

КЛАСИЧНИЙ ПРИВАТНИЙ УНІВЕРСИТЕТ
ІНСТИТУТ ЕКОНОМІКИ ТА ПРАВА
КАФЕДРА ЕКОНОМІКИ ТА СТАТИСТИКИ

СИЛАБУС

навчальної дисципліни
«СТАТИСТИЧНІ АЛГОРИТМИ НАВЧАННЯ»

КОНТАКТНА ІНФОРМАЦІЯ ТА ТЕХНІЧНА ДОПОМОГА
(включаючи електронну пошту, робочий час / місцезнаходження тощо).

Викладач (-і)	Сергеева Людмила Нільсівна
Контактний тел.	+38(061)764-67-50 (внутр. 169)
E-mail:	lud.sergeeva@gmail.com
Сторінка курсу на сайті підтримки навчальних програм КПУ	http://www.zhu.edu.ua/cpu_edu/course/view.php?id=4682
Консультації	<i>Очні консультації:</i> за графіком консультацій викладача, а. 507, головний корпус КПУ <i>Консультації off-line:</i> шляхом повідомлення на сторінці навчальної дисципліни сайту підтримки навчальних програм КПУ http://www.zhu.edu.ua/cpu_edu/message

АНОТАЦІЯ

У період стрімкого розвитку цифрових технологій, статистичні алгоритми навчання стають силою, що дозволяє ефективно використовувати весь величезний потенціал даних. Алгоритми навчання мають чудову здатність вчитися, виявляти закономірності, робити прогнози та автоматизувати прийняття рішень.

Освітній процес з дисципліни здійснюється за такими формами: навчальні заняття; самостійна робота; контрольні заходи. Видами навчальних занять згідно з навчальним планом є: лекції; практичні заняття, а також консультації.

Практичні заняття передбачають: перевірку домашніх завдань; перевірку засвоєння студентами лекційного матеріалу; виконання завдань з побудови нейронних мереж.

Самостійна робота студентів полягає у засвоєнні вивченого навчального матеріалу в час, вільний від обов'язкових навчальних занять, без участі викладача.

Консультації призначені для роз'яснення студентам теоретичних або практичних питань.

Засвоєння навчального матеріалу перевіряється за допомогою поточного контролю, який здійснюється на практичних заняттях у формі перевірки виконаних завдань, самостійних робіт.

Підсумковий (семестровий) контроль після завершення 2 семестру здійснюється у формі екзамену.

ФОРМАТ НАВЧАЛЬНОЇ ДИСЦИПЛІНИ

Загальна кількість годин – 150 год., у т. ч. 56 годин аудиторних занять і 94 годин самостійної роботи студента. Кількість кредитів ECTS – 5.

Всього кредитів	Всього годин	Аудиторних годин	У тому числі			Сам. робота
			Лекц.	Лабор.	Семін. (практ.)	
5	150	56	26	-	30	94

ОЗНАКИ ДИСЦИПЛІНИ

Навчальний рік	Курс (рік навчання)	Семестр	Цикл підготовки	Обов'язкова / вибіркова
2024/2025	1	2	професійна	обов'язкова

МЕТА ТА ЗАВДАННЯ НАВЧАЛЬНОЇ ДИСЦИПЛІНИ

Мета навчальної дисципліни

Метою викладання навчальної дисципліни «Статистичні алгоритми навчання» є оволодіння сучасними методами, теоретичними положеннями та основними застосуваннями статистичних алгоритмів машинного навчання.

Завдання навчальної дисципліни

формування здатності розв'язувати складні спеціалізовані математичні та статистичні задачі, що характеризується комплексністю і невизначеністю умов і передбачає застосування теоретико - ймовірнісних і статистичних алгоритмів; набуття знань, умінь та навичок (компетентностей) на рівні новітніх досягнень у математиці та статистиці.

ЗАПЛАНОВАНІ РЕЗУЛЬТАТИ НАВЧАННЯ

У результаті вивчення дисципліни студенти повинні знати:

- як використовуються нейронні мережі для розв'язку певних задач статистичних алгоритмів навчання,
- обирати їх архітектуру, тлумачити результати навчання з теоретико-ймовірнісної точки зору та статистично порівнювати різні мережі,
- розуміти межі, поза якими мережа навчається гірше, й «пастки», котрих слід уникати.

Після вивчення дисципліни студенти повинні вміти:

- застосовувати ймовірнісно-статистичні моделі та методи;
- використовувати в практичній діяльності спеціалізоване програмне забезпечення;
- застосовувати математичні моделі та методи для розв'язання проблем.

Відповідно до освітньо-професійної програми підготовки магістра галузі знань 11 Математика та статистика за спеціальністю 112 Статистика, освітня програма

Комп'ютерна статистика та аналіз даних вивчення дисципліни «Статистичні алгоритми навчання» сприяє формуванню **компетентностей та програмних результатів навчання:**

Інтегральна компетентність

Здатність розв'язувати складні математичні та статистичні задачі, практичні проблеми у професійній діяльності або у процесі навчання, що передбачає проведення досліджень та здійснення інновацій і характеризується комплексністю та невизначеністю умов.

Загальні компетентності:

ЗК 2. Здатність використовувати у професійній діяльності знання з галузей математичних, природничих, соціально-гуманітарних та економічних наук

ЗК 4. Здатність до пошуку, оброблення й аналізу інформації з різних джерел, необхідної для розв'язування наукових і професійних завдань.

ЗК 7. Здатність до виконання дослідницької роботи з елементами наукової новизни.

Спеціальні (фахові, предметні) компетентності:

СК 1. Знання на рівні новітніх досягнень, необхідні для дослідницької та інноваційної діяльності у сфері статистики та її практичних застосувань.

СК 5. Спроможність розробляти математичну модель ситуації з реального світу та переносити математичні знання у нематематичні контексти.

СК 8. Здатність до розвитку нових та удосконалення існуючих статистичних методів аналізу, моделювання, прогнозування, розв'язування нових проблем у нових галузях знань.

СК 12. Володіння знаннями та здатність ініціювати й проводити наукові дослідження у спеціалізованих областях статистики.

Програмні результати навчання:

РН 1 Знати та розуміти фундаментальні і прикладні аспекти наук у сфері статистики.

РН 2 Відтворювати знання фундаментальних розділів статистики в обсязі, необхідному для володіння математичним та економічним апаратами відповідної галузі знань і використання статистичних методів у обраній професії

РН 3 Володіти основами математичних дисциплін і економічних теорій, зокрема які вивчають моделі природничих і соціальних процесів

РН 4. Володіти знаннями грамотної побудови комунікації в освітньому і науковому процесі, відбору вихідних даних дослідження, складання списку використаних джерел, опису наукових результатів.

РН 5. Уміти використовувати фундаментальні закономірності статистики у професійній діяльності.

РН 6. Читати і розуміти фундаментальні розділи математичної та економічної літератури та демонструвати майстерність їх відтворення в аргументованій усній та письмовій доповіді.

РН 8. Ініціювати і проводити наукові дослідження у спеціалізованій області математичної статистики та розв'язувати задачі в інших галузях знань методами математичного та економічного моделювання.

РН 9. Інтегрувати знання з різних галузей для вирішення теоретичних та практичних задач і проблем.

РН 10. Застосовувати нові підходи для вироблення стратегії прийняття рішень у складних непередбачуваних умовах.

РН 12. Уміти самостійно планувати виконання дослідницького та інноваційного завдання та формулювати висновки за його результатами.

РН 13. Усно й письмово спілкуватися рідною та іноземними мовами в науковій, виробничій та соціально-суспільній сферах діяльності із професійних питань; читати

спеціальну літературу; знаходити, аналізувати та використовувати інформацію з різних джерел.

PH 14. Використовувати раціональні способи пошуку та використання науково-технічної інформації в галузі статистики, включаючи засоби електронних інформаційних мереж; застосовувати інформаційні ресурси, у тому числі електронні, для пошуку відповідних математичних моделей.

PH 15. Дотримуватися норм етичної поведінки стосовно інших людей, адаптуватися та комунікувати.

ПЛАН КУРСУ

Назва змістових модулів та тем	Лекц.	Прак т (сем.)	Завдання для самостійної роботи
ЗМІСТОВИЙ МОДУЛЬ № 1. ТЕОРЕТИЧНІ ОСНОВИ МАШИННОГО НАВЧАННЯ НАВЧАННЯ З ВЧИТЕЛЕМ: ЗАВДАННЯ КЛАСИФІКАЦІЇ			
Тема 1. Вступ до машинного навчання (МН)	2	2	1. Опрацювання лекційного матеріалу, 2. Самостійне опрацювання теоретичних питань: 1) Поняття та сутність кластеризації 3. Підготовка до практичного заняття, опрацювання першоджерел та навчальної літератури.
Тема 2. Навчання простих алгоритмів МН для класифікації	2	2	1. Опрацювання лекційного матеріалу, 2. Самостійне опрацювання теоретичних питань: 1) Класи, які сепарабельні та не сепарабельні. 3. Підготовка до практичного заняття, опрацювання першоджерел та навчальної літератури; виконання індивідуального завдання.
Тема 3. Адаптивні лінійні нейрони та збіжність навчання	2	4	1. Опрацювання лекційного матеріалу, 2. Самостійне опрацювання теоретичних питань: 1) Вплив стандартизації на навчання з градієнтним спуском. 3. Підготовка до практичного заняття, опрацювання першоджерел та навчальної літератури, виконання індивідуального завдання.
Тема 4. Огляд класифікаторів на основі машинного навчання з використанням scikit-learn: логістична регресія	2	2	1. Опрацювання лекційного матеріалу, 2. Самостійне опрацювання теоретичних питань: 1) Порівняння Adaline та логістичної регресії. 3. Підготовка до практичного заняття, опрацювання першоджерел та навчальної літератури; виконання індивідуального завдання.
Тема 5. Огляд класифікаторів на основі машинного навчання з використанням scikit-learn: метод опорних векторів	2	2	1. Опрацювання лекційного матеріалу, 2. Самостійне опрацювання теоретичних питань: 1) Логістична регресія у порівнянні з методами опорних векторів. 3. Підготовка до практичного заняття, опрацювання першоджерел та навчальної літератури; виконання індивідуального завдання.
Тема 6. Огляд класифікаторів на основі машинного навчання з використанням scikit-learn: метод дерева прийняття рішень та метод k найближчих сусідів	2	2	1. Опрацювання лекційного матеріалу, 2. Самостійне опрацювання теоретичних питань: 1) Метод k найближчих сусідів - алгоритм ледачого навчання. 3. Підготовка до практичного заняття, опрацювання першоджерел та навчальної літератури; виконання індивідуального завдання.
ЗМІСТОВИЙ МОДУЛЬ № 2. НАВЧАННЯ З ВЧИТЕЛЕМ: РЕГРЕСІЙНИЙ АНАЛІЗ ТА НАВЧАННЯ БЕЗ ВЧИТЕЛЯ: ЗАВДАННЯ КЛАСТЕРИЗАЦІЇ. НАВЧАННЯ З ПІДКРІПЛЕННЯМ			

Назва змістових модулів та тем	Лекц.	Прак т (сем.)	Завдання для самостійної роботи
Тема 7. Прогнозування значень неперервних цільових змінних за допомогою регресійного аналізу	2	2	1. Опрацювання лекційного матеріалу, 2. Самостійне опрацювання теоретичних питань: 1) Налаштування надійної регресійної моделі з використанням RANSAC. 3. Підготовка до практичного заняття, опрацювання першоджерел та навчальної літератури; виконання індивідуального завдання.
Тема 8. Оцінка ефективності лінійних регресійних моделей. Оцінювання нелінійних залежностей	2	2	1. Опрацювання лекційного матеріалу, 2. Самостійне опрацювання теоретичних питань: 1) Гребенева регресія. 3. Підготовка практичного заняття, опрацювання першоджерел та навчальної літератури; виконання індивідуального завдання
Тема 9. Робота з непоміченими даними - кластерний аналіз	2	2	1. Опрацювання лекційного матеріалу, 2. Самостійне опрацювання теоретичних питань: 1) Кількісна оцінка якості кластеризації через графіки силуетів. 3. Підготовка до практичного заняття, опрацювання першоджерел та навчальної літератури; виконання індивідуального завдання.
Тема 10. Організація кластерів у вигляді ієрархічного дерева	2	4	1. Опрацювання лекційного матеріалу, 2. Самостійне опрацювання теоретичних питань: 1) Чотири фундаментальні категорії алгоритмів кластеризації. 3. Підготовка до практичного заняття, опрацювання першоджерел та навчальної літератури; виконання індивідуального завдання.
Тема 11. Навчання з підкріпленням для прийняття рішень у складних середовищах	2	2	1. Опрацювання лекційного матеріалу, 2. Самостійне опрацювання теоретичних питань: 1) Відмінність між нагородою, віддачею і функцією цінності. 3. Підготовка до семінарського заняття, опрацювання першоджерел та навчальної літератури; виконання індивідуального завдання.
Тема 12. Алгоритми навчання з підкріпленням	1	2	1. Опрацювання лекційного матеріалу, 2. Самостійне опрацювання теоретичних питань: 1) Знаходження оптимальної політики з використанням контролю Монте-Карло. 3. Підготовка до практичного заняття, опрацювання першоджерел та навчальної літератури; виконання індивідуального завдання.
Тема 13. Реалізація алгоритмів навчання із підкріпленням	1	4	1. Опрацювання лекційного матеріалу, 2. Самостійне опрацювання теоретичних питань: 1) Реалізація пам'яті відтворення. 3. Підготовка до практичного заняття, опрацювання першоджерел та навчальної літератури; виконання індивідуального завдання.

ФОРМИ КОНТРОЛЮ ТА КРИТЕРІЇ ОЦІНЮВАННЯ

У процесі вивчення навчальної дисципліни «Статистичні алгоритми навчання» використовуються наступні види контролю:

1. Поточний контроль – здійснюється протягом семестру шляхом перевірки виконання завдань практичних занять, виконання індивідуальних завдань тощо. За

змістом він включає перевірку ступеню засвоєння студентом навчального матеріалу, який охоплюється темою лекційного та практичного заняття, уміння самостійно опрацьовувати навчально-методичну літературу, здатність осмислювати зміст теми, уміння використовувати програмне забезпечення та здійснювати інтерпретацію отриманих результатів.

2. Підсумковий семестровий контроль – здійснюється у формі екзамену відповідно до графіка освітнього процесу.

Для оцінювання студентів використовується система накопичування балів. Згідно з «Положенням про організацію освітнього процесу в КПУ» підсумкова оцінка з дисципліни виставляється за 100-бальною шкалою з наступним переведенням у національну шкалу та шкалу ECTS.

Бали нараховуються за виконання завдань практичних робіт, індивідуальних завдань.

Результати поточного контролю здобувачів вищої освіти є складовими елементами підсумкової оцінки з дисципліни.

Оцінка рівня роботи студента протягом семестру під час навчальних занять та самостійної роботи здійснюється у межах 60 балів. Вага екзамену у підсумковій оцінці складає 40 балів.

Для оцінювання результатів навчання студентів з навчальної дисципліни «Статистичні алгоритми навчання» використовуються наступні критерії:

– рівню «відмінно» відповідає повністю виконане завдання, при якому студент показав всебічне системне знання матеріалу; засвоєння літератури; чітке володіння понятійним апаратом, методами, методиками та інструментами статистичних алгоритмів навчання; уміння використовувати знання для побудови програми за конкретним алгоритмом навчання та надавати змістовну інтерпретацію отриманих результатів;

– рівню «добре» відповідає успішне виконання запропонованого завдання і засвоєння матеріалу літератури без змістовної інтерпретації отриманих результатів;

– рівню «задовільно» відповідає часткове виконання запропонованого завдання і засвоєння матеріалу літератури, при якому студент показав достатній рівень знань з статистичних алгоритмів навчання, але не зміг переконливо аргументувати свою відповідь, помилився у використанні понятійного апарату, показав недостатні знання рекомендованої літератури;

– рівню «незадовільно» відповідає відсутність виконання запропонованого завдання; студент показав незадовільні знання понятійного апарату і літератури чи взагалі нічого не відповів.

РОЗПОДІЛ БАЛІВ ПОТОЧНОГО ТА ПІДСУМКОВОГО (СЕМЕСТРОВОГО) ОЦІНЮВАННЯ

Контрольний захід	Бали
Модуль №1	30
Практичні роботи	20
Поточний контроль	10
Модуль №2	30
Практичні роботи	20
Поточний контроль	10
Екзамен	40
Разом	100

Шкала оцінювання: 100-бальна, національна та ЄКТС

Сума балів за всі види навчальної діяльності	Оцінка за шкалою ЄКТС	Оцінка за національною шкалою для екзамену
---	------------------------------	---

90-100	A	відмінно
82-89	B	добре
75-81	C	
67-74	D	задовільно
60-66	E	
35-59	FX	незадовільно з можливістю повторного складання
0-34	F	незадовільно з обов'язковим повторним вивченням дисципліни

РЕКОМЕНДОВАНА ЛІТЕРАТУРА

1. Інтелектуальний аналіз даних та машинне навчання. Частина 1. Базові методи та засоби аналізу даних / Я. В. Іванчук, В. І. Месюра, А. А. Яровий, О. Д. Манжілевський – Вінниця : ВНТУ, 2021. – 69 с.
2. 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.
3. Новотарський М.А. Лекції з курсу «Статистичні методи машинного навчання» // <https://cloud.comsys.kpi.ua/s/pjw4nAStZFPsDos>
4. Харченко В. О. Основи машинного навчання : навч. посіб. / В. О. Харченко. – Суми : Сумський державний університет, 2023. – 264 с.
5. Sebastian Raschka, Vahid Mirjalili. Machine Learning and Deep Learning with Python, scikit-learn, and TensorFlow 2 3rd ed. Edition, Packt Publishing, 2019, 771 p.
6. Bishop C. M. Pattern Recognition and Machine Learning. — Springer, Series: Information Science and Statistics, 2006. — 740 pp.
7. Hastie T., Tibshirani R., Friedman J. The Elements of Statistical Learning. — Springer, 2001. – 533 pp. <http://http://www-stat.stanford.edu/~tibs/ElemStatLearn>.
8. LeCun Y., Bottou L., Orr G. B., Muller K.-R. Efficient BackProp // Neural Networks: tricks of the trade. – Springer, 1998.

ІНФОРМАЦІЙНІ РЕСУРСИ В ІНТЕРНЕТІ

1. UCI machine learning repository. <https://archive.ics.uci.edu/>
2. Kaggle: Your Machine Learning and Data Science Community <https://www.kaggle.com/>