

Житомирська політехніка	МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ ДЕРЖАВНИЙ УНІВЕРСИТЕТ «ЖИТОМИРСЬКА ПОЛІТЕХНІКА» Система управління якістю відповідає ДСТУ ISO 9001:2015	Ф-22.07- 05.01/121.00.1/ДФ/ ОК7
	Екземпляр № 1	Арк 11 / 1

ЗАТВЕРДЖЕНО

Вченою радою факультету
інформаційно-комп'ютерних
технологій

31 серпня 2022 р., протокол № 8

Голова Вченої ради

Тетяна НІКІТЧУК



РОБОЧА ПРОГРАМА НАВЧАЛЬНОЇ ДИСЦИПЛІНИ «СУЧАСНІ МЕТОДИ ІНТЕЛЕКТУАЛЬНОГО АНАЛІЗУ ДАНИХ»

для здобувачів вищої освіти ступеня доктора філософії
спеціальності 121 «Інженерія програмного забезпечення»
освітньо-наукова програма «Інженерія програмного забезпечення»
факультет інформаційно-комп'ютерних технологій
кафедра інженерії програмного забезпечення

Схвалено на засіданні кафедри
інженерії програмного
забезпечення
26 серпня 2022 р.,
протокол № 6

Розробник: к.т.н., доц., доцент кафедри комп'ютерних наук Інна СУГОНЯК

Житомир
2022 – 2023 н.р.

Житомирська політехніка	МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ ДЕРЖАВНИЙ УНІВЕРСИТЕТ «ЖИТОМИРСЬКА ПОЛІТЕХНІКА» Система управління якістю відповідає ДСТУ ISO 9001:2015	Ф-22.07- 05.01/121.00.1/ДФ/ ОК7
	Екземпляр № 1	Арк 11 / 2

1. ОПИС НАВЧАЛЬНОЇ ДИСЦИПЛІНИ

Найменування показників	Галузь знань, спеціальність, освітній рівень	Характеристика навчальної дисципліни
		денна форма навчання
Кількість кредитів – 5	Галузь знань 12 – Інформаційні технології	нормативна
Модулів – 2	Спеціальність: 121 «Інженерія програмного забезпечення»	Рік підготовки:
Змістових модулів – 2		2-ий
Загальна кількість годин - 150		Семестр
		3-ій
Тижневих годин для денної форми навчання: аудиторних – 4 самостійної роботи – 5,38	Освітньо-науковий ступінь: «доктор філософії»	Лекції
		32 год.
		Практичні
		32 год.
		Лабораторні
		0 год.
		Самостійна робота
86 год.		
	Індивідуальні завдання: -	
	Вид контролю: екзамен	

Житомирська політехніка	МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ ДЕРЖАВНИЙ УНІВЕРСИТЕТ «ЖИТОМИРСЬКА ПОЛІТЕХНІКА» Система управління якістю відповідає ДСТУ ISO 9001:2015	Ф-22.07- 05.01/121.00.1/ДФ/ ОК7
	Екземпляр № 1	Арк 11 / 3

2. МЕТА ТА ЗАВДАННЯ НАВЧАЛЬНОЇ ДИСЦИПЛІНИ

Метою навчальної дисципліни є формування знань, вмінь та навичок, необхідних для розробки систем із штучного інтелекту з використанням сучасних методів і засобів програмування (нейронні мережі, об'єктно-орієнтований підхід, моделі представлення знань) для вирішення проблем інтелектуального аналізу даних.

Завданнями вивчення навчальної дисципліни є:

Вивчення сучасних методів і технологій штучного інтелекту, розроблення та застосування моделей представлення знань для побудови інтелектуальних систем аналізу даних.

У результаті вивчення навчальної дисципліни здобувач повинен знати: особливості застосування штучного інтелекту у комп'ютерних системах; особливості застосування експертних систем; моделі подання знань у експертних системах; методи математичного моделювання та управління виробничими процесами; гібридні інтелектуальні системи.

Результатом вивчення дисципліни є набуття здобувачами вищої освіти таких компетентностей:

Спеціальні компетентності:

СК04. Здатність відстежувати тенденції розвитку інженерії програмного забезпечення та критично переосмислювати наявні технології.

СК05. Здатність до розроблення нових та вдосконалення існуючих моделей, методів, засобів, процесів у сфері інженерії програмного забезпечення, які забезпечують розвиток або надають нові можливості технологіям розробки та супроводження програмного забезпечення.

Програмні результати навчання:

РН05. Застосовувати сучасні інструменти і технології пошуку, оброблення та аналізу інформації, зокрема, статистичні методи аналізу даних великого обсягу та/або складної структури, спеціалізовані бази даних та інформаційні системи для покращення ефективності програмних систем.

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.

Житомирська політехніка	МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ ДЕРЖАВНИЙ УНІВЕРСИТЕТ «ЖИТОМИРСЬКА ПОЛІТЕХНІКА» Система управління якістю відповідає ДСТУ ISO 9001:2015	Ф-22.07- 05.01/121.00.1/ДФ/ ОК7
	<i>Екземпляр № 1</i>	<i>Арк 11 / 4</i>

РН08. Глибоко розуміти загальні принципи та методи інженерії програмного забезпечення, а також методологію наукових досліджень, застосовувати їх у власних дослідженнях та у викладацькій практиці.

РН09. Формулювати та вирішувати задачі оптимізації, адаптації, прогнозування, керування та прийняття рішень щодо процесів, засобів та ресурсів розробки, впровадження, супроводу та експлуатації програмного забезпечення.

РН10. Аналізувати та оцінювати стан і перспективи розвитку інженерії програмного забезпечення та інформаційних технологій у цілому.

Житомирська політехніка	МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ ДЕРЖАВНИЙ УНІВЕРСИТЕТ «ЖИТОМИРСЬКА ПОЛІТЕХНІКА» Система управління якістю відповідає ДСТУ ISO 9001:2015	Ф-22.07- 05.01/121.00.1/ДФ/ ОК7
	Екземпляр № 1	Арк 11 / 5

3. ПРОГРАМА НАВЧАЛЬНОЇ ДИСЦИПЛІНИ

Змістовий модуль 1. Аналіз даних

Тема 1. Інтелектуальний аналіз даних, інструменти та методи.

Що таке аналіз даних. Інтелектуальний аналіз даних як процес.

Інструменти інтелектуального аналізу даних. Задачі інтелектуального аналізу даних. Інтерпретатор Python. Бібліотеки Pandas, NumPy, Matplotlib, SciPy.

Тема 2. Розвідувальний аналіз.

Розвідувальний аналіз. Трактуювання результатів. Парадокс Сімпсона. Імпорт даних та дослідження даних (кількість вимірів, відображення першої та останньої частин об'єкта, імен об'єкта). Генеральна сукупність, вибірка. Репрезентативна та нерепрезентативна вибірки. Центральна тенденція (середнє значення, медіана, мода). Дисперсія та середньоквадратичне відхилення. Квартилі та інтерквартильний розмах. Коробчата діаграма.

Тема 3. Візуальний аналіз даних.

Візуальна аналітика. Бібліотеки Matplotlib і Seaborn. Коробчата діаграма.

Тема 4. Розподіли.

Класичні розподіли. Біноміальний розподіл. Нормальний розподіл. Візуалізація біноміального та нормального розподілів. Розподіл Пуассона. Геометричний розподіл. Рівномірний розподіл. Експоненціальний розподіл.

Тема 5. Кореляційний аналіз.

Коваріація. Коефіцієнт кореляції Пірсона (лінійний коефіцієнт кореляції). Інтерпретація коефіцієнта кореляції. Коефіцієнт кореляції рангів Спірмена

Змістовий модуль 2. Машинне навчання

Тема 6. Регресійний аналіз. Дерева прийняття рішень

Проста лінійна регресія. Метод найменших квадратів. Поліноміальна регресія. Логістична регресія. Дерево рішень в регресії.

Тема 7. Кластеризація.

Базові алгоритми кластеризації. Об'єднання кластерів. Алгоритми ієрархічної кластеризації.

Тема 8. Нейронні мережі.

Розпізнавання рукописних цифр та зображень. MLP (Multi-layer Perceptron). Набір даних MNIST. Згортова нейронна мережа (CNN, Convolutional Neural Networks). Набір даних CIFAR-10. Реалізація і навчання нейронної мережі з використанням Keras.

Тема 9. Асоціативні правила. Математичні методи аналізу текстів

Житомирська політехніка	МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ ДЕРЖАВНИЙ УНІВЕРСИТЕТ «ЖИТОМИРСЬКА ПОЛІТЕХНІКА» Система управління якістю відповідає ДСТУ ISO 9001:2015	Ф-22.07- 05.01/121.00.1/ДФ/ ОК7
	Екземпляр № 1	Арк 11 / 6

Тема 10. Введення в обробку зображень в Python з OpenCV

Установка. Пошук деталей зображення. Поділ зображення на окремі канали. Фільтрація: бінаризація по порозу, вибір; класична фільтрація: Фур'є, ФНЧ, ФВ; вейвлети; кореляція; фільтрації функцій; фільтрації контурів. Логічна обробка результатів фільтрації: морфологія, контурний аналіз, особливі точки.

4. СТРУКТУРА (ТЕМАТИЧНИЙ ПЛАН) НАВЧАЛЬНОЇ ДИСЦИПЛІНИ

Змістові модулі і теми	Кількість годин			
	денна форма			
	усього	лекції	практичні роботи	самостійна робота
Модуль 1				
Змістовий модуль 1. Аналіз даних				
Тема 1. Інтелектуальний аналіз даних, інструменти та методи	6	2	-	4
Тема 2. Розвідувальний аналіз.	14	2	-	12
Тема 3. Візуальний аналіз даних.	14	4	-	10
Тема 4. Розподіли.	10	2	8	-
Тема 5. Кореляційний аналіз.	16	2	8	6
<i>Разом за змістовий модуль 1</i>				
	60	12	16	32
Змістовий модуль 2. Машинне навчання				
Тема 6. Регресійний аналіз. Древа прийняття рішень	16	2	-	14
Тема 7. Кластеризація.	20	4	8	8
Тема 8. Нейронні мережі.	20	4	8	8
Тема 9. Асоціативні правила. Математичні методи аналізу текстів.	18	6	-	12
Тема 10. Введення в обробку зображень в Python з OpenCV.	16	4	-	12
<i>Разом за змістовий модуль 2</i>				
	90	20	16	54
ВСЬОГО				
	150	32	32	86

Житомирська політехніка	МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ ДЕРЖАВНИЙ УНІВЕРСИТЕТ «ЖИТОМИРСЬКА ПОЛІТЕХНІКА» Система управління якістю відповідає ДСТУ ISO 9001:2015	Ф-22.07- 05.01/121.00.1/ДФ/ ОК7
	Екземпляр № 1	Арк 11 / 7

5. ТЕМИ ПРАКТИЧНИХ (ЛАБОРАТОРНИХ) ЗАНЯТЬ

№ з/п	Назва теми	Кількість годин
1	Використання модуля Pandas. Аналіз даних по серцевосудинних захворюваннях. Описова статистика (Descriptive statistics) Python Pandas	8
2	Візуалізації даних в Python з Matplotlib і Seaborn.	8
3	Регресія. Метод найменших квадратів. Кластеризація. Базові алгоритми кластеризації. Ієрархічні алгоритми.	8
4	Розпізнання. Глибоке навчання в Python. Розпізнавання рукописних цифр та зображень. Розпізнавання обличчя на зображенні за допомогою Python та OpenCV.	8
РАЗОМ		32

Житомирська політехніка	МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ ДЕРЖАВНИЙ УНІВЕРСИТЕТ «ЖИТОМИРСЬКА ПОЛІТЕХНІКА» Система управління якістю відповідає ДСТУ ISO 9001:2015	Ф-22.07- 05.01/121.00.1/ДФ/ ОК7
	Екземпляр № 1	Арк 11 / 8

6. ЗАВДАННЯ ДЛЯ САМОСТІЙНОЇ РОБОТИ

Передбачається, що в період вивчення дисциплін аспірант самостійно розв'язує домашнє завдання, вивчає матеріал курсу в процесі підготовки до практичних занять, а також в цілому перед сесією.

№ з/п	Назва теми	Кількість годин
1	Розвідувальний аналіз. Попередній експрес-аналіз даних шляхом їх перетворення та/або представлення у зручному вигляді: графічному, табличному, схем, діаграм тощо.	4
2	Описова статистика. Одновимірний розподіл, двовимірний розподіл або крос-табуляція, міри центральної тенденції. Середнє (Mean). Медіана (Median). Мода (Mode). Нормальний розподіл (Normal distribution). Дисперсія і стандартне відхилення. Бімодальні розподіли. Квартилі	12
3	Візуальний аналіз даних. Візуалізація та її типи. Стандартні елементи візуалізації даних. Графіки та їх типи. Ефективні види графіків. Візуальний аналіз даних	10
4	Розподіли Нормальний розподіл (закон Гаусса). Крива нормального розподілу (крива Гауса). Емпіричний розподіл Кореляційний аналіз Поняття кореляції. Кореляційний момент. Коефіцієнт кореляції. Нормальний закон розподілу на площині	6
5	Регресійний аналіз Кореляційний та регресійний аналіз прямолінійної залежності. Дерева прийняття рішень Дерева рішень.	14
6	Кластеризація, або кластерний аналіз. Задача кластеризації. Два типи кластеризації. Алгоритми кластеризації. Методи кластеризації. Метод k-середніх. Ієрархічна кластеризація	8
7	Нейронні мережі Згорткові нейронні мережі (Convolutional neural networks, CNN). Генеративні нейронні мережі (Generative adversarial network, GAN). Генерація зображень за допомогою нейромереж. Нейронні мережі та розпізнавання зображень. Нейромережа для розпізнавання осіб.	8
8	Аналіз тексту. Текстова аналітика. Класифікація новин. Визначення мови.	12
9	OpenCV Бібліотека комп'ютерного зору. Обробка зображень.	12
РАЗОМ		86

7. МЕТОДИ НАВЧАННЯ

Під час викладання курсу використовуються методи організації та здійснення навчально-пізнавальної діяльності, методи стимулювання й мотивації навчально-пізнавальної діяльності, методи контролю (самоконтролю, взаємоконтролю), корекції (самокорекції, взаємокорекції) за ефективністю навчально-пізнавальної діяльності, інтегровані (універсальні) методи.

Житомирська політехніка	МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ ДЕРЖАВНИЙ УНІВЕРСИТЕТ «ЖИТОМИРСЬКА ПОЛІТЕХНІКА» Система управління якістю відповідає ДСТУ ISO 9001:2015	Ф-22.07- 05.01/121.00.1/ДФ/ ОК7
	Екземпляр № 1	Арк 11 / 9

За джерелами знань використовуються такі методи навчання: словесні – розповідь, пояснення, лекція, інструктаж; наочні – демонстрація, ілюстрація; практичні – лабораторна робота, практична робота.

За характером логіки пізнання використовуються такі методи: аналітичний, синтетичний, аналітико-синтетичний, індуктивний, дедуктивний.

За рівнем самостійної розумової діяльності використовуються методи: проблемний, частково-пошуковий, дослідницький.

Для полегшення засвоєння матеріалу використовуються технічні засоби:

№ з/п	Назва теми	Технічні засоби
1	Тема 1. Інтелектуальний аналіз даних, інструменти та методи	Проектор з комплектом слайдів
2	Тема 2. Розвідувальний аналіз.	Проектор з комплектом слайдів
3	Тема 3. Візуальний аналіз даних.	Проектор з комплектом слайдів
4	Тема 4. Розподіли.	Проектор з комплектом слайдів
5	Тема 5. Кореляційний аналіз.	Проектор з комплектом слайдів
6	Тема 6. Регресійний аналіз. Деревя прийняття рішень	Проектор з комплектом слайдів
7	Тема 7. Кластеризація.	Проектор з комплектом слайдів
8	Тема 8. Нейронні мережі.	Проектор з комплектом слайдів
9	Тема 9. Асоціативні правила. Математичні методи аналізу текстів.	Проектор з комплектом слайдів
10	Тема 10. Введення в обробку зображень в Python з OpenCV.	Проектор з комплектом слайдів

8. МЕТОДИ КОНТРОЛЮ

Під час вивчення дисципліни використовуються наступні методи контролю: поточне та підсумкове тестування за теоретичним матеріалом, захист практичних робіт у формі співбесіди.

Підсумковий контроль – екзамен.

Житомирська політехніка	МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ ДЕРЖАВНИЙ УНІВЕРСИТЕТ «ЖИТОМИРСЬКА ПОЛІТЕХНІКА» Система управління якістю відповідає ДСТУ ISO 9001:2015	Ф-22.07- 05.01/121.00.1/ДФ/ ОК7
	Екземпляр № 1	Арк 11 / 10

9. РОЗПОДІЛ БАЛІВ, ЯКІ ОТРИМУЮТЬ ЗДОБУВАЧІ ВИЩОЇ ОСВІТИ

Поточне тестування та самостійна робота				Сума
модуль 1		модуль 2		
Опитування	Лабораторні роботи	Опитування	Лабораторні роботи	100
10	40 (20 + 20)	10	40 (20 + 20)	

О – опитування; І – індивідуальне завдання; МКР – модульна контрольна

Шкала оцінювання: національна та ECTS

Сума балів за всі види навчальної діяльності	Оцінка ECTS	Оцінка за національною шкалою
		для екзамену, курсового проекту (роботи), практики
90 – 100	A	відмінно
82 – 89	B	добре
74 – 81	C	
64 – 73	D	задовільно
60 – 63	E	
35 – 59	FX	незадовільно
0 – 34	F	незадовільно

10. РЕКОМЕНДОВАНА ЛІТЕРАТУРА

1. Ямпольський Л.С. Нейротехнології та нейрокомп'ютерні системи / Ямпольський Л.С., Лісовиченко О.І., Олійник В.В. // Дорадо-друк, Київ, 2016. - 571 с.
2. Інтелектуальні системи управління: навч. посіб. для студ. / КПІ ім. Ігоря Сікорського; уклад.: Л. Д. Ярошук, В. І. Бородін. – Електронні текстові дані. – Київ: КПІ ім. Ігоря Сікорського, 2018. – 81 с.
2. Гнучкі комп'ютерно-інтегровані системи: планування, моделювання, верифікація, управління / Л.С. Ямпольський, П.П. Мельничук, К.Б. Остапченко, О.І. Лісовиченко. – Житомир: ЖДТУ, 2010.
3. Інтелектуальні системи: навч. посіб. / С. О. Субботін, А. О. Олійник; за ред. С. О. Субботіна. - Запоріжжя: ЗНТУ, 2014. - 219 с.
3. Іванченко Г. Ф. Системи штучного інтелекту: навч. посібник / Г. Ф. Іванченко. – К., 2011. – 382 с
4. Антоненко В. М. Сучасні інформаційні системи і технології: управління знаннями: навчальний посібник / В. М. Антоненко, С. Д. Мамченко, Ю. В. Рогушина. – Ірпінь: Національний університет ДПС України, 2016. – 212 с.

Житомирська політехніка	МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ ДЕРЖАВНИЙ УНІВЕРСИТЕТ «ЖИТОМИРСЬКА ПОЛІТЕХНІКА» Система управління якістю відповідає ДСТУ ISO 9001:2015	Ф-22.07- 05.01/121.00.1/ДФ/ ОК7
	Екземпляр № 1	Арк 11 / 11

Допоміжна література

1. Комп'ютерні системи штучного інтелекту: підручник для студ. Вузів / Доля В.Г. – К. : Університет Україна, 2011. – 296 с.
2. Ямпольський Л.С. Нейротехнології та нейрокомп'ютерні ситеми / Ямпольський Л.С., Лісовиченко О.І., Олійник В.В. // Дорадо-друк, Київ, 2016. - 571 с.
3. Richard S. Sutton, Andrew G. Barto. Reinforcement Learning : An Introduction MIT Press, Adaptive Computation and Machine Learning Ser.: 2018. - 552 p.
4. Руденко О. Г., Бодянський Є. В. Штучні нейронні мережі: Навчальний посібник. – Харків: ТОВ "Компанія СМІТ", 2006. – 404 с.

11. ІНФОРМАЦІЙНІ РЕСУРСИ В ІНТЕРНЕТІ

1. Репозитарій Державного університету «Житомирська політехніка». Доступ до ресурсу: <http://eztuir.ztu.edu.ua/>
2. Національна бібліотека України імені В.І. Вернадського. Доступ до ресурсу: www.nbuv.gov.ua