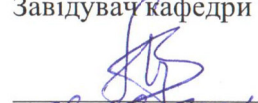


Міністерство освіти і науки України
Черкаський національний університет імені Богдана Хмельницького
ННІ Інформаційних та освітніх технологій

ЗАТВЕРДЖЕНО

Завідувач кафедри прикладної математики та інформатики


28.08.2024

О. В. Піскун

Робоча програма навчальної дисципліни

Аналіз даних

Освітній ступінь	Семестр за навч. планом	Спеціальність	Освітня програма	Обов'язковість дисципліни	Мова навчання	Погодження керівника ГЗ ОП*
бакалавр	5	126 Інформаційні системи та технології	Інтелектуальний аналіз даних	обов'язкова	українська	О. А. Сердюк

* групи забезпечення освітньої програми

Розробники робочої програми

О. В. Піскун	кафедра прикладної математики та інформатики	доцент	к. т. н.	
О. А. Сердюк	кафедра прикладної математики та інформатики	доцент	к.е.н.	

СТРУКТУРА НАВЧАЛЬНОЇ ДИСЦИПЛІНИ

Форма навчання	Семестр у межах дисципліни	Кількість кредитів	Загальна кількість годин	Аудиторна робота				Самостійна робота			Форма підсумкового контролю
денна	перший	6	180 год	60 год / 33 %				120 год / 67 %			іспит
				лекції	лабора- торні	прак- тичні	семінар- ські	розрахункові роботи	індивіду- альні завдання	підготовка до занять	
				30 год	30 год			60 год		60 год	

МЕТА НАВЧАЛЬНОЇ ДИСЦИПЛІНИ

Засвоєння студентами основних теоретичних відомостей та практичних вмінь з курсу. Підготувати студента до ефективного використання сучасних методів інтелектуального аналізу даних та обробки інформації задля створення автоматизованих систем у подальшій професійній діяльності; допомогти набуту навички реалізації програмних засобів для інтелектуального аналізу даних та розв'язання за допомогою них практичних задач.

ПЕРЕДУМОВИ ДЛЯ ВИВЧЕННЯ ДИСЦИПЛІНИ ТА УЗГОДЖЕННЯ З ІНШИМИ ДИСЦИПЛІНАМИ

Курс базується на поняттях, що вивчаються в дисциплінах:

- алгебра та геометрія;
- математичний аналіз;
- теорія ймовірностей та математична статистика;
- програмне забезпечення та інформаційно-комунікаційні технології
- алгоритми та структури даних;
- програмування;
- мови програмування.

ОЧІКУВАНІ РЕЗУЛЬТАТИ НАВЧАННЯ

Знання та практичні навички для використання методів і моделей інтелектуального аналізу даних при розв'язанні задач обробки великих масивів даних. Знання основних понять та визначень інтелектуального аналізу даних; моделей та методів побудови моделей та аналізу залежностей у великих масивах даних; сучасних програмних засобів для проектування і розробки систем інтелектуального аналізу даних; критеріїв порівняння моделей і методів інтелектуального аналізу даних.

Вміння обґрунтовувати й аналізувати вибір конкретного типу моделі та методу інтелектуального аналізу даних при розв'язанні відповідних практичних задач; використовувати сучасні програмні засоби для проектування та дослідження систем інтелектуального аналізу даних; розробляти програми для побудови та використання моделей інтелектуального аналізу даних; аналізувати результати побудови та використання систем інтелектуального аналізу даних при розв'язанні прикладних задач.

Результатом вивчення навчальної дисципліни «Аналіз даних» є формування наступного комплексу компетентностей.

Інтегральна компетентність: здатність розв'язувати складні спеціалізовані задачі та практичні проблеми в області інформаційних систем та технологій, або в процесі навчання, що характеризуються комплексністю та невизначеністю умов, які потребують застосування теорій та методів інформаційних технологій.

Загальні компетентності:

- ЗК 1. Здатність до абстрактного мислення, аналізу та синтезу;
- ЗК 2. Здатність застосовувати знання у практичних ситуаціях;
- ЗК 3. Здатність до розуміння предметної області та професійної діяльності;
- ЗК 4. Здатність спілкуватися іноземною мовою;
- ЗК 5. Здатність вчитися і оволодівати сучасними знаннями;
- ЗК 6. Здатність до пошуку, оброблення та узагальнення інформації з різних джерел.

Фахові компетентності:

- СК 6. Здатність використовувати сучасні інформаційні системи та технології (виробничі, підтримки прийняття рішень, інтелектуального аналізу даних та інші), методики й техніки кібербезпеки під час виконання функціональних завдань та обов'язків;
- СК 11. Здатність до аналізу, синтезу і оптимізації інформаційних систем та технологій з використанням математичних моделей і методів;
- СК 12. Здатність управляти та користуватися сучасними інформаційно-комунікаційними системами та технологіями (у тому числі такими, що базуються на використанні Інтернет);
- СК 13. Здатність проводити обчислювальні експерименти, порівнювати результати експериментальних даних і отриманих рішень;
- СК 16. Здатність до розробки і використання інтелектуальних інформаційних систем, технологій обробки та аналізу (великих) даних, алгоритмів та методів машинного навчання у різних сферах діяльності;
- СК 17. Здатність розробляти модулі обробки даних інформаційних систем, використовувати технічні, програмні засоби та мови програмування для розробки алгоритмів і програм в області інтелектуального аналізу даних, інтелектуальних та інформаційних систем.

ПРОГРАМНІ РЕЗУЛЬТАТИ НАВЧАННЯ

- ПРН 2. Застосовувати знання фундаментальних і природничих наук, системного аналізу та технологій моделювання, стандартних алгоритмів та дискретного аналізу при розв'язанні задач проектування і використання інформаційних систем та технологій.
- ПРН 3. Використовувати базові знання інформатики й сучасних інформаційних систем та технологій, навички програмування, технології безпечної роботи в комп'ютерних мережах, методи створення баз даних та інтернет-ресурсів, технології розроблення алгоритмів і комп'ютерних програм мовами високого рівня із застосуванням об'єктно-орієнтованого програмування для розв'язання задач проектування і використання інформаційних систем та технологій.
- ПРН 4. Проводити системний аналіз об'єктів проектування та обґрунтовувати вибір структури, алгоритмів та способів передачі інформації в інформаційних системах та технологіях.
- ПРН 6. Демонструвати знання сучасного рівня технологій інформаційних систем, практичні навички програмування та використання прикладних і спеціалізованих комп'ютерних систем та середовищ з метою їх запровадження у професійній діяльності.
- ПРН 12. Вміти розробляти програмне забезпечення різного рівня складності, що входить до складу інформаційних систем та технологій, і спрямоване, зокрема, на виконання задач інтелектуального пошуку, аналізу та обробки даних з метою видобутку знань і прогнозування поведінки аналізованого об'єкта при розв'язанні прикладних науково-виробничих задач і задач бізнесу.
- ПРН 13. Вміти будувати стратегії розв'язання практичних і наукових задач у галузі інтелектуального аналізу даних, машинного навчання, штучного інтелекту; обирати найбільш ефективний метод розв'язання та реалізовувати його з використанням сучасних мов програмування та спеціалізованого програмного забезпечення.

КРИТЕРІЇ ОЦІНЮВАННЯ РЕЗУЛЬТАТІВ НАВЧАННЯ

Критерієм успішного проходження здобувачем освіти підсумкового оцінювання є досягнення ним мінімальних порогових рівнів оцінок за кожним запланованим результатом навчання навчальної дисципліни, який визначається до кожного завдання через якісні критерії і трансформується в мінімальну позитивну оцінку обраної для даної дисципліни шкали. Після завершення курсу використана шкала перенормовується у накопичувальну 100-бальну системою з відповідністю у національній («відмінно», «добре», «задовільно», «незадовільно») та ЄКТС (A, B, C, D, E, FX, F) шкалах.

Контроль проводиться у чотири етапи: вхідний (у формі тестів для визначення рівня підготовки студентів з дисциплін, що забезпечують цей курс; за результатами цього контролю розробляються заходи індивідуальної допомоги студентам, коригування освітнього процесу тощо), поточний (при проведенні лекцій, лабораторних занять і має на меті перевірку рівня засвоєння студентом матеріалу курсу та підготовленості студента до виконання конкретних видів навчальної діяльності), проміжний (модульний; при закінченні вивчення логічно завершеної частини навчальної дисципліни) та семестровий контроль.

ЗАСОБИ ОЦІНЮВАННЯ

Екзамен; аналітичні звіти; лістинги програм; тести; розрахункові роботи.

ПРОГРАМА НАВЧАЛЬНОЇ ДИСЦИПЛІНИ

Аудиторна робота «Слухай, читай, обговорюй» (Л – лекція, Лб – лабораторне заняття, Пр – практичне заняття, См – семінарське заняття)

Самостійна робота «Думай, пиши, аналізуй, досліджуй, твори» (Рр – розрахункова робота, Інд – індивідуальне завдання, Пз – підготовка до занять)

Контроль «Захищай, відстоюй, неси відповідальність» (МК – модульний контроль; ПК – підсумковий контроль)

Тема	Тема навчального заняття	К-сть годин за денною/заочною формами навчання	Засіб оцінювання	Максимальна кількість балів за формами навчання
Вхідний контроль			тест	---
Змістовий модуль 1. Вступ до аналізу даних				24 / ---
1.1. Кластеризація та класифікація даних методами kMeans та kNN	Л1 Методи кластеризації kMeans та класифікації kNN. <i>Л-ра: 2</i>	2 / –	---	---
	Лб1 Реалізація методів kMeans та kNN. <i>Л-ра: 2</i>	2 / –	---	---
	Пз1 Реалізація методів kMeans та kNN. <i>Л-ра: 2</i>	4 / –	звіт з лістингами програм	1 / ---
1.2. Генетичні алгоритми	Л2 Генетичні алгоритми. Загальна структура. <i>Л-ра: 7, 10</i>	2 / –	---	---
	Лб2 Реалізація простого генетичного алгоритму. <i>Л-ра: 7, 10</i>	2 / –	---	---
	Пз2 Реалізація простого генетичного алгоритму. <i>Л-ра: 7, 10</i>	4 / –	звіт з лістингами програм	1 / ---
	Л3 Використання генетичних алгоритмів для розв'язування прикладних задач. <i>Л-ра: 7, 10</i>	2 / –	---	---
	Лб3 Пошук наближених розв'язків практичних задач за допомогою генетичних алгоритмів. <i>Л-ра: 7, 10</i>	2 / –	---	---

	Пз3 Пошук наближених розв'язків практичних задач за допомогою генетичних алгоритмів. <i>Л-ра: 7, 10</i>	4 / –	звіт з лістингами програм	1 / ---
	РР1 Розв'язання простої задачі генерації розкладу на основі генетичних алгоритмів. <i>Л-ра: 7</i>	15 / –	аналітичний звіт з лістингами програм	4 / ---
1.3. Логістична регресія	Л4 Програмування логістичної регресії за допомогою методу Ньютона-Рафсона. <i>Л-ра: 9</i>	2 / –	---	---
	Л64 Реалізація методу Ньютона-Рафсона для пошуку параметрів логістичної регресії. <i>Л-ра: 9</i>	2 / –	---	---
	Пз4 Реалізація методу Ньютона-Рафсона для пошуку параметрів логістичної регресії. <i>Л-ра: 9</i>	4 / –	звіт з лістингами програм	1 / ---
	Л5 Навчання методом градієнтного спуску. <i>Л-ра: 9</i>	2 / –	---	---
	Л65 Навчання систем класифікації логістичної регресії за допомогою методу градієнтного спуску. <i>Л-ра: 9</i>	2 / –	---	---
	Пз5 Навчання систем класифікації логістичної регресії за допомогою методу градієнтного спуску. <i>Л-ра: 9</i>	4 / –	звіт з лістингами програм	2 / ---
	Л6 Класифікація на основі багатокласової логістичної регресії. <i>Л-ра: 9</i>	2 / –	---	---
	Л66 Реалізація класифікації на основі багатокласової логістичної регресії. <i>Л-ра: 9</i>	2 / –	---	---
	Пз6 Реалізація класифікації на основі багатокласової логістичної регресії. <i>Л-ра: 9</i>	4 / –	звіт з лістингами програм	2 / ---
1.4. Байєсівська класифікація	Л7 Кластеризація даних за допомогою виведення спрощеного алгоритму Байєса. <i>Л-ра: 2, 8</i>	2 / –	---	---
	Л67 Реалізація кластеризації даних за допомогою виведення спрощеного алгоритму Байєса. <i>Л-ра: 2, 8</i>	2 / –	---	---
	Пз7 Реалізація кластеризації даних за допомогою виведення спрощеного алгоритму Байєса. <i>Л-ра: 2, 8</i>	4 / –	звіт з лістингами програм	1 / ---

	Л8 Наївна Байєсівська класифікація. <i>Л-ра: 2, 8</i>	2 / –	---	---
	Л68 Реалізація програми класифікації зразків за допомогою наївного алгоритма Байєса. <i>Л-ра: 2, 8</i>	2 / –	---	---
	Пз8 Реалізація програми класифікації зразків за допомогою наївного алгоритма Байєса. <i>Л-ра: 2, 8</i>	4 / –	звіт з лістингами програм	2 / ---
	РР2 Класифікація зразків за допомогою багатокласової логістичної регресії та наївного Байєса. <i>Л-ра: 2, 8</i>	15 / –	аналітичний звіт з лістингами програм	4 / ---
Модульний контроль	МК1 Тест за матеріалами змістового модуля 1		тест	5 / ---
Змістовий модуль 2. Штучні нейронні мережі				26 / ---
2.1. Компоненти штучної нейронної мережі	Л9 Основні компоненти штучної нейронної мережі. Навчання нейронної мережі. <i>Л-ра: 1, 3, 4, 6</i>	2 / –	---	---
	Л69 Побудова простої нейронної мережі. <i>Л-ра: 1, 3, 4, 6</i>	2 / –	---	---
	Пз9 Побудова простої нейронної мережі. <i>Л-ра: 1, 3, 4, 6</i>	4 / –	звіт з лістингами програм	1 / ---
2.2. Класифікація зразків	Л10 Штучні нейронні мережі зворотнього розповсюдження помилки. <i>Л-ра: 1, 3, 4, 6</i>	2 / –	---	---
	Л610 Побудова штучної нейронної мережі зі зворотнім розповсюдженням помилки. <i>Л-ра: 1, 3, 4, 6</i>	2 / –	---	---
	Пз10 Побудова штучної нейронної мережі зі зворотнім розповсюдженням помилки. <i>Л-ра: 1, 3, 4, 6</i>	4 / –	звіт з лістингами програм	2 / ---
	Л11 Застосування штучних нейронних мереж зворотнього розповсюдження помилки до класифікації зображень. <i>Л-ра: 1, 3, 4, 6</i>	2 / –	---	---
	Л611 Класифікація зображень за допомогою штучної нейронної мережі зі зворотнім розповсюдженням помилки.	2 / –	---	---

	<i>Л-ра: 1, 3, 4, 6</i>			
	Пз11 Класифікація зображень за допомогою штучної нейронної мережі зі зворотнім розповсюдженням помилки. <i>Л-ра: 1, 3, 4, 6</i>	4 / –	звіт з лістингами програм	2 / ---
	Л12 Застосування штучних нейронних мереж зворотнього розповсюдження помилки до прогнозування часових рядів. <i>Л-ра: 1, 3, 4, 6</i>	2 / –	---	---
	Лб12 Прогнозування часових рядів за допомогою штучної нейронної мережі зі зворотнім розповсюдженням помилки. <i>Л-ра: 1, 3, 4, 6</i>	2 / –	---	---
	Пз12 Прогнозування часових рядів за допомогою штучної нейронної мережі зі зворотнім розповсюдженням помилки. <i>Л-ра: 1, 3, 4, 6</i>	4 / –	звіт з лістингами програм	2 / ---
	РР3 Передбачення погоди за допомогою штучних нейронних мереж. <i>Л-ра: 1, 3, 4, 6</i>	15 / –	аналітичний звіт з лістингами програм	4 / ---
2.3. Кластеризація зразків	Л13 Нейронні мережі SOFM. <i>Л-ра: 1, 3, 4</i>	2 / –	---	---
	Лб13 Застосування нейронних мереж SOFM для розв'язання задач кластеризації. <i>Л-ра: 1, 3, 4</i>	2 / –	---	---
	Пз13 Застосування нейронних мереж SOFM для розв'язання задач кластеризації. <i>Л-ра: 1, 3, 4</i>	4 / –	звіт з лістингами програм	2 / ---
2.4. Асоціація зразків	Л14 Дискретні мережі Хопфілда. <i>Л-ра: 1, 3, 4</i>	2 / –	---	---
	Лб14 Реалізація асоціативної пам'яті. <i>Л-ра: 1, 3, 4</i>	2 / –	---	---
	Пз14 Реалізація асоціативної пам'яті. <i>Л-ра: 1, 3, 4</i>	4 / –	звіт з лістингами програм	2 / ---
2.5. Рекурентні штучні нейронні мережі	Л15 Модульні нейронні мережі. <i>Л-ра: 1, 3, 4</i>	2 / –	---	---

РЕКОМЕНДОВАНІ ДЖЕРЕЛА ІНФОРМАЦІЇ

Основна література

1. Wu X., Kumar V. The top ten algorithms in data mining. – CRC Press, 2009. – 201 p.
2. Сердюк О.А. Розв'язання задачі комівояжера за допомогою генетичного та мурашиного алгоритмів. Методичні вказівки. – 2018. – 43 с.8.

Додаткова література

3. Downey A.B. Think Bayes. Bayesian statistics in Python. – O'Reilly, 2013. – 209 p.
4. Kleinbaum D.G., Klein M. Logistic regression. A self-learning text. – Springer, 2010. – 701 p.
5. Price K.V., Storn R.M., Lampinen J.A. Differential evolution. A practical approach to global optimization. – Springer, 2005. – 538 p