

ІНТЕЛЕКТУАЛЬНИЙ АНАЛІЗ ДАНИХ

СИЛАБУС

Шифр і назва спеціальності	122-Комп'ютерні науки	Інститут / факультет	комп'ютерних наук і програмної інженерії
Назва програми	Комп'ютерні науки та інтелектуальні системи	Кафедра	програмної інженерії та інформаційних технологій управління
Тип програми	Освітньо-професійна	Мова навчання	Українська

Викладач

ПІБ, електронна пошта

ГОЛОСКОКОВА АННА ОЛЕКСАНДРІВНА,
anna.goloskokova@khpі.edu.ua



Науковий ступінь – кандидат технічних наук

Посада – доцент кафедри програмної інженерії та інформаційних технологій управління

Кількість публікацій – більш ніж 25, в тому числі 1 посібник з грифом МОН України та НТУ “ХПІ”, 2 статті у виданні, індексованому у Scopus

<https://scholar.google.com.ua/citations?hl=uk&user=jnDzQRAAAAAAJ>

<https://orcid.org/0000-0001-9012-7889>

<https://www.scopus.com/authid/detail.uri?origin=resultslist&authorId=57190428440&zone=57190428440>

Основні курси – системи штучного інтелекту (бакалаври), моделі та методи м'яких обчислень (бакалаври), вступ до нейронних мереж (бакалаври)

Загальна інформація про курс

Анотація	Дисципліна “Інтелектуальний аналіз даних» є навчальною дисципліною з циклу спеціальної (фахової) підготовки за спеціальністю 122 “Комп'ютерні науки”. Дисципліна викладається у 8 семестрі в обсязі 90 год. (3, кредита ECTS), зокрема: лекції – 20 год., лабораторні – 10 год., самостійна робота – 60 год. У курсі передбачено два змістових модулі та одна підсумкова модульна контрольна робота. Завершується дисципліна диф. заліком.
Цілі курсу	Формування теоретичних і практичних знань студентів в курсі інтелектуального аналізу даних у процесі проектування та впровадження програмних систем, застосування методів інтелектуального аналізу даних для вирішення проблем проектування програмного забезпечення.
Формат	Лекції, лабораторні заняття, самостійна робота. Підсумковий контроль – диф. Залік.
Семестр	8

Програмні
компетентності

ЗК1. Здатність до абстрактного мислення, аналізу та синтезу.

ЗК2. Здатність застосовувати знання у практичних ситуаціях.

ЗК3. Знання та розуміння предметної області та розуміння професійної діяльності.

ЗК6. Здатність вчитися й оволодівати сучасними знаннями.

ЗК7. Здатність до пошуку, оброблення та аналізу інформації з різних джерел.

СК2. Здатність до виявлення статистичних закономірностей недетермінованих явищ, застосування методів обчислювального інтелекту, зокрема статистичної, нейромережевої та нечіткої обробки даних, методів машинного навчання та генетичного програмування тощо.

СК7. Здатність застосовувати теоретичні та практичні основи методології та технології моделювання для дослідження характеристик і поведінки складних об'єктів і систем, проводити обчислювальні експерименти з обробкою й аналізом результатів.

СК11. Здатність до інтелектуального аналізу даних на основі методів обчислювального інтелекту включно з великими та погано структурованими даними, їхньої оперативної обробки та візуалізації результатів аналізу в процесі розв'язування прикладних задач.

СК16. Здатність реалізовувати високопродуктивні обчислення на основі хмарних сервісів і технологій, паралельних і розподілених обчислень при розробці й експлуатації розподілених систем паралельної обробки інформації.

СК17. Здатність застосовувати теоретичні та практичні основи сучасної теорії управління складними організаційно-технічними та соціально-економічними системами для побудови інтелектуальних систем управління, у процесі проектування інтелектуальних систем використовувати сучасні технології обробки інформації та методи обчислювального інтелекту.

Результати навчання

Методи викладання та навчання

Форми оцінювання
(поточне оцінювання CAS, підсумкове оцінювання FAS)

ПР3. Використовувати знання закономірностей випадкових явищ, їх властивостей та операцій над ними, моделей випадкових процесів та сучасних програмних середовищ для розв'язування задач статистичної обробки даних і побудови прогнозних моделей.

У процесі викладання передбачено застосування таких начальних технологій, як: лекції, лабораторні роботи, робота в малих групах, brainstorming, презентації, що розвивають комунікативні та лідерські навички, самостійна робота з літературними джерелами, змішані форми навчання з використанням дистанційних платформ

Письмові індивідуальні завдання до лабораторних робіт (CAS), оцінювання знань на лабораторних заняттях (CAS), експрес-опитування (CAS), онлайн-тести (CAS), підсумковий/семестровий контроль у формі семестрового диф. заліку, відповідно до графіку навчального процесу (FAS)

ПР4. Використовувати методи обчислювального інтелекту, машинного навчання, нейромережевої та нечіткої обробки даних, генетичного та еволюційного програмування для розв'язання задач розпізнавання, прогнозування, класифікації, ідентифікації об'єктів керування тощо.

У процесі викладання передбачено застосування таких начальних технологій, як: лекції, лабораторні роботи, робота в малих групах, brainstorming, презентації, що розвивають комунікативні та лідерські навички, самостійна робота з літературними джерелами, змішані форми навчання з використанням дистанційних платформ

Письмові індивідуальні завдання до лабораторних робіт (CAS), оцінювання знань на лабораторних заняттях (CAS), онлайн -тести (CAS), підсумковий/семестровий контроль у формі семестрового диф. заліку, відповідно до графіку навчального процесу (FAS)

ПР8. Використовувати методологію системного аналізу об'єктів, процесів і систем для задач аналізу, прогнозування, управління та проєктування динамічних процесів в макроекономічних, технічних, технологічних і фінансових об'єктах.

У процесі викладання передбачено застосування таких начальних технологій, як: лекції, лабораторні роботи, робота в малих групах, brainstorming, презентації, що розвивають комунікативні та лідерські навички, самостійна робота з літературними джерелами, змішані форми навчання з використанням дистанційних платформ

Письмові індивідуальні завдання до лабораторних робіт (CAS), оцінювання знань на лабораторних заняттях (CAS), онлайн -тести (CAS), підсумковий/семестровий контроль у формі семестрового диф. заліку, відповідно до графіку навчального процесу (FAS)

ПР12. Застосовувати методи та алгоритми обчислювального інтелекту та інтелектуального аналізу даних в задачах класифікації, прогнозування, кластерного аналізу, пошуку асоціативних правил з використанням програмних інструментів підтримки багатовимірного аналізу даних на основі технологій DataMining, TextMining, WebMining.

У процесі викладання передбачено застосування таких начальних технологій, як: лекції, лабораторні роботи, робота в малих групах, brainstorming, презентації, що розвивають комунікативні та лідерські навички, самостійна робота з літературними джерелами, змішані форми навчання з використанням дистанційних платформ

Письмові індивідуальні завдання до лабораторних робіт (CAS), оцінювання знань на лабораторних заняттях (CAS), онлайн -тести (CAS), підсумковий/семестровий контроль у формі семестрового диф. заліку, відповідно до графіку навчального процесу (FAS)

ПР16. Виконувати паралельні та розподілені обчислення, застосовувати чисельні методи та алгоритми для паралельних структур, мови паралельного програмування при розробці та експлуатації паралельного та розподіленого програмного забезпечення.

У процесі викладання передбачено застосування таких начальних технологій, як: лекції, лабораторні роботи, робота в малих групах, brainstorming, презентації, що розвивають комунікативні та лідерські навички, самостійна робота з літературними джерелами, змішані форми навчання з використанням дистанційних платформ

Письмові індивідуальні завдання до лабораторних робіт (CAS), оцінювання знань на лабораторних заняттях (CAS), онлайн -тести (CAS), підсумковий/семестровий контроль у формі семестрового диф. заліку, відповідно до графіку навчального процесу (FAS)

ПР17. Застосовувати для побудови інтелектуальних систем управління теоретичні та практичні основи сучасної теорії управління, проєктувати інтелектуальні системи з використанням сучасних технологій обробки інформації та методів обчислювального інтелекту.

У процесі викладання передбачено застосування таких начальних технологій, як: лекції, лабораторні роботи, робота в малих групах, brainstorming, презентації, що розвивають комунікативні та лідерські навички, самостійна робота з літературними джерелами, змішані форми навчання з використанням дистанційних платформ

Письмові індивідуальні завдання до лабораторних робіт (CAS), оцінювання знань на лабораторних заняттях (CAS), онлайн -тести (CAS), підсумковий/семестровий контроль у формі семестрового диф. заліку, відповідно до графіку навчального процесу (FAS)

СИСТЕМА ОЦІНЮВАННЯ

Розподіл балів для оцінювання успішності студента	Сума балів за всі види навчальної діяльності	Оцінка ECTS	Оцінка за національною шкалою	Нарахування балів
	90-100	A	відмінно	
	82-89	B	добре	
	74-81	C		
	64-73	D	задовільно	
	60-63	E		
	35-59	FX	незадовільно з можливістю повторного складання	
0-34	F	незадовільно з обов'язковим повторним вивченням дисципліни		

100% підсумкове оцінювання у вигляді диф. заліку (40%) та поточного оцінювання (60%).
40% диф. залік: семестровий диф. залік, відповідно до графіку навчального процесу
60% поточне оцінювання (написання модульних контрольних робіт – 20% семестрової оцінки, лабораторні заняття – 20% семестрової оцінки, самостійна робота – 20% семестрової оцінки)

Політика курсу	<p>Політика навчальної дисципліни визначається системою вимог щодо вивчення дисципліни, неприпустимість пропусків, виконання необхідного мінімуму навчальної роботи; заохочень і стягнень – нарахування або віднімання балів. Участь у лабораторних заняттях вимагає попередньої підготовки та завчасного опрацювання всіх необхідних матеріалів для продуктивних дискусій під час заняття. Письмові завдання повинні бути подані до встановлених строків. Політика навчальної дисципліни базується на урахуванні норм законодавства України щодо академічної доброчесності, Статуту, положень НТУ «ХПІ» .</p> <p>За порушення академічної доброчесності здобувачі освіти можуть бути притягнені до такої академічної відповідальності:</p> <ul style="list-style-type: none"> • зниження результатів оцінювання модульних контрольних робіт, диф. заліку; • повторне проходження оцінювання модульних контрольних робіт, диф. заліку; • повторне проходження оцінювання за лабораторними роботами; • повторне оцінювання за результатами виконання самостійної роботи; • призначення додаткових модульних контрольних індивідуальних завдань, модульних контрольних робіт, тестів.
-----------------------	--

Структура та зміст курсу

Лекція	Зміст лекції	Лабораторна робота	Зміст роботи	Тип роботи	Зміст роботи
Лекція 1	Вступ до інтелектуального аналізу даних. Визначення, потреби та цілі інтелектуального аналізу даних.	Лабораторна робота 1	Попередня обробка даних за допомогою Weka.	Самостійна робота	Задача регресії.
Лекція 2	Процес видобутку даних. Процес CRISP-DM. Процес SEMMA. Застосування процесів інтелектуального аналізу даних	Лабораторна робота 2	Рішення задачі кластеризації за допомогою Weka.		Візуалізація аналізу даних.
Лекція 3	Попередня обробка	Лабораторна робота 3	Обробка даних		Алгоритми видобутку тексту. Вирішення, припинення слів.

	даних. Обробка неінформативних змінних, відсутніх даних, очищення даних.		JSON за допомогою Python.		
Лекція 4	Схеми вибірки для фіксованих наборів даних. Зменшення кількості атрибутів.	Лабораторна робота 4	Обробка XML даних за допомогою Python.		Алгоритми веб -майнінгу. Отримання та аналіз веб -документів.
Лекція 5	Методи міркування на основі пам'яті. Збіг, зважений збіг.	Лабораторна робота 5	Реалізація веб -сканера за допомогою Python.		Агентські системи для реалізації алгоритмів веб -майнінгу.
Лекція 6	Правила асоціації у видобутку знань. Аналіз кошика ринку.				
Лекція 7	Кластеризація. K-means кластеризація.				
Лекція 8	Видобуток тексту. Представлення текстових документів. Побудова моделі векторного простору.				
Лекція 9	Веб -майнінг. Веб -контент, використання Інтернету, видобуток веб -структур.				
Лекція 10	Інтелектуальні методи обробки веб -даних.				

Література

Обов'язкова

1. Kretowski, M. (2019). Evolutionary Decision Trees in Large-Scale Data Mining. Springer.
2. Ranga Suri N. N. R., Murty, N., Athithan, M. G. (2019). Outlier Detection: Techniques and Applications. A Data Mining Perspective. Cham: Springer International Publishing.
3. Anandan, R. (2021). A Closer Look at Big Data Analytics. Nova Science Publishers. Inc.
4. Samanta, D., Banerjee, A. (2021). Computationally Intensive Statistics for Intelligent IoT. Springer.
5. Akerkar, R. (2020). Big Data in Emergency Management: Exploitation Techniques for Social and Mobile Data. Springer.
6. Moreira, J. (2018). Data Analysis Made Easy. New York: Wiley-Interscience.
7. Srinivas, M., Sucharitha, G, Matta, A., Chatterjee, P. (2021). Machine Learning Algorithms and Applications: Theory and Applications. Wiley-Scrivener Publishing.

Додаткова

1. Scitovski, R., Sabo, K., Martínez-Álvarez, F., Ungar, S. (2021). Cluster Analysis and Applications. Springer.
2. Бондаренко, М. Ф., Шабанов-Кушнарченко, Ю. П. Сергиенко, И. В. (Ред.). (2011). Мозгоподобные структуры: Справочное пособие. (Т. 1) Київ: Наукова думка.
3. Tyagi, A. K. (2021). Data Science and Data Analytics: Opportunities and Challenges. New York: Chapman and Hall/CRC.
4. Sakarkar, G., Patil, G., Dutta, P. (2021). Machine Learning Algorithms Using Python Programming. Nova Science Publishers. Inc.
5. Wooldridge, M. J. (2009). An introduction to multiagent systems. John Wiley & Sons. LTD.
6. Бондаренко, М. Ф., Шабанов-Кушнарченко, Ю. П. (2006). Теория интеллекта: Учебник. Харків: Компания СМІТ.

ІНФОРМАЦІЙНІ РЕСУРСИ В ІНТЕРНЕТІ

1. Weka 3: Machine Learning Software in Java. Retrieved from <https://www.cs.waikato.ac.nz/ml/weka>.
2. Data Mining: Practical Machine Learning Tools and Techniques. Retrieved from <https://www.cs.waikato.ac.nz/ml/weka/book.html>.

Норми академічної етики

Студент повинен дотримуватися «Кодексу етики академічних взаємовідносин та доброчесності НТУ «ХПІ»: виявляти дисциплінованість, вихованість, доброзичливість, чесність, відповідальність. Конфліктні ситуації повинні відкрито обговорюватись в навчальних групах з викладачем, а при нерозв'язності конфлікту звертатись до завідувача кафедри програмної інженерії та інформаційних технологій управління

Силабус за змістом повністю відповідає робочій програмі курсу.