

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
ЗАХІДНОУКРАЇНСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ
ЧОРТКІВСЬКИЙ НАВЧАЛЬНО-НАУКОВИЙ ІНСТИТУТ
ПІДПРИЄМНИЦТВА І БІЗНЕСУ

ЗАТВЕРДЖУЮ
Директор ЧННІБ ЗУНУ

Надія КУЛЬЧИЦЬКА

28 08 * 2024 р.

ЗАТВЕРДЖУЮ
Проректор з науково-педагогічної роботи
Віктор ОСТРОВЕРХОВ

30 08 * 2024 р.

РОБОЧА ПРОГРАМА

з дисципліни «Машинне навчання»
ступінь вищої освіти – бакалавр
галузь знань – 01 Освіта/Педагогіка
спеціальність – 015 Професійна освіта (цифрові технології)
освітньо-професійна програма – «Професійна освіта. Цифрові технології»

кафедра фундаментальних та спеціальних дисциплін

Форма навчання	Курс	Семестр	Лекції (год.)	Практ. (год.)	ІРС (год.)	Тренінг (год.)	Самост. робота студ. (год.)	Разом (год.)	Залік (сем.)	Екз. (сем.)
Денна	3	6	32	14	3	6	35	90	-	6

Чортків – ЗУНУ
2024

Робоча програма складена на основі освітньо-професійної програми підготовки бакалавра галузі знань 01 «Освіта/Педагогіка», спеціальності 015 «Професійна освіта (цифрові технології)», затвердженої Вченою радою ЗУНУ (протокол № 9 від 15.06.2022 р.).

Робочу програму склала к.е.н., доцент Ольга ШМИГЕЛЬ

Робоча програма затверджена на засіданні кафедри фундаментальних та спеціальних дисциплін (протокол № 1 від 28.08.2024 р.)

Завідувач кафедри _____ Людмила ДЕРМАНСЬКА

Розглянуто та схвалено групою забезпечення спеціальності 015 «Професійна освіта (цифрові технології)» (протокол № 1 від 30.08.24 р.)

Керівник групи _____ Лілія РЕБУХА
забезпечення спеціальності

Гарант ОПП _____ Ольга ПАВЕЛЧАК-ДАНИЛЮК

СТРУКТУРА РОБОЧОЇ ПРОГРАМИ НАВЧАЛЬНОЇ ДИСЦИПЛІНИ "МАШИННЕ НАВЧАННЯ "

1. Опис дисципліни "Машинне навчання "

Дисципліна «Машинне навчання»	Галузь знань, спеціальність, СВО	Характеристика навчальної дисципліни
Кількість кредитів – 3	Галузь знань – – Освіта/Педагогіка	Статус дисципліни: обов'язкова Мова навчання: Українська
Кількість залікових модулів – 5	Спеціальність – Професійна освіта (цифрові технології)	Рік підготовки: 3 Семестр: <i>Денна</i> – 6
Кількість змістових модулів – 2	Освітньо- професійна програма Професійна освіта. «Цифрові технології»	Лекції: <i>Денна</i> – 32 год. Лабораторні заняття: <i>Денна</i> – 14 год.
Загальна кількість годин – 90	Ступінь вищої освіти – бакалавр	Самостійна робота: <i>Денна</i> – 35 год., Тренінг: <i>Денна</i> – 6 год. Індивідуальна робота: <i>Денна</i> – 3 год.
Тижневих годин – з них аудиторних –		Вид підсумкового контролю – екзамен, 6

2. Мета і завдання дисципліни " Машинне навчання "

2.1. Мета вивчення дисципліни

Мета дисципліни – є навчання майбутніх спеціалістів сучасним методам побудови та аналізу різноманітних моделей машинного навчання, а також практичному застосуванню для вирішення деяких прикладних задач.

2.2. Завдання вивчення дисципліни

Завдання навчальної дисципліни «Машинне навчання»: ознайомити студентів із технологією машинного навчання; поглибити знання у сучасних методах та засобах, що використовуються при навчанні систем із штучним інтелектом; здобути практичні навички побудови моделей.

2.3. Найменування та опис компетентностей, формування яких забезпечує вивчення дисципліни:

СК17. Здатність застосовувати методи та інструментальні засоби з основ теорії і практики машинного навчання.

СК18. Здатність вирішувати складні задачі обробки даних з використанням методів машинного навчання в різних галузях професійної діяльності.

2.4. Передумови для вивчення дисципліни

Курс базується на знаннях, одержаних при вивченні дисциплін «Вища математика», «Дискретна математика», «Чисельні методи та програмування», «Теорія ймовірності і математична статистика», «Основи комп'ютерних наук», «Алгоритми і структури даних», «Бази і сховища даних», «Сучасні парадигми програмування», «Методи та системи штучного інтелекту» та «Інтелектуальний аналіз даних».

2.5. Результати навчання:

ПР17. Застосовувати методи, інструментальні засоби та алгоритми машинного навчання для розв'язку задач класифікації, розпізнавання, прогнозування, кластерного і регресійного аналізу.

ПР18. Розробляти програмні модулі в предметних областях, що використовують парадигми машинного навчання та штучного інтелекту у стратегіях рішення задач.

3. Програма навчальної дисципліни «Машинне навчання»

Змістовий модуль 1 – Основи та методи машинного навчання

Тема 1. Основи машинного навчання.

Основні поняття та історія машинного навчання. Способи нормалізації даних перед навчанням. Підготовка даних для машинного навчання. Види підходів до навчання: Supervised та Unsupervised. Переваги та недоліки кожного підходу

Тема 2. Методи машинного навчання для задач кластеризації.

Визначення задачі кластеризації. Методи кластеризації: DBSCAN, Hierarchical, ін. Оцінка якості кластеризації. Приклади використання кластеризації в реальних задачах.

Тема 3. Регуляризаційні лінійні регресійні моделі навчання.

Лінійна регресія та її обмеження. Гребнева регресія (Ridge Regression). Лассо-регресія (Lasso Regression). Еластична регресія (Elastic Net). Важливість

регуляризації в уникненні перенавчання

Тема 4. Методи опорних векторів (SVM).

Опорні вектори та геометрія SVM. Лінійна класифікація за допомогою SVM. Функція ядра та нелінійна класифікація. Оцінка та налаштування параметрів SVM.

Тема 5. Деревя рішень і правил.

Логічний підхід до класифікації. Алгоритм C4.5 для побудови дерев рішень. Обробка відсутніх значень у C4.5. Алгоритм CART та індекс Джині. Обмеження представлення дерев рішень і правил прийняття рішень.

Тема 6. Ансамблеве навчання.

Ансамблі моделей та їх переваги. Бустінг (Boosting), беггінг (Begging) та стекінг (Stacking) як методи комбінування моделей. Важливість випадкового лісу (Random Forest) у класифікації. Налаштування та використання ансамблів в практиці.

Змістовий модуль 2 – Застосування штучних нейронних мереж

Тема 7. Реалізації нейронних мереж як методів реалізації логічних обчислень.

Введення в нейронні мережі та їх структура. Використання TensorFlow і Keras для розробки нейронних. Тренування та оцінка нейронних мереж. Приклади застосування нейронних мереж у різних областях.

Тема 8. Глибокі нейромережеві архітектури в задачах машинного навчання.

Знайомство з глибоким навчанням та нейромережами. Архітектури глибоких нейромереж: LeNet, AlexNet, GoogLeNet, VGGNet, ResNet, SENet. Приклади застосування глибоких нейромереж в комп'ютерному зорі та обробці зображень. Перспективи розвитку глибокого навчання.

Тема 9. Машинне навчання для обробки природньої мови.

Основи обробки природньої мови (NLP). Бібліотеки для NLP: NLTK та Spacy. Токенізація, лематизація та векторизація тексту. Застосування NLP в аналізі тексту та машинному перекладі.

Тема 10. Навчання з підкріпленням.

Введення в навчання з підкріпленням. Агенти, стани та дії в RL. Алгоритми навчання з підкріпленням: Q-Learning, DDPG, A3C. Застосування RL в гральних середовищах та реальних завданнях.

Тема 11. Масштабування процесу машинного навчання для MLOps.

Основи MLOps та їх значення. Автоматизація процесу навчання та впровадження моделей. Моніторинг та управління моделями в реальному часі. Забезпечення безпеки та якості моделей у MLOps.

4. Структура залікового кредиту з дисципліни «Машинне навчання»

Тема	Кількість годин					
	Лекції	Лабораторні заняття	Індивідуальна робота	Тренінг	Самостійна робота	Контрольні заходи
Змістовий модуль 1 – Основи та методи машинного навчання						
Тема 1. Основи машинного навчання	2	-	2	3	3	Опитування під час заняття
Тема 2. Методи машинного навчання для задач кластеризації	4	2			3	
Тема 3. Регуляризаційні лінійні регресійні моделі навчання	4				3	
Тема 4. Методи опорних векторів (SVM)	4	2			3	
Тема 5. Дерева рішень і правил	2				3	
Тема 6. Ансамблеве навчання	2	2			3	
Змістовий модуль 2 – Застосування штучних нейронних мереж						
Тема 7. Реалізації нейронних мереж як методів реалізації логічних обчислень	4	2	1	3	4	Опитування під час заняття
Тема 8. Глибокі нейромереві архітектури в задачах машинного навчання	4				3	
Тема 9. Машинне навчання для обробки природньої мови	4	2			4	
Тема 10. Навчання з підкріпленням	4	2			3	
Тема 11. Масштабування процесу машинного навчання для MLOps	2	2			3	
Разом	32	14	3	6	35	

5. Тематика лабораторних занять

Лабораторна робота №1. Кластерний аналіз

Мета: Навчитися розробляти модель кластеризації даних та інтерпретувати результати моделювання.

Питання для обговорення:

1. Поняття кластеризації і класифікації даних.
2. Основні етапи процесу кластеризації даних.
3. Кластеризація даних методами: DBSCAN, Hierarchical, ін.

Лабораторна робота №2. Регуляризаційні лінійні регресійні моделі

Мета: Навчитися використовувати регуляризацію в лінійній регресії для покращення моделі та запобігання перенавчанню.

Питання для обговорення:

1. Лінійна регресія та її основні принципи.
2. Проблема перенавчання та її наслідки.
3. Гребнева регресія (Ridge Regression) як метод регуляризації.
4. Лассо-регресія (Lasso Regression) та її особливості.

5. Еластична регресія (Elastic Net) як поєднання гребневої та лассо-регресії.

Лабораторна робота №3. Методи опорних векторів (SVM)

Мета: Навчитися використовувати метод опорних векторів для класифікації даних та розуміти основні концепції цього методу.

Питання для обговорення:

1. Опорні вектори та їх роль у класифікації.
2. Лінійна класифікація за допомогою SVM та її принципи.
3. Функція ядра та нелінійна класифікація з використанням SVM.
4. Важливість налаштування параметрів SVM для досягнення найкращих результатів.
5. Практичні приклади використання SVM у задачах класифікації даних.

Лабораторна робота №4. Дерева рішень

Мета: Навчитися будувати дерево рішень, використовуючи навчальний набір даних та інтерпретувати візуальне представлення елементів дерева.

Питання для обговорення:

1. Поняття дерева рішень.
2. Побудова дерева рішень.

Лабораторна робота №5. Ансамблеве навчання

Мета: Навчитися використовувати Boosting, Bagging і Stacking моделі та інтерпретувати результати моделювання.

Питання для обговорення:

1. Boosting.
2. Bagging.
3. Stacking.

Лабораторна робота №6. Добування ознак із зображень

Мета: Навчитися моделювати удосконаленим проектуванням ознак на основі добування ознак із зображень.

Питання для обговорення:

1. Поняття удосконаленого проектування ознак на основі добування ознак із зображень.
2. Прості ознаки
3. Видобування об'єктів і форм

Лабораторна робота №7. (Заняття 7) Обробка природної мови

Мета: Навчитися передбачення тональності рецензій до фільмів на базі реальних даних.

Питання для обговорення:

1. Вивчення даних і сценарії їх застосування.
2. Побудова вихідної моделі з застосуванням базових NLP ознак і оптимізація параметрів.
3. Витяг більш складних NLP-ознак для збільшення точності моделювання.

Лабораторна робота №7. (Заняття 7) Масштабування процесу машинного навчання

Мета: Навчитися визначити, коли масштабування моделі допомагає збільшити точність і швидкість генерації пророкувань.

Питання для обговорення:

1. Способи масштабування лінійних ML-алгоритмів для роботи з великими навчальними вибірками;

2. Підходи до масштабування нелінійних ML-алгоритмів - зазвичай це набагато складніше завдання;
3. Зменшення затримки і збільшення швидкості генерації прогнозів.

6. Самостійна робота

Самостійна робота студентів в рамках курсу "Машинне навчання" повністю зосереджена на виконанні командного проекту. Кожна команда складається мінімум з 5 осіб. Проект є основним елементом самостійної роботи та передбачає вирішення реальної задачі з використанням методів машинного навчання, які студенти опанували протягом курсу.

Мета проекту: Застосувати знання і навички, отримані під час курсу, для розробки, реалізації та впровадження моделі або системи, що вирішує конкретну задачу з використанням алгоритмів машинного навчання.

Етапи виконання проекту та критерії оцінювання:

Етапи виконання проекту	Опис	Критерії оцінювання командного проекту
Вибір теми проекту	Команди можуть обрати одну з запропонованих тем або запропонувати власну, яка має бути узгоджена з викладачем. Тема повинна бути актуальною, мати практичне застосування та відповідати рівню знань студентів.	Актуальність та складність задачі
Розробка рішення	Студенти займаються збором, обробкою даних, розробкою та впровадженням моделей машинного навчання. Кожен учасник команди відповідає за певний етап проекту, забезпечуючи якісне виконання свого завдання.	Якість технічного виконання (реалізація алгоритмів, вибір архітектури, ефективність рішення)
Аналіз та оптимізація	Команда проводить аналіз результатів роботи моделей, здійснює їх оптимізацію та тестування на нових даних. Важливо не тільки досягти високих показників, але й пояснити, чому саме це рішення є оптимальним.	Аналіз та оптимізація результатів
Підготовка звіту та презентація	Наприкінці проекту команда готує детальний звіт, де описується процес виконання, отримані результати, а також висновки та рекомендації. Проект презентується перед викладачами та одногрупниками.	Оформлення звіту та представлення проекту

Перелік тем:

1. Система прогнозування врожайності сільськогосподарських культур за допомогою машинного навчання. Розробка моделі, яка прогнозує врожайність основних сільськогосподарських культур на основі даних про погоду, типи ґрунтів та історичні показники.
2. Аналіз якості повітря у великих містах України. Використання машинного навчання для аналізу та прогнозування рівня забруднення повітря в мегаполісах України з метою покращення екологічної ситуації.
3. Прогнозування енергоспоживання в житлових будинках. Створення моделі, що прогнозує споживання енергії в житлових будинках з метою оптимізації використання ресурсів та зниження витрат.

4. Аналіз та прогнозування транспортних потоків у великих містах. Використання машинного навчання для аналізу транспортних потоків та розробки моделей прогнозування заторів у містах.
5. Розпізнавання та класифікація аномалій у промисловому обладнанні. Розробка системи, яка аналізує дані з датчиків промислового обладнання та виявляє аномалії, що можуть призвести до аварій або збоїв.
6. Персоналізовані рекомендації щодо харчування та здорового способу життя. Використання машинного навчання для розробки індивідуальних рекомендацій з харчування та фізичної активності на основі аналізу медичних даних.
7. Автоматизоване сортування відходів. Розробка системи машинного навчання для автоматичного сортування відходів на сміттєпереробних заводах, що сприятиме підвищенню ефективності переробки та екологічної ситуації в Україні.
8. Прогнозування руху товарів у логістичних ланцюгах. Модель для оптимізації логістичних процесів, включаючи прогнозування попиту, складування та транспортні маршрути, що важливо для українських виробників і торгових мереж.
9. Аналіз соціальних медіа для виявлення фейкових новин. Розробка алгоритму для виявлення фейкових новин та дезінформації в українських соціальних медіа, що може допомогти у боротьбі з інформаційними загрозами.
10. Система прогнозування ризиків природних катастроф. Використання машинного навчання для прогнозування природних катастроф (повені, зсуви, землетруси) на основі аналізу геологічних і метеорологічних даних в Україні.
11. Автоматизація контролю якості продуктів харчування. Розробка моделі для автоматичного аналізу якості продуктів харчування за допомогою зображень або інших параметрів, що важливо для української харчової промисловості.
12. Персоналізовані системи охорони здоров'я. Створення моделі для прогнозування індивідуальних ризиків захворювань на основі медичних даних пацієнтів, що може бути використано для профілактики та персоналізованого лікування в Україні.
13. Моделі оптимізації водокористування у сільському господарстві. Розробка алгоритму для оптимізації використання водних ресурсів у сільському господарстві на основі кліматичних даних та інформації про ґрунт, що важливо для зрошення полів в Україні.
14. Аналіз і прогнозування тенденцій на ринку праці. Модель для аналізу та прогнозування змін на українському ринку праці на основі даних про вакансії, рівень освіти та демографічні показники, що допоможе в управлінні людськими ресурсами.
15. Прогнозування споживчих тенденцій у ритейлі. Використання машинного навчання для аналізу споживчих поведінкових даних та прогнозування майбутніх тенденцій у роздрібній торгівлі, що може бути корисно для українських торговельних компаній.

7. Тренінг з дисципліни

Завданням тренінгу є розробка програмного забезпечення із застосуванням методів машинного навчання. Тематика проектів буде базуватися на

індивідуальних завданнях самостійної роботи, з особливим акцентом на практичне використання алгоритмів машинного навчання для вирішення реальних задач. Основна мета тренінгу — розробка програмних рішень, що демонструють вміння студентів аналізувати, проектувати та впроваджувати ефективні інструменти машинного навчання у різних прикладних сферах. Нижче наведено основні етапи тренінгу та відповідні завдання:

Етап тренінгу	Завдання
Інструментарії розробки ПЗ для машинного навчання	- Налаштувати середовище розробки та встановити необхідні бібліотеки та інструменти для машинного навчання.
	- Створити базовий проект з підключеними бібліотеками для роботи з даними та реалізації моделей.
Реалізація моделей машинного навчання в програмному рішенні	- Реалізувати програмне рішення з використанням обраного алгоритму машинного навчання (класифікація, регресія, кластеризація тощо).
	- Провести тестування моделі на відповідних наборах даних та оцінити її ефективність.
Презентація та обговорення результатів	- Підготувати презентацію результатів роботи, включаючи аналіз даних, вибір моделі та її оцінку.
	- Захистити проект перед аудиторією, обґрунтувавши вибір методів та отримані результати.

8. Критерії, форми поточного та підсумкового контролю

Підсумковий бал (за 100-бальною шкалою) з дисципліни «Машинне навчання» визначається як середньозважена величина, залежно від питомої ваги кожної складової залікового кредиту:

Модуль 1		Модуль 3	Модуль 4	Модуль 5
20 %	20 %	5%	15%	40%
Поточне оцінювання	Модульний контроль 1	Тренінг.	Самостійна робота.	Екзамен.
	25 тестових питань. - Одне теоретичне питання. - Одна задача			- 10 тестових питань. - Одне теоретичне питання. - Одна задача.

Оцінка за «Поточне оцінювання» визначається як середнє арифметичне з оцінок, отриманих під час занять (кожен здобувач має бути оцінений не рідше як раз на два заняття). Пропуски практичних (лабораторних) занять обов'язково відпрацьовуються в години консультацій, в іншому випадку вони вважаються оцінкою «0» та враховуються при визначенні середнього арифметичного. Для здобувачів, які навчаються за індивідуальним графіком, поточне оцінювання проводиться під час консультацій, та шляхом виконання завдань в системі Moodle.

Модуль «Тренінг» визначається як середнє арифметичне з оцінок, за виконання завдань тренінгу.

Модуль «Самостійна робота» оцінюється як середнє арифметичне з двох оцінок, отриманих під час виконання наскрізного проекту: за якість викладу матеріалу та його представлення, відповідно до критеріїв, визначених у розділі 6 «Самостійна робота» цієї робочої програми.

9. Інструменти, обладнання та програмне забезпечення, використання яких передбачає навчальна дисципліна

№	Найменування	Номер теми
1.	Python	1-7
2.	Мультимедійне обладнання	1-11

РЕКОМЕНДОВАНІ ДЖЕРЕЛА ІНФОРМАЦІЇ

1. Chip Huyen Designing Machine Learning Systems, O'Reilly Media, Inc., 2022. ISBN: 9781098107949.
2. Himansu Das (editor), Jitendra Kumar Rout (editor), Suresh Chandra Moharana (editor), Nilanjan Dey (editor). Applied Intelligent Decision Making in Machine Learning. CRC Press, 2021. ISBN: 9780367503369.
3. Jun Chen; Edward P K Tsang. Detecting Regime Change in Computational Finance, Data Science, Machine Learning and Algorithmic Trading. Chapman & Hall. 2021. ISBN: 9780367540951
4. Hassanien A.E (ed.). Advanced machine learning technologies and applications. AMLTA 2020. Springer, 2021. ISBN: 9789811533822.
5. Saini S., Lata K., Sinha G. R. VLSI and Hardware Implementations Using Modern Machine Learning Methods. Boca Raton : CRC Press, 2021. URL: <https://doi.org/10.1201/9781003201038>.
6. Lee W.-M. Python Machine Learning. Wiley & Sons, Incorporated, John, 2019. 320 p.
7. Liu G. R. Machine Learning with Python. WORLD SCIENTIFIC, 2022. URL: <https://doi.org/10.1142/12774>.
8. Pattanayak S. Quantum Machine Learning with Python. Berkeley, CA : Apress, 2021. URL: <https://doi.org/10.1007/978-1-4842-6522-2>.
9. Bilokon P. A. Python, Data Science and Machine Learning. WORLD SCIENTIFIC, 2021. URL: <https://doi.org/10.1142/11701>.
10. Jonathan Roberge, Michael Castelle. The Cultural Life of Machine Learning: An Incursion into Critical AI Studies. Palgrave Macmillan, 2021. ISBN: 9783030562854,3030562859
11. Chris Fregly, Antje Barth. Data Science on AWS. April 2021 O'Reilly Media, Inc. ISBN: 9781492079392
12. Aboul Ella Hassanien, Ashraf Darwish. Machine Learning and Big Data Analytics Paradigms: Analysis, Applications and Challenges. Springer, Cham. 2021, ISBN: 3030593371.
13. Patel, Govind Singh, editor. Smart agriculture : emerging pedagogies of deep learning, machine learning and Internet of Things / edited by Govind Singh Patel, LPU Phagwara, India, Amrita Rai, UPTU, India, Nripendra Narayan Das, Manipal University Jaipur, India, R.P. Singh, Haramaya University, Diredawa, Ethiopia.

Description: First edition. | Boca Raton : CRC Press/Balkema/ Taylor & Francis Group, (2021), ISBN: 978-0-367-53580-3 (hbk).

Посилання на інформаційні ресурси в Інтернеті, відео-лекції, інше методичне забезпечення

1. Платформа для змагань з аналітики та передбачувального моделювання: <https://www.kaggle.com/>
2. Відео-лекції з машинного навчання: Coursera - Machine Learning by Andrew Ng <https://www.coursera.org/learn/machine-learning>
3. Інтерактивні курси з програмування та машинного навчання: Codecademy - Machine Learning
4. Статті, туторіали та блоги про машинне навчання: Medium - Towards Data Science <https://towardsdatascience.com/>
5. Електронна бібліотека з книгами з машинного навчання та штучного інтелекту: Springer - Machine Learning and AI <https://link.springer.com/>
6. Платформа для спільної роботи над проектами з кодування: GitHub <https://github.com/>
7. Бази даних для практичних завдань з машинного навчання: UCI Machine Learning Repository <https://archive.ics.uci.edu/ml/index.php>
8. Портал відкритих даних України: <https://data.gov.ua/>
9. Європейський портал відкритих даних: <https://data.europa.eu>
10. Світовий банк - Портал відкритих даних: <https://data.worldbank.org>
11. ОпенДатаБот - відкриті дані про Україну: <https://opendatabot.ua>
12. Google Public Data Explorer: <https://publicdata.google.com>
13. Ресурси з відкритими даними для наукових досліджень: Knoema <https://knoema.com/>
14. OECD Data - Портал відкритих даних Організації економічного співробітництва та розвитку: <https://data.oecd.org>
15. UNdata - База даних ООН: <https://data.un.org>
16. DataHub - Спільнота відкритих даних: <https://datahub.io>
17. Portal for Public Health Data: Our World in Data <https://ourworldindata.org/>