МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ

Львівський національний університет імені Івана Франка

Факультет прикладної математики та інформатики

Кафедра кібербезпеки

Затверджено

На засіданні кафедри кібербезпеки
факультету прикладної математики
та інформатики
Львівського національного університету
імені Івана Франка
(Протокол № 15/23 від 29 серпня 2023 р.)

Завідувач кафедри .

Венгерський П.С.

Силабус з навчальної дисципліни

“Великі дані”,

що викладається в межах ОПП Кібербезпека першого (бакалаврського) рівня вищої освіти для здобувачів з спеціальності 125 – кібербезпека та захист інформації

Львів - 2023

Назва дисципліни	Великі дані
Адреса викладання дисципліни	м. Львів, вул. Університетська 1
Факультет та кафедра, за якою закріплена дисципліна	Факультет прикладної математики та інформатики Кафедра кібербезпеки
Галузь знань, шифр та назва спеціальності	12 – інформаційні технології 125 – кібербезпека та захист інформації
Викладачі дисципліни	Винокурова Олена Анатоліївна, професор кафедри кібербезпеки
Контактна інформація викладачів	olena.vynokurova@lnu.edu.ua; https://ami.lnu.edu.ua/employee/vynokurova Головний корпус ЛНУ ім. І. Франка, каб. 260. м. Львів, вул. Університетська, 1
Консультації з питань навчання по дисципліні відбуваються	Консультації в день проведення лекцій/практичних занять (а також за розкладом консультацій кафедри).
Сторінка курсу	https://ami.lnu.edu.ua/admission/specializations
Інформація про дисципліну	Дисципліна “Великі дані” є нормативною дис-ципліною з спеціальності 125 – кібербезпека та захист інформації для освітньої програми Кібербезпека, яка викладається в 6-му семестрі в обсязі 5-ох кредитів (за Європейською Кредитно-Трансферною Системою ECTS).
Коротка анотація дисципліни	Під час навчання студенти отримають теоретичні знання та практичні навички роботи з великими даними, ознайомляться з основними поняттями, методами та засобами обробки великих даних. Навчаться збирати, зберігати, оперувати великими даними, враховуючи їхні особливості.

Мета та цілі дисципліни	Метою викладання навчальної дисципліни є вивчення теоретичних основ та методів обробки великих даних.
Література для вивчення дисципліни	1. Aluru, S., Parthasarathy, S., & Prasanna, V. K. (Eds.). (2020). Big Data Algorithms, Analytics, and Applications. CRC Press. 2. Mayank Bhushan Big Data and Hadoop: Fundamentals, tools, and techniques for data-driven success, BPB Online, 2023 3. Zgurovsky M.Z., Zaychenko Y.P. Big Data: Conceptual Analysis and Applications. Springer, 2020. 4. Shitalkumar R. Sukhdeve, Sandika S. Sukhdeve Google Cloud Platform for Data Science: A Crash Course on Big Data, Machine Learning, and Data Analytics Services, Apress, 2023 5. Amir H. Gandomi; Fang Chen; Laith Abualigah Big Data Analytics Using Artificial Intelligence, MDPI, 2023 6. Jonah Carrio Andersson Learning Microsoft Azure: Cloud Computing and Development Fundamentals O'Reilly Media, 2023 7. T. Ananth Kumar, T. S. Arun Samuel, R. Dinesh Jackson Samuel, M. Niranjanamurthy Privacy and Security Challenges in Cloud Computing. CRC Press Cognitive Approaches in Cloud and Edge Computing, 2022 8. David Santana Cloud Computing Demystified for Aspiring Professionals: Hone your skills in AWS, Azure, and Google cloud computing and boost your career as a cloud engineer Packt Publishing, 2023 Допоміжна 9. https://cloud.google.com/ 10. https://aws.amazon.com/ 11. https://docs.aws.amazon.com/cli/latest/userguide/cli-services- https://azure.microsoft.com/en-us/products/analysis-services
Обсяг курсу	Загальний обсяг: 150 годин. Аудиторних занять: 64 год., з них 32 год. лекцій та 32 год. лабораторних робіт. Самостійної роботи: 86 год.
Очікувані результати навчання	У результаті вивчення навчальної дисципліни студент має набути таких компетентностей: **знати:** - методи та засоби агрегації та інтегрування великих даних різних типів; - методи та засоби проектування сховищ для зберігання великих даних; - методи, засоби та інструменти інтелектуальної обробки великих даних; - методи та засоби візуалізації великих даних; - методи та засоби отримання знання з великих даних.

	вміти: - використовувати інструментальні засоби інтеграції різнорізних великих даних; - правильно обирати та використовувати методи та алгоритми для фільтрації, валідації та зберігання великих даних; - правильно обирати найбільш інформативні способи візуалізації великих даних; - отримувати знання шляхом аналізу великих даних; - проєктувати та розробляти прикладні інформаційні програми для аналізу великих даних та прийняття рішень на основі отриманої інформації. **Курс забезпечує набуття таких компетентностей: ІК, ЗК 1, ЗК 2, ЗК 3, ЗК 5, ФК 2, ФК 3, ФК 9, ФК 11;** **та програмних результатів навчання: ПРН 2, ПРН 3, ПРН 6, ПРН 10, ПРН 11, ПРН 13, ПРН 15, ПРН 17, ПРН 18, ПРН 21, ПРН 24, ПРН 32, ПРН 33, ПРН 35, ПРН 49, ПРН 52.**
Ключові слова	великі дані, методи та засоби їх аналізу та обробки
Формат курсу	Очний Проведення лекцій, лабораторних робіт і консультацій.
Теми	Теми подані у Схемі курсу нижче
Підсумковий контроль, форма	Залік у кінці семестру
Навчальні методи та техніки, які будуть використовуватися під час викладання курсу	Презентації, лекції, лабораторні роботи, індивідуальні завдання, індивідуальні доповіді, опитування теоретичного матеріалу, контрольна робота (модуль) самостійна робота. Лекції та лабораторні: інформаційно-рецептивний метод, репродуктивний метод, евристичний метод, метод проблемного викладу. Самостійна робота: репродуктивний метод, дослідницький метод.
Необхідне обладнання	Комп'ютерний клас із вільно-доступним програмним забезпеченням, локальна комп'ютерна мережа, доступ до Internet мережі.

Критерії оцінювання (окремо для кожного виду навчальної діяльності)	Оцінювання проводиться за 100-бальною шкалою. Бали нараховуються за наступним співвідношенням: • модульний контроль, тестування, усне опитування: 50% семестрової оцінки; максимальна кількість балів 50 • лабораторні роботи: 50% семестрової оцінки; максимальна кількість балів 50 Підсумкова максимальна кількість балів 100. **Академічна доброчесність**: Очікується, що роботи студентів будуть їх оригінальними дослідженнями чи міркуваннями. Відсутність посилань на використані джерела, фабрикування джерел, списування, втручання в роботу інших студентів становлять, але не обмежують, приклади мож-ли-вої академічної недоброчесності. Виявлення ознак академічної недоброчесності в письмовій роботі студента є підставою для її незарахування викладачем, незалежно від масштабів плагіату чи обману. **Відвідання занять** є важливою складовою навчання. Очікується, що всі студенти відвідають усі лекції та практичні зайняття курсу. Студенти повинні інформувати викладача про неможливість відвідати заняття. У будь-якому випадку студенти зобов'язані дотримуватися термінів виз-начених для виконання всіх видів письмових робіт та індивідуальних завдань, передбачених курсом. П**олітика виставлення балів.** Враховуються бали набрані при поточному тестуванні, самостійній роботі та бали підсумкового тестування. При цьому обов'язково враховуються присутність на заняттях та активність студента під час практичного заняття; недопустимість пропусків та запізнень на заняття; користування мобільним телефоном, планшетом чи іншими мо-більними пристроями під час заняття в цілях не пов'язаних з навчанням; списування та плагіат; несвоєчасне виконання поставленого завдання і т. ін. Жодні форми порушення академічної доброчесності не толеруються.
Опитування	Анкету-оцінку з метою оцінювання якості курсу буде надано по завершенню курсу.

Схема курсу

Тиж.	Тема, план, короткі тези	Форма діяльності (заняття)	Література	Завдан-ня, год.	Термін виконання
1	**Тема 1.** Основні поняття обробки даних.	лекція, самостійна робота	[1-11]	2 6	1 тиждень
	Тема 1. Попередня обробка даних.	лаб	[1-11]	2	

2	**Тема 2.** Етапи процесу обробки великих даних.	лекція, самостійна робота	[1-11]	2 6	1 тиждень
	Тема 2. Попередня обробка даних.	лаб.	[1-11]	2	
3	**Тема 3.** Методи та засоби збирання та зберігання даних.	лекція, самостійна робота	[1-11]	2 6	1 тиждень
	Тема 3. Встановлення та налаштування Hadoop.	лаб	[1-11]	2	
4-5	**Тема 4.** Фреймворк Hadoop. Базові складові Hadoop. Додаткові компоненти Hadoop.	лекція, самостійна робота	[1-11]	4 8	2 тижні
	Тема 4. Виконання базових завдань в MapReduce.	лаб.	[1-11]	4	
6	**Тема 5.** Методи машинного навчання для обробки великих даних.	лекція, самостійна робота	[1-11]	2 6	1 тиждень
	Тема 5. Статистична обробка даних.	лаб.	[1-11]	2	
7	**Тема 6.** Методи класифікації та регресії для роботи з великими даними.	лекція, самостійна робота	[1-11]	2 6	1 тиждень
	Тема 6. Алгоритми класифікації та регресії в Mahout.	лаб.	[1-11]	2	
8	**Тема 7.** Методи кластеризації для роботи з великими даними.	лекція, самостійна робота	[1-11]	2 6	1 тиждень
	Тема 7. Алгоритми кластеризації в Mahout.	лаб.	[1-11]	2	
9	**Тема 8.** Глибинне навчання та його паралелізація.	лекція, самостійна робота	[1-11]	2 6	1 тиждень
	Тема 8. Глибинне навчання для великих даних.	лаб.	[1-11]	2	
10-11	**Тема 9.** Системи для аналізу великих даних в реальному часі.	лекція, самостійна робота	[1-11]	4 10	2 тижні
	Тема 9. Обробка великих даних в реальному часі.	лаб.	[1-11]	4	
12	**Тема 10.** Платформи для потокової обробки даних	лекція, самостійна робота	[1-11]	2 6	1 тиждень
	Тема 10. Обробка даних в фреймворку Spark.	лаб.	[1-11]	2	
13-14	**Тема 11.** Управління ресурсами в системах обробки великих даних.	лекція, самостійна робота	[1-11]	4 10	2 тижні

	Тема 11. Обробка даних в фреймворку Spark.	лаб.	[1-11]	4	
15-16	**Тема 12.** Методи та засоби візуалізації великих даних.	лекція, самостійна робота	[1-11]	4 10	2 тижні
	Тема 12. Візуалізація великих даних.	лаб.	[1-11]	4	
	Всього			**150**	