

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИЛЬВІВСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ імені ІВАНА ФРАНКА

Кафедра (циклова комісія) _____ системного проектування _____

“ЗАТВЕРДЖУЮ”

Декан факультету _____

доц. Юрій ФУРГАЛА

“

”

2022 року

РОБОЧА ПРОГРАМА НАВЧАЛЬНОЇ ДИСЦИПЛІНИЗАСОБИ МАШИННОГО НАВЧАННЯ

(шифр і назва навчальної дисципліни)

спеціальність _____ 121 – Інженерія програмного забезпечення _____

(шифр і назва спеціальності)

спеціалізація _____

(назва спеціалізації)

факультет _____ електроніки та комп'ютерних технологій _____

(назва інституту, факультету, відділення)

2022 – 2023 навчальний рік

Робоча програма “Засоби машинного навчання” для студентів
(назва навчальної дисципліни)
 галузі знань “12 – Інформаційні технології”
 за спеціальністю “121 Інженерія програмного забезпечення”

Розробники: Віталій Парубочий (асистент кафедри системного проектування,
факультету електроніки та комп’ютерних технологій)
(вказати авторів, їхні посади, наукові ступені та вчені звання)

Робочу програму схвалено на засіданні кафедри (циклової комісії) _____

 системного проектування

Протокол від “ 30 ” 08 2022 року № 1

Завідувач кафедри _____ системного проектування
 _____ (Роман ШУВАР)

(підпис)

(прізвище та ініціали)

Ухвалено Вченою радою _____ факультету електроніки та комп’ютерних технологій

Протокол від “ 31 ” 08 2022 року № 28/22

1. Опис навчальної дисципліни

Найменування показників	Галузь знань, освітньо-кваліфікаційний рівень	Характеристика навчальної дисципліни	
		денна форма навчання	заочна форма навчання
Кількість кредитів 5,5	Галузь знань <u>12 Інформаційні технології</u> (шифр і назва)	За вибором	
Модулів – 1	Спеціальність: <u>121 Інженерія програмного забезпечення</u>	Рік підготовки	
Змістових модулів – 3		3-й	
Індивідуальне науково-дослідне завдання <u>немає</u> (назва)		Семестр	
Загальна кількість годин – 165		6-й	
Тижневих годин для денної форми навчання: аудиторних – 4 самостійної роботи студента – 6,3	Освітній ступінь <u>бакалавр</u>	Лекції	
		32 год.	
		Практичні, семінарські	
		<i>немає</i>	
		Лабораторні	
		32 год.	
		Самостійна робота	
		101 год.	
		Індивідуальні завдання:	
<i>немає</i>			
Вид контролю:			
<i>залік</i>			

Примітка.

Співвідношення кількості годин аудиторних занять до самостійної і індивідуальної роботи становить

для денної форми навчання – 0,63

для заочної форми навчання – немає

2. Мета та завдання навчальної дисципліни

Мета: надати студентам знання та розуміння термінів машинного навчання, а також навичок та вмінь для вирішення задач регресії, класифікації, зменшення розмірності, кластеризації та прогнозування даних засобами машинного навчання.

Цілі: забезпечити знайомство студентів з загальною термінологією машинного навчання та сформуванню навички їхнього практичного використання; надати можливість студентам отримати знання для аналізу даних та вибору

необхідного методу машинного навчання, який дає змогу вирішити поставлену задачу; надати можливість студентам отримати навички та вміння для аналізу та оцінки отриманих результатів, використовуючи статистичні методи оцінки.

У результаті вивчення навчальної дисципліни студент повинен

знати: основні поняття, визначення і проблеми курсу; вимоги до постановки основних задач машинного навчання; призначення й особливості застосування основних методів машинного навчання;

вміти: володіти математичним апаратом методів машинного навчання; застосовувати методи машинного навчання для розв'язування прикладних задач машинного навчання; виконувати оцінку результатів опрацювання даних методами машинного навчання.

Після вивчення курсу «Засоби машинного навчання» здобувачі набудуть таких Загальних та Фахових компетентностей та Програмних результатів навчання:

ЗК1. Здатність до абстрактного мислення, аналізу та синтезу.

ЗК2. Здатність застосовувати знання у практичних ситуаціях.

ЗК5. Здатність вчитися і оволодівати сучасними знаннями.

ЗК6. Здатність до пошуку, оброблення та аналізу інформації з різних джерел.

ФК16. Здатність ідентифікувати, класифікувати та формулювати вимоги до програмного забезпечення.

ФК17. Здатність брати участь у проектуванні програмного забезпечення, включаючи проведення моделювання (формальний опис) його структури, поведінки та процесів функціонування.

ФК18. Здатність розробляти архітектури, модулі та компоненти програмних систем.

ФК22. Володіння знаннями про інформаційні моделі даних, здатність створювати програмне забезпечення для зберігання, видобування та опрацювання даних.

ФК23. Здатність застосовувати фундаментальні і міждисциплінарні знання для успішного розв'язання завдань інженерії програмного забезпечення.

ФК25. Здатність накопичувати, обробляти та систематизувати професійні знання щодо створення і супроводження програмного забезпечення та визнання важливості навчання протягом всього життя.

ФК27. Здатність здійснювати процес інтеграції системи, застосовувати стандарти і процедури управління змінами для підтримки цілісності, загальної функціональності і надійності програмного забезпечення.

ФК28. Здатність обґрунтовано обирати та освоювати інструментарій з розробки та супроводження програмного забезпечення.

ФК29. Здатність до алгоритмічного та логічного мислення.

ПРН1. Аналізувати, цілеспрямовано шукати і вибирати необхідні для вирішення професійних завдань інформаційно-довідкові ресурси і знання з урахуванням сучасних досягнень науки і техніки.

ПРН8. Вміти розробляти людино-машинний інтерфейс.

ПРН9. Знати та вміти використовувати методи та засоби збору, формулювання та аналізу вимог до програмного забезпечення.

ПРН11. Вибирати вихідні дані для проектування, керуючись формальними методами опису вимог та моделювання.

ПРН13. Знати і застосовувати методи розробки алгоритмів, конструювання програмного забезпечення та структур даних і знань.

ПРН14. Застосовувати на практиці інструментальні програмні засоби доменного аналізу, проектування, тестування, візуалізації, вимірювань та документування програмного забезпечення.

ПРН18. Знати та вміти застосовувати інформаційні технології обробки, зберігання та передачі даних.

ПРН22. Знати та вміти застосовувати методи та засоби управління проектами.

ПРН23. Вміти документувати та презентувати результати розробки програмного забезпечення.

ПРН26. Знати та вміти застосовувати засоби інженерії програмного забезпечення для реалізації проєктів у галузі штучного інтелекту та м'яких обчислень.

ПРН27. Вміти обирати оптимальні алгоритми та технології розробки програмного забезпечення для розв'язання задач аналізу та інженерії даних.

3. Програма навчальної дисципліни

Змістовий модуль 1. Основи машинного навчання.

Тема 1. Машинне навчання. Основні поняття та задачі.

Штучний інтелект як галузь знань. Машинне навчання. Наука про дані. Добування даних. Підходи машинного навчання. Навчання з учителем. Напівавтоматичне навчання. Активне навчання. Навчання з підкріпленням. Навчання без учителя. Задачі машинного навчання. Методи та моделі машинного навчання. Обмеження методів машинного навчання. Оцінювання методів машинного навчання. Програмні та апаратні засоби машинного навчання.

Тема 2. Регресійний аналіз даних.

Поняття регресії та регресійного аналізу. Використання та застосування регресійного аналізу. Загальна регресійна модель. Методи регресійного аналізу. Лінійна регресія. Проста та множинна лінійна регресія. Загальна та узагальнена лінійні моделі. Логістична регресія. Оцінка невідомих параметрів. Узагальнення.

Тема 3. Градієнтний спуск та його застосування в машинному навчанні.

Поняття градієнтного спуску. Умови збіжності градієнтного спуску. Модифікації градієнтного спуску. Стохастичний градієнтний спуск. Модифікації стохастичного градієнтного спуску. Застосування градієнтного спуску.

Змістовий модуль 2. Задачі машинного навчання.

Тема 4. Класифікація даних.

Поняття класифікації. Види класифікації. Методи класифікації. Логістична регресія як метод класифікації. Навчання на дереві рішень. Випадковий ліс.

Метод k-найближчих сусідів. Метод опорних векторів. Оцінювання результатів класифікації.

Тема 5. Зменшення розмірності даних.

Поняття зменшення розмірності даних. Підходи до зменшення розмірності. Методи зменшення розмірності. Лема Джонсона-Лінденштрауса. Аналіз головних компонентів. Факторизація невід'ємної матриці. Лінійний дискримінантний аналіз. Аналіз головних компонентів ядра. Автокодер.

Тема 6. Напіваавтоматичне навчання та його застосування.

Поняття напіваавтоматичного навчання. Методи напіваавтоматичного навчання. Сходові мережі. Рі-модель. Тимчасове ансамблювання. Усереднений вчитель. Варіаційний автокодер. Генеративна змагальна мережа. Методи проксі-міток. Самонавчання. Мета-псевдомітки. Багаторакурсне навчання. MixMatch. ReMixMatch. FixMatch. Методи на основі графів. Поширення міток. Регуляризація.

Тема 7. Кластерний аналіз даних.

Поняття кластеризації. Типи кластеризації. Методи кластеризації. Кластеризація великих наборів даних. Ієрархічна кластеризація. Методи ієрархічної кластеризації. Кластеризація на основі центроїда. Кластеризація k-середніх. Кластеризація на основі розподілу. Кластеризація на основі щільності. DBSCAN. OPTICS. Кластеризація на основі щільності. Кластеризація на основі сітки. Оцінка результатів кластеризації. Застосування кластеризації.

Тема 8. Прогнозування даних.

Поняття прогнозування даних. Класифікація як метод прогнозування. Навчання на дереві рішень. Кластеризація як метод прогнозування. Регресія як метод прогнозування. Властивості часових рядів. Стратегії прогнозування. Якісне та кількісне прогнозування. Методи простої оцінки. Методи часових рядів. Ковзне середнє. ARMA. ARIMA. Груповий метод обробки даних. Рекурентні нейронні мережі. Оцінка точності прогнозування. Застосування прогнозування.

Змістовий модуль 3. Прикладні методи та засоби машинного навчання.

Тема 9. Техніки покращення точності моделей машинного навчання.

Регуляризація. Перенавчання та узагальнення. Класифікація методів регуляризації. Штучні дані та збільшення даних. Зміна вибірки. Регуляризація гіперпараметрів. Автоматичне налаштування гіперпараметрів. Відкидання. Стохастичне об'єднання. Зниження ваги. Рання зупинка. Зменшення розмірності. Підходи до зменшення розмірності. Методи зменшення розмірності. Навчання на дереві рішень. Випадковий ліс. Ансамблеве навчання. Методи ансамблевого навчання. Оптимальний класифікатор Байеса. Початкове агрегування. Методи ансамблевого навчання.

Тема 10. Штучні нейронні мережі та глибинне навчання.

Штучні нейронні мережі. Глибинне навчання. Історія розвитку штучних нейронних мереж. Основні поняття. Багатошаровий перцептрон. Гіперпараметри. Навчання ШНМ. Швидкість навчання та функція витрат. Зворотне поширення. Підходи машинного навчання. Сфери застосування. Програмні та апаратні засоби машинного навчання.

Тема 11. Класифікація штучних нейронних мереж.

Класифікація ШНМ. Нейронна мережа прямого поширення. Згортова нейронна мережа. Мережа радіальних базисних функцій. Регуляторні мережі зворотного зв'язку. Рекурентні нейронні мережі. Довга короткочасна пам'ять. Нейронні мережі з розширеною пам'яттю. Генеративна стохастична штучна нейронна мережа. Модульна нейронна мережа. Спайкова нейронна мережа. Складна ієрархічно-глибока модель. Генеративна змагальна мережа. Динамічна нейронна мережа. Гібридна нейронна мережа. Мережа пам'яті. Фізична нейронна мережа.

Тема 12. Огляд програмних засобів та інфраструктури машинного навчання.

Програмні засоби машинного навчання. Бібліотека scikit-learn. TensorFlow та Keras. Torch та PyTorch. Інші бібліотеки Python. Apache Software Foundation. Бібліотеки C++. Open Neural Network Exchange (ONNX). Хмарні рішення. Отримання даних.

4. Структура навчальної дисципліни

Назви змістових модулів і тем	Кількість годин											
	денна форма						заочна форма					
	усь го	у тому числі					усь ого	у тому числі				
		л	п	лаб.	інд.	с. р.		л	п	лаб	інд.	с. р.
1	2	3	4	5	6	7	8	9	10	11	12	13
Змістовий модуль 1. Основи машинного навчання												
Тема 1. Машинне навчання. Основні поняття та задачі.	21	4		4		13						
Тема 2. Регресійний аналіз даних.	10	2		2		6						
Тема 3. Градієнтний спуск та його застосування в машинному навчанні.	8	2				6						
Разом за змістовим модулем 1	39	8		6		25						
Змістовий модуль 2. Задачі машинного навчання												
Тема 4. Класифікація даних.	10	2		2		6						

Тема 5. Зменшення розмірності даних.	10	2	2	6						
Тема 6. Напівавтоматичне навчання та його застосування.	18	2	4	12						
Тема 7. Кластерний аналіз даних.	13	2	4	7						
Тема 8. Прогнозування даних.	18	2	4	12						
Разом за змістовим модулем 2	69	10	16	43						
Змістовий модуль 3. Прикладні методи та засоби машинного навчання										
Тема 9. Техніки покращення точності моделей машинного навчання.	17	4	4	9						
Тема 10. Штучні нейронні мережі та глибинне навчання.	10	2	2	6						
Тема 11. Класифікація штучних нейронних мереж.	15	4	2	9						
Тема 12. Огляд програмних засобів та інфраструктури машинного навчання.	15	4	2	9						
Разом за змістовим модулем 3	57	14	10	33						
Усього годин	165	32	32	101						

5. Теми семінарських занять

6. Теми практичних занять

7. Теми лабораторних занять

№ з/п	Назва теми	Кількість годин
1	Вступне заняття. Налаштування середовища та необхідного програмного забезпечення	2
2	Лабораторна робота №1. Використання бібліотеки NumPy для роботи з багатовимірними масивами даних	2
3	Лабораторна робота №2. Використання бібліотек pandas та Matplotlib для попередньої обробки та візуалізації даних	2
4	Лабораторна робота №3. Прогнозування даних за	2

	<i>допомогою методів регресійного аналізу</i>	
5	<i>Лабораторна робота №4. Класифікація даних за допомогою методів машинного навчання</i>	2
6	<i>Лабораторна робота №5. Використання зменшення розмірності для класифікації даних за допомогою методів машинного навчання</i>	2
7	<i>Лабораторна робота №6. Використання методів напіваавтоматичного навчання для класифікації даних</i>	4
8	<i>Лабораторна робота №7. Кластеризація даних за допомогою методів машинного навчання</i>	4
9	<i>Лабораторна робота №8. Прогнозування даних за допомогою методів машинного навчання</i>	4
10	<i>Лабораторна робота №9. Дослідження методів машинного навчання для вирішення індивідуальної задачі</i>	4
11	<i>Захисне лабораторне заняття</i>	4
	Разом	32

8. Самостійна робота

№ з/п	Назва теми	Кількість годин
1	<i>Машинне навчання. Основні поняття та задачі</i>	13
2	<i>Регресійний аналіз даних</i>	6
3	<i>Градiєнтний спуск та його застосування в машинному навчанні</i>	6
4	<i>Класифікація даних</i>	6
5	<i>Зменшення розмірності даних</i>	6
6	<i>Напіваавтоматичне навчання та його застосування</i>	12
7	<i>Кластерний аналіз даних</i>	7
8	<i>Прогнозування даних</i>	12
9	<i>Техніки покращення точності моделей машинного навчання</i>	9
10	<i>Штучні нейронні мережі та глибинне навчання</i>	6
11	<i>Класифікація штучних нейронних мереж</i>	9
12	<i>Огляд програмних засобів та інфраструктури машинного навчання</i>	9
	Разом	101

9. Індивідуальні завдання

10. Методи навчання

Інформаційні методи (лекція, бесіда, презентація, ілюстрація, демонстрація); дедуктивні методи на основі узагальнень; евристичні методи (проблемна лекція); інтерактивні методи (дискусія).

11. Методи контролю

Поточний контроль здійснюється шляхом проведення усного опитування та написання письмових звітів по виконаних лабораторних роботах. У кінці курсу проводиться модуль у вигляді письмового есе на задану тему.

12. Розподіл балів, які отримують студенти

Поточне опитування та лабораторні роботи									Модульне завдання	Сума
Змістовий модуль 1			Змістовий модуль 2					Змістовий модуль 3		
Л1	Л2	Л3	Л4	Л5	Л6	Л7	Л8	Л9		
5	5	10	10	10	10	10	10	10	20	100

Л1, Л2 ... Л9 – теми лабораторних робіт.

Шкала оцінювання: національна та ЄКТС

Оцінка ЄКТС	Сума балів за всі види навчальної діяльності	Оцінка за національною шкалою	
		для екзамену, курсового проекту (роботи), практики	для заліку
A	90 – 100	відмінно	зараховано
B	81-89	добре	
C	71-80		
D	61-70		
E	51-60	задовільно	
FX	21-50	незадовільно з можливістю повторного складання	не зараховано з можливістю повторного складання
F	0-20	незадовільно з обов'язковим повторним вивченням дисципліни	не зараховано з обов'язковим повторним вивченням дисципліни

* кількість балів для оцінок «незадовільно» (FX і F) визначається Вченими радами факультетів (педагогічними радами коледжів).

13. Методичне забезпечення

14. Рекомендована література

Основна

- 1) Andreas C. Muller and Sarah Guido., Introduction to Machine Learning with Python.: Published by O'Reilly Media, Inc. – First Edition, 2016. – 392 pp.
- 2) Deep Learning with Python / François Chollet.: Manning Publications Co. 20 Baldwin Road PO Box 761 Shelter Island, 2018. – 386 pp.
- 3) Hands-On Machine Learning with Scikit-Learn and TensorFlow / Aurélien Géron. – Concepts, Tools, and Techniques to Build Intelligent Systems.: Published by O'Reilly Media, Inc. – Second Edition, 2019. – 564 pp.
- 4) Ian Goodfellow, Yoshua Bengio, Aaron Courville. Deep Learning.: The MIT Press Cambridge, 2016. - 800 pp.
- 5) Machine learning: A Probabilistic Perspective / Kevin P. Murphy.: Massachusetts Institute of Technology, 2012. – 1098 pp.
- 6) Python Data Science Handbook / Jake VanderPlas.: O'Reilly Media, Inc., 2016. – 573 pp.
- 7) Xiaojin Zhu, "Semi-Supervised Learning Literature Survey," University of Wisconsin, 2008. 60 pages.
- 8) Chris Piech, K Means. Stanford CS221.
- 9) R.J. Hyndman and G. Athanasopoulos, "Forecasting: principles and practice," 3rd edition, OTexts: Melbourne, Australia, 2021.

Допоміжна

- 1) Hands-On Machine Learning with Scikit-Learn and TensorFlow / Aurélien Géron. – Concepts, Tools, and Techniques to Build Intelligent Systems.: Published by O'Reilly Media, Inc. – First Edition, 2017. – 564 pp.
- 2) Xiaojin Zhu, "Semi-Supervised Learning Literature Survey," University of Wisconsin, 2005. 38 pages.
- 3) Amarnag Subramanya and Partha Pratim Talukdar, "Graph-Based Semi-Supervised Learning," Morgan and Claypool Publishers, 2014. - 126 pages.
- 4) Guoqiang Zhong Ph.D. and Kaizhu Huang Ph.D., "Semi-Supervised Learning: Background, Applications and Future Directions," Nova Science Pub Inc., 2018. - 229 pages.

15. Інформаційні ресурси

- 1) Internet – джерела.
- 2) Наукова бібліотека Львівського національного університету імені Івана Франка (<https://www.lnulibrary.lviv.ua/to-users-2/paid-services/internet/>).
- 3) Львівська національна наукова бібліотека України імені Василя Стефаника (<https://www.lsl.lviv.ua/index.php/uk/elektronni-resursy1/>).