

Чернівецький національний університет імені Юрія Федьковича

(повне найменування закладу вищої освіти)

Факультет математики та інформатики

(назва факультету/навчально-наукового інституту)

Кафедра математичного моделювання

(назва кафедри, що забезпечує викладання)

“ЗАТВЕРДЖУЮ”

**Декан факультету
математики та інформатики**

Ольга МАРТИНЮК



2025 року

РОБОЧА ПРОГРАМА

навчальної дисципліни

Data Mining з використанням мови Python

(назва навчальної дисципліни)

обов'язкова

(вказати: обов'язкова)

Освітньо-професійна програма «Системний аналіз»

(назва програми)

Спеціальність 124 Системний аналіз

(вказати: код, назва)

Галузь знань 12 Інформаційні технології

(вказати: шифр, назва)

Рівень вищої освіти перший (бакалаврський)

(вказати: перший (бакалаврський) / другий (магістерський) / третій (освітньо-науковий))

Факультет математики та інформатики

(назва факультету/ навчально-наукового інституту, на якому здійснюється підготовка фахівців за вказаною освітньою програмою)

Мова навчання українська

(вказати: на якій мові читається дисципліна)

Чернівці 2025 рік

Робоча програма навчальної дисципліни «*Data Mining з використанням мови Python*» складена відповідно до освітньо-професійної програми «Системний аналіз»

Розробник:

Юрченко Ігор Валерійович, доцент кафедри математичного моделювання, кандидат фіз.-мат. наук, доцент

Викладач, що забезпечує читання даної навчальної дисципліни:

Юрченко Ігор Валерійович, доцент кафедри математичного моделювання, кандидат фіз.-мат. наук, доцент

Погоджено з гарантом ОП  Андрій ПЕРЦОВ

Затверджено на засіданні кафедри математичного моделювання
Протокол № 15 від «24» червня, 2025 року

Завідувач кафедри  Ігор ЧЕРЕВКО

Схвалено методичною радою факультету математики та інформатики
Протокол № 12 від «25» червня 2025 року

Голова методичної ради  Віра СІКОРА

Мета навчальної дисципліни: ознайомити студентів із сучасним станом обробки масивів даних, основними задачами DataMining, їх базовими компонентами, методами, що застосовуються при обробці даних, методами використання бібліотек мови Python для розв'язання задач галузі DataMining, навчити студентів розв'язувати прикладні задачі обробки та аналізу даних різної розмірності з використанням комп'ютерних технологій.

Навчальна дисципліна призначена для ознайомлення студентів з основними теоретичними та практичними аспектами обробки масивів даних, інтелектуального аналізу даних (DataMining) за допомогою широко розповсюджених бібліотек мови Python.

Пререквізити. Навчальні дисципліни: “Теорія ймовірностей та математична статистика”, “Бібліотеки мови Python”.

Результати навчання. У результаті вивчення навчальної дисципліни студент повинен

знати: предметну область, основні задачі галузі DataMining; методи та засоби обробки та аналізу даних з використанням бібліотек мови Python,

вміти: застосовувати бібліотеки мови Python для розв'язання прикладних задач обробки та аналізу даних різної розмірності.

Дисципліна формує такі **компетенції** у відповідності до стандарту вищої освіти спеціальності та освітньої програми:

ЗК02. Здатність застосовувати знання у практичних ситуаціях.

ЗК07. Здатність до пошуку, оброблення та аналізу інформації з різних джерел.

ФК1. Здатність використовувати системний аналіз як сучасну міждисциплінарну методологію, що базується на прикладних математичних методах та сучасних інформаційних технологіях і орієнтована на вирішення задач аналізу і синтезу технічних, економічних, соціальних, екологічних та інших складних систем.

ФК6. Здатність до комп'ютерної реалізації математичних моделей реальних систем і процесів; проектувати, застосовувати і супроводжувати програмні засоби моделювання, прийняття рішень, оптимізації, обробки інформації, інтелектуального аналізу даних.

ФК10. Здатність розробляти експериментальні та спостережувальні дослідження і аналізувати дані, отримані в них.

ФК12. Здатність до використання систем штучного інтелекту, алгоритмів машинного навчання для аналізу великих обсягів даних та виявлення закономірностей для прийняття рішень в умовах складних систем

Наведені результати навчання за відповідною дисципліною співвідносяться із такими **програмними результатами навчання:**

ПР12. Застосовувати методи і засоби роботи з даними і знаннями, методи математичного, логіко-семантичного, об'єктного та імітаційного моделювання, технології системного і статистичного аналізу.

ПР13. Проектувати, реалізовувати, тестувати, впроваджувати, супроводжувати, експлуатувати програмні засоби роботи з даними і знаннями в комп'ютерних системах і мережах.

ПР18. Знати та вміти розробляти та застосовувати алгоритми машинного навчання та інші методи штучного інтелекту для аналізу складних систем, в тому числі для прогнозування, оптимізації та прийняття рішень; розуміти етичні та соціальні вимоги до застосування штучного інтелекту.

Опис навчальної дисципліни Загальна інформація

Форма навчання	Рік підготовки	Семестр	Кількість		Кількість годин						Вид підсумкового контролю
			кредитів	годин	лекції	практичні	семінарські	лабораторні	самостійна робота	індивідуальні завдання	
Денна	4	7	3.0	90	30	–	–	30	30	–	залік

Структурний зміст навчальної дисципліни

Назви змістових модулів і тем	Кількість годин													
	денна форма							Заочна форма						
	усього	у тому числі					усього	у тому числі						
		л	п	лаб	інд	с.р.		л	п	лаб	інд	с.р.		
1	2	3	4	5	6	7	8	9	10	11	12	13		
Змістовий модуль 1. Основи аналізу та статистичної обробки даних засобами Python														
Тема 1. Математичне моделювання та аналіз даних програмними засобами Python.	13	4	–	4	–	5								
Тема 2. Методи та засоби візуалізації результатів аналізу даних програмними засобами Python.	17	6	–	6	–	5								
Тема 3. Методи та засоби	15	5	–	5	–	5								

статистичного аналізу даних у бібліотеці SciPy.													
Разом за змістовим модулем 1	45	15	–	15	–	15							
	Змістовий модуль 2. Методи інтелектуального аналізу даних та оцінювання моделей												
Тема 4. Лінійні моделі в задачах інтелектуального аналізу даних.	15	5	–	5	–	5							
Тема 5. Методи визначення якості математичних моделей.	15	5	–	5	–	5							
Тема 6. Ансамблеві математичні моделі.	15	5	–	5	–	5							
Разом за ЗМ 2	45	15	–	15	–	15							
Усього годин	90	30	–	30	–	30							

Тематика лекційних занять з переліком питань

№	Назва теми з основними питаннями
1	Математичне моделювання та аналіз даних програмними засобами Python (NumPy, Matplotlib). Навчальні набори даних для інтелектуального аналізу даних.
2	Методи та засоби візуалізації результатів аналізу даних програмними засобами Python (Matplotlib, Plotly). Експлораційний аналіз даних. Обробка та підготовка даних.
3	Методи та засоби статистичного аналізу даних у бібліотеці SciPy. Моделювання класифікації. Моделювання регресії. Кластерний аналіз. Використання алгоритмів асоціативного аналізу.
4	Використання бібліотек Scikit-learn, TensorFlow та PyTorch для створення та навчання нейронних мереж. Використання методів обробки текстів для аналізу та класифікації документів. Лінійні моделі в задачах інтелектуального аналізу даних.
5	Методи визначення якості математичних моделей.
6	Ансамблеві математичні моделі.

Матеріали до кожної лекції наведено на сайті електронного навчання на сторінці курсу <https://moodle.chnu.edu.ua/course/view.php?id=7576> та у роботі [2].

Тематика лабораторних занять з переліком питань

№	Назва теми (завдання)
1	Засоби реалізації математичного моделювання та аналізу даних DataMining (NumPy, SciPy, Pandas).
2	Методи та засоби візуалізації результатів аналізу даних (Matplotlib, Pandas, Plotly).
3	Застосування статистичних методів та засобів в задачах аналізу даних (NumPy, SciPy).
4	Використання лінійних моделей в задачах інтелектуального аналізу даних (Scikit-learn).
5	Застосування методів визначення якості моделей в задачах інтелектуального аналізу даних (Scikit-learn, Metrics).
6	Використання ансамблевих математичних моделей у задачах інтелектуального аналізу даних (Scikit-learn).

Матеріали до виконання лабораторних робіт наведено на сайті електронного навчання на сторінці курсу <https://moodle.chnu.edu.ua/course/view.php?id=7576> та у роботі [2].

Індивідуальні науково-дослідні завдання (ІНДЗ)

№	Завдання до тем
1	Студент може індивідуально виконувати додаткові завдання навчально-дослідницької спрямованості за завданнями, наданими викладачем.
2	Студенти можуть отримати до 10 балів в рахунок ІНДЗ, якщо самостійно зареєструються на безкоштовних курсах платформи Prometheus або платформи Coursera з DataMining (за попереднім узгодженням тематики курсів з викладачем), пройдуть навчання, отримають відповідний сертифікат і надішлють його на сайт дистанційного навчання викладачу разом з детальним звітом з практичних завдань пройденого курсу (постановки задач, коди виконаних програм, пояснення коду) та скріншотом успішності на курсі. Кількість балів буде виставлена пропорційно до навчальних результатів студента (згідно зі статистикою сайту Prometheus або Coursera).

* ІНДЗ – для змістового модуля, або в цілому для навчальної дисципліни визначається викладачем, з урахуванням специфіки дисципліни.

Завдання для самостійної роботи студентів

Самостійна робота складається з повторення матеріалу, засвоєного на лекціях, самостійного опанування частини теоретичного матеріалу, роботи з контрольними запитаннями та завданнями.

№	Назва теми	Завдання для самостійної роботи	К-сть год.
1	Теми 1-6	підготовка до лекційних занять	10
2	Теми 1-6	підготовка до лабораторних занять	15
3	Теми 1-6	підготовка до підсумкового модуль-контролю	5

Самостійна робота студентів використовується при вивченні наступних тем і передбачає опрацювання теоретичного матеріалу, результати якого застосовуються під час виконання лабораторних робіт.

Опис завдань самостійної роботи для лабораторних занять

До ЛР 1 (2 год.)

Завдання 1. Візуалізація та дослідження набору даних

Опис:

Виберіть публічно доступний табличний набір даних (наприклад, із Kaggle). Використовуючи бібліотеки NumPy, Matplotlib та Pandas, виконайте:

1. Зчитування, попередню обробку (наприклад, пропущені значення, типи змінних).
2. Візуалізації: розподіл змінних, кореляції, гістограми, діаграми розсіяння.
3. Напишіть короткий звіт (1–2 сторінки): що цікавого в даних, які закономірності побачили.

Рекомендовані набори даних:

- На Kaggle — «Find Open Datasets and Machine Learning Projects». ([Kaggle](#))
- Інший варіант: використовувати набір із репозиторію як приклад.

Поради:

- Оберіть набір із не дуже великою кількістю записів (до 10 000) — це полегшить роботу.
- Візуалізації оформляйте підписами, легендами.
- Звіт — короткий, акцент на ваших висновках, а не просто скрині графіків.

До ЛР 2 (2 год.)

Завдання 2. Експлораційний аналіз і підготовка даних

Опис:

Візьміть набір даних із кількома типами змінних (числові + категоріальні). Використовуючи Plotly (або комбінуючи з Matplotlib) зробіть інтерактивні візуалізації (наприклад, коробкові діаграми, гістограми за категоріями, heatmap кореляцій). Попередньо проведіть:

- кодування категоріальних змінних (one-hot, label encoding)
- заповнення або видалення пропущених значень
- масштабування числових змінних (за потреби).

Напишіть код, що робить усе це, і підготуйте короткий звіт: які трансформації зроблені, як змінились розподіли змінних, що можна далі аналізувати.

Поради:

- Оберіть набір із кількома категоріальними змінними (щоб практикувати кодування).
- Збережіть результат підготовки (наприклад, у CSV), щоб його використовувати далі.
- Використання Plotly дає змогу студенту ознайомитись із інтерактивністю, що підвищує якість роботи.

До ЛР 3 (3 год.)

Завдання 3. Статистичний аналіз і класифікація / регресія

Опис:

Виберіть набір даних для задачі класифікації або регресії. Використовуючи SciPy й/або scikit-learn реалізуйте:

- описову статистику (mean, median, стандартне відхилення)
- перевірку гіпотез (наприклад, t-тест чи ANOVA для категоріальних груп)
- побудову моделі класифікації або регресії (наприклад, логістична регресія, лінійна регресія)
- оцінювання моделі (accuracy, confusion matrix, R^2 тощо)
- аналіз важливості змінних, інтерпретацію результатів.

Рекомендовані теми:

- Для класифікації: набір тексту чи числових >2 класів
- Для регресії: набір із числовою цільовою змінною

Поради:

- Документуйте: як ви обирали модель, які передобробки робили.
- Покажіть, які змінні важливі, чи мають сенс, чи потребують трансформації.
- Додайте графіки розподілів, залишків (для регресії) або матрицю помилок.

До ЛР 4 (3 год.)

Завдання 4. Кластерний аналіз і асоціативний аналіз

Опис:

Використовуючи набір без мічених класів (або з невеликою кількістю мічених) застосуйте:

- Кластеризацію (наприклад, алгоритм k-means clustering, ієрархічна, DBSCAN) за допомогою scikit-learn. Документуйте вибір кількості кластерів, візуалізуйте кластери (за допомогою 2D/3D графіків). (scikit-learn.org)
- Асоціативний аналіз: реалізуйте алгоритм «Association Rules» (наприклад, бібліотека mlxtend) на дискретизованих змінних або транзакційному наборі даних. Інтерпретуйте правила: підтримка, достовірність.

Рекомендовані дані:

- Репозиторій з наборами для кластеризації: “Clustering-Datasets” на GitHub. ([GitHub](https://github.com))
- Використати приклади з навчальних матеріалів, наприклад, Iris, Wine, але краще — більш реальний набір із 500+ спостережень.

Поради:

- Для кластеризації: обов’язково стандартизуйте числові змінні.
- Для асоціативного аналізу: покажіть, які правила ви отримали, чи мають вони бізнес-смысл (чи науковий).
- Порівняйте кількість кластерів, спробуйте кілька підходів.

До ЛР 5 (2 год.)

Завдання 5. Обробка текстів: класифікація документів

Опис:

Візьміть публічно доступний текстовий набір даних (наприклад, рецензії, новинні статті). Використовуючи Python:

- Попередня обробка тексту (токенізація, стоп-слова, лематизація)
- Побудова векторного подання (TF-IDF або embeddings)
- Побудова моделі класифікації (наприклад, SVM, RandomForest або простий нейронний підхід)
- Оцінювання моделі (precision, recall, f1-score)
- Короткий звіт: які основні труднощі, що вдалося.

Рекомендовані набори:

- Список 15+ наборів для класифікації текстів. ([Innovatiana](#))
- Репозиторій niderhoff/nlp-datasets з багатьма корисними ресурсами. ([GitHub](#))

Поради:

- Можна обрати дві категорії (наприклад, позитив/негатив) або мультиклас.
- Обов'язково відобразить баланс класів, документи-«важкі» випадки.
- Якщо бажаєте, зробіть просту нейронну мережу (напр., з TensorFlow чи PyTorch), але це необов'язково.

До ЛР 6 (3 год.)

Завдання 6. Лінійні моделі в задачах ІАД (інтелектуального аналізу даних)

Опис:

Оберіть набір даних із числовими змінними та/або категоріальними, реалізуйте лінійну модель (лінійну регресію/логістичну регресію). Виконайте:

- Добір змінних (feature selection)
- Перевірку припущень моделі (нормальність, мультиколінеарність)
- Побудову моделі, інтерпретацію коефіцієнтів
- Оцінювання моделі
- Порівняйте базову лінійну модель із розширеною (наприклад, з поліноміальними ознаками або регуляризацією L1/L2)

Поради:

- Працюйте з набором, де кількість ознак не дуже велика (щоб інтерпретація була зрозумілою).
- Проведіть аналіз залишків, можливо – графік «спостережуване vs передбачене».
- Висновок: чи модель підходить, які обмеження.

Формат подачі та вимоги

1. Надіслати Jupyter Notebook (.ipynb) або .py-файл з коментарями + PDF-звіт (2-4 сторінки).
2. Код має бути читабельним, із коментарями.
3. Звіт має містити: постановку задачі, опис даних, підготовку, результати, висновки.
4. Використовуйте посилання на джерела даних, коротко опишіть, чому вибрали саме цей набір.
5. Студенти повинні самостійно виконати аналіз – не просто скопіювати чи «прогнати» чужий код.

Методи навчання

Методи навчання та викладання: лекції, лабораторні заняття, електронне навчання з використанням системи Moodle, тестування, виконання завдань ІНДЗ.

Методи формування професійної компетентності: розповідь, пояснення, бесіда, демонстрація, візуалізація, дискусія тощо. Методи формування практичних умінь та навичок: розв'язування задач лабораторних робіт, виконання завдань, розробка та аналіз алгоритмів і програмного коду, захист звітів з лабораторних робіт.

Система контролю та оцінювання

Засобами оцінювання та демонстрування результатів навчання є: стандартизовані тести; аналітичні звіти з лабораторних робіт; презентації результатів виконаних завдань та досліджень ІНДЗ, усний контроль у вигляді індивідуального та фронтального опитування на лекціях та лабораторних заняттях.

Формами поточного контролю є усна чи письмова (тестування, лабораторна робота, ІНДЗ) відповідь студента.

Формою підсумкового контролю є залік.

Критерії оцінювання поточного та підсумкового контролю

Критерієм підсумкового оцінювання є досягнення студентом мінімальних порогових рівнів оцінок (балів) за кожним передбаченим результатом навчання.

Система оцінювання рівня навчальних досягнень ґрунтується на принципах ECTS та є накопичувальною.

Студенти можуть отримувати загалом до 10 балів під час відвідування лекційних занять за правильні відповіді на запитання лектора, активне обговорення багатоваріантних підходів до рішення представленої лектором проблеми (для активізації пошукової та дослідної діяльності студентів).

Протягом семестру студенти виконують 6 лабораторних робіт. Кожна лабораторна робота оцінюється в 10 балів.

Звіт з лабораторної роботи, який студенти завантажують на сайт, повинен мати таку структуру:

1) файл у форматі Word, що містить титульний аркуш (назва університету, факультету, кафедри, навчальної дисципліни, хто виконав, хто перевірів, Чернівці – 202_) , назву лабораторної роботи, умову кожного завдання, код програми (не скріншот), скріншоти виконання з демонстрацією різних можливих, в т.ч. особливих, випадків роботи програми;

2) усі файли розробленого студентом проєкту, що містять коди програм;

3) всі вищезазначені файли архівуються в один архів, назва архіву має складатися із прізвища студента латиницею, номера лабораторної та номера варіанту, наприклад: Ivanenko_lab_1_var_21.arj.

У разі невиконання студентами вищезазначених правил оформлення кількість отриманих студентом балів при перевірці буде знижуватися. Якщо студент отримав суттєві зауваження до програмного коду і переробив програму та звіт, надіславши його на повторне оцінювання, максимально можлива кількість отриманих ним балів буде щонайменше на 1 (один) бал меншою, ніж максимально можлива кількість балів за цю роботу.

50% балів, відведених на оцінювання ЛР, студент отримує за працюючий програмний продукт, в якому реалізовано всі завдання ЛР та оформлений звіт. Решта 50% балів виставляється після захисту студентом виконаного звіту. На захисті звіту з ЛР студент має відповісти на питання щодо постановки задачі та розробленого ним алгоритму реалізації кожного із завдань ЛР. При відповіді на теоретичні питання та питання щодо програмної реалізації алгоритму у випадку неістотної помилки знімається 10-20% балів, а у випадку істотної 20-40% балів, якщо ж студент не опанував теоретичний матеріал, плутається в означеннях, наводить логічно невірні твердження, то знімається до 50% балів від усієї суми балів за ЛР.

Максимальна кількість, яку можна набрати на підсумковому модулі (тестування) – 30 балів. Тестування проводиться з використанням сайту дистанційного навчання.

Підсумкова оцінка виставляється за результатами суми балів, набраних на змістових модулях під час семестру та підсумковому модулі згідно з нижче наведеною таблицею.

Критерії оцінювання поточного та підсумкового контролю

Розподіл балів, які отримують студенти

Поточне оцінювання						Кількість балів (модуль-контроль)	Сумарна к-ть балів	
Змістовий модуль №1		Змістовий модуль №2			активність на заняттях			
T1	T2	T3	T4	T5	T6	ЗМ1, ЗМ2	30	100
10	10	10	10	10	10	10		

Шкала оцінювання: національна та ЄКТС

Оцінка за національною шкалою	Оцінка за шкалою ECTS	
	Оцінка (бали)	Пояснення за розширеною шкалою
Відмінно	A (90-100)	відмінно
Добре	B (80-89)	дуже добре
	C (70-79)	добре
Задовільно	D (60-69)	задовільно
	E (50-59)	достатньо
Незадовільно	FX (35-49)	(незадовільно) з можливістю повторного складання
	F (1-34)	(незадовільно) з обов'язковим самостійним опрацюванням освітнього компоненту до перескладання

На оцінку "відмінно" заслуговує студент, який виявив всебічні, систематичні та глибокі знання, здатність самостійно виконувати завдання, передбачені програмою, ознайомлений з основною і додатковою літературою, рекомендованою програмою. Така оцінка передбачає також засвоєння студентом взаємозв'язку основних понять дисципліни та їх значення для набутої професії.

Оцінку "добре" ставлять студентіві, який засвоїв навчально-програмовий матеріал у повному обсязі, успішно виконує передбачені програмою завдання, опрацював основну літературу, рекомендовану програмою, тобто студентіві, який засвідчив систематичний характер знань із дисципліни і здатний до їх самостійного поповнення й оновлення у процесі подальшої навчальної роботи і професійної діяльності.

На оцінку "задовільно" заслуговує студент, який виявив знання основного навчального матеріалу в обсязі, необхідному для подальшого навчання і майбутньої роботи за професією, здатний виконувати завдання, передбачені програмою, ознайомлений з основною літературою, рекомендованою програмою. Як правило, цю оцінку виставляють студентам, які припустилися

огріхів у відповіді на іспиті та при виконанні екзаменаційних завдань, але продемонстрували спроможність усунути їх.

Оцінку "незадовільно" ставлять студентів, у знаннях якого є прогалини, який припустився принципових помилок у виконанні передбачених програмою завдань, тобто студентів, який неспроможний продовжити навчання чи приступити до професійної діяльності після закінчення вищого навчального закладу без додаткових занять з відповідної дисципліни.

Перелік питань для самоконтролю та підсумкового контролю навчальних досягнень студентів

1. Математичне моделювання та аналіз даних програмними засобами Python (NumPy, Matplotlib).
2. Методи та засоби візуалізації результатів аналізу даних програмними засобами Python (Matplotlib, Plotly).
3. Методи та засоби статистичного аналізу даних у бібліотеці SciPy.
4. Навчальні набори даних для інтелектуального аналізу даних.
5. Експлораційний аналіз даних.
6. Обробка та підготовка даних.
7. Моделювання класифікації.
8. Моделювання регресії.
9. Кластерний аналіз.
10. Використання алгоритмів асоціативного аналізу.
11. Використання бібліотек Scikit-learn, TensorFlow та PyTorch для створення та навчання нейронних мереж.
12. Використання методів обробки текстів для аналізу та класифікації документів.
13. Лінійні моделі в задачах інтелектуального аналізу даних.
14. Методи визначення якості математичних моделей.
15. Ансамблеві математичні моделі.

Зарахування результатів неформальної/інформальної освіти

Здобувачі вищої освіти має право на участь у неформальній/інформальній освіті.

У межах поточного контролю можуть визнаватися результати неформальної/інформальної освіти за умови наявності сертифікату або освітньої декларації про результати неформальної/інформальної освіти з питань, що відповідає тематиці курсу («Порядок визнання у Чернівецькому національному університеті імені Юрія Федьковича результатів навчання, здобутих шляхом неформальної та/або інформальної освіти»), <https://www.chnu.edu.ua/media/4g5fzssb/poriadok-vyznannia-rezultativ-navchannia-zdobutykh-shliakhom-neformalnoi-ta-abo-informalnoi-osvity.pdf>).

Студентам можуть бути зараховані додаткові бали, отримані через неформальну освіту, до загальної суми балів, набраної з освітньої компоненти, за умови, що результати з проблеми, за якою відбувалося навчання, відповідають тематиці курсу.

Рекомендована література

Основна

1. Стандарт вищої освіти України перший (бакалаврський) рівень, галузь знань 12 – Інформаційні технології, спеціальність 124 – Системний аналіз // Затверджено і введено в дію наказом Міністерства освіти і науки України від 13.11.2018 р. № 1245.– 23 с.
URL: <https://mon.gov.ua/storage/app/media/vishcha-osvita/zatverdzeni%20standarty/12/21/124-sistemn.analiz-bakalavr-1.pdf>
2. Юрченко І.В. DataMining з використанням Python. Навчальний посібник.– Чернівці: Чернівецький національний університет імені Юрія Федьковича, 2024.– 143 с.
<https://archer.chnu.edu.ua/xmlui/handle/123456789/10067>
3. Інтелектуальний аналіз даних та машинне навчання. Частина 1. Базові методи та засоби аналізу даних / Я. В. Іванчук, В. І. Месюра, А. А. Яровий, О. Д. Манжілевський – Вінниця : ВНТУ, 2021. – 69 с. ISBN 978-966-641-874-9.
4. Ясинський В.К., Юрченко І.В. Прикладний статистичний аналіз. Методичні рекомендації до лабораторних робіт.– Чернівці: Рута, 2008.– 84 с.
5. Юрченко І.В. Прикладний статистичний аналіз з використанням Python.– Чернівці: Технодрук, 2021.– 102 с.
6. Jake VanderPlas. Python Data Science Handbook. Essential Tools for Working with Data.– Beijing, Boston, Farnham, Tokyo: O'Reilly Media, Inc, 2016.– 576 p.– ISBN: 9-781-491-912-058.
7. Davy Cielen, Arno D.B. Meysman, Mohamed Ali. Introducing data science: BigData, Machine Learning, and more, using Python tools.– New York: Manning Publications Co, Shelter Island, 2016.– 322 p.– ISBN: 9-781-633-430-037.
8. Бібліотека Scikit-learn мови Python [Електронний ресурс]. Джерело доступу: <https://scikit-learn.org/stable/>
9. Scikit-learn – Вікі [Електронний ресурс]. Джерело доступу: <https://uk.upwiki.one/wiki/Scikit-learn>

Допоміжна

1. Шевченко О.В. Методи обробки великих масивів даних в пакетному режимі // Магістерська дисертація на здобуття ступеня магістра за освітньо-професійною програмою «Інтелектуальні сервіс-орієнтовані розподілені обчислювання» зі спеціальності 122 «Комп'ютерні науки».– Київ:

- Національний технічний університет України "Київський політехнічний інститут імені Ігоря Сікорського", 2020.– 101 с.
2. Гуцуляк І.В. Застосування методів проектування ознак та кластеризації для машинного навчання // Дипломна робота на здобуття ступеня магістра за освітньою програмою "Інформаційні технології та управління проектами" зі спеціальності 122 "Комп'ютерні науки".– Чернівці: Чернівецький національний університет імені Юрія Федьковича, 2022.– 85 с.

Інформаційні ресурси

<https://moodle.chnu.edu.ua/course/view.php?id=7576>

<https://matplotlib.org/>

<https://plotly.com/>

<https://numpy.org/>

<https://scipy.org/>

<https://scikit-learn.org/stable/>

Політика академічної доброчесності

Дотримання політики щодо академічної доброчесності учасниками освітнього процесу при вивченні навчальної дисципліни регламентовано такими документами:

1. «Етичний кодекс Чернівецького національного університету імені Юрія Федьковича» <https://www.chnu.edu.ua/universytet/normatyvni-dokumenty/etychnyi-kodeks-chernivetsko-ho-natsionalnoho-universytetu-imeni-yurii-a-fedkovycha/>

2. «Положенням про виявлення та запобігання академічного плагіату у Чернівецькому національному університету імені Юрія Федьковича» <https://www.chnu.edu.ua/universytet/normatyvni-dokumenty/polozhennia-pro-vyavlennia-ta-zapobihannia-akademichnomu-plahiatu/>