



МАШИННЕ 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 мультимедійних та інформаційно-пошукових систем</i>
Статус дисципліни	<i>Вибіркова</i>
Форма навчання	<i>Очна (денна)</i>
Рік підготовки, семестр	<i>1 рік навчання, 2 семестр</i>
Обсяг дисципліни	<i>Лекції: 36 год., комп'ютерний практикум: 18 год., самостійна робота: 96 год.</i>
Семестровий контроль/ контрольні заходи	<i>Екзамен, модульна контрольна робота, календарний контроль</i>
Розклад занять	<i>Згідно розкладу поточного навчального року (rozklad.kpi.ua)</i>
Мова викладання	<i>Українська</i>
Інформація про керівника курсу / викладачів	<i>Лектор: к.т.н., доцент, Олещенко Любов Михайлівна, oleshchenkoliubov@gmail.com Комп'ютерний практикум: к.т.н., доцент, Олещенко Любов Михайлівна, oleshchenkoliubov@gmail.com</i>
Розміщення курсу	<i>Google classroom: https://classroom.google.com/c/MzE5MjA3ODU2NjY3</i>

Програма навчальної дисципліни

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

Вивчення дисципліни «Машинне навчання» сприяє формуванню у здобувачів вищої освіти компетенцій, необхідних для розв'язання практичних задач професійної діяльності, пов'язаної з аналізом даних за допомогою технологій машинного навчання.

***Метою** вивчення дисципліни «Машинне навчання» є формування у студентів здатностей використовувати технології машинного навчання для розроблення програмного забезпечення для кластеризації, класифікації даних, розпізнавання зображень та прогнозування часових рядів.*

***Предметом** дисципліни «Машинне навчання» є технології машинного навчання для кластеризації, класифікації даних, розпізнавання зображень та прогнозування часових рядів.*

*Після засвоєння дисципліни «Машинне навчання» **результатами навчання** є:*

знання:

- методів аналізу даних за допомогою технологій машинного навчання.*

уміння:

- створювати програмне забезпечення для аналізу даних за допомогою технологій машинного навчання;
- використовувати можливості мови програмування Python для кластеризації, класифікації даних, розпізнавання зображень та прогнозування часових рядів.

досвід:

- розроблення програмного забезпечення, яке дозволяє виконувати кластеризацію, класифікацію та прогнозування для вирішення практичних завдань в різних галузях людської діяльності.

Вивчення дисципліни «Машинне навчання» сприяє формуванню у здобувачів вищої освіти, які навчаються за освітньою програмою «Інженерія програмного забезпечення мультимедійних та інформаційно-пошукових систем» компетентностей, необхідних для розв'язання практичних задач професійної діяльності, пов'язаної з використанням методів машинного навчання:

ЗК01 Здатність до абстрактного мислення, аналізу та синтезу.

ЗК03 Здатність проводити дослідження на відповідному рівні.

ФК02 Здатність розробляти і реалізовувати наукові та/або прикладні проекти у сфері інженерії програмного забезпечення.

ПРН32 Знати підходи, напрямки, моделі та методи штучного інтелекту, у тому числі машинного навчання; знати технології розроблення програмного забезпечення систем штучного інтелекту, застосовувати методи штучного інтелекту у дослідницькій діяльності та для розв'язання прикладних задач.

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

Успішному вивченню дисципліни «Машинне навчання» передують вивчення дисциплін «Комп'ютерна дискретна математика», «Математичний аналіз», «Лінійна алгебра та аналітична геометрія», «Теорія ймовірностей», «Програмування» навчального плану підготовки бакалаврів за спеціальністю 121 «Інженерія програмного забезпечення».

Дисципліна «Машинне навчання» забезпечує виконання курсових проектів та магістерських дисертацій за спеціальністю 121 «Інженерія програмного забезпечення».

3. Зміст навчальної дисципліни

Дисципліна «Машинне навчання» передбачає вивчення таких тем:

Тема 1. Математичні моделі та програмні інструменти машинного навчання.

Тема 2. Застосування нейронних мереж.

Екзамен.

4. Навчальні матеріали та ресурси

Базова література:

1. Олещенко Л.М. Машинне навчання: комп'ютерний практикум / Л.М. Олещенко; КПІ ім. Ігоря Сікорського. – Електронні текстові дані. – Київ: КПІ ім. Ігоря Сікорського, 2022. – 92 с.
2. Кононова К. Ю. Машинне навчання: методи та моделі: підручник. – Харків: ХНУ ім. В. Н. Каразіна, 2020. – 301 с.
3. Кононюк А.Ю. Нейронні мережі і генетичні алгоритми – К.: «Корнійчук», 2008. – 446 с.
4. Могильний С. Б. Машинне навчання з використанням мікрокомп'ютерів: навч.-метод. посіб. / за ред. О. В. Лісового та ін. – К., 2019. – 226 с.

Додаткова література:

1. Alex Smola and Vishwanathan S.V.N. Introduction to Machine Learning. Електронний ресурс. Режим доступу: <https://alex.smola.org/drafts/thebook.pdf>
2. Dipranjan Sarkar, Raghav Bali, Tushar Sharma. Practical Machine Learning with Python. Електронний ресурс. Режим доступу: <https://link.springer.com/book/10.1007%2F978-1-4842-3207-1>
3. Utilizing Python for Agent-Based Modeling: The Mesa Framework. Електронний ресурс. Режим доступу: https://www.researchgate.net/publication/344675633_Utilizing_Python_for_Agent-Based_Modeling_The_Mesa_Framework
4. Computer Vision: Foundations and Applications / Compiled by Ranjay Krishna, published by Stanford University, 2017. – 213 p.
5. Oleshchenko L.M., Lysenko O.O. Software method for clustering software testing reports using KNN algorithm. Вчені записки Таврійського національного університету імені В. І. Вернадського. Серія: Технічні науки. Том 31 (70) №4. 2020. С. 106-110.
http://www.tech.vernadskyjournals.in.ua/journals/2020/4_2020/17.pdf
6. Dychka, I., Legeza, V., Oleshchenko, L., Bohutskiy, D. Method Simultaneous Using GAN and RNN for Generating Web Page Program Code from Input Image. Advances in Intelligent Systems and Computing. 2021. Vol. 1247. P. 338–349.

Усі навчальні матеріали знаходяться у вільному доступі в мережі Інтернет.

Навчальний контент

5. Методика опанування навчальної дисципліни (освітнього компонента)

№ з/п	Тип навчального заняття	Опис навчального заняття
Тема 1. Математичні моделі та програмні інструменти машинного навчання.		
1	Лекція 1. Історія розвитку машинного навчання.	Історія розвитку машинного навчання. Предиктивний аналіз та задачі машинного навчання. Етапи наукового дослідження. Помилки в предиктивному аналізі. Оцінка результатів моделей машинного навчання. Типи машинного навчання.
2	Лекція 2. Аналіз статистичних залежностей та прогнозні моделі.	Лінійна кореляція та лінійна регресія. Оцінка статистичних залежностей та прогнозні моделі. Поліноміальна регресія. Мультиколінеарність. Логістична регресія.
3	Лекція 3. Використання бібліотек Python для аналізу даних.	Моделювання та прогнозування основних статистичних залежностей. Використання бібліотек Python для аналізу даних.
4	Лекція 4. Програмні методи машинного навчання.	Середовище Jupyter Notebook. Інструменти Python для машинного навчання. Можливості CPU та GPU середовища Google Colab для аналізу та обробки даних, зберігання та поширення наукових результатів досліджень.
5	Лекція 5. Прогнозування часових рядів в Python.	Прогнозування часових рядів в Python. Модель ARIMA.
6	Лекція 6. Методи кластеризації даних.	Методи кластеризації. Приклади реалізації алгоритму K-means у Python.

7	Комп'ютерний практикум 1. Використання алгоритму кластеризації K-Means для аналізу даних у Python.	Завдання: реалізувати алгоритм кластеризації K-Means та алгоритм K-Means ++ для ініціалізації центроїдів у Python, побудувати криву ліктя, щоб визначити потрібну кількість кластерів для набору даних формату csv задачі сегментації споживачів.
8	Лекція 7. Методи асоціації та класифікації даних.	Методи асоціації та класифікації. Метод опорних векторів.
9	Лекція 8. Деревя рішень.	Деревя рішень. Random Forest.
10	Комп'ютерний практикум 2. Прогнозування на основі класифікатора дерева рішень.	Завдання: створити класифікатор дерева рішень, який працюватиме з набором даних, який містить дані про понад 1300 сотень пасажирів у форматі csv, використовуючи бібліотеки Python: pandas, sklearn та застосунок Graphviz.
11	Лекція 9. Агентне моделювання.	Агентне моделювання. Програмне забезпечення для агентного моделювання поведінки населення у надзвичайних ситуаціях.
Тема 2. Застосування нейронних мереж.		
12	Лекція 10. Історія розвитку нейронних мереж.	Нейронні мережі. Історія розвитку нейронних мереж.
13	Лекція 11. Архітектура та алгоритм роботи нейронної мережі.	Моделі та компоненти штучного нейрону. Передатна функція (функція активації). Типи та характеристики функцій активації нейронної мережі. Архітектури з'єднань штучних нейронів. Процес навчання нейронної мережі. Алгоритм роботи нейронної мережі. Епоха навчання НМ. Метод градієнтного спуску. Перенавчання нейронної мережі.
14	Лекція 12. Типи нейронних мереж.	Типи нейронних мереж та їх застосування. Можливості бібліотеки Keras.
15	Лекція 13. Згорткова нейронна мережа CNN.	Згорткова нейронна мережа (Convolutional neural networks, CNN). Приклади застосування згорткових нейронних мереж.
16	Комп'ютерний практикум 3. Використання згорткової нейронної мережі CNN для задачі класифікації зображень.	Завдання: використати бібліотеки Python для завантаження, дослідження та аналізу даних, змінити розмір та масштаб зображень, перетворити мітки у вектори кодування та розділити дані на набори для навчання та перевірки моделі; побудувати модель нейронної мережі: навчити та оцінити модель, побудувати графіки точності та втрат; оцінити нову модель та порівняти результати; зробити прогнози щодо даних випробувань, перетворити ймовірності в мітки класів і побудувати тестові зразки, які модель класифікує правильно, візуалізувати звіт про класифікацію даних.

17	Лекція 14. Рекурентна нейронна мережа RNN.	Рекурентна нейронна мережа (Recurrent Neural Networks, RNN). Приклади застосування рекурентних нейронних мереж.
18	Лекція 15. Довга короткочасна пам'ять LSTM.	Довга короткочасна пам'ять (Long Short Term Memory, LSTM). Приклади застосування довгої короткочасної пам'яті.
19	Лекція 16. Бібліотека програмного забезпечення для машинного навчання TensorFlow.	Призначення і можливості бібліотеки TensorFlow. Sentiment analysis твітів за допомогою TensorFlow. Моделювання логістичної регресії з використанням MNIST у TensorFlow.
20	Лекція 17. Комп'ютерний зір та його застосування. Можливості бібліотеки OpenCV.	Комп'ютерний зір та його застосування. Типові завдання комп'ютерного зору. Специфікації OpenVX. Можливості бібліотеки OpenCV.
21	Лекція 18. Підсумкове заняття.	Модульна контрольна робота.

6. Самостійна робота студента

Дисципліна «Машинне навчання» ґрунтується на самостійній підготовці до аудиторних занять на теоретичні та практичні теми.

№ з/п	Назва теми, що виноситься на самостійне опрацювання	Кількість годин	Література
1	Підготовка до лекції 1	2	2, стор. 8-13.
2	Підготовка до лекції 2	2	2, стор. 15-29.
3	Підготовка до лекції 3	2	3, 4, стор. 10-26.
4	Підготовка до лекції 4	2	1 (дод.), стор. 60-67.
5	Підготовка до лекції 5	2	1 (дод.), стор. 68-81.
6	Підготовка до лекції 6	2	2, стор. 107-118.
7	Підготовка до комп'ютерного практикуму 1	8	1, стор. 5-35. 2, стор. 118-127.
8	Підготовка до лекції 7	2	2, стор. 59-88.
9	Підготовка до лекції 8	2	2, стор. 95-97.
10	Підготовка до комп'ютерного практикуму 2	8	1, стор. 36-56. 2, стор. 97-106.
11	Підготовка до лекції 9	2	2 (дод.), стор. 308-317.
12	Підготовка до лекції 10	2	4, стор. 33-36.
13	Підготовка до лекції 11	2	4, стор. 66-86.
14	Підготовка до лекції 12	2	4, стор. 63-66.
15	Підготовка до лекції 13	2	4, стор. 38-46.
16	Підготовка до комп'ютерного практикуму 3	8	1, стор. 57-92. 2, стор. 210-217.
17	Підготовка до лекції 14	2	2, стор. 201-204.

18	Підготовка до лекції 15	2	2, стор. 205-210.
19	Підготовка до лекції 16	2	1 (дод.), стор. 12-56.
20	Підготовка до лекції 17	2	3 (дод.), стор. 17-81, стор. 161-208.
21	Підготовка до модульної контрольної роботи	8	1, стор. 5-56, 2, стор. 8-127, 210-217, 4, стор. 33-46, 1 (дод.), стор. 12-56, 60-81. 3 (дод.), стор. 17-81, 161-208.
23	Підготовка до екзамену	30	1, стор. 5-92. 2, стор. 8-217, 4, стор. 33-86, 1 (дод.), стор. 12-81, 3 (дод.), стор. 17-81, 161-208.

Політика та контроль

7. Політика навчальної дисципліни (освітнього компонента)

- Відвідування лекційних занять є обов'язковим.
- Відвідування занять комп'ютерного практикуму може бути епізодичним та за потреби захисту робіт комп'ютерного практикуму.
- Правила поведінки на заняттях: активність, повага до присутніх, відключення телефонів.
- Дотримання політики академічної доброчесності.

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

Протягом семестру студенти виконують **3 комп'ютерні практикуми**. Максимальна кількість балів за кожний комп'ютерний практикум: 10 балів.

Бали нараховуються за:

- якість виконання комп'ютерного практикуму: 0-4 бали;
- відповідь під час захисту комп'ютерного практикуму: 0-4 бали;
- своєчасне представлення роботи до захисту: 0-2 бали.

Максимальна кількість балів за виконання та захист комп'ютерних практикумів:
10 балів × 3 комп. практ. = 30 балів.

Завдання на **модульну контрольну роботу** складається з 5 питань – 3 теоретичних та 2 практичних завдань. Відповідь на кожне теоретичне запитання оцінюється 4 балами, відповідь на практичне запитання оцінюється 4 балами.

Критерії оцінювання кожного теоретичного/практичного запитання модульної контрольної роботи:

- 4 бали – відповідь вірна, повна, добре аргументована;
- 3 бали – відповідь вірна, але неповна або погано аргументована;
- 2 бали – у відповіді є незначні помилки;
- 1 бали – у відповіді є суттєві помилки;
- 0 балів – немає відповіді або відповідь невірна.

Максимальна кількість балів за кожну модульну контрольну роботу:
4 бали × 3 теоретичні запитання + 4 бали × 2 практичні запитання = 20 балів.

Семестрова складова рейтингової шкали $R_c = 50$ балів, вона визначається як сума балів, отриманих за виконання та захист комп'ютерних практикумів і результатів модульного контролю.

Екзаменаційна складова рейтингової шкали $R_E = 50$ балів.

Завдання на екзамен складається з 5 питань – 3 теоретичних та 2 практичних завдань.

Відповідь на кожне теоретичне та практичне запитання оцінюється 10 балами.

Критерії оцінювання теоретичного/практичного питання екзаменаційної роботи:

9-10 балів – відповідь вірна, повна, добре аргументована;

7-8 балів – відповідь вірна, але погано аргументована ;

5-6 балів – відповідь вірна, але неповна та погано аргументована;

3-4 бали – у відповіді є незначні помилки;

1-2 бали – у відповіді є суттєві помилки;

0 балів – немає відповіді або відповідь невірна.

Максимальна кількість балів за екзаменаційну роботу:

10 балів × 3 теоретичні запитання + 10 балів × 2 практичні запитання = 50 балів.

Рейтингова шкала з дисципліни за семестр дорівнює:

$R_c = 30$ балів + 20 балів = 50 балів.

За описом: $R = R_{\text{комп.практ}} + R_{\text{МКР}} + R_{\text{екз.}} = 30 + 20 + 50$ балів = 100 балів

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

На першій атестації (8-й тиждень) студент отримує «зараховано», якщо його поточний рейтинг не менше 50 % від максимальної кількості балів, яку може отримати студент до першої атестації (10 балів).

На другій атестації (14-й тиждень) студент отримує «зараховано», якщо його поточний рейтинг не менше 50 % від максимальної кількості балів, яку може отримати студент до другої атестації (15 балів).

Семестровий контроль: **екзамен.**

Умови допуску до семестрового контролю:

При семестровому рейтингу (r_c) не менше 60% (30 балів) та зарахуванні усіх робіт комп'ютерного практикуму.

Необхідною умовою допуску до екзамену є виконання і захист комп'ютерного практикуму.

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)

Перелік питань, які виносяться на семестровий контроль, наведено у Додатку 1.

Робочу програму навчальної дисципліни (силабус):

Складено к.т.н., доц. Олещенко Л.М.

Ухвалено кафедрою ПЗКС (протокол №8 від 25.01.23)

Погоджено Методичною комісією факультету прикладної математики (протокол № 6 від 27.01.2023)

Додаток 1. Перелік питань, які виносяться на семестровий контроль

Модуль 1

- 1. Історія розвитку машинного навчання.*
- 2. Предиктивний аналіз та задачі машинного навчання.*
- 3. Типи машинного навчання.*
- 4. Моделювання та прогнозування основних статистичних залежностей. Бібліотеки Python для аналізу даних.*
- 5. Інструменти Python для машинного навчання.*
- 6. Прогнозування часових рядів в Python. Модель ARIMA.*
- 7. Методи кластеризації. Приклади реалізації алгоритму K-means у Python.*
- 8. Методи асоціації та класифікації. Метод опорних векторів.*
- 9. Дерева рішень. Random Forest.*
- 10. Агентне моделювання. Приклади застосування.*

Модуль 2

- 1. Нейронні мережі. Історія розвитку нейронних мереж.*
- 2. Типи нейронних мереж.*
- 3. Моделі та компоненти штучного нейрону.*
- 4. Передатна функція (функція активації). Типи та характеристики функцій активації нейронної мережі.*
- 5. Архітектури з'єднань штучних нейронів.*
- 6. Процес навчання нейронної мережі.*
- 7. Алгоритм роботи нейронної мережі. Пояснити, що таке епоха навчання НМ.*
- 8. Метод градієнтного спуску в навчанні нейронних мереж, його призначення.*
- 9. Перенавчання нейронної мережі.*
- 10. Convolutional neural networks (CNN).*
- 11. Recurrent Neural Networks (RNN).*
- 12. Long Short Term Memory (LSTM).*
- 13. Призначення TensorFlow. Sentiment analysis twimiv за допомогою TensorFlow.*
- 14. Комп'ютерний зір та його застосування.*
- 15. Типові завдання комп'ютерного зору.*
- 16. Специфікації OpenVX. Можливості бібліотеки OpenCV.*