

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  
року 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 кредити ECTS (120 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”, 2024. – 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 студентів 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 методичної комісії спеціальності 126 “Інформаційні системи та технології” (протокол № 8 від 29.08.2024).

## 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
7 ІНСТРУМЕНТИ, ОБЛАДНАННЯ ТА ПРОГРАМНЕ ЗАБЕЗПЕЧЕННЯ..	8
8 РЕКОМЕНДОВАНІ ДЖЕРЕЛА ІНФОРМАЦІЇ .....	8

## 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 Ф20 “Методи та інформаційні технології обробки великих даних (Big Data)” віднесено такі результати 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 РН	Дисциплінарні результати навчання (ДРН)	
	шифр ДРН	зміст
ПР4	ПР4.1-Ф20	Проводити аналіз об’єктів, які оперують з великими даними, та обґрунтовувати вибір алгоритмів обробки даних
ПР6	ПР6.1-Ф20	Демонструвати знання методів та інформаційних технологій обробки великих даних
ПР9	ПР9.1-Ф20	На основі даних аналізу ІТ-інфраструктури підприємства, запропонувати її елементи, які удосконалюють роботу з великими даними
ПР12	ПР12.1-Ф20	Використовувати сучасні методи розробки програмного забезпечення для обробки великих даних
ПР13	ПР13.1-Ф20	Застосовувати методи регресійного аналізу для рішення задач прогнозування
	ПР13.2-Ф20	Використовувати методи кластеризації та класифікації для обробки великих даних

Шифр РН	Дисциплінарні результати навчання (ДРН)	
	шифр ДРН	зміст
ПР14	ПР14.1-Ф20	Розробляти та використовувати спеціалізоване програмне забезпечення для обробки і візуалізації просторових та багатовимірних даних

### 3 БАЗОВІ ДИСЦИПЛІНИ

Базовими дисциплінами є дисципліни які вивчалися студентами та формують компетентності щодо здатності вивчення дисципліни Ф20 “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, Б3 Теорія ймовірностей та математична статистика Ф11 Комп’ютерна математика	Знати лінійну та векторну алгебру, теорію ймовірностей та математичну статистику в обсязі, необхідному для розробки та використання інформаційних систем.
Ф1 Програмування	Використовувати базові знання інформатики й навички програмування.
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 занять	Обсяг, години	Розподіл за формами навчання, години					
		денна		вечірня		заочна	
		аудиторні заняття	самостійна робота	аудиторні заняття	самостійна робота	аудиторні заняття	самостійна робота
лекційні	58	28	30	-	-	6	52
практичні	-	-	-	-	-	-	-
лабораторні	58	28	30	-	-	6	52

контрольні заходи	4	4	-	-	-	4	-
<b>РАЗОМ</b>	<b>120</b>	<b>60</b>	<b>60</b>	-	-	<b>16</b>	<b>104</b>

## 5 ПРОГРАМА ДИСЦИПЛІНИ ЗА ВИДАМИ НАВЧАЛЬНИХ ЗАНЯТЬ

Шифри ДРН	Види та тематика навчальних занять	Обсяг складових, години
	<b>ЛЕКЦІЇ</b>	<b>58</b>
	<b>Вступ.</b> Мета і завдання дисципліни “ Методи та інформаційні технології обробки великих даних (Big Data)”. Сфери застосування Big Data	2
ПР6.1-Ф20	<b>Тема 1. Подання даних та їх попередня обробка</b> Визначення. Шкали виміру ознак. Життєвий цикл даних. Попередня обробка даних. Алгоритм ZET заповнення пробілів у таблицях даних. Метадані. Життєвий цикл метаданих	6
ПР4.1-Ф20, ПР9.1-Ф20, ПР13.1-Ф20	<b>Тема 2. Методи класифікації даних. Кластерний аналіз</b> Огляд методів класифікації даних. Завдання кластерного аналізу. Міри відстані та збіжності. Основні метрики. Основні методи кластерного аналізу: ієрархічні агломеративні та дивізімні методи, метод К-середніх Мак-Куїна. Нечітка кластеризація, метод С-середніх	8
	<b>Тема 3. Методи класифікації даних. Дерева рішень</b> Структура дерева рішень. Критерії обрання атрибута розбиття дерева. Критерії припинення розбиття. Відсікання гілок. Алгоритм ID3	8
ПР4.1-Ф20, ПР9.1-Ф20, ПР13.2-Ф20	<b>Тема 4. Методи класифікації даних. Регресійний аналіз</b> Багатовимірний лінійний регресійний аналіз. Відбір факторних змінних. Мультиколінеарність. Алгоритм Фаррара-Глобера. Аналіз та прогнозування часових рядів. Автокорреляція часових рядів. AR та MAпроцеси	8
	<b>Тема 5. Методи класифікації даних. Нейронні мережі</b> Елементи нейронної мережі. Функції активації нейрона. Архітектура мереж. Помилка мережі. Навчання мереж. Правила Хебба. Правило Відроу-Хоффа.	8
ПР6.1-Ф20, ПР12.1-Ф20, ПР14.1-Ф20	<b>Тема 6. Програмні інструменти Big Data. Мова Python</b> Статистичні функції мови Python. Аналіз даних засобами мови: регресійний, кластерний, факторний	8
	<b>Тема 7. Програмні інструменти Big Data. Платформа Hadoop</b> Архітектура та компоненти, кластери, їх вузли, файлова система HDFS. Доступ до сервісів Hadoop засобами мови Python	10
	<b>ЛАБОРАТОРНІ РОБОТИ</b>	<b>58</b>
ПР6.1-Ф20	<b>Лабораторна робота 1</b> Імовірнісні розподіли дискретних випадкових величин і точкові оцінки їх параметрів	8
ПР6.1-Ф20	<b>Лабораторна робота 2</b> Імовірнісні розподіли неперервних випадкових величин і точкові оцінки їх параметрів	8
ПР6.1-Ф20	<b>Лабораторна робота 3</b>	

<b>Шифри ДРН</b>	<b>Види та тематика навчальних занять</b>	<b>Обсяг складових, години</b>
	Спеціальні функції математичної статистики, інтервальні оцінки параметрів випадкових величин і тести для перевірки статистичних гіпотез	9
ПР6.1-Ф20	<b>Лабораторна робота 4</b> Кореляційний аналіз	8
ПР13.1-Ф20	<b>Лабораторна робота 5</b> Регресійний аналіз	8
ПР13.2-Ф20 ПР14.1-Ф20	<b>Лабораторна робота 6</b> Кластерний аналіз	9
ПР12.1-Ф20	<b>Лабораторна робота 7</b> Знайомство з пісочницею Apache Hadoop	8
	<b>КОНТРОЛЬНІ ЗАХОДИ</b>	<b>4</b>
	<b>РАЗОМ</b>	<b>120</b>

## **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**

Здобувач вищої освіти може отримати підсумкову оцінку з навчальної дисципліни на підставі поточного оцінювання знань за умови, якщо набрана кількість балів з поточного тестування та самостійної роботи складатиме не менше 60 балів.

Поточна успішність складається з оцінок за лекційну частину курсу та лабораторний практикум. Отримані бали додаються і є підсумковою оцінкою за вивчення навчальної дисципліни. Максимально за поточною успішністю здобувач вищої освіти може набрати 100 балів.

Максимальне оцінювання:

<b>Теоретична частина</b>	<b>Лабораторна частина</b>		<b>Разом при своєчасному складанні</b>
	<b>При своєчасному складанні</b>	<b>При несвоєчасному складанні</b>	
44	56	36	100

Лабораторні роботи приймаються за звітами і контрольними запитаннями до кожної з роботи.

Теоретична частина оцінюється за результатами складання тесту.

### **6.3. Критерії оцінювання теоретичної частини курсу**

Тест містить 27 запитань, які максимально оцінюються в 44 бали. Якщо робота виконується у дистанційному режимі, то видача online-тесту проводиться через систему Moodle у зазначеній викладачем час.

### **6.4. Критерії оцінювання лабораторних робіт**

З кожної лабораторної роботи здобувач вищої освіти отримує 4 запитання з переліку контрольних запитань. Відповідь на питання оцінюється максимально у 2 бали, причому:

- **2 бали** – відповідь вірна;
- **1 бал** – відповідь вірна, але не повна; відповідь вірна, але містить неточності та/або помилки;
- **0 балів** – відповідь невірна.

Максимальна оцінка за лабораторну роботу складає 8 балів. Максимальна оцінка за лабораторний практикум – 56 балів.

## **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 розповсюджене 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. Akerkar R. Models of Computation for Big Data [Текст] / R. Akerkar // Springer, 2018, 110 p.



3. 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.
4. 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.
5. Big Data Fundamentals courses [Електрон. ресурс]. Режим доступа: <https://cognitiveclass.ai/learn/big-data> (дата звернення: 20.08.2024).
6. Big Data Analytics [Електрон. ресурс]. Режим доступа: <https://cognitiveclass.ai/learn/analytics/> (дата звернення: 20.08.2024).

#### **Додаткові:**

1. UCI Machine Learning Repository [Електрон. ресурс]. Режим доступа: <http://archive.ics.uci.edu/ml/index.php/> (дата звернення: 20.08.2024).

Навчальне видання

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

Розробники: 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