

Чернівецький національний університет імені Юрія Федьковича

(повне найменування закладу вищої освіти)

факультет математики та інформатики

(назва факультету/навчально-наукового інституту)

Кафедра

математичного моделювання

(назва кафедри)

“ЗАТВЕРДЖУЮ”

Декан

Мартинюк О.В.

“ 25 ” 06 2024 року



РОБОЧА ПРОГРАМА

навчальної дисципліни

Інтелектуальний аналіз даних та моделювання кризових явищ

(вказіть назву навчальної дисципліни (іноземною, якщо дисципліна викладається іноземною мовою))

обов'язкова

(вказати: обов'язкова чи вибіркова)

Освітньо-професійна програма Системний аналіз

(назва програми)

Спеціальність 124 Системний аналіз

(вказати: код, назва)

Галузь знань 12 Інформаційні технології

(вказати: шифр, назва)

Рівень вищої освіти другий (магістерський)

(вказати: перший бакалаврський/другий магістерський)

факультет математики та інформатики

(назва факультету/навчально-наукового інституту, на якому здійснюється підготовка фахівців за вказаною освітньо-професійною програмою)

Мова навчання українська

(вказати: якою мовою викладається дисципліна)

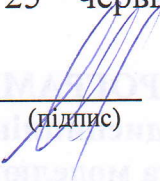
Чернівці 2024 рік

Робоча програма навчальної дисципліни «Інтелектуальний аналіз даних та моделювання кризових явищ» складена відповідно до освітньо-професійної програми «Системний аналіз», затвердженої Вченою радою Чернівецького національного університету імені Юрія Федьковича « 29 » травня 2023 року, протокол № 5.

Розробники: Мельник Василь Сергійович, асистент, канд. фіз.-мат. наук
Виклюк Ярослав Ігорович, професор, доктор техн. наук

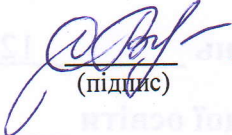
Погоджено з гарантом ОП і затверджено на засіданні кафедри математичного моделювання

Протокол № 18 від “ 25 ” червня 2024 року

Завідувач кафедри  Черевко І.М.
(підпис) (прізвище та ініціали)

Схвалено методичною радою факультету математики та інформатики

Протокол № 11 від “25” червня 2024 року

Голова методичної ради факультету математики та інформатики  Сікора В.С.
(підпис) (прізвище та ініціали)

1. Мета навчальної дисципліни: Мета дисципліни – надати студентам глибокі знання та навички з інтелектуального аналізу даних та генеративних моделей штучного інтелекту.

Основними цілями цієї дисципліни є:

- Засвоєння методів і технологій інтелектуального аналізу даних: Вивчення таких методів, як кластеризація, класифікація, асоціативні правила, дерева рішень, нейронні мережі, методи машинного навчання та їх застосування для виявлення прихованих закономірностей у даних.
- Розвиток навичок збору та обробки даних: Опанування сучасних інструментів для роботи з великими масивами даних (Big Data), розуміння важливості попередньої підготовки, очищення та валідації даних.
- Моделювання кризових явищ: Вивчення способів моделювання й прогнозування кризових подій, таких як економічні кризи, соціальні потрясіння чи природні катаклізми, з використанням інтелектуальних систем.
- Оцінка ризиків та їх мінімізація: Формування вмінь ідентифікувати можливі ризики в реальних сценаріях, моделювати й аналізувати їхній вплив, а також розробляти рекомендації щодо мінімізації наслідків.
- Застосування знань для реальних кейсів: Розвиток вміння застосовувати методи аналізу даних для вирішення практичних завдань, пов'язаних із кризовими ситуаціями у різних галузях (економіка, політика, екологія).

Загалом, дисципліна допомагає студентам оволодіти аналітичним інструментарієм для ефективного аналізу складних систем і сценаріїв, що включають елементи невизначеності, та прийняття рішень в умовах ризику.

Пререквізити. Аналіз даних. Системи штучного інтелекту. Алгебра і геометрія. Математичний аналіз. Дискретна математика.

Постреквізити. Методи Data Science. Технології паралельного програмування. Теоретичні основи прикладних досліджень. Системи та методи прийняття рішень в соціальних та економічних системах. Системи машинного навчання.

2. Результати навчання: охоплюють як теоретичні знання, так і практичні навички, які студенти мають опанувати до кінця курсу. Очікувані результати навчання можна поділити на кілька категорій:

1. Знання та розуміння:

- Методи та алгоритми інтелектуального аналізу даних: Студенти повинні розуміти принципи роботи основних алгоритмів машинного навчання, таких як регресія, класифікація, кластеризація, методи асоціацій та дерева рішень.
- Методи моделювання кризових явищ: Опанування підходів до моделювання складних систем, таких як економічні кризи, соціальні конфлікти чи природні катастрофи. Студенти повинні знати основні методи прогнозування та сценарного аналізу.
- Технології обробки великих даних: Розуміння сучасних технологій і платформ для збору, зберігання та аналізу великих обсягів даних (Big Data), таких як Hadoop, Spark, бази даних NoSQL тощо.
- Основи оцінки ризиків: Студенти повинні знати методи аналізу ризиків та їхнього впливу на різні системи й процеси, розуміти способи мінімізації ризиків.

2. Практичні навички:

- Збір і підготовка даних: Уміння працювати з різними типами даних, їх збирання, очищення та підготовка для аналізу. Студенти мають опанувати інструменти для роботи з неструктурованими даними.

- Аналіз даних: Здатність застосовувати алгоритми інтелектуального аналізу даних для вирішення реальних завдань, зокрема, кластеризацію, класифікацію, аналіз часових рядів і виявлення аномалій.
 - Побудова та оцінка моделей: Вміння створювати моделі для прогнозування кризових явищ, аналізувати результати та робити висновки на основі отриманих даних.
 - Інтерпретація результатів: Здатність правильно інтерпретувати результати аналізу даних, формулювати обґрунтовані висновки та пропозиції для прийняття управлінських рішень у кризових ситуаціях.
3. Комунікаційні вміння:
- Презентація результатів: Здатність чітко і доступно представляти результати аналізу даних для різної аудиторії (керівників, колег, замовників), у тому числі у формі звітів, презентацій та візуалізацій даних.
 - Робота в команді: Уміння працювати в команді над спільними проєктами, координувати дії з іншими аналітиками, економістами та фахівцями з інших галузей.
4. Критичне мислення та прийняття рішень:
- Оцінка альтернативних підходів: Здатність критично оцінювати різні методи аналізу та моделювання, порівнювати їх ефективність і вибирати найбільш відповідні для конкретної задачі.
 - Прийняття рішень в умовах невизначеності: Вміння приймати рішення на основі аналізу даних, враховуючи неповну або неточну інформацію, а також оцінювати можливі наслідки і ризики.
5. Цифрова грамотність та технічні навички:
- Використання програмних засобів: Опанування сучасних інструментів та платформ для аналізу даних (Python, NumPy, Pandas, Scikit-learn), візуалізації даних (Matplotlib) та роботи з великими даними.
 - Автоматизація процесів: Навички автоматизації збору та обробки даних за допомогою програмування та відповідних інструментів для підвищення ефективності процесів.
 - Таким чином, після завершення курсу студенти будуть готові застосовувати знання інтелектуального аналізу даних та моделювання в умовах реальних кризових явищ, що сприятиме їх професійному розвитку у сфері аналітики, управління ризиками та прогнозування.

Дисципліна формує такі компетенції у відповідності до стандарту вищої освіти спеціальності 124 – Системний аналіз та освітньої програми:

ЗК1. Здатність до абстрактного мислення, аналізу та синтезу.

ЗК2. Здатність спілкуватися іноземною мовою.

ЗК3. Здатність вчитися та оволодівати сучасними знаннями.

ЗК4. Здатність до пошуку, оброблення та аналізу інформації з різних джерел.

ЗК5. Здатність генерувати нові ідеї (креативність)

ЗК6. Здатність спілкуватися з представниками інших професійних груп різного рівня (з експертами з інших галузей знань/видів економічної діяльності).

ЗК7. Здатність виявляти та вирішувати проблеми на основі обґрунтованих рішень.

ЗК9. Здатність діяти на основі етичних міркувань.

СК1. Здатність інтегрувати знання та здійснювати системні дослідження, застосовувати методи математичного та інформаційного моделювання складних систем та процесів різної природи.

СК3. Здатність розробляти системи підтримки прийняття рішень та рекомендаційні системи.

СК4. Здатність оцінювати ризики, розробляти алгоритми управління ризиками в складних системах різної природи.

СК6. Здатність застосовувати теорію і методи Data Science для здійснення інтелектуального аналізу даних з метою виявлення нових властивостей та генерації нових знань про складні системи.

СК8. Здатність розробляти і реалізовувати наукові та прикладні проекти в галузі інформаційних технологій та дотичні до неї міждисциплінарні проекти.

СК10. Здатність до самоосвіти та професійного розвитку.

СК11. Здатність ефективно проводити системний аналіз, здійснювати вибір концептуальної моделі середовища інформаційної системи на основі математичних моделей і методів прийняття рішень, параметризацію компонентів інтелектуальної системи підтримки прийняття рішень

Наведені результати навчання за відповідною дисципліною співвідносяться із такими програмними результатами навчання:

РН2. Будувати та досліджувати моделі складних систем і процесів застосовуючи методи системного аналізу, математичного, комп'ютерного та інформаційного моделювання.

РН3. Застосовувати методи розкриття невизначеностей в задачах системного аналізу, розкривати ситуаційні невизначеності та невизначеності в задачах взаємодії, протидії та конфлікту стратегій, знаходити компроміс при розкритті концептуальної невизначеності.

РН5. Використовувати міри оцінювання ризиків та застосовувати їх при аналізі багатофакторних ризиків в складних системах.

РН6. Застосовувати методи машинного навчання та інтелектуального аналізу даних, математичний апарат нечіткої логіки, теорії ігор та розподіленого штучного інтелекту для розв'язання складних задач системного аналізу.

РН7. Розробляти інтелектуальні системи в умовах слабо структурованих даних різної природи.

3. Опис навчальної дисципліни

3.1 Загальна інформація

Форма навчання	Рік підготовки	Семестр	Кількість		Кількість годин						Вид підсумкового контролю
			кредитів	годин	лекції	практичні	семінарські	лабораторні	самостійна робота	індивідуальні завдання	
Денна	1	1	4	120	15	-	-	15	90	-	залік

3.2 Структура змісту навчальної дисципліни

Назви змістових модулів і тем навчальних занять	Кількість годин												
	усього	денна форма						заочна форма					
		у тому числі						у тому числі					
		лекц.	прак./семін.	лаб.	інд.	с.р.	усього	лекц.	прак./сем.	лаб	інд	с.р.	
1	2	3	4	5	6	7	8	9	10	11	12	13	
Теми навчальних занять	Змістовий модуль 1. Основи інтелектуального аналізу даних												
Тема 1. Основні поняття інтелектуального	15	2		2		10							

аналізу даних. Завдання аналізу даних. CRISP-DM та підготовка даних.												
Тема 2. Збір даних. Обробка даних в Pandas. Візуалізація даних в Matplotlib. Реляційні дані.	15	2		2		10						
Тема 3. Вектори, матриці та лінійна алгебра засобами NumPy. Робота з графами.	15	2		2		10						
Тема 4. Конструювання промптів та Обробка природної мови. Модель «Bag of words». Моделі з урахуванням семантики. Токенізація. Техніки векторизації. Генеративні моделі.	15	2		2		10						
Разом за ЗМ1	60	8		8		40						
Теми навчальних занять	Змістовий модуль 2. Стратегії інтелектуального аналізу даних та їх реалізації.											
Тема 1. Основи машинного навчання. Лінійна регресія. Лінійна класифікація.	15	1		1		10						
Тема 2. Нелінійне моделювання. Крос-валідація. Наївний баєсів класифікатор.	15	2		2		10						
Тема 3. Тестування гіпотез. Рекомендаційні системи. Колаборативна фільтрація.	15	2		2		10						
Тема 4. Методи моделювання кризових явищ. Використання нейронних мереж для прогнозування кризових явищ. НМ в задачах апроксимації та прогнозування. НМ в задачах класифікації та кластеризації, карти Кохонена. Глибоке навчання.	15	2		2		20						
Разом за ЗМ 2	60	7		7		45						
Усього годин	120	15		15		90						

3.3 Тематика лабораторних занять

№	Назва теми (завдання)	Кількість годин
1	Обробка даних з використанням NumPy та Pandas. Веб-скрапінг та візуалізація даних.	4
2	Парсинг даних з використанням NLP. Конструювання промптів	4
3	Задачі кластеризації та класифікації.	4
4	Використання глибокого навчання.	3

3.4. Зміст завдань для самостійної роботи

Самостійна робота складає 90 годин. Розподіл самостійної роботи за видами навчальних робіт:

- підготовка до лекційних занять – 30 годин;
- підготовка до лабораторних занять – 40 годин;
- підготовка до підсумкового модуль-контролю – 20 годин.

Самостійна робота складається з повторення матеріалу, засвоєного на лекціях, самостійного опанування частини теоретичного матеріалу, роботи з контрольними запитаннями та завданнями.

4. Освітні технології, методи навчання і викладання навчальної дисципліни

Методи навчання та викладання: лекції, лабораторні заняття, електронне навчання з використанням системи Moodle, тестування, виконання завдань ІНДЗ.

5. Критерії та засоби оцінювання результатів навчання з навчальної дисципліни

5.1. Критерієм підсумкового оцінювання Система оцінювання рівня навчальних досягнень ґрунтується на принципах ECTS та є накопичувальною. Протягом семестру студенти виконують 4 лабораторних робіт. Кожна лабораторна робота оцінюється в 15 балів. Виконуючи завдання лабораторної роботи, студент повинен оформити і завантажити для подальшої перевірки на сайт електронного навчання звіт разом із працездатними файлами програмної реалізації завдань ЛР (правила оформлення наведені на сторінці навчальної дисципліни на сайті). 50% балів, відведених на оцінювання ЛР, студент отримує за працюючий програмний продукт, в якому реалізовано всі завдання ЛР та оформлений звіт. Решта 50% балів виставляється після захисту студентом виконаного звіту. На захисті звіту з ЛР студент має відповісти на питання щодо постановки задачі та розробленого ним алгоритму реалізації кожного із завдань ЛР. При відповіді на теоретичні питання та питання щодо програмної реалізації алгоритму у випадку неістотної помилки знімається 10-20% балів, а у випадку істотної 20-40% балів, якщо ж студент не опанував теоретичний матеріал, плутається в означеннях, наводить логічно невірні твердження, то знімається до 50% балів від усієї суми балів за ЛР. Максимальна кількість, яку можна набрати на підсумковому модулі (тестування) – 40 балів. Підсумкова оцінка виставляється за результатами суми балів, набраних на змістових модулях під час семестру та підсумковому модулі згідно з нижченаведеною таблицею.

5.2. Шкала оцінювання: національна та ECTS

Оцінка за національною шкалою	Оцінка за шкалою ECTS	
	Оцінка (бали)	Пояснення за розширеною шкалою
Зараховано	A (90-100)	відмінно
	B (80-89)	дуже добре

	C (70-79)	добре
	D (60-69)	задовільно
	E (50-59)	достатньо
Не зараховано	FX (35-49)	(незадовільно)з можливістю повторного складання
	F (1-34)	(незадовільно) з обов'язковим повторним курсом

5.3. Засоби оцінювання

Засобами оцінювання та демонстрування результатів навчання можуть бути:

- контрольні роботи; стандартизовані тести; проекти (наскрізні проекти; індивідуальні та командні проекти; дослідницько-творчі та ін.); аналітичні звіти; реферати; есе; розрахункові, графічні, розрахунково-графічні роботи; презентації результатів виконаних завдань та досліджень; студентські презентації та виступи на наукових заходах; контрольні роботи; завдання налабораторному обладнанні, тренажерах, реальних об'єктах тощо; інші види індивідуальних та групових завдань.

6. Форми поточного та підсумкового контролю

Формами поточного контролю є усна чи письмова (тестування, есе, реферат, творча робота, лабораторна робота) відповідь студента.

Формами підсумкового контролю є залік.

7. Рекомендована література

1.1. Фахова (основна) – українських і зарубіжних видань.

1. Han J. Data Mining: Concepts and Techniques (Second Edition) / J. Han, M. Kamber – Morgan Kaufmann Publishers, 2006. – 800 p. 10
2. Witten, I. H. Data mining : practical machine learning tools and techniques. / Ian H. Witten, Frank Eibe, Mark A. Hall. – 3rd ed. – organ Kaufmann Publishers, 2011. – 630 p.
3. Ситник В.Ф., Краснюк М.Т. Інтелектуальний аналіз даних (дейтамайнінг): Навч. посібник. – К: КНЕУ, 2007. – 376 с.
4. McKinney. Python for Data Analysis: Data Wrangling with pandas, NumPy, and Jupyter. O'Reilly Media. September 20, 2022. - 621 p
5. Witten I. Data Mining: Practical Machine Learning Tools and Techniques / Ian Witten, Eibe Frank, Mark A. Hall. – Morgan Kaufmann, 2016. – 654p.
6. Inmon W.H. Building the Data Warehouse, 4th Edition. – Hoboken, NJ:Wiley, 2005. – 576 p.

1.2. Допоміжна.

1. Ranga Suri N.N.R., Murty N., Athithan M.G. Outlier Detection: Techniques and Applications. A Data Mining Perspective / N.N.R. Ranga Suri, N. Murty, M.G. Athithan. – Cham: Springer International Publishing, 2019, 214 p.

8. Інформаційні ресурси

- <https://moodle.chnu.edu.ua/course/view.php?id=7756>
<https://training.epam.ua/ua/blog/592>
<http://www.the-data-mine.com>
<https://www.coursera.org/learn/data-analysis-with-python>

Політика щодо академічної доброчесності

Усі учасники освітнього процесу повинні дотримуватись вимог чинного законодавства України, Статуту і Правил внутрішнього розпорядку ЧНУ, загальноприйнятих моральних принципів, правил поведінки та корпоративної культури; підтримувати атмосферу доброзичливості, відповідальності, порядності й толерантності.

Під час навчання та викладання учасники освітнього процесу зобов'язані дотримуватися академічної доброчесності.

Дотримання академічної доброчесності здобувачами передбачає: самостійне виконання навчальних завдань, завдань поточного та підсумкового контролю. За необхідністю з метою з'ясування всіх незрозумілих під час самостійної та індивідуальної роботи питань студент може відвідувати консультації викладача. Виконаний студентом не свій варіант завдання не оцінюється.

Складання (перескладання) заліку проводиться за встановленим деканатом розкладом.