

Базова інформація про дисципліну	
Назва дисципліни	Основи машинного навчання
Рівень вищої освіти / фахової передвищої освіти	Фахова передвища
Семестр	II семестр
Кафедра/циклова комісія	Циклова комісія програмування
Анотація курсу	<p>Даний курс формує практичні навички використання бібліотечних інструментів мови програмування Python для побудови проєктів з машинним навчанням. У якості інструментальної бази використовуються мова програмування Python, бібліотеки numpy, pandas, scipy, scikit-learn.</p> <p>Машинне навчання представляє собою множину математичних, статистичних та обчислювальних методів для розробки алгоритмів, здатних розв'язувати задачу на основі пошуку закономірностей в різноманітних вхідних даних. Програмні проєкти з машинним навчанням мають власний життєвий цикл, заснований на аналізі та підготовці даних, а також моделюванні та оцінці отриманого рішення. Таким чином, курс розбитий на змістові модулі, які включають використання Python-бібліотек для завантаження даних та їх попередньої обробки, розгляд різних моделей (алгоритмів) навчання та оцінки отриманих результатів.</p>
Сторінка курсу в MOODLE	http://78.137.2.119:1919/m72/course/view.php?id=104
Мова викладання	Українська

Лектор курсу	Марченко С.В., викладач II категорії канали комунікації: СДН «Moodle»: повідомлення в чаті E-mail: numosmanius@gmail.com
Місце дисципліни в освітній програмі	
Перелік загальних компетентностей (ЗК)	Здатність до абстрактного мислення, аналізу і синтезу. Здатність застосовувати знання у практичних ситуаціях.
Перелік спеціальних компетентностей (СК)	Здатність створювати програмне забезпечення для збору, опрацювання та зберігання даних. Здатність накопичувати, обробляти та систематизувати професійні знання щодо створення і супроводу програмного забезпечення та визнання важливості навчання протягом усього життя. Здатність застосовувати фундаментальні та міждисциплінарні знання для успішного розв'язання завдань інженерії програмного забезпечення.
Перелік програмних результатів навчання	Вміти знаходити аналогії та застосовувати знання, вміння та навички з суміжних дисциплін для формування та вирішення професійних завдань. Знати основні інструментальні засоби для розробки та супроводу програмного забезпечення та вміти застосовувати їх на практиці з урахуванням специфіки отриманого завдання та вимог користувача. Знати основні підходи до видобування, зберігання, обробки даних та вміти застосовувати їх для створення відповідного програмного забезпечення.
Опис дисципліни	
Структура навантаження на студента	Загальна кількість годин – 90 Кількість кредитів – 3 Кількість лекційних годин – 0 Кількість практичних занять – 34 Кількість годин для самостійної роботи

	студентів – 56 Форма підсумкового контролю – залік
Методи навчання	За подачею навчального матеріалу: методи готових знань, дослідницький метод. З огляду на мету навчання: методи здобуття нових знань, метод формування умінь і навичок, метод застосування знань на практиці, методи закріплення знань, умінь і навичок, методи перевірки і оцінювання знань, умінь і навичок.
Зміст дисципліни	
Тема 1. Математичні основи машинного навчання. Бібліотека numpy	Бібліотека numpy: вектори, матриці, масиви. Елементи роботи з зображеннями. Статистичні функції numpy.
Тема 2. Завантаження та обробка даних. Пакет pandas	Базове влаштування та організація доступу до даних. Завантаження та нормалізація даних. Базові перетворення даних у pandas. Фільтрування даних. Групування та агрегація даних.
Тема 3. Основи візуалізації даних	Поняття візуалізації даних. Типи візуалізації. Побудова графіків для різних задач засобами бібліотеки Matplotlib. Управління кольором при програмній візуалізації даних.
Тема 4. Навчання на основі регресійного аналізу	Проекти машинного навчання та бібліотека scikit-learn. Проста лінійна регресія. Регресія та метод k найближчих сусідів (kNN). Багатовимірні лінійні регресії. Поліноміальна регресія.
Тема 5. Навчання з учителем. Задача класифікації	Бінарна класифікація на базі логістичної регресії. Наївний баєсівський класифікатор. Метод опорних векторів та SVM-класифікатори.
Тема 6. Навчання без учителя. Задача кластеризації	Кластеризація за допомогою k-середніх. Кластеризація методом зсуву до середнього. Кластеризація методом DBSCAN. Кластеризація методом ієрархічного злиття.
Тема 7. Деревя прийняття рішень та випадкові ліси	Класифікаційні дерева прийняття рішень. Регресійні дерева прийняття рішень. Класифікаційні випадкові ліси. Регресійні

	випадкові ліси.
Тема 8. Ансамблеві методи навчання	Методи збору ансамблів алгоритмів машинного навчання. Прості ансамблеві техніки: класифікатори з голосуванням, усереднення, зважене усереднення. Беггінг і вставка. Бустинг. Стекінг (укладка).

Політика дисципліни

Політика відвідування	Регулярне відвідування всіх видів занять, своєчасність виконання самостійної роботи. За об'єктивних причин (наприклад, хвороба, міжнародне стажування) навчання зорганізується в он-лайн формі за погодженням із керівником курсу.
Політика щодо дедлайнів та перескладання	Роботи, які здаються із порушенням термінів без поважних причин, оцінюються на нижчу оцінку.
Академічна доброчесність	У випадку недотримання політики академічної доброчесності (плагіат, самоплагіат, фабрикація, фальсифікація, списування, обман, хабарництво) передбачено повторне проходження оцінювання.

Система оцінювання

Поточний контроль здійснюється протягом семестру під час проведення практичних, семінарських та інших видів занять і оцінюється сумою набраних балів (максимальна сума – 100 балів; мінімальна сума, що дозволяє студенту отримати атестацію з предмету – 60 балів); підсумковий/ семестровий контроль, проводиться у формі заліку або іспиту, відповідно до графіку навчального процесу.

Підсумкова оцінка за умови заліку виставляється як загальна сума балів, набраних за результатами поточного контролю. Підсумкова оцінка за умови іспиту виставляється як загальна сума балів набраних за результатами поточного (70%) та підсумкового контролю.

Накопичування рейтингових балів з навчальної дисципліни

Види навчальної роботи	Мах кількість балів
Практичні завдання (8 тем)	56
Тестування (2 тести)	14
Захист практичного завдання до самостійної роботи	30
Разом	100

Шкала оцінювання		
ECTS	Бали	Зміст
A	90-100	Бездоганна підготовка в широкому контексті
B	80-89	Повні знання, міцні вміння
C	70-79	Хороші знання та вміння
D	65-69	Задовільні знання, стереотипні вміння
E	60-64	Виконання мінімальних вимог діяльності в стандартних умовах
FX	35-59	Слабкі знання, відсутність умінь
F	1-34	Необхідний повторний курс

Список рекомендованих джерел

1. Харрисон М. Как устроен Python. Гид для разработчиков, программистов и интересующихся. СПб.: Питер, 2019. 272 с.
2. Лутц М. Изучаем Python, том 1, 5-е издание. СПб.: Символ-Плюс, 2019. 832 с.
3. Stepanek H. Thinking in Pandas: How to Use the Python Data Analysis Library the Right Way. Portland: Apress, 2020. – 188с.
4. Harrison M., Petrou T. Pandas 1.x Cookbook: Practical recipes for scientific computing, time series analysis, and exploratory data analysis using Python, Second Edition. Birmingham: Packt Publishing, 2020. – 606с.
5. Элбон К. Машинное обучение с использованием Python. Сборник рецептов: Пер. с англ. СПб.: БХВ-Петербург, 2019. 384 с.: ил.
6. Жерон О. Прикладное машинное обучение с помощью Scikit-Learn, Keras и TensorFlow: концепции, инструменты и техники для создания интеллектуальных систем, 2-е изд.: Пер. С англ. – СПб.: ООО “Диалектика”, 2020. – 1040с.: ил.
7. Liu Y. Python Machine Learning By Example: Build intelligent systems using Python, TensorFlow 2, PyTorch, and scikit-learn, Third Edition. Birmingham: Packt Publishing, 2020. – 504с.