



Теорія та методи машинного навчання

Робоча програма навчальної дисципліни (Силабус)

Реквізити навчальної дисципліни

Рівень вищої освіти	<i>Другий (магістерський)</i>
Галузь знань	<i>12 Інформаційні технології</i>
Спеціальність	<i>121 Інженерія програмного забезпечення</i>
Освітня програма	<i>Інженерія програмного забезпечення інформаційних систем</i>
Статус дисципліни	<i>Нормативна</i>
Форма навчання	<i>очна(денна)</i>
Рік підготовки, семестр	<i>1 курс, осінній семестр</i>
Обсяг дисципліни	<i>5 кредитів (150 годин, з них 36 годин лекцій, 36 годин лабораторних робіт, 78 годин СРС)</i>
Семестровий контроль/ контрольні заходи	<i>залік/модульна контрольна робота/захист лабораторних робіт</i>
Розклад занять	<i>1 лекція (2 години) 1 раз на тиждень 1 лабораторна робота (2 години) 1 раз на тиждень Детальний розклад: https://schedule.kpi.ua</i>
Мова викладання	<i>Українська</i>
Інформація про керівника курсу / викладачів	<i>Лектор: ст. викл. Сарнацький Владислав Віталійович, sarnatskyi.vladyslav@iit.kpi.ua Лабораторні роботи: ст. викл. Сарнацький Владислав Віталійович, sarnatskyi.vladyslav@iit.kpi.ua</i>
Розміщення курсу	<i>https://classroom.google.com/u/2/c/NzEwMzM1MTIyNTkx</i>

Програма навчальної дисципліни

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

Метою навчальної дисципліни є отримання студентами знань з теоретичних та практичних методів машинного навчання; навичок вирішення задач регресії, класифікації, кластеризації, пониження розмірності.

Предмет навчальної дисципліни – теоретичні та практичні методи машинного навчання.

Дисципліна формує та закріплює у здобувачів такі загальні та спеціальні (фахові) компетентності:

- ЗК 01 Здатність до абстрактного мислення, аналізу та синтезу;

- ФК 14 Здатність використовувати методи машинного навчання при розробленні програмного забезпечення інформаційних систем.

Після засвоєння дисципліни студенти мають продемонструвати такі **результати навчання**:

Програмні результати навчання:

- ПРН 19 Формулювати, експериментально перевіряти, обґрунтовувати і застосовувати на практиці в процесі розроблення програмного забезпечення інноваційні методи та конкурентоспроможні технології розв'язання професійних, науково-технічних задач у мультидисциплінарних контекстах;
- ПРН 21 Знання методів машинного навчання.

Досвід з аналізу поставлених задач, вибору алгоритмів машинного навчання для їх вирішення, проектування та побудови програмних компонентів з використанням технологій машинного навчання, оцінка якості роботи систем машинного навчання.

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

Для успішного засвоєння дисципліни студент повинен володіти базовими знаннями з програмування, математичного аналізу, лінійної алгебри, теорії ймовірностей та математичної статистики, аналізу даних.

Знання, одержані студентами при вивченні дисципліни будуть використовуватися при вивченні наступних дисциплін:

ПО 2 Науково-дослідна практика

ПО 3 Виконання магістерської дисертації

3. Зміст навчальної дисципліни

ВСТУП

Вступ до дисципліни

Розділ 1 Алгоритми машинного навчання з вчителем

Тема 1.1 Лінійна та поліноміальна регресія

Тема 1.2 Логістична регресія

Тема 1.3 Метод опорних векторів

Тема 1.4 Методи на основі дерев прийняття рішень

Тема 1.5 Ансамбльові методи

Розділ 2 Алгоритми машинного навчання без вчителя

Тема 2.1 Кластеризація

Тема 2.2 Пониження розмірності

Тема 2.3 Пошук аномалій

Розділ 3 Штучні нейронні мережі

Тема 3.1 Просунуті методи математичної оптимізації

Тема 3.2 Штучні нейронні мережі

Тема 3.3 Оптимізація гіперпараметрів

Розділ 4 Проектування систем машинного навчання

Тема 4.1 Життєвий цикл систем машинного навчання

4. Навчальні матеріали та ресурси

Базова література

1. Дистанційний курс «Теорія та методи машинного навчання» для магістрів 1-го курсу спеціальності 121 Інженерія програмного забезпечення. Адреса розміщення: <https://classroom.google.com/u/2/c/NzEwMzM1MTIyNTkx>;

Додаткова література

1. Goodfellow I., Bengio Y., Courville A. Deep learning. – MIT press, 2016. Адреса розміщення: <https://www.deeplearningbook.org/>.
2. Géron A. Hands-on machine learning with Scikit-Learn, Keras, and TensorFlow: Concepts, tools, and techniques to build intelligent systems. – « O'Reilly Media, Inc.», 2019 Адреса розміщення: <https://www.rasa-ai.com/wp-content/uploads/2022/02/Aur%C3%A9lien-G%C3%A9ron-Hands-On-Machine-Learning-with-Scikit-Learn-Keras-and-Tensorflow-Concepts-Tools-and-Techniques-to-Build-Intelligent-Systems-O%E2%80%99Reilly-Media-2019.pdf>.

Навчальний контент

5. Методика опанування навчальної дисципліни

5.1 Тематика лекцій

Теми лекцій та перелік основних питань наведені в таблиці 1.

Таблиця 1 Тематика лекцій

№	Назва теми лекції та перелік основних питань	К-ть годин
1	Вступ до дисципліни. Знайомство з основними розділами дисципліни та видами робіт, що передбачено під час вивчення дисципліни. Введення базових понять.	2
2	Лінійна регресія Задача регресії. Лінійна регресія. Метод градієнтного спуску.	2
3	Поліноміальна регресія Поліноміальна регресія. Моделювання нелінійних даних.	2
4	Логістична регресія Задача класифікації. Логістична регресія.	2
5	Метод опорних векторів Метод опорних векторів. Метод ядра.	2
6	Дерево прийняття рішень Вирішення задачі класифікації методом дерева прийняття рішень.	2
7	Ансамблі Ансамблювання методів. Ліс прийняття рішень. Adaboost.	4
8	Задача кластеризації Метод к-середніх. Ієрархічна кластеризація. DBSCAN.	2
9	Задача пониження розмірності Аналіз головних компонент. Алгоритми TSNE, UMAP.	2
10	Задача пошуку аномалій Ймовірнісні методи. Ізолюючий ліс.	2
11	Методи математичної оптимізації Алгоритми SGD, RmsProp, Adam, BFGS	2

№	Назва теми лекції та перелік основних питань	К-ть годин
12	Штучні нейронні мережі-1 Означення. Алгоритм прямого розповсюдження. Алгоритм зворотнього розповсюдження.	4
13	Штучні нейронні мережі-2 Програмування нейронних мереж. Фрейворки TensorFlow, PyTorch.	2
14	Мета-оптимізація. Пошук гіперпараметрів Методи оптимізації гіперпараметрів.	2
15	Життєвий цикл систем машинного навчання Особливості проектування, розробки, розгортання та підтримування	2
16	Залік	2

5.2 Тематика лабораторних та практичних занять

Таблиця 2 Лабораторні заняття

№	Назва	Кількість ауд. годин
1	Знайомство з бібліотекою NumPy	6
2	Лінійна та поліноміальна регресія	6
3	Логістична регресія	6
4	Навчання без вчителя	8
5	Програмування штучних нейронних мереж	8
6	Виконання модульної контрольної роботи	2

6. Самостійна робота студента

У рамках самостійної роботи, студенти опрацьовують матеріал щодо деталей реалізацій методів машинного навчання, найкращих практик щодо організації коду підготовки даних, організації коду роботи алгоритмів. Усі навчальні матеріали (тексти лекцій, методичні вказівки до виконання практичних робіт) розміщені в електронному вигляді на сайті (<https://classroom.google.com>). Навчальний контент доступний із будь-якого місця в мережі Інтернет.

Таблиця 3. Питання, які виносяться на самостійне опрацювання

№	Назва теми, що виносяться на самостійне опрацювання	К-ть годин
	Тема 1.1 Програмні інструменти матричних обчислень	8
	Тема 1.2 Багатокласова логістична регресія	8
	Тема 1.4 Візуалізація та аналіз моделей на основі дерев	8
	Тема 1.5 Комбіновані методи ансамблювання	8

Тема 2.1 Просунуті методи кластеризації	8
Тема 2.2 Просунуті методи пониження розмірності	8
Тема 3.1 Методи математичної оптимізації другого порядку	6
Тема 3.2 Штучні нейронні мережі для обробки зображень	8
Тема 3.3 Програмні інструменти оптимізації гіперпараметрів	8
Тема 4.1 Основи MLOps	8

Розподіл годин за видами самостійної роботи наступний:

№ з/п	Вид самостійної роботи	Кількість годин СРС
1	Підготовка до лекційних занять	28
2	Підготовка до лабораторних занять	36
3	Підготовка до модульної контрольної роботи	8
4	Підготовка до заліку	6
	ВСЬОГО	78

Політика та контроль

7. Політика навчальної дисципліни (освітнього компонента)

Як викладач, так і студент зобов'язані дотримуватись [Кодексу честі Національного технічного університету України «Київський політехнічний інститут»](#).

Основні положення політики:

- впродовж занять студенти можуть задавати питання стосовно матеріалу, що викладається; студенти мають можливість підняти будь-яке питання, яке стосується процедури проведення та оцінювання контрольних заходів;
- студенти мають право оскаржити результати контрольних заходів, аргументовано пояснивши з яким критерієм не погоджуються відповідно до оціночного листа та/або зауважень;
- у випадку виявлення факту академічної недобросовісності робота не зараховується;
- заохочувальні бали виставляються за: активну участь на лекціях, кількість заохочуваних балів на більше 10;
- штрафні бали за невчасне виконання та здачу лабораторних робіт відсутні, студент має право здавати та захищати лабораторні роботи не пізніше останнього практичного заняття семестру.

Визнання результатів навчання, набутих у неформальній/інформальній освіті.

Порядок визнання таких результатів регламентується Положенням про визнання результатів навчання, набутих у неформальній / інформальній освіті (<https://osvita.kpi.ua/index.php/node/179>). Можуть бути зараховані окремі змістовні модулі або теми дисципліни. В такому разі здобувач звільняється від виконання відповідних завдань, отримуючи за них максимальний бал відповідно до рейтингової системи оцінювання.

Політика щодо академічної доброчесності. Обов'язковою умовою виконання завдань з освітньої компоненти є дотримання політики та принципів академічної доброчесності (<https://kpi.ua/academic-integrity>), які, у тому числі, викладено у Кодексі честі Національного технічного університету України «Київський політехнічний інститут імені Ігоря Сікорського» (<https://kpi.ua/code>), Положенні про систему запобігання академічному плагіату в КПІ ім. Ігоря Сікорського (<https://osvita.kpi.ua/node/47>). У разі виявлення дублювання робіт, плагіату роботи здобувачі отримують нульовий рейтинг.

Політика використання штучного інтелекту. Використання штучного інтелекту (далі, ШІ) регламентується «Політикою використання штучного інтелекту для академічної діяльності в КПІ ім. Ігоря Сікорського» (<https://osvita.kpi.ua/node/1225>). Усі навчальні завдання з дисципліни мають бути результатом власної оригінальної роботи здобувача.

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

8.1 Поточний контроль

Поточний контроль успішності засвоєння знань виконується шляхом виконання ними: лабораторних робіт, МКР, завдань, що видані для самостійної проробки. Таким чином, семестровий рейтинг студента з дисципліни складається з балів, що він отримує за:

- лабораторні роботи;
- одну модульну контрольну роботу;
- самостійну роботу;
- відповідь на заліку (при необхідності).

8.2 Система рейтингових (вагових) балів та критерії оцінювання

Лабораторні роботи

Вагові бали кожної лабораторної роботи наведені у таблиці 1. Сумарний ваговий бал за даний контрольний захід складає **60 балів**.

Критерії оцінювання лабораторних робіт включають якість її виконання, захисту та оформлення звіту (таблиця 4).

Таблиця 4 – Вагові бали та критерії оцінювання лабораторних робіт

Назва роботи	Бали			
	Виконання	Захист	Звіт	Сума
Знайомство з бібліотекою NumPy	6	6		12
Лінійна та поліноміальна регресія	6	6		12
Логістична регресія	6	6		12
Навчання без вчителя	6	6		12
Програмування штучних нейронних мереж	6	6		12
Разом	30	30		60

Критерії оцінювання лабораторних робіт 1-5:

“відмінно” – робота виконана та захищена без зауважень, максимальний балів;

“добре” – достатньо повне виконання роботи з деякими похибками, 80% від максимальної кількості балів;

“задовільно” – неповна відповідь (не менше 60% потрібної інформації), 60% від максимальної кількості балів;

“незадовільно” – при виконанні або під час захисту роботи були виявлені помилки, -0 балів.

Модульна контрольна робота

Модульна контрольна робота складається з 2-х частин: перелік тестових питань та завдання на розробку системи машинного навчання з подальшим захистом та виконується студентами протягом 2-х аудиторних годин.

Ваговий бал кожної частини МКР – 20 балів. Ваговий бал за даний контрольний захід для однієї МКР – 20×2 бали = 40 балів.

Критерії оцінювання першої частини МКР:

Кожне з 20 тестових питань оцінюється на 1 бал за вірну відповідь і 0 за невірну.

Критерії оцінювання другої частини МКР:

- “відмінно”, робоча система і повна відповідь (не менше 90% потрібної інформації) –20 балів;
- “добре”, робоча система і достатньо повна відповідь (не менше 75% потрібної інформації), або повна відповідь з незначними помилками – 15-19,5 балів;
- “задовільно”, робоча система і неповна відповідь (не менше 60% потрібної інформації) та незначні помилки – 12-15 балів;
- “незадовільно”, неробоча система – 0 балів.

Заохочувальні бали за:

- виконання додаткових завдань – 10 балів.

Умови позитивної проміжної атестації

За результатами навчальної роботи за перші 7 тижнів «ідеальний студент» має набрати 24-бали. На першій атестації (8-й тиждень) студент отримує «зараховано», якщо його поточний рейтинг не менше 13 балів.

За результатами 13 тижнів навчання «ідеальний студент» має набрати 48-бал. На другій атестації (14-й тиждень) студент отримує «зараховано», якщо його поточний рейтинг не менше 25 балів.

Розрахунок шкали рейтингу R :

Максимальна сума вагових балів контрольних заходів протягом семестру складає:

$$R_c = 60 + 40 = 100 \text{ балів.}$$

Необхідною умовою отримання заліку є:

- виконання 5 лабораторних робіт на оцінку не нижче ніж “задовільно”;
- виконання МКР не нижче ніж на оцінку “задовільно” (тобто не менше 20 балів);
- стартовий рейтинг (r_c) не менше 60% від R , тобто 60 балів.

Студенти, які наприкінці семестру мають рейтинг менше 60 балів, а також ті, хто хоче підвищити оцінку, виконують залікову контрольну роботу. При цьому існуючі бали анулюються і отримана рейтингова оцінка за залікову контрольну роботу стає остаточною. Завдання залікової контрольної роботи складаються з одного практичного завдання з різних розділів дисципліни.

Критерії оцінювання залікової контрольної роботи наведені в таблиці 5. Критерії застосовуються, якщо програмне забезпечення компілюється та запускається. Інакше -0

Таблиця 5. Критерії оцінювання залікової контрольної роботи

Вирішення задачі класифікації з використанням методів машинного навчання	
Організовані дані для навчання	20
Виконана оптимізація гіперпараметрів	20
Самостійно реалізована модель	20
Застосований метод ансамблювання	20
Досягнена точність >80%	20

Для отримання студентом відповідних оцінок за університетською шкалою його рейтингова оцінка R_D переводиться згідно з таблицею 6.

Таблиця 6

Бали	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)

Перелік питань, які виносяться на семестровий контроль розміщений в системі «Електронний кампус» та оприлюднюються в Google Classroom.

Робочу програму навчальної дисципліни (силабус):

Складено ст. викл. Сарнацький Владислав Віталійович

Ухвалено кафедрою ІПІ (протокол №16 від 29.05.2024)

Погоджено Методичною комісією факультету (протокол № 10 від 21.06.2024)