

СИЛАБУС НАВЧАЛЬНОЇ ДИСЦИПЛІНИ

«Машинне навчання»



Ступінь освіти	Магістр
Галузь знань	F Інформаційні технології
Спеціальність	F1 Інженерія програмного забезпечення; F4 системний аналіз та наука про дані
Тривалість викладання	1 семестр
Заняття:	весняний семестр (3 та 4 чверть)
Лекції	2 год./тижд.
Практичні	1 год./тижд.
Мова викладання	українська

Передумови для вивчення: вивчення дисципліни «Машинне навчання» у встановлених відповідною робочою програмою обсягах передбачає розуміння математичного аналізу, алгебри, аналізу даних.

Сторінка курсу в СДО НТУ «ДП»:

<https://do.nmu.org.ua/course/view.php?id=6019>

Інші додаткові ресурси:

Консультації: за окремим розкладом, погодженим зі здобувачами вищої освіти

Онлайн-консультації: MS Teams, електронна пошта.

Інформація про викладача



Викладач:

Владико Олександр Борисович

Кандидат технічних наук, доцент кафедри САУ

Посилання на профілі:

[Сторінка кафедри САУ](#)

[Scopus iD](#)

[ORCID iD](#)

[Google Scholar](#)

1. Анотація до курсу

Особливістю методів машинного навчання є не прямий розв'язок задачі, а навчання на множині подібних прикладів, що дозволяє використовувати ці методи для обробки великих обсягів даних та виявляти в них нові, нетривіальні, корисні та доступні для інтерпретації знання. Окрім методів штучного інтелекту, під час розробки моделей машинного навчання в якості допоміжних використовуються

засоби математичної статистики, чисельних методів, методів оптимізації, теорії ймовірностей, теорії графів, різні техніки роботи з даними в цифровій формі. Машинне навчання застосовують в ряді обчислювальних задач, в яких розробка та програмування явних алгоритмів з доброю продуктивністю є складною або нездійсненною задачею. Для практичного засвоєння навчальних матеріалів ряд тем дисципліни поглиблено вивчається на лабораторних заняттях.

2. Мета та завдання навчальної дисципліни

Мета дисципліни – сформувати у магістрів навички застосування та використання сучасних технологій обробки даних для вирішення задач класифікації, регресійного аналізу, прогнозування та ухвалення рішень.

Завдання курсу:

- теоретичних основ базових технологій машинного навчання;
- методів, алгоритмів роботи, засобів реалізації та технологій налагодження систем обробки даних з використанням технологій обчислювального інтелекту;
- сучасних програмних засобів реалізації технологій машинного навчання.

3. Результати навчання

1. Знати теоретичні основи базових технологій машинного навчання; методів, алгоритмів роботи, засобів реалізації та технологій налагодження систем обробки даних з використанням технологій обчислювального інтелекту; сучасних програмних засобів реалізації технологій машинного навчання.
2. Вміти обґрунтовано вибирати конкретні технології та алгоритми машинного навчання при розв'язанні відповідних практичних задач; здійснювати підготовку та первинну обробку даних для побудови моделей систем методами машинного навчання; вирішувати задачі автоматизації підтримки рішень, розпізнавання образів, діагностики, класифікації, оптимізації та аналізу даних методами машинного навчання; використовувати сучасні програмні засоби для реалізації технологій машинного навчання.
3. Вміти проводити попередню обробку даних різної природи та інтерполювати невідомі значення

4. Структура курсу

Види та тематика навчальних занять	Внесок в загальну оцінку, %
ЛЕКЦІЇ	30
1. Машинне навчання. Вступ. Постановка задачі машинного навчання. Лінійні моделі. Поняття перенавчання. Способи оцінювання ступеня перенавчання моделей.	6
2. Задачі бінарної класифікації. Рівняння гіперплощини у задачах бінарної класифікації. Розв'язання простого завдання бінарної класифікації. Функції втрат у задачах лінійної бінарної класифікації.	8

3. Формула Байєса. Формула Байєса під час вирішення конкретних завдань. Байєсовський висновок. Наївна байєсівська класифікація. Гаусівський байєсівський класифікатор. Лінійний дискримінант Фішера.	8
4. Методи SVM, PCA. Введення метод опорних векторів (SVM) Реалізація методу опорних векторів (SVM) Метод опорних векторів (SVM) із нелінійними ядрами. Імовірна оцінка якості моделей.	8
5. Методи головних компонентів. Показники precision та recall. F-міра. Метрики якості ранжирування. ROC-крива. Метод головних компонентів (Principal Component Analysis). Скорочення розмірності простору ознак за допомогою PCA. Сингулярне розкладання та його зв'язок з PCA.	8
5. Методи класифікації. Багатокласова класифікація. Методи one-vs-all та all-vs-all. Метричні методи класифікації. Метод k найближчих сусідів. Методи парзенівського вікна та потенційних функцій. Метричні регресійні методи. Формула Надарая-Ватсона.	8
6. Алгоритми кластеризації. Завдання кластеризації. Постановка задачі. Алгоритм кластеризації Ллойда (K-середніх, K-means). Алгоритм кластеризації DBSCAN. Агломеративна ієрархічна кластеризація. Дендограма.	8
7. Вирішальні дерева. Логічні методи класифікації. Критерії якості для побудови вирішальних дерев. Побудова вирішальних дерев жадібним алгоритмом ID3. Усічення (pruning) дерева, обробка перепусток та категоріальних ознак. Вирішальні дерева у задачах регресії. Алгоритм CART. Випадкові дерева та випадковий ліс. Бутстреп та бегінг.	8
8. Бустінг. Введення у бустінг (boosting). Алгоритм AdaBoost під час класифікації. Алгоритм AdaBoost у задачах регресії. Градієнтний бустінг та стохастичний градієнтний бустінг. Нейронні сітки. Короткий вступ до теорії. Навчання нейронної мережі. Алгоритм back propagation.	8
ПРАКТИЧНІ ЗАНЯТТЯ	70
1. Завдання лінійної бінарної класифікації. Мета: у середовищі Python, створити алгоритм бінарної класифікації.	10
2. Реалізація гаусівського байєсівського класифікатора. Мета: у середовищі Python, створити та вирішити алгоритм гаусівського байєсовського класифікатора.	12
3. Апроксимація даних ядерним згладжуванням. Мета: у середовищі Python, провести апроксимацію даних ядерним згладжуванням.	12
4. Алгоритм кластеризації Ллойда (K-середніх). Мета: у середовищі Python, створити алгоритм кластеризації Ллойда.	12
5. Алгоритм класифікації CART на вирішальних деревах. Мета: у середовищі Python, створити алгоритм класифікації CART на вирішальних деревах.	12
6. Алгоритм класифікації AdaBoost на вирішальних деревах. Мета: у середовищі Python, створити Алгоритм класифікації AdaBoost.	12
Разом	100

5. Технічне обладнання та/або програмне забезпечення

Технічні засоби навчання: мультимедійні та комп'ютерні пристрої.

Засоби дистанційної освіти: Moodle, MS Teams.

Пакети прикладних програм: Python, бібліотеки: Keras, TensorFlow (без-коштовні).

6. Система оцінювання та вимоги

6.1. Навчальні досягнення здобувачів вищої освіти за результатами вивчення курсу оцінюватимуться за шкалою, що наведена нижче:

Рейтингова шкала	Інституційна шкала
90 – 100	відмінно
75-89	добре
60-74	задовільно
0-59	незадовільно

6.2. Здобувач ступеня освіти «Магістр» може отримати підсумкову оцінку з навчальної дисципліни на підставі поточного оцінювання знань за умови, якщо набрана кількість балів з поточного тестування та самостійної роботи складатиме не менше 60 балів. Поточна успішність складається з успішності за теоретичну частину курсу (максимум – 36 балів) та оцінок за виконання практичних робіт (максимум 8 балів за кожну роботу та максимальною сумарною оцінкою за всі роботи – 64 бали). Отримані бали за теоретичну частину курсу та практичні роботи додаються і є підсумковою оцінкою за вивчення навчальної дисципліни. Максимально за поточною успішністю здобувач вищої освіти може набрати 100 балів.

Шкала оцінювання (зазначено максимально можливі бали):

Теоретична частина	Практичні роботи		Разом
	При своєчасному складанні	При несвоєчасному складанні	
30	40	30	100

6.3. Критерії оцінювання поточного та підсумкового контролю:

– підсумкове оцінювання відбувається у формі диференційованого заліку у форматі тесту, який складається з 16 завдань (15 запитань із вибором варіанту відповіді – 2 бали за правильну відповідь; 1 завдання у формі задачі – максимум 6 балів, якщо надано повністю правильну і обґрунтовану відповідь);

– поточне оцінювання практичних робіт відбувається шляхом захисту звіту з відповідної роботи (максимальний бал – 8, який формується наступним чином: 50 % – правильність і повнота викладення матеріалу в звіті, 50 % – захист індивідуальної роботи шляхом відповіді на контрольні питання).

6.4. Критерії оцінювання практичної роботи

З кожної практичної роботи здобувач вищої освіти отримує 5 запитань з переліку контрольних запитань. Кількість вірних відповідей визначають кількість отриманих балів.

7. Політика курсу

7.1. Політика щодо академічної доброчесності. Академічна доброчесність здобувачів вищої освіти є важливою умовою для опанування результатами навчання за дисципліною і отримання задовільної оцінки з поточного та підсумкового контролів. Академічна доброчесність базується на засудженні практик списування (виконання письмових робіт із залученням зовнішніх джерел інформації, крім дозволених для використання), плагіату (відтворення опублікованих текстів інших авторів без зазначення авторства), фабрикації (вигадування даних чи фактів, що використовуються в освітньому процесі). Політика щодо академічної доброчесності регламентується положенням «Положення про систему запобігання та виявлення плагіату у Національному технічному університеті «Дніпровська політехніка».

У разі порушення здобувачем вищої освіти академічної доброчесності (списування, плагіат, фабрикація), робота оцінюється незадовільно та має бути виконана повторно. При цьому викладач залишає за собою право змінити тему завдання.

7.2. Комунікаційна політика. Здобувачі вищої освіти повинні мати активовану університетську пошту.

Усі письмові запитання до викладачів стосовно курсу мають надсилатися на університетську електронну пошту.

7.3. Політика щодо перескладання. Роботи, які здаються із порушенням термінів без поважних причин оцінюються на нижчу оцінку. Перескладання відбувається із дозволу деканату за наявності поважних причин (наприклад, лікарняний).

7.4. Відвідування занять. Для студентів денної форми відвідування занять є обов'язковим. Поважними причинами для неявки на заняття є хвороба, участь в університетських заходах, відрядження, які необхідно підтверджувати документами у разі тривалої (два тижні) відсутності. Про відсутність на занятті та причини відсутності студент має повідомити викладача або особисто, або через старосту. Якщо студент захворів, ми рекомендуємо залишатися вдома і навчатися за допомогою дистанційної платформи. Студентам, чий стан здоров'я є незадовільним і може вплинути на здоров'я інших студентів, буде пропонуватися залишити заняття (така відсутність вважатиметься пропуском з причини хвороби). Лабораторні заняття не проводяться повторно, ці оцінки неможливо отримати під час консультації.

За об'єктивних причин (наприклад, міжнародна мобільність) навчання може відбуватись в онлайн формі за погодженням з керівником курсу.

7.5. Участь в анкетуванні. Наприкінці вивчення курсу та перед початком сесії студентам буде запропоновано анонімно заповнити електронні анкети (MS Office 365), які буде розіслано на ваші університетські поштові скриньки. Заповнення анкет є важливою складовою вашої навчальної активності, що дозволить оцінити дієвість застосованих методів викладання та врахувати ваші пропозиції стосовно покращення змісту навчальної дисципліни.

8. РЕКОМЕНДОВАНІ ДЖЕРЕЛА ІНФОРМАЦІЇ

Базові

1. Murphy, K. P. (2012). Machine Learning: a Probabilistic Perspective. MIT Press, Cambridge, MA, USA
2. Farabet, C., LeCun, Y., Kavukcuoglu, K., Culurciello, E., Martini, B., Akselrod, P., and Talay, S. (2011). Large-scale FPGA-based convolutional networks. In R. Bekkerman, M. Bilenko, and J. Langford, editors, Scaling up Machine Learning: Parallel and Distributed Approaches. Cambridge University Press.
3. Машинне навчання [Електронний ресурс]: конспект лекцій для здобувачів ступеня магістра зі спеціальності 124 Системний аналіз / уклад.: Т.А. Желдак, О.Б. Владико ; М-во освіти і науки України, Нац. техн. ун-т «Дніпровська політехніка». – Дніпро : НТУ «ДП», 2024. – 162 с
4. Машинне навчання [Електронний ресурс] : методичні рекомендації до виконання практичних робіт для здобувачів ступеня магістра зі спеціальності 124 Системний аналіз / уклад.: Т.А. Желдак, О.Б. Владико, А.В. Малієнко, Д.М. Гаранжа ; М-во освіти і науки України, Нац. техн. ун-т «Дніпровська політехніка». – Дніпро : НТУ «ДП», 2024. – 51 с

Додаткові

1. А. И. Черняк, П. В. Захарченко, Интеллектуальный анализ данных: Підручник / О.І. Черняк, Київський національний університет імені Т. Шевченко. К.: Знання, 2014. 599 с.
2. Kevin P. Murphy [«Machine Learning: A Probabilistic Perspective»](#), 2012
3. Кононова К. Ю. Машинне навчання: методи та моделі: підручник для бакалаврів, магістрів та докторів філософії спеціальності 051 «Економіка» / К. Ю. Кононова. – Харків: ХНУ імені В. Н. Каразіна, 2020. – 301 с.