439_Mathematical_modeling_clustering_algorithms_and_applications
22. Peter Trebuña. Jana Halčinová. Mathematical Tools of Cluster Analysis. https://www.researchgate.net/publication/276492929_Mathematical_Tools_of_Cluster_Analysis
23. Joaquín Pérez Ortega. The K-Means Algorithm Evolution. https://www.researchgate.net/publication/332340076_The_K-Means_Algorithm_Evolution
24. Dimitris Dineff. Clustering using k-means algorithm in multivariate dependent models with factor structure. <https://uu.diva-portal.org/smash/get/diva2>