



Методи бінарної класифікації

Робоча програма навчальної дисципліни (Силабус)

Реквізити навчальної дисципліни

Рівень вищої освіти	<i>Перший (бакалаврський)</i>
Галузь знань	<i>12 Інформаційні технології</i>
Спеціальність	<i>122 Комп'ютерні науки 124 Системний аналіз</i>
Освітня програма	<i>Системи і методи штучного інтелекту Системний аналіз і управління</i>
Статус дисципліни	<i>Вибіркова</i>
Форма навчання	<i>очна(денна)</i>
Рік підготовки, семестр	<i>4 курс, весняний семестр</i>
Обсяг дисципліни	<i>120 годин/ 4 кредитів ЕКТС (лекції – 36 год., практичні заняття – 18 год., СРС – 66 год.)</i>
Семестровий контроль/ контрольні заходи	<i>Залік, МКР</i>
Розклад занять	<i>https://schedule.kpi.ua/ (4 год. лекційних та 2 год. практичних занять на тиждень)</i>
Мова викладання	<i>Українська</i>
Інформація про керівника курсу / викладачів	<i>Лектор: д.ф.-м.н., професор, Купенко Ольга Петрівна, kupenko.olga@gmail.com Практичні: д.ф.-м.н., професор, Купенко Ольга Петрівна, kupenko.olga@gmail.com</i>
Розміщення курсу	<i>Платформа дистанційного навчання «Сікорський», додаткові ресурси за посиланням https://drive.google.com/drive/folders/1_LkkgulrHca-HI-19qblvNcxZYMnDKxz3RXL9IN4d8Fs16kj6EPdgyM9wgEKXppmn0VBnEyM</i>

Програма навчальної дисципліни

1. Опис навчальної дисципліни, її мета, предмет вивчення та результати навчання

Цілі дисципліни: сформувати навички щодо коректного формулювання різних типів задач машинного навчання в якості задач бінарної класифікації, розуміння особливостей базових алгоритмів бінарної класифікації та їх поєднання, оцінювання якості побудованих моделей та інтерпретації отриманих результатів.

Компетентності:

Здатність використовувати математичні методи для аналізу моделей формалізованої предметної області

Здатність збирати і аналізувати дані (включно з великими), для забезпечення якості прийняття проектних рішень

Здатність застосовувати існуючі і розробляти нові алгоритми розв'язування задач у галузі комп'ютерних наук

Програмні результати навчання:

Розробляти математичні моделі та методи аналізу даних (включно з великими)

Розробляти алгоритмічне та програмне забезпечення для аналізу даних (включно з великими)

Використовувати технології обчислювального інтелекту при розробці систем прийняття рішень та інтелектуальних інформаційних систем

2. Пререквізити та постреквізити дисципліни (місце в структурно-логічній схемі навчання за відповідною освітньою програмою)

Курс “Методи бінарної класифікації” є одним із завершальних курсів професійної підготовки бакалаврів спеціальності “Комп'ютерні науки”.

Цей курс підсумовує раніше прочитані спеціальні дисципліни в напрямку теорії та систем прийняття рішень і дає систематизоване детальне викладання основ теорії, методів та засобів побудови систем обчислювального інтелекту та їх застосування в системах прийняття рішень в економіці, бізнесі та фінансовій сфері. Тому ця дисципліна має глибокі логічні зв'язки з попередніми дисциплінами навчального плану підготовки, зокрема з курсами “Дослідження операцій”, “Теорія прийняття рішень в складних системах”, “Моделювання економічних систем”, “Статистичний аналіз економічних процесів”.

Проходження курсу дозволить застосовувати набуті знання та уміння для грамотного формулювання конкретних задач машинного навчання у вигляді задач бінарної класифікації у відповідь на завдання бізнесу, розробки відповідних моделей машинного навчання та оцінювання якості їх роботи.

Матеріали курсу широко використовуються при підготовці бакалаврської кваліфікаційної роботи

3. Зміст навчальної дисципліни

Освітній компонент охоплює такі теми:

ЛЕКЦІЇ

Тема 1. Формулювання різних типів задач машинного навчання як задач бінарної класифікації

Тема 2. Функції втрат

Тема 3. Метрики класифікації. Точність, повнота, F-міра

Тема 4. Крива ROC для оцінки якості класифікаційних моделей

Тема 5 Класичні, ітераційні та евристичні методи класифікації

Тема 6. Базові моделі бінарної класифікації. Логістична регресія

Тема 7. Базові моделі бінарної класифікації. Наївний баєсів класифікатор

Тема 8. Базові моделі бінарної класифікації. Метод опорних векторів

Тема 8. Базові моделі бінарної класифікації. Дерева рішень

Тема 9. Базові моделі бінарної класифікації. Нейро-мережеві моделі

Тема 10. Підходи та методи до класифікації з незбалансованими даними.

Активне навчання. Метрики для активного навчання.

Тема 11. Ансамблі базових моделей та мета-навчання

Тема 12. Калібрування ймовірностей

ПРАКТИЧНІ ЗАНЯТТЯ

<u>№ п/п</u>	<u>Найменування практичного заняття</u>	<u>Кількість годин</u>
<u>Пр. №1</u>	<u>Порівняння моделей логістичної регресії на базового класифікатора на конкретних даних. Оцінка точності моделей</u>	<u>2</u>
<u>Пр. №2</u>	<u>Порівняння дерева рішень та методу опорних векторів на конкретних даних. Оцінка точності моделей</u>	<u>4</u>
<u>Пр. №3</u>	<u>Нейромережі для бінарної класифікації</u>	<u>4</u>
<u>Пр. №4</u>	<u>Ансамблі базових алгоритмів та моделі мета-навчання. Оцінка підвищення точності. Калібрування ймовірностей</u>	<u>2</u>
<u>Пр. №5</u>	<u>Класифікація на незбалансованих даних. Залучення спеціальних метрик, оцінка інтерпретабельності моделей</u>	<u>4</u>
<u>Пр. №6</u>	<u>Розв'язання задачі прогнозування методами бінарної класифікації</u>	<u>2</u>

ДИДАКТИЧНІ МЕТОДИ:

На лекційних заняттях: Лекція, пояснення, мозковий штурм, проблемні завдання.

На практичних заняттях: Завдання до виконання, опитування та тестування студентів

4. Навчальні матеріали та ресурси

1. Базова література

1. M. Zgurovsky, Yu. Zaychenko. Fundamentals of computational intelligence- System approach. Springer, 2016.-275 p. (НТБ ім. Г.І. Денисенка)
 2. Yuriy Zaychenko. Problem of Fuzzy portfolio optimization and its solution with application of forecasting methods. Scholar Press. – 2015, 54 p. url: <http://journal.iasa.kpi.ua/article/view/65695>
 3. Інтелектуальний аналіз даних: навчальний посібник / укладачі: Я.М. Дрінь, І.В. Малик, Ю.А. Літвінчук // ЧНУ імені Ю. Федьковича, 2020. - 95 с. (НТБ ім. Г.І. Денисенка)
 4. Черняк О.І. Інтелектуальний аналіз даних: Підручник / Черняк О.І., Захарченко П.В./ К.: Знання, 2014 р. - 599 с. url: https://moodle.znu.edu.ua/pluginfile.php/767346/mod_resource/content/1/%D0%9F%D1%96%D0%B4%D1%80%D1%83%D1%87%D0%BD%D0%B8%D0%BA.pdf
- Гороховатський В. О. Методи інтелектуального аналізу та оброблення даних : навч. посіб. / В. О. Гороховатський, І. С. Творошенко ; М-во освіти і науки України, Харків. нац. ун-т радіоелектроніки. – Харків : ХНУРЕ, 2021. – 92 с. (НТБ ім. Г.І. Денисенка)

2. Допоміжна література

1. Zgurovsky M., Zaychenko Yu. Big Data: Conceptual Analysis and Applications. Springer Nature Switzerland AG. – 2020. – 275 p. (НТБ ім. Г.І. Денисенка)

2. Землянський О.М., Пашинська Н.М., Снитюк В.Є. Інформаційно-аналітичні технології прогнозування наслідків хімічних аварій, Київ, ВПЦ “Київський університет”, 2017, 167 с. (за запитом викладачу)
3. Huyen C. Designing Machine Learning Systems. - O'Reilly Media. USA – 2022. – 463 p. (за запитом викладачу)
4. Kantarzic M. Data Mining. Concepts, Models, Methods and Algorithms / M. Kantarzic, 3rd Ed. – Publisher : Wiley, 2019. – 672 p (за запитом викладачу)
5. Інтелектуальний аналіз даних: Комп'ютерний практикум [Електронний ресурс] : навч. посіб. для студ. спеціальності 122 «Комп'ютерні науки та інформаційні технології», спеціалізацій «Інформаційні системи та технології проектування», «Системне проектування сервісів» / О. О. Сергеев-Горчинський, Г. В. Іщенко ; КПІ ім. Ігоря Сікорського. – Електронні текстові данні (1файл: 1,72 Мбайт). – Київ : КПІ ім. Ігоря Сікорського, 2018. – 73 с. (НТБ ім. Г.І. Денисенка)

Навчальний контент

5. Методика опанування навчальної дисципліни (освітнього компонента)

У межах дисципліни заплановано такі види занять:

- лекції
- практичні заняття (комп'ютерні практикуми)
- самостійна робота

Теми дисципліни взаємопов'язані, матеріал вивчається в логічній послідовності. На лекціях розкриваються найбільш суттєві теоретичні питання, які дозволяють забезпечити студентам можливість глибокого самостійного вивчення всього програмного матеріалу. Теми та порядок виконання практичних завдань сформовано в логічній послідовності і повністю узгоджується і підтримується лекційним матеріалом. Теоретичні та практичні знання та навички поглиблюються шляхом самостійної роботи з використанням рекомендованої літератури та мережі Internet.

6. Самостійна робота здобувача вищої освіти

Важливим аспектом для якісного засвоєння матеріалу, відпрацювання прийомів і алгоритмів вирішення основних завдань дисципліни є самостійна робота, що дає змогу перетворити отримані знання в об'єкт власної діяльності. Самостійна робота включає в себе читання літератури, огляд літератури по темі, виконання звітів з комп'ютерних практикумів, підготовку до їх захисту та підготовку до іспиту.

Політика та контроль

7. Політика навчальної дисципліни (освітнього компонента)

Викладач повинен звернути увагу студентів на те, що дисципліна Інтелектуальний аналіз даних - це дисципліна, що займається розробкою і застосуванням методів та технологій AI в прикладних задачах розпізнавання образів, класифікації, кластерного аналізу в різних областях людської діяльності в умовах неповноти та невизначеності.

Рекомендовані методи навчання: проектний метод, імітаційні вправи, презентація та опитування студентів.

Студенту рекомендується вести докладний конспект лекцій і фіксувати основні результати практичних занять.

8. Види контролю та рейтингова система оцінювання результатів навчання (PCO)

Поточний контроль: фронтальний (усний, письмовий), лабораторні роботи.

Календарний контроль: проводиться двічі на семестр як моніторинг поточного стану виконання вимог силабусу.

Рейтингова система оцінювання включає всі види тестування: контрольні роботи, активність на лабораторних заняттях та якість захисту лабораторних робіт.. Кожен студент отримує свій підсумковий рейтинг по дисципліні.

Рейтинг студента з кредитного модуля у сьомому семестрі складається з балів, які він отримує за:

- написання модульної контрольної роботи;*
- робота на лабораторних заняттях та захист лабораторних робіт;*
- відвідування лекцій та написання конспекту під час лекцій;*
- відповіді на екзамені.*

Система рейтингових (вагових) балів та критеріїв оцінювання:

<i>Метод оцінювання</i>	<i>Кількість</i>	<i>Мінімальна оцінка в балах</i>	<i>Максимальна оцінка в балах</i>
<i>Лабораторні роботи</i>	<i>6</i>	<i>2</i>	<i>10</i>
<i>Модульна контрольна робота</i>	<i>1</i>	<i>4</i>	<i>20</i>
<i>Стартовий рейтинг</i>		<i>16</i>	<i>80</i>
<i>Залік</i>	<i>1</i>		<i>40</i>
<i>Підсумковий рейтинг</i>		<i>60</i>	<i>100</i>

Таблиця відповідності рейтингових балів оцінкам за університетською шкалою:

<i>Кількість балів</i>	<i>Оцінка</i>
<i>100-95</i>	<i>Відмінно</i>
<i>94-85</i>	<i>Дуже добре</i>
<i>84-75</i>	<i>Добре</i>
<i>74-65</i>	<i>Задовільно</i>
<i>64-60</i>	<i>Достатньо</i>
<i>Менше 60</i>	<i>Незадовільно</i>
<i>Не виконані умови допуску</i>	<i>Не допущено</i>

9. Додаткова інформація з дисципліни (освітнього компонента)

можливість зарахування сертифікатів проходження дистанційних чи онлайн курсів за відповідною тематикою.

Робочу програму навчальної дисципліни (силабус):

Складено професор кафедри ШІ, професор, д.ф.-м.н., Купенко Ольга Петрівна

Ухвалено кафедрою штучного інтелекту (протокол № 14 від 14.06.2024)

Погоджено Методичною комісією ННІПСА (протокол № 10 від 24.06.2024)