



Силабус навчальної дисципліни
«Використання програмних засобів для
обробки великих обсягів даних (BigData)»
Галузь знань: 12 Інформаційні технології
Спеціальність: 123 Комп'ютерна інженерія
Освітньо-професійна програма: «Комп'ютерна
інженерія»

Освітньо-професійний ступінь	фаховий молодший бакалавр
Статус дисципліни	вибіркова навчальна дисципліна
Форма навчання	Очна (денна)
Обсяг дисципліни, кредити ЕКТС/ загальна кількість годин	3 кредити / 90 годин
Мова викладання	Українська
Анотація дисципліни	Дисципліна «Використання програмних засобів для обробки великих обсягів даних (Big Data)» знайомить студентів із сучасними методами та технологіями роботи з великими даними. У рамках курсу розглядаються основи зберігання, обробки та аналізу даних, використовуються популярні інструменти та платформи, зокрема Hadoop, Spark, NoSQL, Apache Kafka, Power BI та MS Excel. Студенти навчатися застосовувати алгоритми машинного навчання для аналізу великих масивів даних, працювати з розподіленими обчисленнями та оптимізувати процеси обробки інформації. Особлива увага приділяється візуалізації та аналітиці даних у Power BI та MS Excel, що дозволяє ефективно працювати з бізнес-аналітикою та приймати обґрунтовані рішення. Дисципліна орієнтована на набуття практичних навичок, необхідних для ефективної роботи з Big Data у реальних проєктах. Курс буде корисним для студентів, які цікавляться аналізом даних, штучним інтелектом, Data Science та інформаційними технологіями.
Що буде вивчатися (предмет навчання)	Основні теми курсу Вступ до Big Data -Основні поняття та принципи роботи з великими даними - Виклики та можливості Big Data - Сучасні тенденції та сфери застосування Технології зберігання та обробки великих даних -Реляційні та нереляційні бази даних (SQL vs NoSQL) - Hadoop: архітектура, HDFS, MapReduce -Apache Spark: основи, моделі обробки, Spark SQL Інструменти для обробки потокових даних -Apache Kafka: принципи роботи та застосування

	- Інші платформи потокової обробки (Flink, Storm) 4. MS Excel у роботі з великими даними - Обробка та аналіз великих наборів даних - Використання зведених таблиць, Power Query - Автоматизація аналізу даних (макроси, VBA) 5. Power BI: бізнес-аналітика та візуалізація даних - Основи роботи з Power BI: імпорт та підготовка даних - Створення дашбордів та аналітичних звітів - DAX-функції та моделювання даних 6. Машинне навчання та аналітика великих даних - Основи Data Science та машинного навчання - Використання Python/R для аналізу Big Data - Практичні кейси аналізу даних 7. Практичні кейси та застосування Big Data - Реальні проєкти в аналізі великих даних - Оптимізація процесів бізнес-аналітики - Тренди та перспективи розвитку Big Data
Чому це цікаво/потрібно вивчати (мета)	Дисципліна «Використання програмних засобів для обробки великих обсягів даних (Big Data)» спрямована на формування у студентів теоретичних знань і практичних навичок у роботі з великими масивами даних. У сучасному світі компанії генерують величезні обсяги інформації, які необхідно швидко обробляти, аналізувати та використовувати для ухвалення ефективних рішень. Тому володіння навичками роботи з Big Data є критично важливим для фахівців у галузі бізнес-аналітики, Data Science, IT та цифрових технологій. Цей курс допоможе студентам: ✓ Розібратися в принципах роботи з великими даними та їх значенні в сучасному бізнесі та науці. ✓ Опанувати ключові технології та інструменти для зберігання, обробки, аналізу та візуалізації даних (Hadoop, Spark, Apache Kafka, NoSQL, Power BI, MS Excel). ✓ Навчитися працювати з потоковими даними та використовувати алгоритми машинного навчання для аналізу Big Data. ✓ Отримати практичні навички створення інтерактивних аналітичних звітів і дашбордів для підтримки ухвалення рішень. ✓ Зрозуміти, як застосовувати Big Data у реальних проєктах у різних сферах: фінанси, медицина, маркетинг, логістика тощо.

	Курс орієнтований на студентів, які хочуть стати конкурентоспроможними фахівцями у сфері аналізу даних, а також тих, хто прагне отримати навички роботи з сучасними інструментами обробки інформації.
Чому можна навчитися (результати навчання)	Після завершення курсу «Використання програмних засобів для обробки великих обсягів даних (Big Data)» студенти набудуть таких знань, умінь і навичок: 1. Теоретичні знання: - Розуміння основних концепцій та технологій Big Data. - Знання архітектури та принципів роботи з розподіленими системами обробки даних (Hadoop, Spark). - Ознайомлення з потоковою обробкою даних (Apache Kafka). - Розуміння принципів роботи з реляційними (SQL) та нереляційними (NoSQL) базами даних. - Основи машинного навчання та його застосування для аналізу великих даних. 2. Практичні навички: - Використання MS Excel для обробки великих масивів даних, застосування зведених таблиць, Power Query, макросів. - Робота з Power BI: підключення до джерел даних, створення інтерактивних дашбордів, використання DAX-функцій. - Аналіз та обробка великих обсягів даних у Hadoop та Spark. - Робота з Apache Kafka для потокової обробки даних. - Використання SQL та NoSQL для роботи з базами даних. - Застосування Python/R для аналізу великих даних. 3. Аналітичні та проєктні компетенції: - Уміння працювати з великими наборами даних та знаходити в них корисні закономірності. - Розробка аналітичних звітів та інтерактивних візуалізацій для ухвалення управлінських рішень. - Оптимізація процесів обробки даних та підвищення продуктивності аналітичних систем. - Використання інструментів Big Data у бізнесі, фінансах,

	маркетингу, медицині та інших сферах. - Робота над реальними кейсами та застосування отриманих знань у практичних проєктах. Курс надасть студентам актуальні навички, які затребувані на ринку праці та відкривають можливості для роботи у сфері Data Science, аналітики, машинного навчання, ІТ та бізнес-інтелекту (BI).
Як можна користуватися набутими знаннями і уміннями (компетентності)	ЗК3. Здатність до абстрактного мислення, аналізу і синтезу. ЗК 4 Здатність застосовувати знання у практичних ситуаціях. ЗК7. Здатність працювати в команді. ЗК8. Здатність вчитися і оволодівати сучасними знаннями. ЗК9. Здатність використовувати інформаційні та комунікаційні технології. ЗК10. Здатність генерувати нові ідеї (креативність). СК3. Здатність вільно користуватись сучасними комп'ютерними та інформаційними технологіями, прикладними та спеціалізованими комп'ютерно-інтегрованими середовищами для розробки, впровадження та обслуговування апаратних та програмних засобів комп'ютерної інженерії. СК4. Здатність розробляти системне та прикладне програмне забезпечення засобів комп'ютерної інженерії з використанням ефективних алгоритмів, сучасних методів і мов програмування. СК5. Здатність забезпечувати захист інформації в комп'ютерних системах та мережах з метою реалізації встановленої політики інформаційної безпеки. СК10. Здатність аргументувати вибір методів розв'язування спеціалізованих задач, критично оцінювати отримані результати, обґрунтовувати прийняті рішення. СК13. Здатність ідентифікувати, класифікувати та описувати роботу програмно-технічних засобів, комп'ютерних систем, мереж та їх компонентів шляхом використання аналітичних методів і методів моделювання. СК14. Здатність оцінювати і враховувати економічні, соціальні, технологічні та екологічні чинники, що впливають на сферу професійної діяльності. РН8. Вміти застосовувати знання для формулювання і розв'язування технічних задач спеціальності, використовуючи методи, що є найбільш придатними для досягнення поставлених цілей. РН9. Вміти використовувати методи аналізу та синтезу при розробці апаратних та програмних засобів комп'ютерної інженерії. РН10. Вміти системно мислити та застосовувати творчі здібності до формування нових та нестандартних рішень при розв'язуванні задач комп'ютерної інженерії.
Навчальна логістика	Модуль 1. Вступ до Big Data та основні концепції Тема 1. Основні поняття та значення Big Data.

	Тем 2. Джерела та типи великих даних. Тема 3. Виклики та можливості обробки великих даних. Тема 4. Основні підходи та технології обробки (Hadoop, Spark, NoSQL). Тема 5. Основи роботи з Apache Spark: структура, обробка потокових даних.
Пререквізити	Модуль 5. Аналіз даних у MS Excel та Power BI Тема 1. Робота з великими наборами даних у Excel. Тема 2. Використання зведених таблиць та функцій для аналізу даних. Тема 3. Інструменти Power Query для обробки та трансформації даних. Тема 4. Автоматизація обробки даних (макроси, VBA). Тема 5. Моделювання даних, використання DAX-функцій. Для успішного засвоєння матеріалу курсу «Використання програмних засобів для обробки великих обсягів даних (Big Data)» бажано мати базові знання та навички у таких сферах: Основи інформатики та програмування ✓ Знання основ алгоритмізації та структур даних. ✓ Базові навички програмування (Python, R, або SQL бажані, але не обов'язкові). Бази даних та SQL ✓ Основи реляційних баз даних та знання SQL-запитів (SELECT, JOIN, GROUP BY, агрегатні функції). ✓ Бажано мати базове уявлення про NoSQL (MongoDB, Cassandra). Робота з MS Excel ✓ Основи роботи з таблицями, формулами та функціями. ✓ Досвід використання зведених таблиць буде перевагою. Математика та статистика ✓ Базові знання математичної статистики (середні, медіана, дисперсія, кореляція). ✓ Розуміння основ аналітики та ймовірності. Основи аналізу даних ✓ Вміння працювати з великими наборами даних та їх обробкою. ✓ Розуміння концепцій бізнес-аналітики та основ візуалізації даних. Курс розрахований на студентів, які вже мають базові технічні знання, але спеціальна підготовка з Big Data не є обов'язковою. Усі необхідні інструменти та концепції будуть пояснені протягом курсу.

Постреквізити	Після завершення курсу студент буде здатний розуміти основи концепції Big Data, зокрема різні типи даних, їх характеристики, такі як обсяг, швидкість та різноманітність, а також зможе визначати підходи до зберігання даних, працюючи з різними типами, зокрема структурованими, неструктурованими та напівструктурованими даними. Студент набуде навичок використання основних технологій для обробки великих даних, таких як Hadoop та Apache Spark, та зможе вибирати правильні інструменти для зберігання даних, наприклад, NoSQL бази даних, як MongoDB або Cassandra. Також студент зможе писати програми на таких мовах, як Python, R або Scala, для обробки великих обсягів даних, застосовуючи ці мови для реалізації паралельних обчислень у середовищах, таких як Apache Spark. Курс дозволить йому освоїти застосування алгоритмів машинного навчання до великих даних, таких як класифікація, регресія та кластеризація, а також дасть змогу оптимізувати ці алгоритми для роботи з великими обсягами даних. Студент навчиться проводити аналіз даних і візуалізувати результати, створюючи графіки, діаграми та дашборди, використовуючи такі інструменти, як Tableau або Power BI. Також він зможе застосовувати різні архітектури обробки великих даних, такі як Batch Processing та Stream Processing, і реалізовувати практичні проекти за допомогою інструментів для обробки поточкових даних, наприклад, Apache Kafka. У результаті курсу студент зможе інтегрувати інструменти Big Data в бізнес-процеси, розуміти, як використовувати ці технології для покращення рішень у таких сферах, як фінанси, маркетинг, медицина чи електронна комерція. Також він буде готовий працювати з хмарними платформами, такими як Google Cloud, AWS або Azure, для ефективної обробки та аналізу великих даних.	
Рекомендована література	1. Олещенко Л. М. Технології оброблення великих даних: конспект лекцій. Запоріжжя: ЗНУ, 2021. 162 с. URL: https://files.znu.edu.ua/files/Bibliobooks/Inshi72/0053281.pdf 2. Черняк О. І. Інтелектуальний аналіз даних: підручник. Київ: КНЕУ, 2017. 304 с. 3. Сидоренко В. В., Константинова Л. В., Смірнов С. А. Організація баз даних: навчальний посібник. Кропивницький: ЦНТУ, 2020. 210 с. URL: 4. Aurelien Géron – <i>Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow</i> (2022)	
Інформаційні ресурси	1. Coursera – Big Data Specialization Навчальний курс від Каліфорнійського університету в Сан-Дієго, що охоплює основи обробки великих даних. URL: https://www.coursera.org/specializations/big-data 2. Harvard University – Data Science: Big Data Безкоштовний курс із аналізу великих даних на платформі edX. URL: https://cs50.harvard.edu/x/	
Формат та обсяг курсу	Вид занять	Кількість годин
	Лекції	18
	Практичні	36
	Самостійна робота	66

Розподіл балів, форма контролю	Форми контролю		Максимальна кількість балів
	Модуль 1		25
	Модуль 2		25
	Залік		50
Циклова комісія	Прикладної математики та інформаційних технологій		
Відділення	Природниче відділення		
Шкала оцінювання, національна та ЄКТС	Сума балів за всі види навчальної діяльності	Оцінка ЄКТС	Оцінка за національною шкалою
	90-100	A	Зараховано
	80-89	B	
	70-79	C	
	60-69	D	
	50-59	E	
	35-49	FX	Не зараховано (з можливістю повторного складання)
	0-34	F	Не зараховано (з обов'язковим повторним вивченням дисципліни)
Викладач(і)	ЛУЦЮК ЮРІЙ ВІТАЛІЙОВИЧ Посада: викладач Категорія: вища кваліфікаційна категорія Педагогічне звання: викладач Науковий ступінь: – Вчене звання: – Профайл викладача (-ів): https://amit-college.chnu.edu.ua/tsyklova-komisiia/sklad-tsyklovoi-komisii/lustiuk-yurii-vitaliiovych/ E-mail: y.lustiuk@chnu.edu.ua		