

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ

ЧОРНОМОРСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ ІМЕНІ ПЕТРА МОГИЛИ

ФАКУЛЬТЕТ КОМП'ЮТЕРНИХ НАУК

КАФЕДРА ІНТЕЛЕКТУАЛЬНИХ ІНФОРМАЦІЙНИХ СИСТЕМ



2020 року

РОБОЧА ПРОГРАМА НАВЧАЛЬНОЇ ДИСЦИПЛІНИ

«МЕТОДИ ТА СИСТЕМИ МАШИННОГО НАВЧАННЯ»

Спеціальність: 122 Комп'ютерні науки

Розробник

Гожий О.П.

Завідувач розробника

Кондратенко Ю.П.

Завідувач кафедри спеціальності

Кондратенко Ю.П.

Гарант освітньої програми КН

Гожий О.П.

Декан факультету Комп'ютерних наук

Бойко А.П.

Начальник НМВ

Шкірчак С.І.

Миколаїв – 2020 рік

1. Опис навчальної дисципліни

Найменування показника	Характеристика дисципліни	
Найменування дисципліни	Методи та системи машинного навчання	
Шифр курсу в освітній програмі	ОПП5	
Галузь знань	12 Інформаційні технології	
Спеціальність	Комп'ютерні науки	
Спеціалізація (якщо є)	-	
Освітня програма	Інтелектуальні інформаційні системи	
Рівень вищої освіти	Магістр	
Статус дисципліни	Нормативна	
Курс навчання	V	
Навчальний рік	2020-2021	
Номер семестрів:	Денна форма	Заочна форма
	10	-
Загальна кількість кредитів ЄКТС/годин	5,5 кредитів / 165 годин	
Структура курсу: – лекції – практичні – годин самостійної роботи студентів	Денна форма	Заочна форма
	18	-
	36	-
	111	-
Відсоток аудиторного навантаження	33%	-
Мова викладання	Українська	
Форма проміжного контролю	Тестування	
Форма підсумкового контролю	Іспит; Курсова робота	

2. Мета, завдання та результати вивчення дисципліни

Концепція викладання дисципліни:

Дисципліна «Методи та системи машинного навчання» присвячена розгляду актуальних питань використання сучасних інформаційних технологій для вирішення інтелектуальних задач і застосуванню інтелектуальних методів при вирішенні різноманітних проблем, зокрема проблем побудовання систем машинного навчання, вирішенню попередньої обробки даних, задач нормалізації, задач класифікації та кластеризації, проблем побудови алгоритмів для вирішення завдань машинного навчання та прогнозування, розв'язання завдань розпізнавання образів та завдань глибинного навчання. Постійний розвиток технологій машинного навчання та інформаційних інтелектуальних систем, розвиток та оновлення програмного забезпечення у сфері інтелектуалізації обробки і представлення даних та знань створює передумови для застосування нових методів та інформаційних технологій вирішення інтелектуальних задач у різних галузях людської діяльності. У зв'язку із цим виникає необхідність викладати у програмі підготовки бакалаврів з навчального напрямку 122 - "Комп'ютерні науки" дисципліну «Методи та системи та машинного навчання».

Мета курсу: формування у студентів теоретичних знань з сучасних інформаційних технологій машинного навчання при вирішенні різноманітних задач, та ефективного використання інтелектуальних методів та їх реалізацій для розв'язування практичних завдань, та придбання навичок їх ефективного використання.

«Методи та системи машинного навчання» є *нормативно* дисципліною при підготовці магістрів з галузі знань «Інформаційні технології». Для студентів спеціальності 122 - "Комп'ютерні науки" викладається у 10 семестрі в обсязі 5,5 кредитів (4 розділів), в тому числі 54 годин аудиторних занять; з них 18 годин лекцій, 36 годин практичних занять; 111 години самостійної роботи. Згідно навчального плану по дисципліні «Методи та системи та машинного навчання» передбачено виконання курсової роботи.

Передумови для вивчення дисципліни.

Вивчення дисципліни передбачає наявність першого бакалаврського рівня. Таким чином викладання курсу «Методи та системи машинного навчання» базується на знаннях, які студенти одержали при вивченні дисциплін "Вища математика", "Теорія алгоритмів", "Теорія ймовірностей, імовірнісні процеси і математична статистика", "Методи та системи штучного інтелекту" та інших. Студент повинен мати стійкі навички роботи на персональному комп'ютері, знати типові структури даних та основи програмування.

З дисципліни "Вища математика" використовуються знання і навички з лінійної алгебри та матричного обчислення.

З дисципліни "Теорія алгоритмів" використовуються знання і навички з розробки алгоритмів.

З дисципліни "Теорія ймовірностей, імовірнісні процеси і математична статистика" використовуються знання стосовно ймовірності, випадкової величини, випадкової функції, випадкового процесу, функції розподілу, щільності ймовірності, статистичних моментів їх оцінок.

З дисципліни "Методи та системи штучного інтелекту" використовуються знання і навички з розділу "Нейронні мережі".

В результаті вивчення дисципліни у студентів повинні бути сформовані **уявлення** про:

- Цілі і завдання досліджень в галузі машинного навчання, системах машинного навчання, принципи їх побудови та різних областях застосування.
- Проблеми і способи побудови нейронних мереж.
- Постановку задач класифікації, передбачення, розпізнавання образів і шляхи їх розв'язання.

Після вивчення дисципліни студенти повинні **знати**:

- Основні етапи та алгоритми, та процедури навчання та обробки даних.
- Основні алгоритми класифікації та кластеризації.
- Основні моделі нейронних мереж, методів і алгоритмів їх навчання.
- Основні поняття про навчання з підкріпленням.
- Структури та методи побудови систем МН та їх архітектурних особливостей залежно від особливостей розв'язуваної задачі. Етапи побудови систем МН.

В результаті вивчення дисципліни студенти повинні придбати **вміння** та **навички**:

- Будувати та застосовувати основні моделі нейронних мереж, класифікаторів та кластеризаторів для вирішення прикладних завдань МН.
- Ставити завдання побудови системи МН для вирішення різноманітних задач в різних предметних областях.
- Розробляти алгоритми для вирішення завдань класифікації, кластеризації та систем навчання за допомогою нейронних мереж в різних предметних областях.
- Розробляти компоненти систем машинного навчання, використовуючи сучасні програмні та інструментальні засоби.

Знання, які студенти набудуть при вивченні курсу "Методи машинного навчання" будуть необхідними при подальшому навчанні та освоєнні фахових та спеціальних дисциплін, а також у виробничій діяльності зі спеціальності.

В результаті вивчення дисципліни студент отримує:

Загальні компетентності:

- ЗК4. Здатність до пошуку, оброблення та аналізу інформації з різних джерел.
- ЗК5. Вміння виявляти, ставити та вирішувати проблеми.

Спеціальні (фахові, предметні) компетентності:

- ФК1. Виконання проектних робіт на усіх стадіях проектування інтелектуальних інформаційних систем і технологій (*обстеження, техніко-економічне обґрунтування, розробка технічного завдання, ескізне, технічне та робоче проектування, випробування, експлуатація*)

- ФК2. Розробка алгоритмів та алгоритмічних моделей інтелектуальних інформаційних систем і технологій.
- ФК3. Розробка програмного забезпечення окремих функціональних задач для інтелектуальних інформаційних систем.
- ФК5. Контроль працездатності та характеристик програм за галуззю інформаційні технології
- ФК9. Здатність розробляти математичні моделі і алгоритми для вирішення наукових та практичних задач прийняття оптимальних рішень і проектування систем, керування системами, процесами, аналізу та обробки даних, інтелектуального пошуку та видобування знань
- ФК10. Готовність до розробки та впровадження нових інтелектуальних інформаційних технологій і програмного забезпечення для управління, проектування, прийняття рішень, пошуку, аналізу і обробки даних.

Програмні результати навчання:

- ПР1. Знати і розуміти особливості та можливості сучасних інтелектуальних технологій та їх застосування у наукових дослідженнях.
- ПР2. Вміння організовувати самостійну, індивідуальну роботу, здійснювати комплексні дослідження та прийняття рішень в міждисциплінарних областях.
- ПР3. Знати та розуміти сучасні методи дослідження математичних моделей та алгоритмів інтелектуального аналізу даних, машинного навчання, аналізу сигналів, зображень та текстів, інформаційного пошуку та видобування знань, захисту інформації.
- ПР6. Вміння організовувати і проводити наукові дослідження та виконувати інноваційні розробки в галузі інтелектуальних інформаційних технологій.
- ПР7. Вміти розробляти математичні моделі і алгоритми для вирішення наукових та практичних задач прийняття оптимальних рішень і проектування систем, керування системами, процесами та проектами, аналізу та обробки даних, інтелектуального пошуку та видобування знань.
- ПР8. Знати методи розробки та впровадження нових інформаційних технологій і програмного забезпечення для управління, проектування, прийняття рішень, пошуку, аналізу і обробки даних.
- ПР9. Вміти розробляти математичні моделі та алгоритми для аналізу процесів і станів технічних систем.

3. Програма навчальної дисципліни

Денна форма:

№	Теми	Лекції	Практичні (групові)	Самостійна робота
<i>РОЗДІЛ 1. ВСТУП. Предметна область та основні поняття машинного навчання.</i>				
1	Тема 1. Вступ в машинне навчання. Головні розділи. Основні поняття.	1	-	10
2	Тема 2. Типи систем машинного навчання. Головні завдання та проблеми машинного навчання.	1	-	10
<i>РОЗДІЛ 2. Етапи розробки проекту машинного навчання.</i>				
3	Тема 3. Етапи проекту МН.	1	-	10
4	Тема 4. Вибір та навчання моделі даних. Налаштування моделей даних.	1	-	10
<i>РОЗДІЛ 3. Методи класифікації та кластеризації</i>				
5	Тема 5. Методи класифікації. Одновимірний та багатовимірні методи. Методи нормалізації даних.	2	8	10

6	Тема 6. Навчання моделей. Зменшення розмірності. (Узагальнення).	2	-	10
7	Тема 7. Методи кластеризації.	2	6	10
8	Тема 8. Дерева прийняття рішень. Ансамблеві методи МН.	2	4	10
РОЗДІЛ 4. Вступ до нейромережових технологій МН.				
9	Тема 9. Основи нейронних мереж.	2	8	10
10	Тема 10. Глибоке навчання та згорточні нейронні мережі.	2	4	10
11	Тема 11. Рекурентні мережі (RNN). Типи рекурентних нейронних мереж . Бібліотеки <i>TensorFlow</i> , <i>Keras</i> , <i>Scikit-Learn</i> .	2	6	11
РАЗОМ		18	36	111

4. Зміст навчальної дисципліни

4.1. План лекцій

№	Тема заняття / план
1	Тема 1: Вступ в машинне навчання. Головні розділи. Основні поняття. 1) Предмет та зміст дисципліни. 2) Два типи навчання: навчання з вчителем, навчання без вчителя. Навчання з підкріпленням. Тема 2: Типи систем машинного навчання. Головні завдання та проблеми машинного навчання. 1) Головна схема МН. 2) Етапи МН. Особливість кожного етапу. 3) Приклади завдань МН.
2	Тема 3: Етапи проекту МН. 1) Робота з реальними даними. 2) Отримання даних. 3) Підготовка даних для алгоритмів МН (нормалізація та агрегація даних). Тема 4: Вибір та навчання моделі даних. Налаштування моделей даних. 1) Методи попередньої обробки даних. 2) Налаштування моделей даних. 3) Приклади вирішення прикладних завдань.
3	Тема 5: Вступ до методів класифікації. 1) Двійковий класифікатор. 2) Байєсівський класифікатор. 3) Показники якості класифікації. 4) Багатокласова та багатозначна класифікація.
4	Тема 6: Навчання моделей. 1) Матриця неточностей. Методи побудови МН. 2) Лінійна регресія. 3) поліноміальна регресія. 4) логістична регресія.
5	Тема 7: Методи кластеризації. 1) Методи кластеризації. 2) Методи опорних векторів. 3) Лінійна класифікація SVM. 4) Регресія SVM.

6	Тема 8: Дерева прийняття рішень. Ансамблеві методи МН. 1) Дерева прийняття рішень. 2) Ансамблеві методи МН. 3) Бустінг, Беггінг.
7	Тема 9: Основи нейронних мереж. 1) Біологічний нейрон та модель штучного нейрону. 2) Класифікація НМ. 3) Перцептрон. 4) Багатошаровий перцептрон. 5) Алгоритми навчання. Приклад побудови системи розпізнавання символів.
8	Тема 10: Глибоке навчання та згорточні нейронні мережі. 1) Завдання глибокого навчання. 2) Архітектура CNN. 3) Типи шарів CNN. Створення класифікаторів зображень на основі CNN.
9	Тема 11: Рекурентні мережі (RNN). Типи рекурентних нейронних мереж . Бібліотеки <i>TensorFlow, Keras, Scikit-Learn</i>. 1) Базові рекурентні нейронні мережі в бібліотеках <i>TensorFlow</i> . 2) Навчання рекурентних мереж. 3) Навчання RNN для прогнозування часових рядів.

4.2 План практичних (семінарських, лабораторних, півгрупових) занять

№	Тема заняття / план
1	Тема 5. Вступ до методів класифікації. <i>Робота 1.</i> Наївний Байєсівський класифікатор. Мета роботи: Вивчити принцип побудови бінарної класифікації за допомогою бібліотек мови <i>Python</i> . Завдання: Виконати бінарну класифікацію для обраного набору даних за варіантом. Завдання виконати за допомогою бібліотек <i>numpy, sklearn.naive_bayes</i> мови <i>Python</i> .
2	Тема 5. Вступ до методів класифікації. <i>Робота 1.</i> Наївний Байєсівський класифікатор (<i>продовження</i>). Мета роботи: Вивчити принцип побудови бінарної класифікації за допомогою бібліотек мови <i>Python</i> . Завдання: Розробити та протестувати програмне забезпечення для розв'язання завдання класифікації за допомогою наївного Байєсівського класифікатора. До звіту додати скрін-шоти кроків алгоритму.
3	Тема 5. Вступ до методів класифікації. <i>Робота 2.</i> Логістичний класифікатор. Мета роботи: Вивчити принцип побудови логістичного класифікатора за допомогою бібліотек мови <i>Python</i> . Завдання: Побудувати логістичний класифікатор для обраного набору даних за варіантом. Виконати за допомогою бібліотек <i>numpy, sklearn</i> мови <i>Python</i> .
4	Тема 5. Вступ до методів класифікації. <i>Робота 2.</i> Логістичний класифікатор (<i>продовження</i>). Мета роботи: Вивчити принцип побудови логістичного класифікатора за допомогою бібліотек мови <i>Python</i> . Завдання: Розробити та протестувати програмне забезпечення для розв'язання завдання класифікації за допомогою логістичного класифікатора. До звіту додати скрін-шоти кроків алгоритму.

№	Тема заняття / план
5	Тема 7. Методи кластеризації. <i>Робота 3.</i> Машина опорних векторів (SVM). Мета роботи: Вивчити принцип класифікацію даних за допомогою машин опорних векторів. Завдання: Виконати класифікацію даних за допомогою машин опорних векторів. До звіту додати скрін-шоти кроків алгоритму.
6	Тема 7. Методи кластеризації. <i>Робота 4.</i> Створення одновимірних та багатовимірних регресорів. Мета роботи: Вивчити принципи побудови одновимірних та багатовимірних регресорів за допомогою бібліотек мови <i>Python</i>. Завдання: Побудувати регресійну модель для обраного набору даних за варіантом. Виконати за допомогою бібліотек <i>numpy, sklearn</i> мови <i>Python</i>.
7	Тема 7. Методи кластеризації. <i>Робота 4.</i> Створення одновимірних та багатовимірних регресорів (<i>продовження</i>). Мета роботи: Вивчити принципи побудови одновимірних та багатовимірних регресорів за допомогою бібліотек мови <i>Python</i>. Завдання: Розробити та протестувати програмне забезпечення для розв'язання завдання кластеризації за допомогою одновимірних та багатовимірних регресорів. До звіту додати скрін-шоти кроків алгоритму.
8	Тема 8. Дерева прийняття рішень. Ансамблеві методи МН. <i>Робота 5.</i> Навчання та візуалізація дерева прийняття рішень. Алгоритм CART. Завдання: Побудувати дерева прийняття рішень для регресії. Розробити алгоритм CART для вирішення завдань регресії. Розробити програмне забезпечення для розв'язання завдань регресії.
9	Тема 8. Дерева прийняття рішень. Ансамблеві методи МН. <i>Робота 5.</i> Навчання та візуалізація дерева прийняття рішень. Алгоритм CART (<i>продовження</i>). Завдання: Дослідити чутливість набору даних для навчання. Використовувати для виконання роботи бібліотеку <i>Scikit-Learn</i>.
10	Тема 9. Основи нейронних мереж. <i>Робота 6.</i> Використання методу зворотного розповсюдження помилки для навчання нейронної мережі. Мета роботи: Вивчити основні принципи функціонування нейронних мереж, реалізувати алгоритм зворотного розповсюдження помилки для навчання мережі, дослідити можливості застосування нейронних мереж для вирішення задач теорії штучного інтелекту. Завдання: 1. Скласти програму, яка реалізує функціонування нейронної мережі (див. рис.) для задачі розпізнавання (<i>класифікації</i>) чисел. Представити операцію навчання мережі на прикладах для навчання (див. рис.). Передбачити для користувача можливість зміни таких параметрів: 1) початкового діапазону розподілу ваги для зв'язків нейронної мережі $A[a_1, a_2]$; 2) кількості прихованих шарів нейронної мережі t; 3) кількості нейронів кожного прихованого шару $q_i, i=[1..t]$; 4) коефіцієнту навчання η; 5) коефіцієнту інерції α.

№	Тема заняття / план
11	Тема 9. Основи нейронних мереж. <i>Робота 6.</i> Використання методу зворотного розповсюдження помилки для навчання нейронної мережі (<i>продовження</i>). Мета роботи: Вивчити основні принципи функціонування нейронних мереж, реалізувати алгоритм зворотного розповсюдження помилки для навчання мережі, дослідити можливості застосування нейронних мереж для вирішення задач теорії штучного інтелекту. Завдання: 1. Протестувати програму, що була розроблена, для контрольних прикладів (див. рис.). 2. Дослідити вплив змінних параметрів на швидкість навчання мережі та якість розпізнавання образів. 3. В звіті представити: завдання: розпізнати символи, які складають прізвища студента; теоретичні відомості; покрокові результати навчання НМ; результати розпізнавання; текст програми; висновки.
12	Тема 9. Основи нейронних мереж. <i>Робота 7.</i> Використання теорії адаптивного резонансу (ТАР). Мета роботи: Вивчити принципи та алгоритм функціонування ТАР. Завдання: 1. Реалізувати структуру ТАР для задачі класифікації образів. 2. Дослідити вплив коефіцієнту порівняння ρ на кількість класів образів, що створюються в блоці розпізнавання при послідовному подаванні на вхід системи. 3. Для роботи програми обрати $\rho=0,8$ (ступінь порівняння); $\beta=1$ (коефіцієнт навчання); $N = 63$ (кількість біт вхідного вектора); $M = 25$ (кількість нейронів для запам'ятовування класів образів).
13	Тема 9. Основи нейронних мереж. <i>Робота 7.</i> Використання теорії адаптивного резонансу (<i>продовження</i>). Мета роботи: Вивчити принципи та алгоритм функціонування теорії адаптивного резонансу. Завдання: 1. Відповідно до алгоритму роботи ТАР подати на вхід 10 еталонних образів, та 10 образів для розпізнавання. 2. Передбачити можливість послідовно виводу на екран всіх класів образів, що збережені в шарі порівняння. 3. В звіті представити: завдання; теоретичні відомості (з блок-схемою алгоритму); покрокові результати навчання НМ; результати розпізнавання; текст програми; висновки; перелік літератури (не менш 10 посилань).
14	Тема 10. Глибоке навчання та згорточні нейронні мережі. <i>Робота 8.</i> Побудова лінійного регресора на основі перцептрона. Мета роботи: Вивчення алгоритму побудови лінійної регресійної моделі на основі перцептронів з використанням бібліотеки засобів глибокого навчання <i>TensorFlow</i>. Завдання: Побудувати алгоритм навчання нейронної мережі за заданими вхідними даними. Скласти програмний код алгоритму побудови лінійного регресора на основі перцептрона.
15	Тема 10. Глибоке навчання та згорточні нейронні мережі. <i>Робота 8.</i> Побудова лінійного регресора на основі перцептрона (<i>продовження</i>). Мета роботи: Вивчення алгоритму побудови лінійної регресійної моделі на основі перцептронів з використанням бібліотеки засобів глибокого навчання <i>TensorFlow</i>. Завдання: Протестувати програмний код алгоритму побудови лінійного регресора на основі перцептрона. Визначити складність алгоритму. Визначити швидкодію алгоритму. Зробити висновки. До звіту додати скрін-шоти кроків процесу навчання нейронної мережі.
16	Тема 11. Рекурентні мережі (RNN). Типи рекурентних нейронних мереж. Бібліотеки <i>TensorFlow, Keras, Scikit-Learn</i>. <i>Робота 9.</i> Рекурентні мережі (RNN). Типи рекурентних нейронних мереж. Мета роботи: Вивчення структури RNN та її різновидів. Завдання: Побудувати рекурентну нейронну мережу в середовищі <i>TensorFlow</i>. Навчити RNN для прогнозування часових рядів. До звіту додати скрін-шоти кроків процесу навчання нейронної мережі. Зробити висновки.

№	Тема заняття / план
17	Тема 11. Рекурентні мережі (RNN). Типи рекурентних нейронних мереж. Бібліотеки <i>TensorFlow, Keras, Scikit-Learn</i>. <i>Робота 9.</i> Рекурентні мережі (RNN). Типи рекурентних нейронних мереж (<i>продовження</i>). Мета роботи: Вивчення структури RNN та її різновидів. Завдання: Побудувати рекурентну нейронну мережу в середовищі <i>Keras</i>. Навчити RNN для прогнозування часових рядів. До звіту додати скрін-шоти кроків процесу навчання нейронної мережі. Зробить висновки.
18	Тема 11. Рекурентні мережі (RNN). Типи рекурентних нейронних мереж. Бібліотеки <i>TensorFlow, Keras, Scikit-Learn</i>. <i>Робота 9.</i> Рекурентні мережі (RNN). Типи рекурентних нейронних мереж (<i>продовження</i>). Мета роботи: Вивчення структури RNN та її різновидів. Завдання: Побудувати рекурентну нейронну мережу в середовищі <i>Scikit-Learn</i>. Навчити RNN для прогнозування часових рядів. До звіту додати скрін-шоти кроків процесу навчання нейронної мережі. Зробить висновки.

4.3. Завдання для самостійної роботи

4.3.1. Загальні положення

Одним з основних напрямів успішного засвоєння матеріалів навчальної дисципліни є самостійна робота студентів над основною й додатковою літературою з вивчення й використання сучасних комп'ютерних технологій при рішенні вимірювальних задач.

Основними видами самостійної роботи є:

1. Вивчення лекційного матеріалу.
2. Вивчення рекомендованої літератури.
3. Вивчення термінів і основних понять з тем навчальної дисципліни.
4. Підготовка до лабораторних занять і розробка ескізів документів з кожної лабораторної роботи.
5. Підготовка до тестового контролю з навчальної дисципліни.
6. Робота з опрацювання та вивчення рекомендованої літератури.
7. Систематизація вивченого матеріалу перед іспитом.

4.3.2. Обов'язкові види самостійної роботи

Фіксований перелік тем для виконання індивідуальних з дисципліни у семестрі студентам не пропонується. Теми обираються студентами самостійно та є засобом поглиблення знань про алгоритми та засоби їх побудови, які розглядаються в межах дисципліни. Крім того, можуть бути розглянутими деякі специфічні використання методів та алгоритмів.

Теми індивідуальних занять узгоджуються з викладачем протягом семестру, до початку залікового тижня.

Теми інформаційних повідомлень співпадають з темами та основними питаннями, які розглядаються на лекціях. В інформаційних повідомленнях також можуть розглядатись новітні засоби та методи в сфері обчислювальних алгоритмів.

4.3.3. Додаткові теми для самостійної роботи

1. Застосування систем МН для інформаційно-аналітичного забезпечення прийняття рішень (на основі обраного прикладу).
2. Основні форми подання й способи пояснення (інтерпретації) результатів МН.
3. Основні етапи аналізу й інтерпретації даних.
4. Стратегії введення, подання й організації вихідних даних в інформаційних системах з МН.
5. Вимірювальні шкали, подання змінних, введення й редагування даних.
6. Трансформація даних і файлів у методах МН (на обраному методі, прикладі рішення завдання).

7. Основні способи завантаження даних у системи МН з різних джерел, у тому числі на основі ODBC.
8. Об'єднання й агрегування (файлів) даних у інформаційно-аналітичних системах із застосуванням засобів МН.
9. Базові розвідницькі методи аналізу даних.
10. Табличні способи подання результатів МН.
11. Застосування методів і засобів картографії для аналізу даних й інтерпретації його результатів.
12. Побудова й інтерпретації карт візуалізації залежностей (на обраному прикладі, у тому числі - аналізу категоріальних даних).
13. Побудова, застосування та інтерпретації карт категоріальних даних.
14. Підготовка звітів та експорт результатів аналізу із систем МН.
15. Класичні методи багатомірного статистичного аналізу.
16. Роль і місце методів класичного статистичного аналізу для розв'язання основних завдань МН.
17. Застосування методів класифікації й кластеризації для сегментації даних.
18. Використання дерев рішень у завданнях класифікації (на обраному прикладі).
19. Компоненти й моделі часових рядів (тренда, сезонності, циклу, випадкового компонента).
20. Введення в нейронні мережі: основні принципи організації, побудови й застосування.
21. Рішення завдань класифікації нейронними мережами (на обраному прикладі).
22. Застосування нейронних мереж до завдань регресійного аналізу (на обраному прикладі).
23. Застосування нейронних мереж до завдань прогнозування часових рядів (на обраному прикладі).
24. Древа рішень, принципи їхньої побудови, практичного застосування й інтерпретації результатів (характеристика дерев).
25. Аналіз за допомогою дерева рішень: звичайні завдання аналізу за допомогою дерева рішень, області додатка аналізу за допомогою дерева рішень.
26. Застосування дерева рішень для класифікації (на обраному прикладі).
27. Застосування дерева рішень для прогнозування (на обраному прикладі).
28. Порівняльний аналіз засобів (систем) аналізу й прогнозування на основі дерев рішень (на прикладі не менше трьох систем).
29. Автоматизація виконання повторюваних завдань у системах МН.
30. Об'єднання й агрегування (файлів) даних в інформаційно-аналітичних системах із застосуванням засобів МН.
31. Створення звітів й експорт результатів із систем МН у різні додатки (на прикладі обраної системи).
32. Вплив розмірності й обсягів вихідних даних на вибір і застосування засобів МН.
33. Методи аналізу різнотипних експериментальних даних.
34. Застосування генетичних алгоритмів для аналізу даних (на прикладі пакета SNN).
35. Організація й реалізація оптимального вибору структури й параметрів моделей в МН (на обраному прикладі).

4.3.4. Вибіркові види самостійної роботи

Студентам пропонується виконання творчих завдань для самостійного опрацювання (одне – за вибором студента). Виконання творчих завдань не є обов'язковим, але може бути зараховане як залікове завдання, що надає шанс набрати бажану кількість балів до сесії.

Кожне творче завдання оцінюється в 10 балів та являє собою практичне завдання з виконання певних алгоритмів або окремих програмних модулів.

За результатами виконання творчого завдання студент повинен оформити звіт, в якому будуть задокументовані на скрін-шотах послідовно всі дії з виконання завдання, до отримання кінцевого результату.

Планом навчального процесу передбачено виконання студентами курсової роботи (КР). Навчальний посібник з рекомендаціями щодо виконання КР: Гожий О.П., Калініна І.О. Методи та системи та машинного навчання: Навчальний посібник. – Миколаїв: Вид-во МДГУ ім. Петра Могили, 2020 знаходиться в Moodle3 та в локальній мережі за посиланням: \\main\Documents\Computer Science\Гожий Олександр Петрович\401-402_Методи_та_системи_машинного навчання

**Приклад завдання на курсову роботу
ЗАВДАННЯ
на курсову роботу
з дисципліни «Методи та системи машинного навчання»
студенту**

_____ групи _____
Варіант № _____

1. Тема роботи

Розробка моделі класифікації на основі структурованих даних.

2. Вхідні дані для роботи

Значення признаков та векторів задається в залежності від структурованої вибірки для кожного варіанту окремо.

а) визначення формату структурованих даних, признаки $N_1, N_2, \dots, N_p$.

б) вектори (об'єкти): $M_1, M_2, \dots, M_s$;

в) значення признаку по кожному вектору.

*Приклад вхідних даних для завдання класифікації
Passenger*

Позначення	Роз'яснення
<i>pclass</i>	клас пасажира (1 – перший, 2 – другий, 3 – третій)
<i>survival</i>	виживання (0 – ні, 1 – так)
<i>name</i>	ім'я
<i>sex</i>	стать
<i>age</i>	вік
<i>sibsp</i>	кількість братів, сестер або подружжя на борту
<i>parch</i>	кількість батьків або дітей на борту
<i>ticket</i>	номер квитка
<i>fare</i>	тариф пасажира
<i>cabin</i>	каюта

<i>embarked</i>	точка посадки (С – Шербур, Q – Квінстаун, S – Саутгемптон)
<i>boat</i>	рятувальна шлюпка
<i>body</i>	ідентифікаційний номер тіла
<i>home.dest</i>	дім або місце призначення

3. Етапи обробки даних

Збір даних, очистка даних, створення признаков, вибірка даних, заміщення даних, нормалізація даних, рефакторинг, створення моделі, оцінка моделі, оптимізація моделі, матриця неточностей, крива ROC, крива навчання, розгортання моделі.

Зауваження! Для розв'язку задачі необхідно використання мови *Python* та бібліотеки *pandas*.

4. Зміст пояснювальної записки

- Анотація.
- Завдання на курсову роботу.
- Постановка задачі.
- Розв'язання завдання класифікації на основі методів машинного навчання (задача класифікації об'єктів при певній ситуації).
- Результат класифікації, його обґрунтування.
- Висновки.
- Список використаних джерел літератури (15-20 найменувань).

5. Перелік графічного матеріалу

Пояснювальна записка містить рисунки, що забезпечують ілюстрацію результатів роботи, а також Power-Point презентацію (10-12 слайдів).

Зауваження! Для захисту курсової роботи перед комісією кожен студент готує текст доповіді (до 5-7 хвилин).

4.4. Забезпечення освітнього процесу

- Ноутбук, проектор, екран.
- Комплект слайд-презентацій по курсу.
- Програмне забезпечення для демонстрацій слайд-презентацій.

5. Підсумковий контроль

Кожне залікове завдання складається з теоретичної та практичної частини. Перелік теоретичних питань наведений нижче:

1. Рання архітектура ШНМ.
2. Взаємозв'язок перцептрона і байесовського класифікатора в гауссовій середовищі.
3. Алгоритм мінімізації середньоквадратичної помилки.
4. Лінійний фільтр, побудований по методу найменших квадратів.
5. Переваги і обмеження навчання методом зворотного поширення.
6. Навчання з вчителем як задача оптимізації.
7. Порівняння мереж RBF і багатошарових перцептронів.
8. Асоціативні машини.
9. Машини опорних векторів.
10. Карти самоорганізації.
11. Нейродинамічне програмування.
12. Архітектура рекурентних мереж.
13. Мережа Хопфілда.
14. Мережа Хеммінга.
15. Мережа Кохонена.
16. Мережі зустрічного поширення.
17. Стохастичні мережі.

18. Радіально-базисна мережа. Архітектура RBF. Навчання RBF.
19. Когнітрон і неокогнітрон.
20. Мережі на основі адаптивного теорії адаптивного резонансу.
21. Генетичні алгоритми і ШНМ.
22. Каскадно-кореляційні мережі.
23. Мережа з елементами затримки сигналу.
24. Методи спрощення структури мереж. Методи видалення вагів. Методи видалення нейронів.
25. Асоціативна пам'ять. Алгоритми навчання асоціативної пам'яті.
26. Методи навчання ШНМ.
27. Програмні пакети для реалізації і використання ШНМ.
28. Мережі з перетворенням вхідного простору.
29. Еволюційне програмування та еволюційні стратегії.
30. Нейро-фаззі системи.
31. Система нейромереженого моніторингу активності користувачів у комп'ютерних системах.
32. Стандартні модифікації методу зворотного поширення.
33. Алгоритми евристичного пошуку.
34. Агенти, інтелект та еволюція.
35. Інтелектуальні агенти.
36. Стохастичний підхід до опису невизначеності.
37. Штучні імуні системи та їх використання.
38. Нечіткі моделі очікувального значення.
39. Навчання нейромереж за допомогою генетичних алгоритмів.
40. Основні поняття і принципи машинного навчання.
41. Дайте визначення інтелектуальних обчислень.
42. Коротка історична довідка про розвиток теорії штучних нейронних мереж.
43. Будова біологічного нейрона.
44. Штучний нейрон та активаційні функції. Наведіть приклади різних функцій активацій.
45. Пояснить можливості кожного засобу навчання НМ.
46. Чим відрізняється навчання між собою засоби навчання НМ.
47. Які бувають ШНМ, відносно їх топології? Наведіть приклади.
48. Будова та алгоритм функціонування бінарного перцептрона.
49. Проблема лінійної роздільності досліджуваної області.
50. В чому полягає зміст алгоритму навчання перцептрона Розенблатта?
51. Наведіть приклади застосування одношарового перцептрона.
52. У чому полягає зміст алгоритму оберненого поширення похибки?
53. За яким оптимізаційним методом функціонує алгоритм?
54. Яке значення для алгоритму має функція похибки?
55. Для розв'язку яких задач застосовується даний алгоритм?
56. Які переваги та недоліки *back propagation*?
57. Чим відрізняється навчання з учителем від навчання без учителя?
58. Які існують методи навчання нейронних мереж?
59. До якого класу навчання відноситься навчання методом оберненого поширення похибки - до навчання без учителя чи з учителем?
60. Чому порядок пред'явлення прикладів в навчальній вибірці може впливати на якість навчання?
61. Як впливає зменшення кількості вхідних нейронів на функціонування
62. мережі?
63. Які властивості повинна мати функція активації при використанні алгоритма оберненого поширення похибки?
64. У чому полягає особливість навчання за алгоритмом *Delta-Bar-Delta*?
65. Які евристики використовуються в алгоритмах *Delta-Bar-Delta* та *Enhanced Delta-Bar-Delta*?
66. Прокоментуйте алгоритм пружного поширення.
67. У чому полягає основна ідея алгоритму *Quick Propagation*?
68. Охарактеризуйте загальну процедуру еволюційних обчислень.

69. Охарактеризуйте методи еволюційного програмування та еволюційних стратегій.

Типові тестові завдання

1. Назвіть елементи математичної моделі формального нейрона.
 - a) Суматор
 - b) Синапс
 - c) Множник
 - d) Дільник
 - e) Зв'язок
2. Який компонент входить до складу нейросетевого інтелектуального блоку?
 - a) Вчитель
 - b) Синапс
 - c) Суматор
 - d) Контрастера
 - e) Мережа
3. Назвіть метод формування значень вихідних параметрів нейромережі.
 - a) Оптимізація
 - b) Експертна
 - c) Статистичний
 - d) Нелінійного програмування
 - e) Найменших квадратів
4. Назвіть методи попередньої обробки даних при використанні нейронних мереж для вирішення задач.
 - a) Фільтрація
 - b) Заповнення пропусків у даних
 - c) Проріджування
 - d) Підсумовування
 - e) Усереднення
5. Назвіть метод навчання багат шарових нейронних мереж
 - a) Центру невизначеності
 - b) Симплекс-метод
 - c) Зворотного поширення помилки
 - d) Найменших квадратів
 - e) Дисперсійний аналіз

Типові задачі для розв'язання

Практична робота № 0: Алгоритми класифікації. Алгоритм створення байесівського класифікатора

Вступ. Наївний байесівський: класифікатор (*Naive Bayes classifier*) - це простий класифікатор, заснований на використанні теореми Байеса, яка описує ймовірність події з урахуванням пов'язаних з ним умов. Такий класифікатор створюється за допомогою присвоювання міток класів екземплярів задачі. Останні подаються у вигляді векторів значень ознак. При цьому вважається, що значення будь-якого заданої ознаки не залежить від значень інших ознак. Його припущення про незалежність розглянутих ознак і становить наївну частину байесівського класифікатора.

Послідовність виконання практичної роботи.

Для створення наївного байесівського класифікатора необхідно:

1. Створити новий файл *Python* і імпортувати наступні пакети.

```
import numpy as np
import matplotlib.pyplot as plt
from sklearn.naive_bayes import GaussianNB
from sklearn import cross_validation
```

```
from utilities import visualize_classifier
```

2. Завантажити вхідні данні з файлу `data_multivar_nb(N).txt` За відповідним варіантом (N-номер варіанта за журналом).

3. Створити байесівський класифікатор. В даному випадку використовується гауссовський наївний байесовський класифікатор, в якому передбачається, що значення, асоційовані з кожним класом, діють за законом розподілу Гаусса.

```
# Создание наивного байесовского классификатора
classifier = GaussianNB()
```

4. Провести навчання класифікатора, використовуючи тренувальні дані.

Запустимо класифікатор на тренувальних даних і спрогнозуємо результати.

```
# Прогнозирование значений для тренировочных данных
y_pred = classifier.predict(X)
```

5. Обчислити якість (ассигасу) 1 класифікатора, порівнявши передбачені значення з істинними мітками, а потім візуалізувати результат.

```
# Вычисление качества классификатора
accuracy = 100.0 * (y== y_pred) .sum() / X.shape[0] print("Accuracy of Naive Bayes
classifier =", round(accuracy, 2), "%")

#Визуализация результатов работы классификатора
visualize_classifier(classifier, X, y)
```

6. Підготувати звіт.

«0» варіант іспитового білету з зазначенням максимальної кількості балів за кожне виконане завдання:

№	Питання білету	Максимальна кількість балів
1	Головні завдання та проблеми машинного навчання.	10
2	Дерева прийняття рішень. Ансамблеві методи МН. Наведіть приклади.	15
3	Байесівський класифікатор. Приклад використання.	15
	Всього	40

6. Критерії оцінювання та засоби діагностики результатів навчання

а) для денної форми навчання:

№	Вид діяльності (завдання)	Максимальна кількість балів
1	Практична робота №1	5
2	Практична робота №2	5
3	Практична робота №3	5
4	Практична робота №4	5
5	Практична робота №5	5
6	Практична робота №6	5
7	Практична робота №7	5
8	Практична робота №8	5
9	Практична робота №9	5
10	Виконання контрольного тестового завдання	5
11	Самостійна робота студента	10
12	Разом за семестр	60
13	Залік	40
14	Всього	100

Критерії оцінювання завдань для досягнення максимальної кількості балів при виконанні практичних робіт

Максимальна кількість балів (відповідно до попередньої таблиці) – студент з високою якістю самостійно виконав весь обсяг робіт, відповідає на всі питання, пов'язані з виконаними роботами, та робить додаткові розрахунки, які йому пропонує викладач. У викладача немає претензій щодо

реалізації та вимог до виконання роботи.

4 бали – студент з достатньою якістю виконав всі завдання, але в процесі роботи він робив деякі помилки, які, після вказування на них викладачем, самостійно виправляв. На деякі питання він відповідає з похибкою. Запропоновані викладачем додаткові розрахунки робить з деякою потугою. Не всі вимоги до виконання роботи дотримані.

3 бали – студент самостійно виконав всі роботи, але якість реалізації недостатня (помилки при розрахунках, не всі вимоги до роботи дотримані). На питання щодо виконання робіт відповідає не зовсім чітко. Є помилки при відповідях.

2-1 бали – студент самостійно виконав не всі роботи, при цьому якість реалізації недостатня (помилки при розрахунках, не дотримується вимог до оформлення роботи). На питання щодо виконання робіт відповідає не чітко. Є грубі помилки при відповідях.

0 балів – студент не виконав весь обсяг робіт, або виконав з грубими помилками. Він має проблеми з розрахунками, не знає теоретичного матеріалу, програмна реалізація не відповідає поставленим вимогам.

При отриманні незадовільної оцінки студент має право виправити всі помилки або виконати нові варіанти завдань, якщо викладач невпевнений, що студент виконав їх самостійно. Такий варіант пропонується, коли студент має багато пропусків занять.

Критерії оцінювання для досягнення максимальної кількості балів при виконанні контрольного тестового завдання

Оцінювання контрольного тестового завдання студентів по дисципліні «Методи та системи машинного навчання» проводиться по 5-бальній шкалі.

Тести допомагають отримати об'єктивніші оцінки рівня знань, умінь, навиків, перевірити відповідність вимог до підготовки студентів заданим стандартам, виявити пропуски в підготовці студентів.

Виходячи з технологічності процедури тестування відповіді кодуються двійковим кодом: **1** – істинно і **0** – помилково, і у такому вигляді можуть поступати в сучасні автоматизовані системи обробки інформації. Також може використовуватись відсоткове кодування відповіді (згідно з кількістю правильних відповідей у разі можливості надання декількох відповідей) відповідно до вагового коефіцієнту кожної відповіді з усіх можливих.

Тестування з дисципліни «Методи та системи машинного навчання» студенти можуть проходити як одноразово, так і необмежену кількість разів, поки не досягнуть результату, який відповідає їх уяві про власні знання. Чим буде більше наполегливим та зацікавленим студент - тим скоріше він досягне найвищого рівня професійних знань та вмінь. Така наполегливість також є складовою мети дисципліни.

Тестові бали переводяться в традиційну систему оцінок. Наприклад, якщо випробовуваний виконав більше 90 % завдань, то він отримує оцінку “відмінно” (**5 балів**), що вирішив від 75 до 90 % завдань “добре” (**4 бали**), від 50 до 75 % – “задовільно” (**3 бали**). Якщо студент виконав менш ніж 50% завдань, то він обов’язково повинен перескласти тестові завдання для отримання допуску до заліку.

Критерії оцінювання для досягнення максимальної кількості балів при виконанні курсової роботи

Оцінювання курсової роботи студентів по дисципліні «Методи та системи машинного навчання» проводиться на основі рейтингової системи оцінювання знань по 100-бальній системі.

Таблиця. Шкала оцінювання курсової роботи з дисципліни «Методи та системи машинного навчання»

Оцінка в балах	Оцінка за національною шкалою	Оцінка за шкалою ECTS	
		Оцінка	Пояснення
90-100	Відмінно	A	<i>Відмінно</i> студент з високою якістю самостійно виконав весь обсяг завдань курсової роботи, відповідає на всі питання, пов'язані з виконаними роботами. При захисті курсової роботи допускається лише незначна кількість помилок не принципового характеру. У викладача (та членів комісії) немає претензій щодо реалізації, оформленню та представленню курсової роботи.
82-89	Добре	B	<i>Дуже добре</i> студент виконав всі завдання курсової роботи, але в процесі роботи він робив деякі помилки, які, після вказування на них викладачем, самостійно виправляв. На деякі питання комісії він відповідає з похибкою або не дуже впевнено. Не всі вимоги до оформлення курсової роботи дотримані.
75-81	Добре	C	<i>Добре</i> студент виконував завдання курсової роботи вище середнього рівня з кількома суттєвими помилками, але після вказування на них викладачем, самостійно виправляв їх. На деякі питання комісії він відповідає з похибкою або з потугою. Не всі вимоги до виконання або оформлення курсової роботи дотримані.
67-74	Задовільно	D	<i>Задовільно</i> студент самостійно виконав всі завдання курсової роботи, але якість реалізації недостатня (помилки при розрахунках, не всі вимоги до роботи дотримані). На питання щодо виконання завдань відповідає не зовсім чітко. Є помилки при відповідях.
60-66	Задовільно	E	<i>Достатньо</i> студент самостійно виконав не всі завдання курсової роботи, при цьому якість задовольняє мінімальним критеріям (помилки при розрахунках, не дотримується вимог до оформлення роботи). На питання щодо виконання завдань відповідає не чітко. Є грубі помилки при відповідях.
35-59	Незадовільно	FX	<i>Незадовільно</i> (з можливістю обрання іншого варіанта завдання, його виконання та процедури захисту) студент виконав менш ніж 50% обсягу завдань курсової роботи, або виконав з грубими помилками. Він має проблеми з розрахунками, не знає теоретичного матеріалу, програмна реалізація не відповідає поставленим вимогам.

Оцінка в балах	Оцінка за національною шкалою	Оцінка за шкалою ECTS	
		Оцінка	Пояснення
1-34	Незадовільно	F	<i>Незадовільно</i> (з можливістю обрання іншого варіанта завдання, його виконання та процедури захисту) студент виконав менш ніж 25% обсягу завдань курсової роботи, або виконав з грубими помилками. Він має проблеми з розрахунками, не знає теоретичного матеріалу, програмна реалізація не відповідає поставленим вимогам, або вона відсутня.

Студент, який за виконання та захист курсової роботи набрав менше, ніж 60 балів, обирає інший варіант завдання для курсової роботи та виконує його.

7. Рекомендовані джерела інформації

7.1. Основні:

1. Бодянский Є.В., Пелешко Д.Д., Винокурова О.А. Машталір С.В., Іванов Ю.С. Гібридна нейронна мережа з глибинним навчанням в задачах динамічного інтелектуального аналізу потоків даних. - Монографія – Львів: 2016. – 267 с.
2. Глибовець М.М. Штучний інтелект. - К.:НУКМА, 2002.
3. Тарик Рашид Создаем нейронную сеть. – М.: Диалектика, 2018, 272 с.
4. Хайкин С. Нейронные сети – М.: Вильямс, 2006.
5. Hastie T., Tibshirani R, Friedman J. The Elements of Statistical Learning (2nd edition). Springer, 2009.
6. Bishop C. M. Pattern Recognition and Machine Learning. Springer, 2006.
7. Mohri M., Rostamizadeh A., Talwalkar A. Foundations of Machine Learning. MIT Press, 2012.
8. Murphy K. Machine Learning: A Probabilistic Perspective. MIT Press, 2012.
9. Mohammed J. Zaki, Wagner Meira Jr. Data Mining and Analysis. Fundamental Concepts and Algorithms. Cambridge University Press, 2014.
10. HastieT., TibshiraniR., FriedmanJ. The Elementsof Statistical Learning. Springer, 2014. — 739 p.
11. Bishop C.M. Pattern Recognitionand Machine Learning. — Springer, 2006. — 738 p.
12. Kohonen T. Self-Organizing Maps / T. Kohonen. - Berlin: Springer-Verlag, 1995. - 362 p.
13. Jang J.-S.R. Neuro-fuzzy and soft computing / J.-S.R. Jang, C.-T. Sun, E. Mizutani. - Prentice Hall, Upper Saddle River, NJ. - 1997. - 640 p.
14. Han, J. Data Mining: Concepts and Techniques / J. Han, M. Kamber. – Amsterdam e.a.: Morgan Kaufmann Publishers, 2006. – 754 p.
15. Gan, G. Data Clustering: Theory, Algorithms and Applications / G. Gan, Ch. Ma, J. Wu. – Philadelphia, Pennsylvania: SIAM, 2007. – 455 p.
16. Abonyi J. Cluster analysis for data mining and systems identification / J. Abonyi, B. Feil. - BaselBoston-Berlin: Birkheueser Verlag AG. - 2007. - 303 p.
17. Lughofer E. Evolving fuzzy systems: methodologies, advanced concepts and applications / E. Lughofer. - Springer. - 2011. - 454 p.

7.2 Додаткові:

1. Люгер Дж.Ф. Искусственный интеллект. – М.: Вильямс, 2003.
2. Руденко О.Г. Основы теории искусственных нейронных сетей / О.Г. Руденко, Е.В. Бодянский. - Харьков: ТЕЛІТЕХ. - 2002. - 317с.
3. Haykin S. Neural networks. A comprehensive foundation / S. Haykin. - Upper Saddle River, NJ: Prentice Hall. - 1999. - 842 p.

Студентам для вивчення навчального матеріалу надається конспект лекцій з надлишком навчального матеріалу для самостійного опрацювання, а також перелік літератури для засвоєння теоретичного матеріалу.

