

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

Затверджено

На засіданні кафедри радіоелектронних і
комп'ютерних систем
факультету електроніки та комп'ютерних
технологій
Львівського національного університету
імені Івана Франка
(протокол № 1/24 від 28.08 2023 р.)

Завідувач кафедри:

 Ігор ОЛЕНИЧ

Силабус з навчальної дисципліни
“Машинне навчання”,
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	Машинне навчання
Адреса викладання дисципліни	Корпус факультету електроніки та комп'ютерних технологій, Львівський національний університет імені Івана Франка, вул. Драгоманова 50, м. Львів, 79005, вул. Ген. Тарнавського 107, м. Львів, 79011
Факультет та кафедра, за якою закріплена дисципліна	Факультет електроніки та комп'ютерних технологій, кафедра радіоелектронних і комп'ютерних систем
Галузь знань, шифр та назва спеціальності	12 - Інформаційні технології, 122 — Комп'ютерні науки
Викладачі дисципліни	Сінкевич Олег Олександрович, доцент кафедри радіоелектронних і комп'ютерних систем, доктор філософії
Контактна інформація викладачів	oleh.sinkevych@lnu.edu.ua, https://electronics.lnu.edu.ua/employee/o_sinkevych
Консультації з питань навчання по дисципліні відбуваються	Консультації в день проведення лекційних занять (за попередньою домовленістю). Також можливі онлайн консультації через MS Teams та Telegram (у текстовому режимі). Для погодження часу онлайн консультацій слід писати на електронну пошту викладача, або в особистий Telegram.
Сторінка дисципліни	https://electronics.lnu.edu.ua/course/mashynne-navchannia-ist https://e-learning.lnu.edu.ua/course/view.php?id=4417
Інформація про дисципліну	Дисципліна «Машинне навчання» є нормативною дисципліною з спеціальності 122 Комп'ютерні науки для освітньої програми «Комп'ютерні науки», яка викладається в 7 семестрі в обсязі 4 кредитів (за Європейською Кредитно-Трансферною Системою ECTS).
Коротка анотація дисципліни	Навчальну дисципліну розроблено таким чином, щоб надати студентам необхідні знання на практичні навички, обов'язкові для того, щоб вміти реалізовувати та застосовувати типові алгоритми та відповідні їм структури даних для вирішення задач машинного навчання. Тому у дисципліні представлено як теоретичні та математичні основи моделей машинного навчання (лінійна та логістична регресії, нейронні мережі, метод головних компонент, k-середніми та ієрархічної кластеризації), так і конкретні реалізації та бібліотеки, які необхідні для створення прототипів та прикладних програмних інтерфейсів (API). Зокрема, детально розглянуті кроки для інженерії ознак, визначення ефективності моделей та підготовки моделей до їх веб-розгортання.
Мета та цілі дисципліни	<i>Мета:</i> вивчення нормативної дисципліни «Машинне навчання» є одержання студентами знань про базові моделі машинного навчання; типи машинного навчання; процес підготовки та обробки даних; виділення значущих ознак даних, їх перетворення; процес розгортання моделей машинного навчання та їх огортання в прикладний програмний інтерфейс; ознайомлення з бібліотеками машинного навчання для мови Python 3: numpy, scikit-learn, pandas, PyTorch. <i>Цілі:</i> забезпечити знайомство студентів з загальною теорією методів машинного навчання, сформувати навички їх практичного використання; навчити студента вибирати і обґрунтувати методи машинного навчання та обробки даних, визначати доцільність та допустимість вибраних методів під відповідний клас задач, програмно реалізувати, або модифікувати методи під потребу системи.
Література для вивчення дисципліни	Основна література: 1. Andreas C. Müller, Sarah Guido. Introduction to Machine Learning with

	<p>Python. O'Reilly Media, 2017. — 377 с.</p> <ol style="list-style-type: none"> 2. Sebastian Raschka, Vahid Mirjalili. Python Machine Learning, 2 edition. Packt Publishing. — 2017. — 622 с. 3. Luis Pedro Coelho, Willi Richert. Building Machine Learning Systems with Python, 2 edition. Packt Publishing. — 2015. — 326 с. 4. Jones A. The Unsupervised Learning Workshop / A. Jones, K. Christopher, B. Johnston., 2020. – 549 с. – (Packt Publishing). 5. Pramod Singh. Deploy Machine Learning Models to Production: With Flask, Streamlit, Docker, and Kubernetes on Google Cloud Platform. Apress. — 2021. — 161 с. 6. Rokach L. Data mining with decision trees. Theory and Applications / L. Rokach, O. Maimon. – Singapore: World Scientific Publishing Co. Pte. Ltd., 2015. – 328 с. 7. Goodfellow, I., Bengio, Y., & Courville, A. (2016). Deep Learning. MIT Press. – 800 с. 8. C. Aggarwal. Neural Networks and Deep Learning A Textbook. 2023. 529 с. 9. Foster D. Generative Deep Learning: Teaching Machines to Paint, Write, Compose, and Play. O'Reilly Media, Incorporated, 2023. 10. Feature Engineering / S. Sumathi та ін. Machine Learning for Decision Sciences with Case Studies in Python. Boca Raton, 2022. С. 351–371. URL: https://doi.org/10.1201/9781003258803-7. <p>Додаткова література (інтернет-ресурси):</p> <ol style="list-style-type: none"> 11. Machine Learning Mastery. MachineLearningMastery.com. URL: https://machinelearningmastery.com/ 12. Machine Learning. Google for Developers. Google for Developers. URL: https://developers.google.com/machine-learning.
<p>Обсяг курсу</p>	<p>Загальний обсяг: 120 годин. Аудиторних занять: 64 год., з них 32 год. лекційних та 32 годин лабораторних робіт. Самостійної роботи: 56 год.</p>
<p>Очікувані результати навчання</p>	<p>Після завершення цього курсу студент буде:</p> <ul style="list-style-type: none"> - Знати процес створення “поточної лінії” машинного навчання - від обробки даних до розгортання моделі; основи обробки даних, їх перетворення та підготовки для моделей машинного навчання; типові алгоритми виявлення та оцінки ознак даних; основи навчання моделей та оцінки їх ефективності; типові програмні засоби машинного навчання; теоретичний та практичний матеріал згідно програми курсу: математичні основи моделей машинного навчання, програмні реалізації тощо. - Вміти застосовувати засвоєний матеріал для створення прототипів моделей машинного навчання; використовувати Python для програмної реалізації моделей машинного навчання, їх тестування та оцінки, а також визначати тип задачі машинного навчання: регресії, класифікації та кластеризації і розв’язувати поставлену проблему на основі здобутих знань та навиків. <p>Після вивчення даного курсу «Машинне навчання» здобувачі набудуть таких Загальних (ЗК), Спеціальних/Фахових (СК) компетентностей та Програмних результатів навчання (ПР):</p> <p>ЗК1. Здатність до абстрактного мислення, аналізу та синтезу. ЗК7. Здатність до пошуку, оброблення та аналізу інформації з різних джерел. ЗК8. Здатність генерувати нові ідеї (креативність).</p>

	<p>СК 2. Здатність до виявлення статистичних закономірностей недетермінованих явищ, застосування методів обчислювального інтелекту, зокрема статистичної, нейромережевої та нечіткої обробки даних, методів машинного навчання та генетичного програмування тощо.</p> <p>СК 4. Здатність використовувати сучасні методи математичного моделювання об'єктів, процесів і явищ, розробляти моделі й алгоритми чисельного розв'язування задач математичного моделювання, враховувати похибки наближеного чисельного розв'язування професійних задач.</p> <p>СК 11. Здатність до інтелектуального аналізу даних на основі методів обчислювального інтелекту включно з великими та погано структурованими даними, їхньої оперативної обробки та візуалізації результатів аналізу в процесі розв'язування прикладних задач.</p> <p>СК 12. Здатність забезпечити організацію обчислювальних процесів в інформаційних системах різного призначення з урахуванням архітектури, конфігурування, показників результативності функціонування операційних систем і системного програмного забезпечення.</p> <p>ПР 2. Використовувати сучасний математичний апарат неперервного та дискретного аналізу, лінійної алгебри, аналітичної геометрії, в професійній діяльності для розв'язання задач теоретичного та прикладного характеру в процесі проектування та реалізації об'єктів інформатизації.</p> <p>ПР 4 Використовувати методи обчислювального інтелекту, машинного навчання, нейромережевої та нечіткої обробки даних, генетичного та еволюційного програмування для розв'язання задач розпізнавання, прогнозування, класифікації, ідентифікації об'єктів керування тощо.</p> <p>ПР 8. Використовувати методологію системного аналізу об'єктів, процесів і систем для задач аналізу, прогнозування, управління та проектування динамічних процесів в макроекономічних, технічних, технологічних і фінансових об'єктах.</p> <p>ПР 12. Застосовувати методи та алгоритми обчислювального інтелекту та інтелектуального аналізу даних в задачах класифікації, прогнозування, кластерного аналізу, пошуку асоціативних правил з використанням програмних інструментів підтримки багатовимірного аналізу даних на основі технологій DataMining, TextMining, WebMining.</p>
Ключові слова	Python, машинне навчання, логістична регресія, кластеризація, класифікація, нейронні мережі, PyTorch.
Формат курсу	Очний
	Проведення лекцій, лабораторних робіт та консультації для кращого розуміння тем.
Теми	Див. СХЕМА КУРСУ
Підсумковий контроль, форма	Залік в кінці семестру
Пререквізити	Для вивчення курсу студенти потребують базових знань з дисциплін «Алгоритми та структури даних», «Основи програмування», «Дискретна математика», «Вища математика».
Навчальні методи та техніки, які будуть використовуватися під час викладання курсу	Інформаційні методи (лекції, презентації, лабораторні роботи, написання рефератів, виконання індивідуальних завдань, робота у групі, командна робота, обговорення, консультації для поглибленого розуміння тем, бесіда, ілюстрація, демонстрація), дедуктивні методи на основі узагальнень, евристичні методи (проблемна лекція),

	інтерактивні методи (дискусія).
Необхідне обладнання	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, доступ до мережі Internet, 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. Необхідне програмне забезпечення включає в себе ОС Windows 10, або Ubuntu 22.04 LTS, середовище розробки MS Visual Studio/PyCharm/Vim/VS Code, компілятор мови програмування Python.
Критерії оцінювання (окремо для кожного виду навчальної діяльності)	<p>Оцінювання проводиться за 100-бальною шкалою. Бали нараховуються за наступним співвідношенням:</p> <ul style="list-style-type: none"> • лабораторні роботи: 60% семестрової оцінки; максимальна кількість балів 60. • змістові модулі (2 модулі): 40% семестрової оцінки; максимальна кількість балів 40. <p>Загалом упродовж семестру 100 балів.</p> <p>Змістовий модуль – самостійна робота студента оформлена у вигляді робочої програми та звіту – робота друкованим текстом, рекомендованим обсягом до 5 сторінок (шрифт Times New Roman, 14). Звіт включає в себе детальний розгляд завдання до модуля, приведення прикладів та лістингів коду програм, огляду технологій. Код програм повинен бути обов'язково прокоментований та пояснений, необхідно також продемонструвати його роботу у разі, якщо в якості прикладу наводяться не окремі елементи технології, а суцільна програма. Фінальна версія звіту у .pdf форматі разом з кодом завантажується на віддалений Git-репозиторій.</p> <p>Академічна доброчесність: Очікується, що роботи студентів будуть оригінальними дослідженнями чи міркуваннями. Списування та втручання в роботу інших студентів становлять, але не обмежують, приклади можливої академічної недоброчесності. Виявлення ознак академічної недоброчесності в написанні завдань є підставою для її незарахування викладачем, незалежно від масштабів плагіату чи обману. Жодні форми порушення академічної доброчесності не толеруються.</p> <p>Відвідання занять є важливою складовою навчання. Очікується, що всі студенти відвідають усі лекції та лабораторні заняття курсу. Студенти повинні інформувати викладача про неможливість відвідати заняття. У будь-якому випадку студенти зобов'язані дотримуватися термінів визначених для виконання всіх 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</p>

використання також й іншої літератури та джерел, яких немає серед рекомендованих.

Політика виставлення балів. Враховуються бали, набрані при поточному контролі та бали за виконання лабораторних робіт. При цьому обов'язково враховуються присутність на заняттях та активність студента під час практичного заняття; недопустимість пропусків та запізень 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.

Оцінювання лабораторних робіт (8 лабораторних робіт, максимальна кількість балів: 60) відбувається шляхом оцінки роботи студента під час проведення лабораторної роботи в аудиторії та захисту звіту по виконаній лабораторній роботі (0-7.5 балів за одну роботу).

Бали оцінювання лабораторних робіт нараховуються за наступним співвідношенням:

7.5 – 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;

7 - 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, проте код не є правильно оформлений згідно стандартів;

6 – студент достатньо розуміє розглянутий матеріал та принципи написаного ним коду програми, присутні неточності та незначні помилки у відповідях на запитання по темі, код програми функціонує відповідно до завдання (або з несуттєвими недоліками);

5 - студент достатньо розуміє розглянутий матеріал та принципи написаного ним коду програми, присутні неточності та незначні помилки у відповідях на запитання по темі, код програми функціонує відповідно до завдання, проте з помірними недоліками;

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;

3 – 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 вагомими недоліками;

2 – 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>1 - студент погано розуміє розглянутий матеріал та написаний ним код програми, код програми не функціонує належним чином;</p> <p>0 - студент зовсім не засвоїв розглянутий матеріал, написаний ним код програми не відповідає темі/не функціонує взагалі.</p> <p>Оцінювання змістових модулів (2 змістових модулів, 20 балів за кожний) — за результатами написаних студентом звіту та коду програми.</p> <p>Бали оцінювання змістових модулів нараховуються за наступним співвідношенням:</p> <p>20-16 - розглянута тема відтворюється в повному обсязі, правильно, обґрунтовано, логічно, містить аналіз і систематизацію, аргументовані висновки. Засвідчено глибоке володіння матеріалом. Наведений код повністю робочий та відповідає темі. Можуть бути присутні несуттєві помилки та невідповідності;</p> <p>16-12 - відтворюється значна частина розглянутої теми. Виявлено знання і розуміння основних положень навчальної дисципліни, проте присутні неточності та/або невідповідності основній темі. Наведений код частково робочий, проте в загальному відповідають темі;</p> <p>12-8 - відстежується загальне розуміння розглянутої теми. Виявлені множинні неточності та невідповідності, пояснення наведеного коду відсутні, код функціонує із значними неточностями (або відсутні приклади запуску коду на виконання взагалі);</p> <p>8-4 – студент погано розуміє розглянуту тему. Виявлені суттєві неточності та невідповідності. Наведені приклади коду з суттєвими недоліками, або не відповідають темі;</p> <p>4 – 0 – студент взагалі не розуміє розглянуту тему. Тему не розкрито, кількість викладеного матеріалу не відповідає загальним нормам обраного виду роботи. Наведений код не робочий, або відсутній як такий.</p> <p>Критерії оцінювання результатів неформальної освіти:</p> <p>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>Питання до модульного контролю</p>	<p>Категорії машинного навчання. Навчання з учителем, без учителя та навчання з підкріпленням. Виявлення патернів у даних та їх використання у розробці алгоритмів</p>

	<p>прогнозування. Pipeline проєктів машинного навчання. Параметричні та “жадібні” моделі.</p> <p>Поняття навчання моделі з точки зору її параметрів. Непараметричний метод k-найближчих сусідів для задачі класифікації. Переваги та недоліки. Поняття регресії проти класифікації.</p> <p>Виведення рівняння лінійної регресії. Робота з бібліотекою NumPy для матричних і векторних обчислень. Векторизація задачі лінійної регресії. Визначення методу найменших квадратів та його застосування у машинному навчанні. Поняття функції втрат.</p> <p>Задача класифікації методом лінійної регресії. Сигмоїд. Представлення задачі з використанням NumPy. Ймовірнісна інтерпретація логістичної регресії. Алгоритм реалізації. Зведення до задачі з багатозначною цільовою змінною.</p> <p>Оцінка ефективності навчання моделі логістичної регресії. Візуалізація навчальної кривої. Поняття перенавчання моделей. Перехресне затвердження та k-кратне перехресне затвердження. Регуляризація моделі.</p> <p>Вибір та інженерії ознак даних у машинному навчанні. Критерії значущості ознак. Алгоритми вибору ознак для задач машинного навчання. Стискання вимірності та метод головних компонент.</p> <p>Визначення штучної нейронної мережі. Матричне представлення. Поняття функції активації, їх типи. Вибір функції втрат: бінарна та багато класова класифікація. Різновиди сучасних нейронних мереж та їх застосування.</p> <p>Навчання нейронних мереж: алгоритм градієнтного спуску та його модифікації. Покроковий алгоритм методу зворотного поширення помилки. Початкове налаштування гіпер-параметрів та ініціалізація ваг. Функція активації у процесі навчання.</p> <p>Багатошарові нейронні мережі та глибоке навчання. Поняття рекурентної нейронної мережі та довгої короткочасної пам’яті. Трансформери та масштабована скалярно-добуткова увага. Гіпер-параметри, проблеми та алгоритми їх оптимізація.</p> <p>Огортання моделей машинного навчання у Flask/Fastapi. Створення Restapi для моделей машинного навчання. Структурування коду. Реалізація простого full pipeline засобами scikit-learn, MongoDB та Flask/Fastapi.</p>
<p>Опитування</p>	<p>Анкету-оцінку з метою оцінювання якості курсу буде надано по завершенню курсу.</p>

**Схема курсу “Машинне навчання”
для студентів спеціальності 122 – Комп’ютерні науки**

Тиж.	Тема, план, короткі тези	Форма діяльності (заняття)	Література. Ресурси в Інтернеті	Завдання (лабораторна робота), год	Термін виконання
1	Основи машинного навчання. Типи машинного навчання. Інтелектуальний аналіз даних. Основні поняття. Категорії машинного навчання. Навчання з учителем, без учителя та навчання з підкріпленням. Виявлення патернів у даних та їх використання у розробці алгоритмів прогнозування. Pipeline проєктів машинного навчання.	Лекція	2, 4, 5, 10	2	кінець поточного тижня
2	Параметричні та непараметричні моделі. Алгоритм k-найближчих сусідів. Поняття класифікації та регресії. Параметричні та “жадібні” моделі. Поняття навчання моделі з точки зору її параметрів. Непараметричний метод k-найближчих сусідів для задачі класифікації. Переваги та недоліки. Поняття регресії проти класифікації.	Лекція	2, 4, 6, 10	Реалізація алгоритму k-найближчих сусідів з використанням numpy та pandas, 2 год.	кінець поточного тижня
3	Лінійна регресія. Основи лінійної алгебри для векторизації та роботи з матрицями в алгоритмах машинного навчання. Метод найменших квадратів. Функція втрат. Виведення рівняння лінійної регресії. Робота з бібліотекою Numru для матричних і векторних обчислень. Векторизація задачі лінійної регресії. Визначення методу найменших квадратів та його застосування у машинному навчанні. Поняття функції втрат.	Лекція	4, 10	Створення лінійно-регресійної моделі для багатьох ознак з використанням пакету numpy, 2 год.	кінець поточного тижня
4	Логістична регресія (ЛР). Представлення гіпотези. Мультикласова класифікація. Ймовірнісна інтерпретація ЛР. Задача класифікації методом лінійної регресії. Сигмоїд. Представлення задачі з використанням Numru. Ймовірнісна інтерпретація логістичної регресії. Алгоритм реалізації. Зведення до задачі з багатозначною цільовою змінною.	Лекція	2, 4, 5, 6, 10	Реалізація логістичної регресії, 2 год.	кінець поточного тижня
5	Аналіз ефективності моделей ЛР. Навчальна крива. Основи регуляризації. Оцінка ефективності навчання моделі логістичної регресії. Візуалізація навчальної кривої.	Лекція	2, 4, 5, 7, 10	2	кінець поточного тижня

	Поняття перенавчання моделей. Перехресне затвердження та кратне перехресне затвердження. Регуляризація моделі.				
6	Інженерія ознак. Вибір ознак та аналіз їх значущості. Метод головних компонент. Вибір та інженерії ознак даних у машинному навчанні. Критерії значущості ознак. Алгоритми вибору ознак для задач машинного навчання. Стискання вимірності та метод головних компонент.	Лекція	2, 3, 10	Відбір та виявлення важливості ознак. Метод головних компонент, 2 год.	кінець поточного тижня
7	Нейронні мережі (НМ). Математична основа НМ. Матричне представлення. Функції активація та втрат. Визначення штучної нейронної мережі. Матричне представлення з використанням NumPy. Поняття функції активації, їх типи. Вибір функції втрат: бінарна та багато класова класифікація. Різновиди сучасних нейронних мереж та їх застосування.	Лекція	1, 7	2	кінець поточного тижня
8	Навчання нейронних мереж. Метод зворотного поширення помилки. Особливості ініціалізації ваг. Навчання нейронних мереж: алгоритм градієнтного спуску та його модифікації. Покроковий алгоритм методу зворотного поширення помилки. Початкове налаштування гіпер-параметрів та ініціалізація ваг. Функція активації у процесі навчання	Лекція	1, 7, 10	Реалізація штучної нейронної мережі з використанням numpy. Навчання методом зворотного поширення помилки, 2 год.	кінець поточного тижня
9	Основи глибокого навчання. Рекурентні (CNN), LSTM, Трансформери. Проблема оптимізації гіпер-параметрів. Багатошарові нейронні мережі та глибоке навчання. Поняття рекурентної нейронної мережі та довгої короткочасної пам'яті. Трансформери та масштабована скалярно-добуткова увага. Гіпер-параметри, проблеми та алгоритми їх оптимізація.	Лекція	1, 6, 7, 10	2	кінець поточного тижня
10	Дерева прийняття рішень. Індекс gini. Побудова класифікаційних і регресійних дерев. Метод випадкового лісу. Алгоритм дерева прийняття рішень. Поняття індексу gini. Алгоритм CART. Ітераційна побудова дерева для задач класифікації та регресії.	Лекція	10, 12	2	кінець поточного тижня

	Перенавчання дерева прийняття рішень та ансамблевий метод випадкового лісу. Огляд алгоритму випадкового лісу.				
11	Основи підсилювання. Мета-алгоритми. Adaboost та xgboost методи. Поняття градієнтного підсилення та мета-алгоритмів. Слабкі класифікатори та їх навчання. Алгоритми adaboost. Бутстреп-агрегування (бегінг). Бібліотека xgboost та його використання у Python-інфраструктурі машинного навчання.	Лекція	10, 12	Дослідження дерев прийняття рішень. Порівняння з методами випадкового лісу та градієнтного підсилення, 4 год.	кінець поточного тижня
12	Навчання без учителя. Кластеризація. Метод k-середніх. Оцінка моделі та зниження розмірності даних. Основи навчання без учителя та його види: кластеризація та визначення асоціацій. Алгоритм методи k-середніх: підготовка даних (нормалізація/стандартизація), реалізація алгоритму та оцінка оптимальності. Застосування методів головних компонент та t-SNE.	Лекція	8, 10, 11	Реалізація кластеризації методом k-середніх. Визначення оптимального параметру k, 2 год.	кінець поточного тижня
13	Ієрархічна кластеризація. Інтеграція моделей з MongoDB. MongoDB pipeline. Ієрархічна кластеризація. Побудова дендрограми та оцінка кількості кластерів. Типи зчеплення (linkage). Агломеративна та розділююча кластеризація. Переваги та недоліки. Збереження даних у NoSQL базі даних MongoDB.	Лекція	8, 10, 11	2	
14, 15	Основи розгортання моделей. Flask і Fastapi. Огортання моделей машинного навчання в API. Огортання моделей машинного навчання у Flask/Fastapi. Створення Restapi для моделей машинного навчання. Структурування коду. Реалізація простого full pipeline засобами scikit-learn, MongoDB та Flask/Fastapi.	Лекція	3, 9, 10	Розгортання моделі з використанням мікро-фреймворків Flask, або Fastapi API, 2 год.	кінець поточного тижня
16	Підсумкове заняття. Підведення підсумків. Наставочі питання.	Лекція	1, 9, 10	2	кінець поточного тижня