МІНІСТЕРСТВО ОСВІТИ ТА НАУКИ УКРАЇНИ
НАЦІОНАЛЬНИЙ ТРАНСПОРТНИЙ УНІВЕРСИТЕТ
ФАКУЛЬТЕТ ТРАНСПОРТНИХ ТА ІНФОРМАЦІЙНИХ ТЕХНОЛОГІЙ
КАФЕДРА ІНФОРМАЦІЙНИХ СИСТЕМ І ТЕХНОЛОГІЙ

ЗАТВЕРДЖУЮ
Завідувач кафедри
інформаційних систем і
технологій
_____ проф. Гавриленко В.В.
«28» серпня 2023 року

РОБОЧА ПРОГРАМА НАВЧАЛЬНОЇ ДИСЦИПЛІНИ

ПП.2.6. ТЕХНОЛОГІЯ BIG DATA
(шифр і назва навчальної дисципліни)

галузі знань **12 «Інформаційні технології»**
(шифр і назва галузі знань)

спеціальність **121 «Інженерія програмного забезпечення»**
(шифр і назва спеціальності)

форма навчання **денна**
(денна, вечірня, заочна (дистанційна), екстернат)

факультет **Транспортних та інформаційних технологій**
(назва інституту, факультету, відділення)

Київ НТУ 2023

Робоча програма дисципліни «**Технологія Big Data**» для студентів за спеціальністю **121 «Інженерія програмного забезпечення»** розроблена на онові освітньої програми підготовки магістра з **інженерії програмного забезпечення.**

Розробники: пофесор кафедри ІСТ, к.ф.-м.н., доцент, Вітер М.Б.; старший викладач кафедри ІСТ, Огарков А.В.

Робочу програму схвалено на засіданні Ради факультету транспортних та інформаційних технологій

Протокол від «30» серпня 2023 року №1

Робочу програму схвалено на засіданні кафедри інформаційних систем і технологій

Протокол від «28» серпня 2023 року №1

Завідуючий кафедри
інформаційних систем
і технологій, професор _____ В.В.Гавриленко

© Вітер М.Б., 2023 рік
©Огарков А.В., 2023 рік
© НТУ, 2023 рік

1.Опис навчальної дисципліни

Найменування показників	Галузь знань, напрям підготовки, рівень вищої освіти, ступінь	Характеристика навчальної дисципліни
		денна форма навчання
Кількість кредитів – 3,00	Галузь знань 12 «Інформаційні технології»	нормативна
Модулів – 2	Спеціальність: 121 «Інженерія програмного забезпечення»	Рік підготовки
Змістовних модулів – 4		1-й
Індивідуальне комплексне завдання –		Семестр
Загальна кількість годин – 90		1-й
		Лекції
Тижневих годин для денної форми навчання: аудиторних – 3,00 самостійної роботи студента – 3,00	Другий (магістерський) рівень вищої освіти: ступінь – магістра	15 годин
		Практичні
		-
		Лабораторні
		30 годин
		Самостійна робота
		45 годин
		Індивідуальні завдання:
		-
		Вид контролю:
		залік

Примітка.

Співвідношення кількості годин аудиторних занять до самостійної роботи і індивідуальної роботи становить (%):
Для денної форми навчання – 100%

2.Мета та завдання навчальної дисципліни

Мета викладання дисципліни**:** полягає у набутті теоретичних і практичних знань з основ технологій що використовується для побудови архітектури Big Data.

Завдання вивчення дисципліни: навчити студентів базовим поняттям технологій, що використовуються у Big Data; набути практичних навичок у використанні відповідних інструментів в Big Data.

В освітній програмі підготовки магістра з інженерії програмного забезпечення студенти в результаті вивчення дисципліни «Технологія Big Data» набувають такі компетентності: ЗК01, ЗК03, ЗК05, СК01, СК02, СК04, СК05, СК10 та досягають наступні програмні результати: РН01-РН03, РН06, РН07, РН09, РН13, РН14, РН16, РН18.

Загальні компетентності:

ЗК01	Здатність до абстрактного мислення, аналізу та синтезу.

ЗК03	Здатність проводити дослідження на відповідному рівні.

ЗК05	Здатність генерувати нові ідеї (креативність).

Спеціальні (фахові, предметні) компетентності:

СК01	Здатність аналізувати предметні області, формувати, класифікувати вимоги до програмного забезпечення.

СК02	Здатність розробляти і реалізовувати наукові та/або прикладні проєкти у сфері інженерії програмного забезпечення.

СК04	Здатність розвивати і реалізовувати нові конкурентоспроможні ідеї в інженерії програмного забезпечення.

СК05	Здатність розробляти, аналізувати та застосовувати специфікації, стандарти, правила і рекомендації в сфері інженерії програмного забезпечення.

СК10	Здатність до аналізу і застосування сучасних інформаційних технологій при моделюванні транспортних процесів.

Професійні компетентності:

РН01	Знати і застосовувати сучасні професійні стандарти і інші нормативно-правові документи з інженерії програмного забезпечення

РН02	Оцінювати і вибирати ефективні методи і моделі розроблення, впровадження, супроводу програмного забезпечення та управління відповідними процесами на всіх етапах життєвого циклу.

РН03	Будувати і досліджувати моделі інформаційних процесів у прикладній області.

РН06	Розробляти і оцінювати стратегії проєктування програмних засобів; обґрунтовувати, аналізувати і оцінювати варіанти проектних рішень з точки зору якості кінцевого програмного продукту, ресурсних обмежень та інших факторів.

РН07	Аналізувати, оцінювати і застосовувати на системному рівні сучасні програмні та апаратні платформи для розв’язання складних задач інженерії програмного забезпечення.

РН09	Обґрунтовано вибирати парадигми і мови програмування для розроблення програмного забезпечення; застосовувати на практиці сучасні засоби розроблення програмного забезпечення.

РН13	Конфігурувати програмне забезпечення, керувати його змінами та розробленням програмної документації на всіх етапах життєвого циклу.

РН14	Прогнозувати розвиток програмних систем та інформаційних технологій.

РН16	Планувати, організовувати та здійснювати тестування, верифікацію та валідацію програмного забезпечення.

РН18	Оцінювати і вибирати ефективні методи сучасних інформаційних технологій при моделюванні транспортних процесів.

3. Програма навчальної дисципліни

Модуль 1. Теоретичне підґрунтя технологій великих даних

Змістовний модуль 1. Вступ та архітектура великих даних

Тема 1. Вступ до Big Data
1. Контекст використання технологій Big Data.
2. Зростання складності роботи з даними на прикладі тенденцій розвитку e-commerce.
3. Основні характеристики великих даних.
4. Огляд навичок та інструментів, необхідних для роботи з технологіями Big Data.

Тема 2. Архітектура великих даних
1. Принципи та властивості архітектури.
2. Схема та компоненти архітектури великих даних.
3. Різноманіття моделей баз даних.
4. Властивості баз даних та CAP теорема.

Змістовний модуль 2. Інфраструктура великих даних

Тема 3. Складові інфраструктури великих даних
1. Огляд варіантів розміщення інфраструктури.
2. Віртуалізація.
3. Контейнеризація.

Тема 4. Хмарні технології
1. Властивості та перевари хмарної інфраструктури.
2. Основні типи послуг хмарних обчислень.
3. Огляд типових продуктів найбільших IaaS провайдерів.

Модуль 2. Інструменти Big Data

Змістовний модуль 3. Налаштування роботи з різними типами даних

Тема 5. Використання Docker Engine для контейнеризації
1. Основні команди управління контейнерами.
2. Налаштування програмного підключення до контейнеру з SQL базою даних.
3. Створення власного образу системи на основі Dockerfile.

Тема 6. Робота з NoSQL базами даних
1. Огляд характеристик та приклади основних типів NoSQL баз даних.
2. Налаштування програмного підключення до контейнеру з NoSQL базою даних.

Змістовний модуль 4. Робота з реальними даними

Тема 7. Налаштування роботи з відкритими даними
1. Вступ до відкритих даних.
2. Огляд джерел та ініціатив з відкритих даних на рівні країн та міст.
3. Використання API для отримання даних з веб-платформ та сервісів.

Тема 8. Аналіз та візуалізація даних
1. Попередня обробка даних.
2. Методи та засоби аналізу даних.
3. Засоби візуалізації даних.

4. Структура навчальної дисципліни

При вивченні даної дисципліни студент має ознайомитися з програмою дисципліни, з її структурою, формами та методами навчання, видами та методами контролю знань.

Тематичний план навчальної дисципліни складається з двох модулів, що логічно пов’язують кілька навчальних елементів дисципліни за змістом та взаємозв’язками.

Навчальний процес здійснюється у таких формах: лекційні, лабораторні заняття та самостійна робота студента.

Структура залікового кредиту з дисципліни наведено у табл.

Назви змістових модулів і тем	**Кількість годин**					
	денна форма					
	усього	у тому числі				
		лекції	практичні	лабораторні	Індивідуальні завдання	Самостійна робота

1	*2*	*3*	*4*	*5*	*6*	*7*
Семестр 1.						
Модуль 1. Теоретичне підґрунтя технологій великих даних						
Змістовний модуль 1. Вступ та архітектура великих даних						
Тема 1. Вступ до Big Data	6	2	-	2	-	2
Тема 2. Архітектура великих даних	9	4	-	2	-	3
Разом за змістовним модулем 1	**15**	**6**	**-**	**4**	**-**	**5**
Змістовний модуль 2. Інфраструктура великих даних						
Тема 3. Огляд варіантів розміщення інфраструктури.	7	2	-	2	-	3
Тема 4. Хмарні технології	8	2	-	1	-	5
Разом за змістовним модулем 2	**15**	**4**	**-**	**3**	**-**	**8**
Підготовка до МК 1	**2**					**2**
Разом за модулем 1	**32**	**10**	**-**	**7**	**-**	**15**
Модуль 2. Інструменти Big Data						
Змістовний модуль 3. Налаштування роботи з різними типами даних						
Тема 5. Використання Docker Engine для контейнеризації	9	1	-	2	-	6
Тема 6. Робота з NoSQL базами даних	17	1	-	9	-	7
Разом за змістовним модулем 3	**26**	**2**	**-**	**11**	**-**	**13**
Змістовний модуль 4. Робота з реальними даними						
Тема 7. Налаштування роботи з відкритими даними	14	1	-	6	-	7
Тема 8. Аналіз та візуалізація даних.	12	2	-	6	-	4
Разом за змістовним модулем 4	**26**	**3**	**-**	**12**		**11**
Підготовка до МК 2	**2**					**2**
Разом за модулем 2	**54**	**5**	**-**	**23**	**-**	**26**
Підготовка до заліку	**4**					**4**
Усього годин за рік	**90**	**15**	**-**	**30**	**-**	**45**

<u>Примітка:</u>

Розподіл самостійної роботи міститься в розділі 8 робочої програми, яка включає в себе:

1. ***ПМК*** *– підготовка до модульного контролю;*
2. ***ПЛ*** *– підготовка до лабораторних занять;*
3. ***ППК*** *– підготовка до підсумкового контролю (заліку/екзамену);*
4. ***ІКЗ*** *– індивідуальне комплексне завдання або;*
5. ***ІНДЗ*** *– індивідуальне навчально-дослідне завдання (курсова робота);*

Самостійна робота включає в себе суму ПЛ (підготовка до лабораторних занять)+ПМК (підготовка до модульного контролю)+ППК (підготовка до підсумкового контролю (заліку/екзамену)).

5. Теми семінарських занять

Семінарські заняття навчальним планом дисципліни «Технологія Big Data» не передбачені.

6. Теми практичних занять

Практичні заняття навчальним планом дисципліни «Технологія Big Data» не передбачені.

7. Теми лабораторних занять

На кожному лабораторному занятті до виконання лабораторної роботи студент має відповісти на контрольні питання, які відображають його готовність до виконання лабораторної роботи, зокрема оволодіння необхідними теоретичними знаннями та усвідомлення мети роботи. По закінченні виконання лабораторної роботи викладач оцінює ступінь оволодіння навичками та досягнення мети даної роботи.

Для здачі лабораторної роботи студенту необхідно оформити індивідуальний звіт, у якому повинна бути: постановка завдання, роздруковані основні результати роботи, аналіз розрахунків та чіткі висновки.

Підсумкові оцінки за виконання кожної лабораторної роботи вносяться у відповідний журнал. Отримані студентом оцінки за лабораторні роботи враховуються при виставленні підсумкової оцінки з даної навчальної дисципліни.

№ з/п	Назва теми	Кількість годин
1	Налаштування Docker контейнера з SQL базою даних.	6
2	Робота з NoSQL базами даних з використанням Docker.	6
3	Створення та інтеграція інформаційної системи з SQL та NoSQL базами даних розміщеними в Docker.	6
4	Інтеграція з відкритими API.	6
5	Проведення аналізу та візуалізація отриманих результатів.	6
	Всього	**30**

8. Самостійна робота

Для опанування матеріалу дисципліни «Технологія Big Data» окрім лекційних і лабораторних занять, тобто аудиторної роботи, значну увагу необхідно приділяти самостійній роботі.

Основні види самостійної роботи студента:

1. Вивчення додаткової літератури та відкритих джерел інформації.
2. Підготовка до лабораторних занять.
3. Підготовка до проміжного та підсумкового контролю.

Розподіл годин самостійної роботи

1 семестр	
Всього годин - 45	
ПМК –підготовка до модульного контролю	**4** *(4 годин на семестр)*
ПЛ – підготовка до лабораторних занять	**16** *(до 2 годин на пару)*
ППК- підготовка до підсумкового контролю (заліку/екзамену)	**4**
ІКЗ – індивідуальне комплексне завдання **ІНДЗ -** Індивідуальне навчально-дослідне завдання (-)	**21**

Розподіл годин самостійної роботи за темами

№ з/п	Назва теми	Кількість годин	
		ПЛ	**ІКЗ**
1	Вступ до Big Data.	1	1
2	Архітектура великих даних.	1	2
3	Складові інфраструктури великих даних.	1	2
4	Хмарні технології.	1	4
5	Використання Docker Engine для контейнеризації.	4	2
6	Робота з NoSQL базами даних.	3	4
7	Налаштування роботи з відкритими даними.	3	4
8	Аналіз та візуалізація даних.	2	2
	Всього	**16**	**21**
	Підготовка до модульного контролю №1	2	-
	Підготовка до модульного контролю №2	2	-
	Підготовка до підсумкового контролю (заліку/екзамену)	4	-
	Всього за семестр	**24**	**21**

9. Індивідуальні завдання для самостійної роботи

Індивідуальні комплексні завдання навчальним планом дисципліни «Технологія Big Data» не передбачені.

10. Методи навчання

При вивченні курсу «Технологія Big Data» застосовуються 3 групи методів навчання:

– методи організації і здійснення навчально-пізнавальної діяльності;
– методи стимулювання і мотивації навчально-пізнавальної діяльності;
– методи контролю і самоконтролю за ефективністю навчально-пізнавальної діяльності.

Перша група охоплює вербальні методи передачі і сприймання навчальної інформації (розповідь, лекція); наочні (ілюстрація, презентація); практичні (вправи, групові та індивідуальні завдання). В межах самостійної роботи – робота з книгами, методичними матеріалами, Інтернет-джерелами, творчі завдання.

При вивченні курсу активно використовуються інтерактивні методи (при веденні лекцій та лабораторних занять) та проблемно-пошукові методи навчання (як при веденні аудиторних занять, так і при організації самостійної роботи студентів).

11. Методи контролю

Методи поточного контролю: поточне тестування, індивідуальне опитування, фронтальне опитування, перевірка домашніх завдань, перевірка індивідуальних завдань.

Методи модульного контролю: письмова контрольна робота, підсумкове тестування.

Методи підсумкового контролю: екзамен.

12. Розподіл балів, які отримують студенти

Вид оцінювання	**Модулі**				**Підсумковий контроль (екзамен)**	**Сума**
Кількість балів за модуль	**60**					
Змістовні модулі	**ЗМ 1**	**ЗМ 2**	**ЗМ 3**	**ЗМ 4**		
Кількість балів за змістовні модулі та	**10**	**10**	**10**	**10**	**40**	**100**

модульний контроль	**МК1 10**	**МК2 10**		

Шкала оцінювання: національна та ECTS

Сума балів за всі види навчальної діяльності	ОцінкаECTS	Оцінка за національною шкалою	
		для екзамену, курсового проекту (роботи), практики	для заліку
90 – 100	**A**	відмінно	зараховано
82-89	**B**	добре	
74-81	**C**		
64-73	**D**	задовільно	
60-63	**E**		
35-59	**FX**	незадовільно з можливістю повторного складання	не зараховано з можливістю повторного складання
0-34	**F**	незадовільно з обов'язковим повторним вивченням дисципліни	не зараховано з обов'язковим повторним вивченням дисципліни

Критерії оцінювання:

«Відмінно» - A ***(90-100 балів)***–виставляється студенту, який глибоко та міцно засвоїв матеріал, відмінно справляється з задачами та питаннями, показує знайомство з фаховою літературою, володіє різносторонніми навичками та прийомами виконання практичних завдань, вміє добре орієнтуватись у виробничих ситуаціях.

«Добре» - BC*(74-89 балів)*–виставляється студенту, який твердо знає програмний матеріал, правильно застосовує теоретичні знання при рішенні практичних завдань, володіє необхідними навичками та прийомами їх виконання.

«Задовільно» - DE*(64-73балів)*–виставляється студенту, який має знання тільки основного матеріалу, але не засвоїв його деталей, допускає неточності, неправильне тлумачення окремих елементів завдання та відчуває труднощі при виконанні практичних завдань.

«Незадовільно» - FX*(35-59 балів)* **-**виставляється студенту, який дає необґрунтовані відповіді на запитання, допускає суттєві помилки у використанні понятійного апарату. Не простежується логічність та послідовність думки. Формулювання хаотичні та не усвідомлені.

«Незадовільно» - F*(1-34 балів)* **-**виставляється студенту, який не засвоїв зміст дисципліни, вміння та навички не набуті.

13. Методичне забезпечення

1. Програма навчальної дисципліни.
2. Робоча програма навчальної дисципліни.
3. Курс лекцій з дисципліни.
4. Методичні вказівки до виконання лабораторних робіт.
5. Питання на залік/екзамен.

14. Рекомендована література

Базова

1. Zgurovsky M.Z., Zaychenko Y.P. Big Data: Conceptual Analysis and Applications. Springer, 2020. – 298 p.
2. Robert Slane. Big Data Essentials. — 2018. — 356p.
3. Rajkumar Buyya. Big Data. Principles and Paradigms. — Elsevier, 2016. – 496p.
4. Олійник А. О. Інтелектуальний аналіз даних : навчальний посібник. — Запоріжжя : ЗНТУ, 2012. — 278 с.

Додаткова література

1. Phil Simon. Too Big to Ignore: The Business Case for Big Data. — Wiley, 2019. — 257p.
2. Thomas Erl. Big Data Fundamentals. Concepts, Drivers & Techniques. — Prentice Hall, 2016. — 235p.

15. Інформаційні ресурси

1. IoT Fundamentals: Big Data & Analytics // Електронний ресурс. Режим доступу: https://www.netacad.com/courses/iot/big-data-analytics
2. Containerization // Електронний ресурс. Режим доступу: https://www.ibm.com/cloud/learn/containerization
3. Офіційна документація Docker // Електронний ресурс. Режим доступу: https://docs.docker.com
4. Connecting to a NoSQL Database with Spring Boot // Електронний ресурс. Режим доступу: https://www.baeldung.com/spring-boot-nosql-database
5. Kaggle // Електронний ресурс. Режим доступу: https://www.kaggle.com
6. Портал відкритих даних України // Електронний ресурс. Режим доступу: https://data.gov.ua