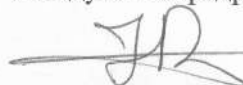


МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
Львівський національний університет імені Івана Франка
Факультет електроніки та комп'ютерних технологій
Кафедра радіофізики та комп'ютерних технологій

Затверджено

На засіданні кафедри радіофізики та
комп'ютерних технологій
факультету електроніки та комп'ютерних
технологій
Львівського національного університету
імені Івана Франка
(протокол № 15/23 від 29.03 2023 р.)

Завідувач кафедри:



Іван КАРБОВНИК

Силабус з навчальної дисципліни
“Аналіз даних”,
що викладається в межах ОПП “Комп'ютерні науки”
першого (бакалаврського) рівня вищої освіти для здобувачів з
спеціальності 122 – Комп'ютерні науки

Львів 2023 р.

Назва дисципліни	Аналіз даних
Адреса викладання дисципліни	Корпус факультету електроніки та комп'ютерних технологій, Львівський національний університет імені Івана Франка, вул. Ген.Тарнавського 107, м. Львів,79011
Факультет та кафедра, за якою закріплена дисципліна	Факультет електроніки та комп'ютерних технологій, кафедра радіофізики та комп'ютерних технологій
Галузь знань, шифр та назва спеціальності	12 Інформаційні технології, 122 Комп'ютерні науки
Викладачі дисципліни	Любунь Зіновій Михайлович, канд. тех. наук, доцент, доцент
Контактна інформація викладачів	zinoviy.lyubun@lnu.edu.ua
Консультації з питань навчання по дисципліні відбуваються	Консультації в день проведення лекційних занять (за попередньою домовленістю): ауд.312, корпус факультету електроніки та комп'ютерних технологій, вул. Ген.Тарнавського, 107 Також можливі он-лайн консультації через MS Teams. Для погодження часу он-лайн консультацій слід писати на електронну пошту викладача
Сторінка дисципліни	https://teams.microsoft.com/l/team/19%3a0-I021yR8IrGdNSleXZ_IYAQqlrvMWdIKA59ctWULPI1%40thread.tacv2/conversations?groupId=4b8739f0-33cb-47f2-b821-b9a502930564&tenantId=70a28522-969b-451f-bdb2-abfea3aaa5bf
Інформація про дисципліну	Дисципліна «Аналіз даних» є вибірковою дисципліною, яка викладається в межах ОПП «Комп'ютерні науки» першого (бакалаврського) рівня вищої освіти для здобувачів з спеціальності 122 «Комп'ютерні науки» в 6 семестрі, в обсязі 5 кредитів (за Європейською Кредитно-Трансферною Системою ECTS).
Коротка анотація дисципліни	Навчальну дисципліну розроблено таким чином, щоб надати учасникам необхідні знання, для реалізації основних алгоритмів аналізу даних, в тому числі, з використанням нейронних мереж.
Мета та цілі дисципліни	Мета викладання навчальної дисципліни “Аналіз даних ” полягає у одержанні студентами знань для вивчення та реалізації базових алгоритмів аналізу даних.
Література для вивчення дисципліни	Основна література: <ol style="list-style-type: none"> 1. І. А. Терейковський, Д. А. Бушуєв, Л. О. Терейковська. Штучні нейронні мережі: базові положення: навчальний посібник – Київ : КПІ ім. Ігоря Сікорського, 2022. – 123 с. 2. Processing Sensor Signal Under Low Values of Signal to Noise Ratio Liubun, Z., Bryk, B., Mandziy, V., Kalivoshka, B., Velhosh, S. 2023 IEEE 13th International Conference on Electronics and Information Technologies, ELIT 2023 - Proceedings, 2023, pp. 73–76. 3. Любунь З. М., Рабик В. Г, Карбовник І. Д. Інтелектуальний аналіз даних. Методичні вказівки до виконання лабораторних робіт для студентів за напрямом підготовки 6.050101

	<p>“Комп’ютерні науки”.– Львів: Видавничий центр ЛНУ імені Івана Франка, 2015. –70 с.</p> <p>4. Python for Data Analysis by Wes McKinney Copyright © 2013 Wes McKinney. All rights reserved. Printed in the United States of America. Published by O’Reilly Media, Inc., 1005 Gravenstein Highway North, Sebastopol, CA 95472, 2020.</p> <p>5. Гладун Ф.Я., Рогушина Ю.В. Data Mining: пошук знань в даних: підручник. Київ: ТОВ «ВД «АДЕФ-Україна», 2016. 452 с.</p> <p>6. І. А. Терейковський, Д. А. Бушуєв, Л. О. Терейковська. Штучні нейронні мережі: базові положення: навчальний посібник – Київ : КПІ ім. Ігоря Сікорського, 2022. – 123 с. Додаткова література.</p> <p>7. Liubun Z. Hover Signal-Profile Detection / Liubun, V. Mandziy, H. Klein, O. Karpin, V. Rabyk // Proceedings of the XV International Scientific and Technical Conference “Computer Science and Information Technologies” – 2020. P. 7 – 10.(Scopus)</p> <p>8. Karpin O. Method of Neural Network Training with Integer Weights / O. Karpin, Mandziy, Z. Liubun, V. Rabyk // Proceedings of the XIth International Scientific and Practical Conference “Electronics and Information Technologies” (ELIT – 2019), September 16 – 18, 2019, Lviv, Ukraine. P. 168 – 172. doi: 10.1109/ELIT.2019.8893349.</p> <p>9. Цемко А.,Любунь З. Створення та реалізація нейронних мереж для прогнозування ціни віртуальних активів. / А. Цемко, З. Любунь // Збірник наукових праць „Електроніка та інформаційні технології”. – 2022.</p> <p>10. Любунь З. М. Основи теорії нейромереж / З. М. Любунь /: Текст лекцій. – Львів: Видавничий центр ЛНУ імені Івана Франка, 2007. –142 с.</p>
Обсяг курсу	Загальний обсяг:150 годин. 64 години аудиторних занять. З них 32 години лекцій, 32 годин лабораторних занять та 86 годин самостійної роботи
Очікувані результати навчання	На основі отриманих знань студент повинен вміти провести аналіз отриманих даних в результаті моделювання чи реальних вимірювань згідно поставленої задачі використовуючи відповідні базові алгоритми аналізу Також повинен вміти провести візуалізацію отриманих результатів та зробити висновки.
Ключові слова	Статистичний аналіз даних. Кореляційний аналіз, Data Mining, нейромережеві методи аналізу даних. Аналіз даних. Машинне навчання та Штучний інтелект. Візуалізація даних. Бізнес-аналітика та прийняття рішень.
Формат курсу	Очний
Теми	Див. СХЕМА КУРСУ
Підсумковий контроль, форма	Залік в кінці семестру
Пререквізити	Для вивчення курсу студенти потребують базових знань з дисциплін «Вища математика», "Об’єктно – орієнтоване програмування", " Алгоритмізація і програмування ", «Мікропроцесорні системи».
Навчальні методи та техніки, які будуть використовуватис	Презентація, лекції, лабораторні заняття, обговорення, дискусія.

<p>я під час викладання курсу</p>	
<p>Необхідне обладнання</p>	<p>Для проведення лекційних занять: комп'ютер (мінімальні характеристики: процесор Intel Core i3(4 ядра/8 потоків), 8ГБ оперативної пам'яті, 50ГБ вільного місця на диску, відеокарта Nvidia GEFORCE GT1030 2048MB), засоби мультимедіа (в т.ч. проектор). Для проведення лабораторних занять: Комп'ютер (мінімальні характеристики: процесор Intel Core i3(4 ядра/8 потоків), 8ГБ оперативної пам'яті, 50ГБ вільного місця на диску, відеокарта Nvidia GEFORCE GT1030 2048MB). Мова програмування Python або аналогічні безкоштовні мови програмування. Мультимедіа, платформа MS Teams, комп'ютерний клас з встановленим ліцензійним програмним забезпеченням.</p>
<p>Критерії оцінювання (окремо для кожного виду навчальної діяльності)</p>	<p>Оцінювання проводиться упродовж семестру за 100-бальною шкалою. Бали нараховуються за такими видами робіт з наступним співвідношенням:</p> <ul style="list-style-type: none"> • лабораторні заняття: 72% семестрової оцінки; максимальна кількість балів 72. • контрольна робота: 28% семестрової оцінки; максимальна кількість балів 28 <p>Загалом упродовж семестру 100 балів.</p> <p>Оцінювання лабораторних робіт (12 лабораторних робіт, максимальна кількість балів: 72) відбувається шляхом оцінки роботи студента під час проведення лабораторної роботи в аудиторії (4 бали за одну роботу) та захисту звіту по виконаній лабораторній роботі (2 бали за одну роботу)</p> <p>Бали оцінювання лабораторних робіт нараховуються за наступним співвідношенням:</p> <p>6 – студент в повному обсязі володіє навчальним матеріалом, має повне розуміння розглянутої теми, надає правильні відповіді на запитання по темі, код програми функціонує відповідно до завдання;</p> <p>4 – студент не досить добре розуміє розглянутий матеріал та написаний ним код програми, вагається та надає неточні/не конкретні відповіді на запитання по темі, код програми функціонує неточно, або з помірними недоліками;</p> <p>1 – студент погано розуміє розглянутий матеріал та написаний ним код програми, студент в більшості надає помилкові відповіді на питання по темі, код програми функціонує з суттєвими недоліками;</p> <p>0 - студент зовсім не засвоїв розглянутий матеріал, написаний ним код програми не відповідає темі/не функціонує взагалі.</p> <p>Оцінювання підсумкової контрольної роботи :</p> <p>25-28 - розглянута тема відтворюється в повному обсязі, правильно, обґрунтовано, логічно, містить аналіз і систематизацію, аргументовані висновки. Засвідчено глибоке володіння матеріалом.</p> <p>20-25 - відтворюється значна частина розглянутої теми. Виявлено знання і розуміння основних положень навчальної дисципліни, проте присутні неточності та/або невідповідності основній темі.</p> <p>10-20 - відстежується загальне розуміння розглянутої теми. 10-5 – студент погано розуміє розглянуту тему. Виявлені суттєві неточності та невідповідності.</p> <p>0 – 5 – студент взагалі не розуміє розглянуту тему. Тему не розкрито, кількість викладеного матеріалу не відповідає загальним</p>

	<p>нормам обраного виду роботи.</p> <p>Критерії оцінювання результатів неформальної освіти: Нарахування балів відбувається за написання студентом тез доповідей на конференціях, наукових статей, участь у діяльності наукових гуртків, участь у наукових семінарах та круглих столах, конкурсах, участь у заходах неформальної освіти за отримання сертифікатів про проходження навчання на різних освітніх платформах (Coursera, Prometheus тощо), курсах на провідних ІТ компаніях за тематикою навчальної дисципліни. Кількість балів визначається відсотком покриття результатів відповідної активності до вимог результатів навчання з навчальної дисципліни.</p> <p>Контрольні заміри проводяться у формі опитування та захисту лабораторних завдань та написання контрольної роботи. Академічна доброчесність: Очікується, що лабораторні та контрольні 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 можливої академічної недоброчесності. Виявлення ознак академічної недоброчесності в роботі студента є підставою для її незарахування викладачем, незалежно від масштабів плагіату чи обману.</p> <p>Відвідання занять є важливою складовою навчання. Очікується, що всі студенти відвідають усі лекції і лабораторні заняття курсу. Студенти мають інформувати викладача про неможливість відвідати заняття. Студенти зобов'язані дотримуватися усіх термінів визначених для виконання усіх видів робіт, передбачених курсом.</p> <p>Література. Уся література, яку студенти не зможуть знайти самостійно, буде надана викладачем виключно в освітніх цілях без права її передачі третім особам. Студенти заохочуються до використання також й іншої літератури та джерел, яких немає серед рекомендованих.</p> <p>Політика виставлення балів. Враховуються бали набрані при виконанні лабораторних робіт та виконання контрольних робіт. При цьому обов'язково враховуються присутність на заняттях та активність студента під час лабораторного заняття; недопустимість пропусків та запізнь на заняття; користування мобільним телефоном, планшетом чи іншими мобільними пристроями під час заняття в цілях не пов'язаних з навчанням; списування та плагіат; несвоєчасне виконання поставленого завдання і т. ін.</p> <p>Жодні форми порушення академічної доброчесності не толеруються.</p>
<p>Питання до модульного контролю</p>	<p>Перелік питань та завдань для проведення підсумкової оцінки знань певних тем до контрольних робіт:</p> <ul style="list-style-type: none"> • Задачі аналізу даних • Data Mining • Збір даних та підготовка даних для аналізу • Методи первісної обробки даних • Методи згладжування та фільтрації даних. • Визначення числових характеристик статистичного розподілу одновимірної послідовності випадкових величин • Кореляційний і регресійний аналіз даних

	<ul style="list-style-type: none"> • Методи кластеризації даних • Прогнозування часових послідовностей (наївні моделі прогнозу) • Прогнозування часових послідовностей (Метод Хольта-Брауна) • Прогнозування часових послідовностей (лінійний та параболічний тренди) • Вирішення задачі ІАД за допомогою нейронних мереж • Генетичні алгоритми. • Дерева рішень. • Алгоритми пошуку асоціативних правил
Опитування	Анкету-оцінку з метою оцінювання якості курсу буде надано по завершенню курсу.

СХЕМА КУРСУ

Тиж.	Тема, план, короткі тези	Форма діяльності (заняття)	Література. Ресурси в Інтернеті	Лабораторні заняття, год.	Термін виконання
1	Задачі аналізу даних та етапи їх вирішення. Методи аналізу даних. Data Mining. Сфера застосування технологій інтелектуальних обчислень. Бізнес-застосування Data Mining . Популярні інструменти аналізу даних: Python (pandas, numpy), R, SQL, Excel. Лекція 1. 2 год.	Лекція	1, 3, 4,5	Лабораторна робота № 1. Реалізація елементарних алгоритмів аналізу даних. 2 год.	1-й тиж. семестру
2	Методи первісної обробки даних. Збір даних. Очистка даних (Data Cleaning). Згладжування та фільтрація даних. Нормування даних. Візуалізація даних (Exploratory Data Analysis, EDA). Лекція 2. 4 год.	Лекція	1, 2,4,5,9	Лабораторна робота № 2. Методи первісної обробки даних. 4 год.	2-й тиж. семестру
3	Статистичні методи аналізу даних. Обчислення математичних очікувань та дисперсій для одновимірних послідовностей випадкових величин. Кореляційний і регресійний аналіз даних. Лекція 3. 2 год.	Лекція	1, 3,5	Лабораторна робота № 3. Використання методу найменших квадратів для згладжування даних. 2 год.	3-й тиж. семестру
4	Кластерний аналіз. Ітераційні методи кластеризації. Ієрархічна	Лекція	1, 3, 4,5	Лабораторна робота № 4. Візуалізація даних. Автоматичне	5й тиж. семестру

	кластеризація. Лекція 4. 4 год.			групування даних. 2 год.	
5	Прогнозування часових послідовностей (наївні моделі прогнозу). Прогнозування часових послідовностей (Метод Хольта-Брауна) Прогнозування часових послідовностей (лінійний та параболічний тренди) Лекція 5. 2 год.	Лекція	3, 4, 7	Лабораторна робота № 5. Визначення числових характеристик статистичного розподілу одновимірної послідовності випадкових величин. 2 год.	6-й тиж. семестру
6	Асоціаційні правила. Послідовне відображення шаблонів даних. Метод Аргіогі, побудова FP-дерев пошуку шаблонів даних. Min-max асоціації у базах даних. Лекція 6. 2 год.	Лекція	2, 3, 4, 5	Лабораторна робота № 6. Кластеризація даних. 4 год.	7-й тиж. семестру
7	Аналітичні обчислення на основі відомих закономірностей. Методи апроксимації та екстраполяції залежностей заданих аналітично. Метод найменших квадратів з точки зору виявлення закономірностей. Загальні поняття про методи дослідження операцій. Гradientні методи знаходження екстремумів функцій заданих аналітично. Лекція 7. 2 год.	Лекція	2, 3, 6	Лабораторна робота № 7. Прогнозування часових послідовностей (наївні моделі прогнозу). 4 год.	8-й тиж. семестру
8	Технологія використання дерев прийняття рішень. Типи задач прийняття рішень, формальні означення дерева рішень і системи прийняття рішень. Приклади дерев рішень та їх використання для вирішення прикладних аналізу даних. Алгоритм. Методика проектування бази знань на основі використання дерев рішень із застосуванням алгоритму ID3. Лекція 8. 2 год.		3,7	Лабораторна робота № 8. Прогнозування часових послідовностей (Метод Хольта-Брауна). 2 год.	9-й тиж. семестру
9	Використання нейромереж для вирішення задач аналізу даних. Приклади	Лекція	1,2,8,9.	Лабораторна робота № 9. Прогнозування часових послідовностей	11 тиж. семестру

	використання нейромереж для вирішення прикладних задач аналізу даних. Лекція 9. 4 год.			(лінійний та параболічний тренди). 2 год.	
10	Виявлення закономірностей у багатомірному потоку даних за допомогою нейромереж. Навчання без вчителя. Пониження розмірності даних. Карти Кохонена. Кластеризація даних. Приклади використання нейромереж для вирішення прикладних задач аналізу даних. Лекція 11. 2 год.	Лекція	1,2,8,9.	Лабораторна робота № 10. Методи одно та багатовимірної оптимізації . 2 год.	12 тиж. семестру
11	Генетичні алгоритми і традиційні методи оптимізації Основні поняття генетичних алгоритмів. Класичний генетичний алгоритм. Модифікації класичного ГА. Приклади використання генетичних алгоритмів. Лекція 11. 4год.	Лекція	1, 2 5,6	Лабораторна робота № 10. Вирішення задачі класифікації за допомогою однеї нейронної мережі. 2 год.	14-й тиж. семестру
12	Нейромережеві методи аналізу даних в системах реального часу. Спайкові нейронні мережі. Лекція 12. 2 год	Лекція	7,8	Лабораторна робота № 12. Вирішення задачі кластеризації за допомогою нейронної мережі. 2 год.	15-й тиж. семестру
13	Нові напрямки та технології аналізу даних. Обробка природної мови. Глибинні нейронні мережі. Аналіз даних в реальному часі.		1,3	Підсумкова контрольна робота. 2 год.	16-й тиж. семестру