

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ

**ХАРКІВСЬКИЙ НАЦІОНАЛЬНИЙ
УНІВЕРСИТЕТ РАДІОЕЛЕКТРОНІКИ**

Кафедра Програмної інженерії



ЗАТВЕРДЖУЮ
Декан факультету КН

А.Л.Єрохін
(підпис, ініціали, прізвище)

15.09.2020 р.

РОБОЧА ПРОГРАМА НАВЧАЛЬНОЇ ДИСЦИПЛІНИ

МЕТОДИ ПРОГРАМНОЇ ІНЖЕНЕРІЇ У BIG DATA

рівень вищої освіти _____ магістерський _____
(бакалаврський, магістерський, освітньо-науковий)

спеціальність 121 Інженерія програмного забезпечення

освітньо-наукова програма - _____ Інженерія програмного забезпечення _____

освітньо-професійна програма - _____ Інженерія програмного забезпечення _____

Харків, 2020

Розробники: К.С. Смеляков професор кафедри ПІ, д.т.н., професор

(ініціали, прізвище, посада, науковий ступінь, вчене звання)

Робочу програму схвалено на засіданні кафедри Програмної інженерії

Протокол № 1 від «1» 09 2020 року

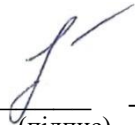
Завідувач кафедри


(підпис)

Дудар З.В.

(ініціали, прізвище)

Керівник групи забезпечення якості


(підпис)

Лєсна Н.С.

(ініціали, прізвище)

Схвалено методичною комісією факультету КН.

Протокол від «15» 09 2020 р. № 2

Голова методичної комісії


(підпис)

Лановий О.Ф.

(ініціали, прізвище)

1. ОПИС НАВЧАЛЬНОЇ ДИСЦИПЛІНИ

Найменування показників	Характеристика навчальної дисципліни	
	денна форма навчання	заочна форма навчання
Кількість кредитів ЄКТС – <u>5</u>	Професійна та практична підготовка (обов'язкова, норм. дисц. – ПЗ; дисц. сам. вибору НЗ - ПЗС)	
Модулів – 5	Рік підготовки	
Змістових модулів – 5	2019	
Індивідуальних завдання: РГЗ та КР	Семестр	
Загальна кількість годин – 150	1-й	
	Навчальні заняття:	
Тижневих годин для денної форми навчання: аудиторних – 2,4 самостійної роботи студента – 4,8 Тижневих годин для заочної форми навчання: аудиторних – 0,9 самостійної роботи студента – 6,3	1) лекції, год	
	30	
	2) практичні, год	
	4	
	3) лабораторні, год	
	16	
	4) консультації, год.	
	10	18
	Самостійна робота, год	
	90	132
в тому числі: 1) РГЗ та КР.		
-	-	
Вид контролю: комбінований іспит.		

2. МЕТА ДИСЦИПЛІНИ ТА ОЧІКУВАНІ РЕЗУЛЬТАТИ З ЇЇ ВИВЧЕННЯ

Метою дисципліни є визначення підходів і методології при розробці інноваційних Big Data додатків, які зорієнтовано на роботу із потоками даних, базами даних, вітринами даних, сховищами даних і озерами даних у реальному масштабі часу.

Визначення та аналіз основних моделей даних, методології обробки великих даних у реальному масштабі часу, а також інформаційних технологій обробки великих даних провідних світових виробників програмного забезпечення.

Виконання індивідуальних завдань практичної спрямованості для отримання відповідних навичок з дисципліни.

Наукова складова дисципліни – дослідження моделей і алгоритмів паралельної обробки даних.

Інформаційні технології: Python, Apache HTTP Server, OpenMP, MPI, CUDA и OpenCL.

Завдання дисципліни: за результатом вивчання дисципліни студенти мають володіти загальними / фаховими компетентностями та програмними результатами відповідно до змісту *освітньо-наукової програми Інженерія програмного забезпечення підготовки фахівців 2-го (магістерського) рівня вищої освіти спеціальності 121 Інженерія програмного забезпечення, та освітньо-професійної програми Програмне забезпечення систем підготовки фахівців 2-го (магістерського) рівня вищої освіти спеціальності 121 Інженерія програмного забезпечення, а саме:*

А) Загальними компетентностями:

1. Здатністю спілкуватися іноземною мовою для оцінювання джерел й отримання інформації з різноманітних джерел даних щодо створення Big Data програмного забезпечення (ПЗ).
2. Здатністю проведення теоретичних та прикладних досліджень щодо створення інноваційного програмного забезпечення за напрямом Big Data.
3. Здатністю удосконалювати свої навички на основі аналізу попереднього досвіду.
4. Здатністю генерувати нові ідеї відповідно до класифікації даних / джерел даних та форм їх використання щодо створення інноваційного ПЗ.

Б) Фаховими компетентностями:

1. Здатністю аналізувати предметні області, формувати, аналізувати та моделювати вимоги до інноваційного Big Data ПЗ.
2. Здатністю розвивати і реалізовувати нові конкурентоспроможні інноваційні ідеї та інноваційні моделі для створення Big Data стартапів в ІТ та в інженерії ПЗ.
3. Здатністю оцінювати ступінь обґрунтованості застосування специфікацій, стандартів, правил і рекомендацій в професійній галузі та дотримуватися їх при реалізації процесів життєвого циклу Big Data цифрових продуктів та послуг.
4. Здатністю ефективно обирати інструменти та ресурси при плануванні / проектуванні сучасного програмного забезпечення у Big Data додатках з метою забезпечення ефективності аналізу не тільки статичних, але й поточних даних.
5. Здатністю забезпечити виконання норм законодавства України відносно інтелектуальної власності та свідомо її застосовувати для захисту прав та економічних інтересів на інноваційне Big Data ПЗ.
6. Здатністю обґрунтовувати та вибирати критерії оцінки технічного стану основних й економічних показників при виробництві та експлуатації інноваційного Big Data ПЗ.

В) Програмні результати:

1. Знати і системно застосовувати методи аналізу та моделювання прикладної області, виявлення інформаційних потреб і збору вихідних даних для проектування інноваційного Big Data ПЗ.
2. Знати й вміти застосовувати сучасні технології розробки і тестування Big Data систем і сервісів.
3. Розробляти і оцінювати стратегії проектування інноваційного Big Data ПЗ; враховуючи обґрунтування вибору інструментів, сервісів та обладнання з точки зору якості кінцевого Big Data цифрового продукту.
4. Вміти приймати організаційно-управлінські рішення в умовах невизначеності для ефективності процесів та механізм розробки інноваційних Big Data проектів.
5. Набувати нові наукові і професійні знання, вдосконалювати навички, прогнозувати тенденції ринку ІТ стосовно потреб та інновацій у сфері Big Data рішень.
6. Формулювати, експериментально підтверджувати, обґрунтовувати і застосовувати на практиці в процесі розробки ПЗ конкурентоспроможні ідеї, методи, технології вирішення професійних, науково-технічних завдань в умовах невизначеності. А також в складних умовах функціонування Big Data сервісів.

3. ПРОГРАМА НАВЧАЛЬНОЇ ДИСЦИПЛІНИ

Змістовий модуль 1. Основні поняття та визначення.

Тема 1. Основні поняття та визначення Big Data.

Тема 2. Big Data у фактах та цифрах.

Тема 3. Особливості реалізації і імплементації Big Data проектів.

Змістовий модуль 2. Основні моделі і методи аналізу даних

Тема 4. Сховища даних, їх архітектура і особливості роботи з ними.

Тема 5. Отримання і інтеграція даних. Потоки даних.

Тема 6. Візуалізація і основні методи аналізу даних.

Змістовий модуль 3. Програмні додатки Apache Software Foundation

Тема 7. Додаток Apache Hadoop.

Тема 8. Додаток Hadoop MapReduce.

Тема 9. Додаток Apache Spark.

Змістовий модуль 4. Програмні додатки SAP

Тема 10. SAP Big Data рішення.

Тема 11. Додаток SAP HANA.

Тема 12. Додаток SAP Lumira.

Змістовий модуль 5. Прогнозування

Тема 13. Статистичні моделі.

Тема 14. Штучний інтелект.

Тема 15. Система координат Big Data & Data Science.

4. СТРУКТУРА НАВЧАЛЬНОЇ ДИСЦИПЛІНИ

Назви змістових модулів і тем	Кількість годин						Кількість годин					
	денна форма						заочна форма					
	усього	у тому числі					усього	у тому числі				
лк		пз	лаб.	конс	с. р.	конс		пз	лаб.	конс	с.р.	
1	2	3	4	5	6	7	8	9	10	11	12	13
Змістовий модуль 1.												
Тема 1. Основні поняття та визначення Big Data.	8	2				6	10	1				9
Тема 2. Big Data у фактах та цифрах.	10	2	2			6	10	1				9
Тема 3. Особливості реалізації і імплементації Big Data проєктів.	12	2	2		2	6	10	1				9
Разом за змістовим модулем 1	30	6	4			2	18	3				27
Змістовий модуль 2.												
Тема 4. Сховища даних, їх архітектура і особливості роботи з ними.	8	2				6	9	1				8
Тема 5. Отримання і інтеграція даних. Поток даних.	8	2				6	10	1				9
Тема 6. Візуалізація і основні методи аналізу даних.	14	2		4	2	6	11	2				9
Разом за змістовим модулем 2	30	6		4	2	18	30	4				26
Змістовий модуль 3.												
Тема 7. Додаток Apache Hadoop.	8	2				6	10	1				9
Тема 8. Додаток Hadoop MapReduce.	8	2				6	10	1				9
Тема 9. Додаток Apache Spark.	14	2		4	2	6	10	1				9
Разом за змістовим модулем 3	30	6		4	2	18	30	3				27
Змістовий модуль 4.												
Тема 10. SAP Big Data рішення.	8	2				6	9	1				8
Тема 11. Додаток SAP HANA.	8	2				6	10	1				9
Тема 12. Додаток SAP Lumira.	14	2		4	2	6	11	2				9
Разом за змістовим модулем 4	30	6		4	2	18	30	4				26
Змістовий модуль 5.												
Тема 13. Статистичні моделі.	8	2				6	9	1				8
Тема 14. Штучний інтелект.	8	2				6	10	1				9
Тема 15. Система координат Big Data & Data Science.	14	2		4	2	6	11	2				9
Разом за змістовим модулем 5	30	6		4	2	18	30	4				26
Усього годин	150	30	4	16	10	90	150	18				132

5. ТЕМИ ПРАКТИЧНИХ ЗАНЯТЬ

№ з/п	Теми занять	Обсяг, год.
1	Технології обробки даних. Частина 1.	2
2	Технології обробки даних. Частина 2.	2
	Загальна кількість	4

6. ТЕМИ ЛАБОРАТОРНИХ ЗАНЯТЬ

№ з/п	Теми занять	Обсяг, год.
1	Попередня обробка даних.	4
2	Сервіси Apache Software. Частина 1.	4
3	Сервіси Apache Software. Частина 2.	4
4	Сервіси SAP.	4
	Загальна кількість	16

7. САМОСТІЙНА РОБОТА

№ з/п	Вид самостійної роботи	Форма навчання	
		Денна обсяг, год	Заочна обсяг, год
1	Тема 1. Основні джерела Big Data.	6	9
2	Тема 2. Big Data у фактах та цифрах.	6	9
3	Тема 3. Засоби імплементації Big Data проектів.	6	9
4	Тема 4. Класифікація сховищ даних.	6	8
5	Тема 5. Засоби отримання статичних і потокових даних.	6	9
6	Тема 6. Засоби візуалізації багатовимірних даних.	6	9
7	Тема 7. Загальна характеристика Apache Software Foundation стосовно проектів Big Data.	6	9
8	Тема 8. Особливості використання Hadoop MapReduce.	6	9
9	Тема 9. Додатки Apache Software Foundation для роботи с базами даних.	6	9
10	Тема 10. Класифікація і характеристика програмних додатків SAP.	6	8
11	Тема 11. Інтеграція Big Data додатків на платформі SAP.	6	9
12	Тема 12. Візуалізація рішень у додатках SAP.	6	9
13	Тема 13. Класифікація методів прогнозування.	6	8
14	Тема 14. Методи прогнозування на основі штучного інтелекту.	6	9
15	Тема 15. Методи прогнозування у багатовимірних просторах.	6	9
	Загальна кількість	90	132

8. ІНДИВІДУАЛЬНІ ЗАВДАННЯ

8.1 Розрахунково-графічні завдання (РГЗ) та контрольні роботи (КР) – не передбачені планом.

8.2 Курсова робота – не передбачена планом.

9. МЕТОДИ НАВЧАННЯ ТА ЗАСОБИ ОЦІНЮВАННЯ

Методи навчання є системними об'єктами, які об'єднують низку взаємопов'язаних дій викладача й студентів, спрямованих на виконання освітньої, розвивальної, виховної і контрольної функцій.

Методи навчання за джерелами знань:

- словесні (лекція, консультація)
- наочні (ілюстрації, демонстраційні приклади).

Вивчення дисципліни здійснюється традиційними методами із застосуванням новітніх інформаційних технологій, слайд лекцій.

Теоретичні знання, що викладаються під час лекцій, використовуються на практичних і лабораторних роботах, що проводяться у комп'ютерних аудиторіях, які обладнані сучасними комп'ютерними засобами.

1. Методи організації та здійснення навчально-пізнавальної діяльності за допомогою слайд-лекцій, пояснень на багатьох практичних прикладів.
2. Методи стимулювання й мотивації навчально-пізнавальної діяльності студентів у виконанні власних проектів з практичної реалізації завдань дисципліни.
3. Методи контролю (самоконтролю, корекції (самокорекції)), за ефективністю навчально-пізнавальної діяльності студента ці методи спрямовані на самостійну, творчу пізнавальну діяльність студентів, особливо при створенні власних проектів.
4. Універсальні методи поєднують самостійну роботу студентів під час практичних занять з інструктуванням, допомогою викладача, у результаті чого студенти набувають навичок самостійності та самостійною роботою студентів поза аудиторного навантаження. Крім цього студент має у своєму розпорядженні слайд-лекції, приклади розв'язання задач з роз'ясненнями – усе це поєднується в наочно-ілюстративно-практичний комплект матеріалів для навчання.

10. МЕТОДИ КОНТРОЛЮ ТА РЕЙТИНГОВА ОЦІНКА ЗА ДИСЦИПЛІНОЮ

Методи контролю – це способи діагностичної діяльності, які дозволяють здійснювати зворотний зв'язок у процесі навчання з метою отримання даних про успішність навчання, ефективність навчального процесу. Вони мають забезпечувати систематичне, повне, точне і оперативне отримання інформації про навчальний процес.

Виділяють наступні **методи контролю**: методи *усного контролю*, методи *письмового контролю*, методи *практичного контролю*, *тести*.

Методи усного контролю – виступ перед аудиторією за означеною проблематикою.

Усне опитування допомагає контролювати не лише знання, а й вербальні вміння, сприяє виправленню мовних помилок. Відтворення студентом раніше вивченого матеріалу сприяє кращому запам'ятовуванню, активному використанню наукових понять, що неможливо без достатнього застосування їх у мові.

Усне опитування може бути індивідуальним і фронтальним. За фронтального опитування студенти відповідають з місця, доповнюючи один одного. Частковим випадком фронтального опитування є групове опитування – 5-6 осіб одночасно. Індивідуальне опитування здійснюється у процесі проведення співбесіди під час практичних занять.

Запитання для усної перевірки знань поділяють на основні, додаткові і допоміжні. Основні запитання передбачають самостійну розгорнуту відповідь (наприклад, запитання щодо змісту лабораторного заняття). Додаткові - спрямовані на уточнення того, як студент розуміє певне питання, формулювання, формулу та ін. Допоміжні запитання мають за мету виправлення помилок та неточностей, якщо такі мали місце у відповіді студента. Усі запитання - логічні, чіткі, зрозумілі, а їх сукупність – послідовна і системна.

Письмовий контроль (аудиторна контрольна робота у вигляді тесту або відповіді на запитання) забезпечує глибоку і всебічну перевірку засвоєння, оскільки вимагає комплексу знань і умінь студента.

Виконання практичних завдань, домашньої контрольної роботи. У письмовій роботі студенту необхідно показати і теоретичні знання, і вміння застосовувати їх для розв'язування конкретних задач, проблем, крім того, виявляється ступінь оволодіння письмовою мовою, уміння логічно, адекватно проблемі скласти свій текст, давати оцінку, експерименту, проблемі.

Письмовий контроль можна здійснюється у вигляді відповідей на запитання, розв'язання задач під час виконання практичних робіт. Письмові роботи допомагають за короткий час з'ясувати рівень засвоєння матеріалу у великій кількості студентів. Результати письмових робіт можна проаналізувати і з'ясувати деталі і неточності у відповідях та діагностувати їх причини.

Практичний контроль – передбачає виявлення вмінь і навичок студентів, що набуті під час практичної діяльності (практичних занять, робота над власним проектом). Така перевірка дає змогу виявити, на якому рівні студент засвоїв теоретичні основи цих дій.

Програмований контроль здійснюється за допомогою бланкових тестів або в системі Open Test.

Поточний контроль дозволяє отримати достатньо повні дані про студента: і рівень його знань, умінь з предмету, і ставлення до навчання, ступінь його пізнавальної активності, свідомості, і вміння мислити, розв'язувати самостійно різноманітні завдання.

Іспит є найбільш активною перевіркою знань за визначений період навчання.

10.1 Розподіл балів, які отримують студенти (Кількісні критерії оцінювання)

Min/ max рейтингова оцінка	Вид заняття / контрольний захід										
	ПЗ№ 1-2	АКР 1 / 1 КТ	ЛР№ 1	АКР 2 / 2 КТ	ЛР№ 2	АКР 3 / 3 КТ	ЛР№ 3	АКР 4 / 4 КТ	ЛР№ 4	АКР 4 / 5 КТ	Разом за 1-у і 2-гу КТ
Відпрацювання та захист Min	3	9	3	9	3	9	3	9	3	9	60
Відпрацювання та захист Max	5	15	5	15	5	15	5	15	5	15	100

10.2 Якісні критерії оцінювання

Необхідний обсяг знань для одержання позитивної оцінки.

- Знати основні поняття та визначення Big Data.
- Знати особливості реалізації і імплементації Big Data проектів.
- Вміти працювати зі сховищами даних.
- Вміти застосовувати інструменти для отримання і інтеграції даних; потоків даних.
- Вміти використовувати сучасні засоби візуалізації даних.
- Вміти використовувати основні програмні додатки Apache Software Foundation для обробки й аналізу Big Data (Apache Hadoop, Hadoop MapReduce, Apache Spark).
- Вміти використовувати основні програмні додатки SAP для обробки й аналізу Big Data (SAP HANA, SAP Lumira, та інші).
- Знати основні методи й програмні додатки прогнозування.
- Володіти статистичними методами прогнозування.
- Володіти методами прогнозування на основі використання моделей штучного інтелекту (перш за все – нейронні мережі, дерева рішень).
- Володіти методами прогнозування у багатовимірних просторах. Володіти методами й технологіями підготовки даних (фільтрація, очищення, заповнення пропусків), методом головних компонент, а також методами розбиття багатовимірного простору.

Необхідний обсяг умінь для одержання позитивної оцінки.

- Вміти застосовувати основні додатки Apache Software Foundation.
- Вміти застосовувати основні додатки SAP.
- Вміти використовувати основні моделі прогнозування.

10.3 Критерії оцінювання роботи студента протягом семестру.

Як форма підсумкового контролю для дисципліни використовується письмовий (комбінований) іспит. При цьому виді контролю підсумкова оцінка P_{Π} обчислюється за формулою: $P_{\Pi}=0.6 \cdot O_{\text{сем}}+0.4 \cdot O_{\text{ісп}}$, де $O_{\text{сем}}$ – оцінка за семестр у 100-бальній системі, $O_{\text{ісп}}$ – оцінка за іспит у 100-бальній системі.

Задовільно, D, E (60-74). Відпрацювати та захистити усі лабораторні роботи, АКР. Знати основні поняття та визначення Big Data; основні технології збору й обробки Big Data; обробку статичних і потокових даних; основні технології Apache Software Foundation та SAP.

Добре, C (75-89). Відпрацювати та захистити усі лабораторні роботи, АКР. Знати основні поняття та визначення Big Data; основні технології збору й обробки Big Data; обробку статичних і потокових даних; основні технології Apache Software Foundation та SAP. Крім того, знати й уміти застосовувати інструменти Predictive Analytics зазначених виробників ПЗ.

Відмінно, A, B (90-100). Відпрацювати та захистити усі лабораторні роботи, АКР. Знати основні поняття та визначення Big Data; основні технології збору й обробки Big Data; обробку статичних і потокових даних; основні технології Apache Software Foundation та SAP. Крім того, знати й уміти застосовувати інструменти Predictive Analytics зазначених виробників ПЗ. Досконально знати прогностичні моделі, оцінки ефективності їх застосування, оцінки похибки.

Критерії оцінювання знань та вмінь студента на комбінованому іспиті.

Задовільно, D, E (60-74). Висвітлити необхідний мінімум теоретичних знань. Знати методи розв'язання практичної задачі.

Добре, C (75-89). Показати добрі теоретичні знання матеріалу. Розв'язати практичне завдання.

Відмінно, A, B (90-100). Показати бездоганні знання теоретичного матеріалу та додаткового матеріалу. Розв'язати практичне завдання без помилок та вміти обґрунтувати своє рішення.

Білет для письмового (комбінованого) іспиту складається з двох теоретичних запитань та практичного завдання. Теоретичні запитання оцінюються за 100-бальною шкалою в 25 балів кожне, а завдання в 50 балів.

Розподіл балів, які отримують студенти.

Шкала оцінювання: національна та ЄКТС

Сума балів за всі види навчальної діяльності	Оцінка ЄКТС	Оцінка за національною шкалою	
		для іспиту, курсового проекту (роботи), практики	для заліку
96-100	A	відмінно добре задовільно	зараховано
90-95	B		
75-89	C		
66-74	D		
60-65	E	незадовільно з можливістю повторного складання	не зараховано з можливістю повторного складання
35-59	FX		
0-34	F	незадовільно з обов'язковим повторним вивченням дисципліни	не зараховано з обов'язковим повторним вивченням дисципліни

11. МЕТОДИЧНЕ ЗАБЕЗПЕЧЕННЯ ТА РЕКОМЕНДОВАНА ЛІТЕРАТУРА

11.1. Базова

1. Брендан Тірні, Джон Келлехер Наука про дані. - Видавництво Альпіна Діджітал, 2020. - 175с.

2. Віктор Майєр-Шенбергер, Кеннет Кук'єр Великі дані. Революція, яка змінить те, як ми живемо, працюємо та мислимо. - Видавництво Манн, Іванов та Фербер, 2013. - 240с.
3. Конспект лекцій з дисципліни.
4. Андреас Вайгенд BIG DATA. Уся технологія в одній книзі. - Видавництво Ексмо, 2018. - 414с.
5. <https://open.sap.com> (Driving Business Results with Big Data)
6. <https://hadoop.apache.org/>

12. МАТЕРІАЛЬНО-ТЕХНІЧНЕ, ІНФОРМАЦІЙНЕ ЗАБЕЗПЕЧЕННЯ

Інформаційні технології: Python, Apache HTTP Server, OpenMP, MPI, CUDA и OpenCL.