



Національна академія наук України
Інститут фізіології ім. О.О. Богомольця
Силабус навчальної дисципліни

"ВИКОРИСТАННЯ МЕТОДІВ МАШИННОГО НАВЧАННЯ В
БІОМЕДИЧНИХ ДОСЛІДЖЕННЯХ"

МВК30

Галузь знань І Охорона здоров'я та соціальне
забезпечення (22 Охорона здоров'я)

Спеціальність І2 (222) Медицина

Ступінь освіти Доктор філософії
Освітньо-наукова Медицина (Патологічна фізіологія)
програма

Статус Навчальна дисципліна вибіркового
компонента з фахового переліку

Форма навчання Денна / заочна

Семестровий контроль Диференційований залік

Курс	1
Семестр	2

ECTS	3
Годин	90

Розподіл годин

Аудиторні години		Самостійна робота
Лекції	Практичні/Семінари	
6	20	64

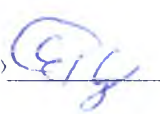
Інформація про викладача

	Лекція	Практичні/семінарські
ПІБ	Меженський Олег Русланович	Меженський Олег Русланович
Вчене звання	—	—
Науковий ступінь	Доктор філософії	Доктор філософії
Профіль викладача	https://orcid.org/0009-0003-1151-3100	https://orcid.org/0009-0003-1151-3100
e-mail	olegmezh@biph.kiev.ua	olegmezh@biph.kiev.ua

Розробник доктор філософії Меженський О.Р.

Гарант ОНП
д.м.н. Моїсєнко Є.В.

Поточна редакція від «11» вересня 2025 р.

Зав. відділу
«Випускова кафедра»  К.В. Розова



Анотація навчальної дисципліни

Навчальна дисципліна "Використання методів машинного навчання в біомедичних дослідженнях" охоплює вивчення основних методів та технологій, аналізу біологічних даних з використанням технології машинного навчання та штучного інтелекту. Аспіранти матимуть змогу ознайомитись з основами програмування та data science. Детально будуть розглянуті такі методи як лінійна регресія, кластеринг, дерева рішень, та симуляції Монте-Карло, а також основи методів глибинного навчання. Крім того, аспіранти дізнаються про використання просунутих методів візуалізації даних. Після завершення курсу аспіранти матимуть достатні знання про основні методи машинного навчання та матимуть змогу використовувати їх у власній роботі.

Загальна мета курсу – забезпечити аспірантів знаннями та практичними навичками, необхідними для використання машинного навчання у сучасних біомедичних дослідженнях.

Місце навчальної дисципліни в програмі навчання

Навчальна дисципліна «Використання методів машинного навчання в біомедичних дослідженнях» є дисципліною за вибором, необхідною для тих аспірантів, чий науковий проєкти пов'язані із аналізом експериментальних даних.

Необхідні навички

Використання методів машинного навчання в біомедичних дослідженнях - складний курс, який вимагає від аспірантів певних навичок та знань:

Базові знання з біології, анатомії, та біохімії, щодо структур й функції біомолекул, а також механізмів захворювань.

Базові навички програмування: знання мов програмування, таких як Python або R.

Рівень набуття знань

В результаті вивчення теоретичного курсу та виконання практичних завдань за обраною тематикою аспірант повинен **знати** та **вміти**:

1. Робота з алгоритмами машинного навчання та обробки біомедичних даних.
2. Здатність працювати з великими наборами даних (геномними, протеомними, медичними зображеннями) і знання інструментів для їх аналізу, таких як Pandas, NumPy, або спеціалізовані бібліотеки для біоінформатики.
3. Знання основних методів і алгоритмів машинного навчання: лінійна регресія, дерева рішень, нейронні мережі, а також їх застосування в біомедичних дослідженнях.
4. Попередня обробка даних: очищення, нормалізація та візуалізація даних для якісного моделювання.
5. Аналіз та оцінка ефективності моделей, інтерпретація результатів, формування висновків для біомедичних контекстів.
6. Планування та виконання проєктів, здатність розробляти та реалізовувати дослідницькі проєкти, які включають збір, аналіз даних та використання результатів для розв'язання наукових питань.

Соціальні/загальні компетентності (ЗК)

ЗК5 Інформаційна грамотність і робота з доказовою базою

ЗК6 Самонавчання, самоорганізація, професійний розвиток

Фахові/дослідницькі компетентності (ФК):

ФК4 Аналіз даних, статистика, біоінформатика

ФК8 Організація лабораторії/групи, менеджмент досліджень

Програмні результати навчання (ПРН)

ПРН5 Вибирати, застосовувати та оптимізувати методи дослідження патологічних процесів на різних рівнях біологічної організації, оцінювати їх ефективність і обмеження.

ПРН6 Обробляти, аналізувати та інтерпретувати експериментальні й клінічні дані із застосуванням статистичних методів та біоінформатичних інструментів.

ПРН7 Формулювати наукові гіпотези, розробляти дизайн досліджень і планувати експерименти відповідно до принципів наукової доброчесності та біоетики.

Перелік тем, завдань та терміни виконання

4. Структура навчальної дисципліни

"Використання методів машинного навчання в біомедичних дослідженнях"

№	Тема	Кількість годин					
		Очне відділення			Заочне відділення		
		Л	Пр/С	СР	Л	Пр/С	СР
Змістовій модуль 1. Використання методів машинного навчання в біомедичних дослідженнях							
1	Вступ до машинного навчання - Структуровані та неструктуровані дані. - Типи змінних (безперервні, категорійні) та їх попередня обробка.	0	2	6	2	0	6
2	Підготовка даних - Очищення, нормалізація, шкалювання та трансформація змінних. - Виявлення та обробка пропущених даних. - Методи візуалізації даних	2	2	10	2	2	10

3	Лінійна регресія • Принципи та припущення лінійної регресії. • Методи оцінки параметрів (метод найменших квадратів).	2	2	6	2	2	6
4	Логістична регресія • Використання для класифікації. • Інтерпретація коефіцієнтів логістичної регресії. • Реалізація	0	2	6	2	0	6
5	Деревоподібні методи • Рішення дерев (Decision Trees): структура, принцип розбиття вузлів. • Випадковий ліс (Random Forest): використання ансамблів дерев для класифікації та регресії. • Реалізація	0	2	6	2	0	6
Змістовий модуль 2. Використання методів машинного навчання в біомедичних дослідженнях							
6	Крос-валідація та оцінка моделей • Розбиття даних на тренувальні та тестові набори. • Методологія крос-валідації (LOOCV, K-fold). • Реалізація	0	2	6	2	0	6
7	Метрики оцінки ефективності моделей • R^2, MSE для регресії. • Accuracy, Precision, Recall, F1-Score для класифікації. • Візуалізація ROC-кривих	2	2	6	2	2	6
8	Методи підвищення продуктивності моделей • Гіперпараметрична оптимізація (Grid Search, Random Search). • Розуміння принципів регуляризації (L1, L2). • Реалізація регуляризації	0	2	6	2	0	6
9	Кластеризація та PCA • Використання PCA для зниження вимірності даних. • Методи кластеризації (K-means, Hierarchical Clustering). • Використання tSNE та UMAP для візуалізації.	0	2	6	2	0	6

10	Виконання практичного проєкту	0	0	8	2	0	6
Всього годин:		6	20	64	20	6	64

Л – Лекції

Пр/С – Практичні / Семінари

СР – Самостійна робота

Теми семінарських/практичних занять

№ з/п	Назва теми	Кількість годин
1	Вступ до машинного навчання	2
2	Підготовка даних	2
3	Лінійна регресія	2
4	Логістична регресія	2
5	Деревоподібні методи	2
6	Крос-валідація та оцінка моделей	2
7	Метрики оцінки ефективності моделей	2
8	Методи підвищення продуктивності моделей	2
9	Кластеризація та PCA	2

6. Самостійна робота

№ з/п	Назва теми	Кількість годин
1	Вступ до машинного навчання <i>Читання наукової літератури та відео матеріалів, пов'язаних з тематикою курсу, з метою поглиблення знань та розширення уявлень про тему.</i>	6
2	Підготовка даних <i>Читання наукової літератури та відео матеріалів, пов'язаних з тематикою курсу, з метою поглиблення знань та розширення уявлень про тему.</i>	10
3	Лінійна регресія <i>Читання наукової літератури та відео матеріалів, пов'язаних з тематикою курсу, з метою поглиблення знань та розширення уявлень про тему.</i>	6
4	Логістична регресія <i>Читання наукової літератури та відео матеріалів, пов'язаних з тематикою курсу, з метою поглиблення знань та розширення уявлень про тему.</i>	6
5	Деревоподібні методи <i>Читання наукової літератури та відео матеріалів, пов'язаних з тематикою курсу, з метою поглиблення знань та розширення уявлень про тему.</i>	6

6	Крос-валідація та оцінка моделей <i>Читання наукової літератури та відео матеріалів, пов'язаних з тематикою курсу, з метою поглиблення знань та розширення уявлень про тему.</i>	6
7	Метрики оцінки ефективності моделей <i>Читання наукової літератури та відео матеріалів, пов'язаних з тематикою курсу, з метою поглиблення знань та розширення уявлень про тему.</i>	6
8	Методи підвищення продуктивності моделей <i>Читання наукової літератури та відео матеріалів, пов'язаних з тематикою курсу, з метою поглиблення знань та розширення уявлень про тему.</i>	8
9	Кластеризація та PCA <i>Читання наукової літератури та відео матеріалів, пов'язаних з тематикою курсу, з метою поглиблення знань та розширення уявлень про тему.</i>	8
Всього годин		64

ПИТАННЯ ДЛЯ ПОТОЧНОГО І ПІДСУМКОВОГО КОНТРОЛЮ

1. Алгоритми машинного навчання та їх застосування в біомедичних дослідженнях.
2. Методи збору, очищення, оброблення та візуалізації дані, такі як геномні дані, зображення медичних обстежень, результати лабораторних досліджень.
3. Методи класифікації, регресії, кластеризації, а також нейронні мережі для вирішення біомедичних задач.
4. Методи прогнозування та діагностики, створення моделей, які можуть прогнозувати розвиток захворювань або допомагати у їх діагностиці на основі великих наборів даних.
5. Аналіз високорозмірних даних (наприклад, генетичних, протеомних, метаболомних) та їх інтеграції для отримання нових наукових висновків.
6. Методи оптимізації та автоматизації для поліпшення моделей та автоматизації процесів аналізу.
7. Інструменти машинного навчання: бібліотеки TensorFlow, PyTorch, Scikit-learn, та їхнє використання в проєктах.
8. Етичні принципи у роботі з біомедичними даними, таких як забезпечення конфіденційності, інформована згода та уникнення упередженості у моделях.
9. Оцінка ефективності моделей, інтерпретація показників ефективності моделей: точність, чутливість, специфічність, тощо.
10. Програмного забезпечення для аналізу даних, структурування коду та його документування.
11. Практична розробка проєктів із інтеграцією машинного навчання у дослідження захворювань, створення нових діагностичних методів або підвищення ефективності лікування.

Система оцінювання

Усне опитування на семінарських/практичних заняттях (1-10 балів), обов'язкові модульні опитування за тестовою системою (0-30 балів за модуль). Заохочуючі бали (1-5 балів) можуть застосовуватись при експрес-опитуванні в процесі лекції (на розуміння її суті), за присутність на лекції 1 бал. Підсумковий тест (залік) – 40 балів. Сумарна оцінка за курс формується, виходячи з максимальної кількості балів - 100. Отримана в такий спосіб оцінка є підсумковою заліковою.

Розподіл балів, які отримують аспіранти

Поточний контроль та самостійна робота										Підсумковий тест (залік)	Сума
Змістовий модуль 1					Змістовий модуль 2						100
T1	T2	T3	T4	T5	T6	T7	T8	T9	T10	T1÷T10	
6	6	6	6	6	6	6	6	6	6	40	100

Семестрова атестація аспірантів

Шкала оцінювання: національна та ECTS

Сума балів за всі види навчальної діяльності	Оцінка ECTS	Оцінка за національною шкалою
		для заліку
90 – 100	A	зараховано
85-89	B	
75-84	C	
70-74	D	
60-69	E	
35-59	FX	не зараховано з можливістю повторного складання
0-34	F	не зараховано з обов'язковим повторним вивченням дисципліни

Засвоєння аспірантом програмного матеріалу змістового модуля вважається успішним, якщо рейтингова оцінка його становить не менше, ніж 60 балів за 100-бальною шкалою.

Додаткові умови допуску до заліку:

У разі виникнення спірних питань щодо не допуску аспірантів до семестрової атестації, вони вирішуються лектором дисципліни спільно із завідувачем кафедри.

Політика навчальної дисципліни

Відвідування занять є обов'язковим для всіх аспірантів.

Пропущені контрольні заходи можна перескласти у визначений викладачем час з дозволу завідувача відділу «Випускова кафедра». Аспіранти, які в поточному семестрі мали пропуски занять і до початку екзаменаційної сесії не засвоїли матеріал пропущених тем і розділів змістових модулів навчальної дисципліни та не подали обґрунтоване письмове пояснення причин пропущених занять, до семестрової атестації з відповідної дисципліни не допускаються.

Академічна доброчесність. Норми етичної поведінки

Політика та принципи академічної доброчесності визначені Законами України.

Норми етичної поведінки аспірантів і працівників визначені Статутом, відповідними законами, підзаконними актами України та відповідними положеннями Інституту.

Процедура оскарження результатів контрольних заходів

Аспіранти мають можливість підняти будь-яке питання, яке стосується процедури контрольних заходів та очікувати, що воно буде розглянуто згідно із наперед визначеними процедурами.

ЛІТЕРАТУРА ТА ІНФОРМАЦІЙНІ РЕСУРСИ

Python

- **Базовий рівень:** *Byte of Python* (Swaroop C.H.)

Дуже дружній для початківців вступ до Python, що охоплює синтаксис, типи даних і базові концепції з чіткими прикладами ([українська версія PDF](#)).

- **Середній рівень:** *Automate the Boring Stuff with Python* (Al Sweigart)

Практичний посібник із застосування Python для щоденних задач — робота з файлами, веб-скрапінг, автоматизація Excel тощо.

[Посилання](#)

- **Просунутий рівень:** *Fluent Python* (Luciano Ramalho)

Зосереджений на написанні ідіоматичного, ефективного коду Python, включно з просунутими структурами даних, метапрограмуванням і багатопотоковістю.

[Посилання](#)

- **Machine Learning:** *Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow* (Aurélien Géron)

Покрокове введення в машинне та глибинне навчання в Python із практичними прикладами та проектами.

[Посилання 1](#)

[Посилання 2](#)

- **Machine Learning 2:** *The Orange Book of Machine Learning* (Carl McBride Ellis)

Основи побудови моделей для передбачення за допомогою регресії та класифікації для табличних даних.

[Посилання](#)

R

- **Базовий рівень:** *R for Data Science* (Hadley Wickham & Garrett Grolemund)

Просте й практичне введення в R, зосереджене на обробці даних, візуалізації та філософії tidyverse.

Посилання

- **Середній рівень:** *Discovering Statistics using R* (Andy Field)

Захопливе та доступне введення в статистику з великою кількістю прикладів та гумору.

Посилання

- **Просунутий рівень:** *Advanced R* (Hadley Wickham)

Розглядає глибокі внутрішні механізми середовища R, функціональне програмування та об'єктно-орієнтовані системи — для досвідчених програмістів.

Посилання

- **Застосування статистики в біології:** *Modern Statistics for Modern Biology* (W. Huber)

Складний, але доступний виклад сучасної статистики, адаптований до біологічних і медичних наук.

Посилання

- **Матеріали CSAMA школи:** Посилання
-

Курси

Python

- CS50's Introduction to Programming with Python

Безкоштовний курс Гарварду — основи Python, розв'язування задач та проєкти.

- Introduction to Machine Learning with Python

Офіційні туторіали scikit-learn (з Jupyter-ноутбуками) з навчання під наглядом і без.

- DS100 Principles and Techniques of Data Science

Курс, який охоплює базові та деякі просунуті теми машинного навчання й науки про дані.

- TensorFlow Bootcamp

Основи нейронних мереж з TensorFlow та вступ до комп'ютерного зору.

R

- Data Science: R Basics

Чітке безкоштовне введення в синтаксис R, структури даних і логіку програмування.

- Introduction to Statistical Learning (ISLR)

Безкоштовна книга + відеолекції; практичне машинне навчання та статистика в R.

Bash

Introduction to Linux (LFS101)

Практичне введення в команди Bash та написання скриптів.