

Міністерство освіти і науки України
Національний технічний університет
“Дніпровська політехніка”

Кафедра інформаційних систем та комп’ютерної інженерії



«ЗАТВЕРДЖЕНО»

завідувач кафедри

Гнатушенко В.В.

« 29 »__серпня__2024 р.

протокол № 8

РОБОЧА ПРОГРАМА НАВЧАЛЬНОЇ ДИСЦИПЛІНИ
“Методи та інформаційні технології обробки великих даних (Big Data)”

Галузь знань	12 Інформаційні технології
Спеціальність	126 Інформаційні системи та технології
Рівень вищої освіти	перший (бакалаврський)
Ступінь	бакалавр
Освітня програма	Інформаційні системи та технології
Спеціалізація	
Статус.....	нормативна
Загальний обсяг	4,5 кредити ECTS (135 годин)
Форма підсумкового контролю	іспит
Термін викладання	8-й семестр
Мова викладання	українська

Викладач: доц. Кожевников А.В.

Пролонговано: на 20__/20__ н.р. _____ (_____) “__”__ 20__р.
(підпис, ПІБ, дата)

на 20__/20__ н.р. _____ (_____) “__”__ 20__р.
(підпис, ПІБ, дата)

Дніпро
НТУ “ДП”
2024

Робоча програма навчальної дисципліни “Методи та інформаційні технології обробки великих даних (Big Data)” для бакалаврів спеціальності 126 “Інформаційні системи та технології” / Нац. техн. ун-т. “Дніпровська політехніка”, каф. ІТКІ. – Д. : НТУ “ДП”, 2025. – 10 с.

Розробники: Кожевников А.В. – доцент, кандидат технічних наук, доцент кафедри інформаційних систем та комп’ютерної інженерії.

- мету дисципліни;
- дисциплінарні результати навчання, сформовані на основі трансформації очікуваних результатів навчання освітньої програми;
- базові дисципліни;
- обсяг і розподіл за формами організації освітнього процесу та видами навчальних занять;
- програму дисципліни (тематичний план за видами навчальних занять);
- алгоритм оцінювання рівня досягнення дисциплінарних результатів навчання (шкали, засоби, процедури та критерії оцінювання);
- інструменти, обладнання та програмне забезпечення;
- рекомендовані джерела інформації.

Робоча програма призначена для реалізації компетентнісного підходу під час планування освітнього процесу, викладання дисципліни, підготовки здобувачів вищої освіти до контрольних заходів, контролю провадження освітньої діяльності, внутрішнього та зовнішнього контролю забезпечення якості вищої освіти, акредитації освітніх програм у межах спеціальності.

Робоча програма буде в пригоді для формування змісту підвищення кваліфікації науково-педагогічних працівників кафедр університету.

Погоджено рішенням науково-методичної комісії спеціальності 126 «Інформаційні системи та технології» (протокол № _8_ від _29_. _08_.2024 р.).

ЗМІСТ

1. МЕТА НАВЧАЛЬНОЇ ДИСЦИПЛІНИ	4
2. ОЧІКУВАНІ ДИСЦИПЛІНАРНІ РЕЗУЛЬТАТИ НАВЧАННЯ	4
3. БАЗОВІ ДИСЦИПЛІНИ	5
4. ОБСЯГ І РОЗПОДІЛ ЗА ФОРМАМИ ОРГАНІЗАЦІЇ ОСВІТНЬОГО ПРОЦЕСУ ТА ВИДАМИ НАВЧАЛЬНИХ ЗАНЯТЬ	5
5. ПРОГРАМА ДИСЦИПЛІНИ ЗА ВИДАМИ НАВЧАЛЬНИХ ЗАНЯТЬ	6
6. ОЦІНЮВАННЯ РЕЗУЛЬТАТІВ НАВЧАННЯ	7
6.1. Шкали оцінювання	7
6.2. Підсумкова оцінка	7
6.3. Критерії оцінювання теоретичної частини курсу	8
6.4. Критерії оцінювання лабораторних робіт	8
6.5. Критерії оцінювання самостійної роботи	8
7. ІНСТРУМЕНТИ, ОБЛАДНАННЯ ТА ПРОГРАМНЕ ЗАБЕЗПЕЧЕННЯ	9
8. РЕКОМЕНДОВАНІ ДЖЕРЕЛА ІНФОРМАЦІЇ	9

1. МЕТА НАВЧАЛЬНОЇ ДИСЦИПЛІНИ

В освітньо-професійній програмі Національного технічного університету “Дніпровська політехніка” спеціальності 126 “Інформаційні системи та технології” здійснено розподіл програмних результатів навчання (ПРН) за організаційними формами освітнього процесу. Зокрема, до дисципліни Ф19 “Методи та інформаційні технології обробки великих даних (Big Data)” віднесено такі результати навчання:

Шифр ПРН	Зміст результату навчання
ПР4	Проводити системний аналіз об’єктів проектування та обґрунтовувати вибір структури, алгоритмів та способів передачі інформації в інформаційних системах та технологіях
ПР6	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
ПР9	Здійснювати системний аналіз архітектури підприємства та його ІТ-інфраструктури, проводити розроблення та вдосконалення її елементної бази і структури
ПР12	Використовувати сучасні методи розробки програмного забезпечення, яке є складовою інформаційних систем та технологій
ПР13	Застосовувати методи штучного інтелекту для рішення різноманітних задач виробництва та суспільства
ПР14	Розробляти комп’ютерні системи розпізнавання та обробки зображень, зокрема аерокосмічних зображень, обробки просторових даних, а також програмне забезпечення різноманітної спрямованості

Мета дисципліни – формування умінь та компетенцій щодо методів та інформаційні технології обробки даних, розмір яких перевищує можливості звичайних програмних аналітичних платформ та баз даних по аналізу, зберіганню, і управлінню інформацією, або великих даних (Big Data). Реалізація мети вимагає трансформації програмних результатів навчання в дисциплінарні та адекватний відбір змісту навчальної дисципліни за цим критерієм.

2. ОЧІКУВАНІ ДИСЦИПЛІНАРНІ РЕЗУЛЬТАТИ НАВЧАННЯ

Шифр ПРН	Дисциплінарні результати навчання (ДРН)	
	Шифр ДРН	Зміст
ПР4	ПР4.1-Ф19	Проводити аналіз об’єктів, які оперують з великими даними, та обґрунтовувати вибір алгоритмів обробки даних
ПР6	ПР6.1-Ф19	Демонструвати знання методів та інформаційних технологій обробки великих даних
ПР9	ПР9.1-Ф19	На основі даних аналізу ІТ-інфраструктури підприємства, запропоновувати її складові, які удосконалюють роботу з великими даними
ПР12	ПР12.1-Ф19	Використовувати сучасні методи розробки програмного забезпечення для обробки великих даних
ПР13	ПР13.1-Ф19	Застосовувати методи регресійного аналізу для рішення задач прогнозування

Шифр РН	Дисциплінарні результати навчання (ДРН)	
	Шифр ДРН	Зміст
	ПР13.2-Ф19	Використовувати методи кластеризації та класифікації для обробки великих даних
ПР14	ПР14.1-Ф19	Розробляти та використовувати спеціалізоване програмне забезпечення для обробки і візуалізації просторових та багатовимірних даних

3. БАЗОВІ ДИСЦИПЛІНИ

Базовими дисциплінами є дисципліни які вивчалися здобувачами вищої освіти та формують компетентності щодо здатності вивчення дисципліни Ф19 “Методи та інформаційні технології обробки великих даних (Big Data)” та ініціативності, відповідальності, навичок безпечної діяльності відповідно майбутнього профілю роботи.

Назва дисципліни	Здобуті результати навчання
Б1 Вища математика, Б2 Фізика, Б3 Теорія ймовірностей та математична статистика	Застосовувати знання фундаментальних наук при розв’язанні задач використання інформаційних систем та технологій
Ф1 Програмування Ф11 Комп’ютерна математика	Використовувати базові знання інформатики й сучасних інформаційних систем та технологій, навички програмування для розв’язання задач використання інформаційних систем та технологій
Ф3 Алгоритми та структури даних Ф6 Бази даних в інформаційних системах	Використовувати 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
Ф15 Проектування інформаційних систем	Аргументувати вибір програмних та технічних засобів для створення інформаційних систем та технологій на основі аналізу їх властивостей, призначення і технічних характеристик з урахуванням вимог до системи і експлуатаційних умов
Ф16 Системи штучного інтелекту	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

4. ОБСЯГ І РОЗПОДІЛ ЗА ФОРМАМИ ОРГАНІЗАЦІЇ ОСВІТНЬОГО ПРОЦЕСУ ТА ВИДАМИ НАВЧАЛЬНИХ ЗАНЯТЬ

Вид навчальних	яг, год	Розподіл за формами навчання, години		
		денна	вечірня	заочна

заняць		аудиторні заняття	самостійна робота	аудиторні заняття	самостійна робота	аудиторні заняття	самостійна робота
лекційні	58	28	30	-	-	6	56
практичні	-	-	-	-	-	-	-
лабораторні	73	28	45	-	-	6	67
контрольні заходи	4	4	-	-	-	-	-
РАЗОМ	135	60	75	-	-	12	123

5. ПРОГРАМА ДИСЦИПЛІНИ ЗА ВИДАМИ НАВЧАЛЬНИХ ЗАНЯТЬ

Шифри ДРН	Види та тематика навчальних занять	Обсяг складових, години
	ЛЕКЦІЇ	58
ПР6.1-Ф19	Вступ. Мета і завдання дисципліни “ Методи та інформаційні технології обробки великих даних (Big Data)”. Сфери застосування Big Data	2
	Тема 1. Подання даних та їх попередня обробка Визначення. Шкали виміру ознак. Життєвий цикл даних. Попередня обробка даних. Алгоритм ZET заповнення пробілів у таблицях даних. Метадані. Життєвий цикл метаданих	6
ПР4.1-Ф19, ПР9.1-Ф19, ПР13.1-Ф19	Тема 2. Методи класифікації даних. Кластерний аналіз Огляд методів класифікації даних. Завдання кластерного аналізу. Міри відстані та збіжності. Основні метрики. Основні методи кластерного аналізу: ієрархічні агломеративні та дивізимні методи, метод К-середніх Мак-Куїна. Нечітка кластеризація, метод С-середніх	8
	Тема 3. Методи класифікації даних. Дерева рішень Структура дерева рішень. Критерії обрання атрибута розбиття дерева. Критерії припинення розбиття. Відсікання гілок. Алгоритм ID3	8
ПР4.1-Ф19, ПР9.1-Ф19, ПР13.2-Ф19	Тема 4. Методи класифікації даних. Регресійний аналіз Багатовимірний лінійний регресійний аналіз. Відбір факторних змінних. Мультиколінеарність. Алгоритм Фаррара-Глобера. Аналіз та прогнозування часових рядів. Автокорреляція часових рядів. AR та МАРпроцеси	8
	Тема 5. Методи класифікації даних. Нейронні мережі Елементи нейронної мережі. Функції активації нейрона. Архітектура мереж. Помилка мережі. Навчання мереж. Правила Хебба. Правило Відроу-Хоффа.	8
ПР6.1-Ф19, ПР12.1-Ф19, ПР14.1-Ф19	Тема 6. Програмні інструменти Big Data. Мова Python Статистичні функції мови Python. Аналіз даних засобами мови: регресійний, кластерний, факторний	8
	Тема 7. Програмні інструменти Big Data. Платформа Hadoop Архітектура та компоненти, кластери, їх вузли, файлова система HDFS. Доступ до сервісів Hadoop засобами мови Python	10
	ЛАБОРАТОРНІ РОБОТИ	73
ПР6.1-Ф19	Лабораторна робота 1 Імовірнісні розподіли дискретних випадкових величин і точкові	10

Шифри ДРН	Види та тематика навчальних занять	Обсяг складових, години
	оцінки їх параметрів	
ПР6.1-Ф19	Лабораторна робота 2 Імовірнісні розподіли неперервних випадкових величин і точкові оцінки їх параметрів	10
ПР6.1-Ф19	Лабораторна робота 3 Спеціальні функції математичної статистики, інтервальні оцінки параметрів випадкових величин і тести для перевірки статистичних гіпотез	10
ПР6.1-Ф19	Лабораторна робота 4 Кореляційний аналіз	10
ПР13.1-Ф19	Лабораторна робота 5 Регресійний аналіз	11
ПР13.2-Ф19 ПР14.1-Ф19	Лабораторна робота 6 Кластерний аналіз	11
ПР12.1-Ф19	Лабораторна робота 7 Знайомство з пісочницею Apache Hadoop	11
	КОНТРОЛЬНІ ЗАХОДИ	4
	РАЗОМ	135

6. ОЦІНЮВАННЯ РЕЗУЛЬТАТІВ НАВЧАННЯ

6.1. Шкали оцінювання

Навчальні досягнення здобувачів вищої освіти за результатами вивчення курсу оцінюватимуться за шкалою, що наведена нижче:

Рейтингова шкала	Інституційна шкала
90 – 100	відмінно
75 – 89	добре
60 – 74	задовільно
0 – 59	незадовільно

6.2. Підсумкова оцінка

Здобувачі вищої освіти можуть отримати підсумкову оцінку з навчальної дисципліни на підставі поточного оцінювання знань за умови, якщо набрана кількість балів складатиме не менше як 60 балів.

Максимальне оцінювання становить:

Теоретична частина	Лабораторні роботи	Самостійна робота з підготовки до лабораторних робіт	Разом
51	42	7	100

Теоретична частина оцінюється за результатами online тесту, який містить 27 запитань.

6.3. Критерії оцінювання теоретичної частини курсу

Для кожного запитання тесту потрібно вибрати єдину правильну відповідь. Кількість варіантів відповіді в залежності від характеру запитання може складати від 5-ти до 10-ти. Кількість балів, які нараховуються за правильну відповідь залежності від складності запитання може складати від 1-го до 4-х. У разі неправильної відповіді на запитання, здобувач вищої освіти отримує за нього 0 балів. Максимальна кількість балів, яку можна набрати за виконання тесту становить 51.

6.4. Критерії оцінювання лабораторних робіт

За кожен лабораторну роботу здобувач вищої освіти може отримати до 6 балів (усього 42 бали), а саме:

6 балів: програма обробки даних правильно функціонує, супроводжується достатньою кількістю коментарів, звіт з роботи оформлений згідно з методичними рекомендаціями і містить повні, інформативні та обґрунтовані висновки.

4 бали: програма обробки даних правильно функціонує, але містить незначні помилки, які суттєво не впливають на отримувані результати, або кількість коментарів недостатня, або звіт з роботи оформлений із незначними відхиленнями від методичних рекомендацій або висновки не повністю задовольняють вимогам повноти, інформативності та обґрунтованості.

2 бали: програма обробки даних функціонує, але містить помилки, які суттєво впливають на отримувані результати, або коментарі відсутні, або звіт з роботи оформлений із суттєвими відхиленнями від методичних рекомендацій, або висновки не задовольняють вимогам повноти, інформативності та обґрунтованості.

0 балів: програма обробки даних не функціонує, або звіт з роботи відсутній, або не містить висновків, або висновки протирічають фактично отриманим результатам.

У разі отримання позитивної оцінки здобувач вищої освіти може підвищити її шляхом очної співбесіди з викладачем. При цьому за кожен правильну відповідь на запропоновані запитання нараховується 1 бал.

6.5. Критерії оцінювання самостійної роботи

За підготовку вхідних даних згідно з варіантом завдання до кожної лабораторної роботи здобувач вищої освіти отримує:

1 бал (разом до 7 балів).

0 балів: підготовку вхідних даних не виконано, дані розраховані невірно або не відповідають варіанту завдання.

7. ІНСТРУМЕНТИ, ОБЛАДНАННЯ ТА ПРОГРАМНЕ ЗАБЕЗПЕЧЕННЯ

Використовуються лабораторна та інструментальна бази випускової кафедри інформаційних технологій та комп'ютерної інженерії, а також комп'ютерне та мультимедійне обладнання:

1. Персональний комп'ютер або ноутбук зі сталим доступом до мережі Інтернет
2. Активованій акаунт університетської пошти (student.i.p.@nmu.one) на Офіс365.
3. Активний обліковий запис у системі дистанційної освіти Moodle.
4. Програмне забезпечення:
 - 4.1. Операційна система Windows 10 або Windows 11.
 - 4.2. Office 365.
 - 4.3. Вільно розповсюджуване середовище математичних розрахунків Scilab.
 - 4.4. Середовище програмування Python на базі оболонки Anaconda.
 - 4.5. Віртуальні машини Oracle VM VirtualBox та HDP Sandbox.

8. РЕКОМЕНДОВАНІ ДЖЕРЕЛА ІНФОРМАЦІЇ

Основні:

1. Zgurovsky M.Z. Big Data: Conceptual Analysis and Applications. [Текст] / M.Z. Zgurovsky, Y.P. Zaychenko // Springer, 2021, 298 p.
2. Кожевніков А.В. Методи та інформаційні технології обробки великих даних (Big Data) [Електронний ресурс] : методичні рекомендації до виконання лабораторних робіт для здобувачів ступеня бакалавра освітньо-професійної програми «Інформаційні системи та технології» зі спеціальності 126 Інформаційні системи та технології / А. В. Кожевніков ; М-во освіти і науки України, Нац. техн. ун-т «Дніпровська політехніка». – Дніпро : НТУ «ДП», 2025. – 89 с.
3. Akerkar R. Models of Computation for Big Data [Текст] / R. Akerkar // Springer, 2018, 110 p.
4. Ghavami P. Big Data Governance: Modern Data Management Principles for Hadoop, NoSQL & Big Data Analytics [Текст] / P. Ghavami // CreateSpace Independent Publishing Platform, 2016, 204 p.
5. Feeney K. Engineering Agile Big-Data Systems [Текст] / K. Feeney, J. Davies, J. Welch, S. Hellmann, C. Dirschl, A. Koller, P. Francois, A. Marciniak // River Publishers, 2018, 436 p.
6. Big Data Fundamentals courses [Електрон. ресурс]. Режим доступа: <https://cognitiveclass.ai/learn/big-data> (дата звернення: 20.08.2024).
7. Big Data Analytics [Електрон. ресурс]. Режим доступа: <https://cognitiveclass.ai/learn/analytics/> (дата звернення: 20.08.2024).

Додаткові:

1. Aziukovskyi O., Udovyk I., Kozhevnykov A., Powroźnik T. Creating using the mathcad system of laboratory experimentation on the subject «intelligent data analysis» [Текст] / Матеріали XVI міжнародної конференції "Проблеми використання інформаційних технологій у сфері освіти, науки та промисловості" – Дніпро: НТУ «Дніпровська політехніка», 2021. – С. 21–26.
2. Кожевников А.В., Удовик І.М. Створення відкритої нейронної мережі бінарного класифікатора засобами системи Mathcad [Текст] / Матеріали XVI міжнародної конференції "Проблеми використання інформаційних технологій у сфері освіти, науки та промисловості" – Дніпро: НТУ «Дніпровська політехніка», 2021. – С. 27–32.
3. Кожевников А.В., Удовик І.М. Створення відкритої нейронної мережі предиктора лінійного часового ряду засобами системи Mathcad [Текст] / Матеріали XVI міжнародної конференції "Проблеми використання інформаційних технологій у сфері освіти, науки та промисловості" – Дніпро: НТУ «Дніпровська політехніка», 2021. – С. 45–49.
4. UCI Machine Learning Repository [Електрон. ресурс]. Режим доступу: <http://archive.ics.uci.edu/ml/index.php/> (дата звернення: 21.10.2024).
5. HDFS Commands Guide – Apache Hadoop 3.4.1 [Електрон. ресурс]. Режим доступу: <https://hadoop.apache.org/docs/r3.4.1/hadoop-project-dist/hadoop-hdfs/HDFSCommands.html> (дата звернення: 21.10.2024).

Навчальне видання

РОБОЧА ПРОГРАМА НАВЧАЛЬНОЇ ДИСЦИПЛІНИ
“Методи та інформаційні технології обробки великих даних (Big Data)”
для бакалаврів спеціальності 126 “Інформаційні технології”

Розробник: доц. Кожевников А.В.

В редакції автора

Підготовлено до виходу в світ
у Національному технічному університеті
“Дніпровська політехніка”.
Свідоцтво про внесення до Державного реєстру ДК № 1842
49005, м. Дніпро, просп. Д. Яворницького, 19