



# 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 – 40 год., практичні заняття – 20 год., СРС – 60 год.)</i>
Семестровий контроль/ контрольні заходи	<i>Залік, МКР</i>
Розклад занять	<i><a href="https://schedule.kpi.ua/">https://schedule.kpi.ua/</a> (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, <a href="mailto:kupenko.olga@gmail.com">kupenko.olga@gmail.com</a> Практичні: д.ф.-м.н., професор, Купенко Ольга Петрівна, <a href="mailto:kupenko.olga@gmail.com">kupenko.olga@gmail.com</a></i>
Розміщення курсу	<i>Платформа дистанційного навчання «Сікорський», додаткові ресурси за посиланням <a href="https://drive.google.com/drive/folders/1_LkkgulrHca-HI-19qblvNcxZYMnDKxz3RXL9IN4d8Fs16kj6EPdgyM9wgEKXppm0VBnEyM">https://drive.google.com/drive/folders/1_LkkgulrHca-HI-19qblvNcxZYMnDKxz3RXL9IN4d8Fs16kj6EPdgyM9wgEKXppm0VBnEyM</a></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:*

**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 басів класифікатор**

**Тема 8. Базові моделі бінарної класифікації. Метод опорних векторів**

**Тема 8. Базові моделі бінарної класифікації. Деревя 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.**

#### 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](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. Kantarizic M. Data Mining. Concepts, Models, Methods and Algorithms / M. Kantarizic, 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. Сергєєв-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.

## **ЛЕКЦІЇ**

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

Лекція 2. Функції втрат

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

Лекції 4-5. Крива ROC для оцінки якості класифікаційних моделей

Лекції 6-7. Класичні, ітераційні та евристичні методи класифікації

Лекції 8-9. Базові моделі бінарної класифікації. Логістична регресія

Лекції 10-11. Базові моделі бінарної класифікації. Наївний баєсів класифікатор

Лекції 12-13. Базові моделі бінарної класифікації. Метод опорних векторів

Лекція 14. Базові моделі бінарної класифікації. Деревя рішень

Лекція 15. Базові моделі бінарної класифікації. Нейро-мережеві моделі

Лекція 16. Підходи та методи до класифікації з незбалансованими даними. Активне навчання. Метрики для активного навчання.

Лекція 17. Ансамблі базових моделей та мета-навчання

Лекція 18. Калібрування ймовірностей

## **ПРАКТИЧНІ ЗАНЯТТЯ**

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

## **ДИДАКТИЧНІ МЕТОДИ:**

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

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

## 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), комп'ютерні практикуми.

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

Рейтингова система оцінювання включає всі види тестування: контрольні роботи, активність на практичних заняттях та якість захисту комп'ютерних практикумів. 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 практичних заняттях та захист комп'ютерних практикумів;
- написання підсумкової роботи.

Система рейтингових (вагових) балів та критеріїв оцінювання:

Метод оцінювання	Кількість	Мінімальна оцінка в балах	Максимальна оцінка в балах
Комп'ютерні практикуми	6	6	10
Модульна контрольна робота	1		20
Підсумкова робота	1		40
Підсумковий рейтинг			100

Умовами допуску до заліку є виконання всіх комп'ютерних практикумів і рейтингова оцінка не нижче 40 балів.

Студенти, які виконали всі умови допуску до заліку та мають рейтингову оцінку 60 і більше балів, отримують відповідну до набраного рейтингу оцінку без додаткових випробувань («автоматом»).

Студенти, які наприкінці семестру мають рейтинг менше 60 балів, але виконали умови допуску до заліку, а також ті, хто хоче підвищити оцінку, виконують залікову контрольну роботу. При цьому попередній рейтинг з дисципліни скасовується і до залікової відомості заносяться бали за залікову контрольну роботу («жорстка» РСО). Ця оцінка є остаточною.

Залік оцінюється у 100 балів. Завдання залікової контрольної роботи складається з двох теоретичних питань різних розділів робочої програми і двох практичних завдань. Кожне теоретичне питання і практичне завдання залікової контрольної роботи оцінюється у 25 балів:

Критерії оцінювання:

- «відмінно», повна відповідь (не менше 95% потрібної інформації) – 24-25 балів;
- «добре», достатньо повна відповідь (не менше 75% потрібної інформації) – 19-23 бали;
- «задовільно», неповна відповідь (не менше 60% потрібної інформації) – 15-18 балів;
- «незадовільно», незадовільна відповідь – 0 балів.

Таблиця відповідності рейтингових балів оцінкам за університетською шкалою:

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

## 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)