



Силабус освітнього компонента Програма навчальної дисципліни



Інструменти Python для AI та Data Science

Шифр та назва спеціальності
122 – Комп'ютерні науки

Інститут
ННІ Комп'ютерних наук та інформаційних
технологій

Освітня програма
Комп'ютерні науки. Штучний інтелект та
управління проєктами

Кафедра
Системного аналізу та інформаційно-
аналітичних технологій

Рівень освіти
Бакалавр

Тип дисципліни
Профільна підготовка. Вибіркова

Семестр
3

Мова викладання
Українська

Викладачі, розробники



Коваленко Світлана Миколаївна

Svitlana.Kovalenko@khpі.edu.ua

канд. техн. наук, доцент, доцент кафедри системного аналізу та
інформаційно-аналітичних технологій (за сумісництвом)

Google Scholar:

<https://scholar.google.com/citations?user=jeD1w74AAAAJ&hl>

ORCID: <https://orcid.org/0000-0001-6770-6778>

Scopus: <https://www.scopus.com/authid/detail.uri?authorId=57212035934>

Web of Science: <https://www.webofscience.com/wos/author/record/F-8252-2017>

[Детальніше про викладача на сайті кафедри](#)

Загальна інформація

Анотація

Дисципліна спрямована на ознайомлення студентів з ключовими інструментами, бібліотеками та концепціями, які необхідні для роботи в галузі штучного інтелекту (AI) та науки про дані (Data Science) з використанням мови програмування Python. Вона допомагає студентам оволодіти основними інструментами та технологіями, які використовуються в сферах аналізу даних. Також дисципліна надає студентам практичні навички та досвід роботи з реальними наборами даних та вирішення реальних задач у галузі аналізу даних та штучного інтелекту для прийняття обґрунтованих рішень.

Мета та цілі дисципліни

Мета викладання дисципліни полягає у формуванні системи теоретичних знань і практичних навичок щодо ефективної роботи з даними з використанням мови програмування Python, розуміння методів обробки, аналізу та візуалізації даних. Завдання дисципліни визначаються змістом і специфікою її предмета і спрямоване на ознайомлення студентів з ключовими інструментами та технологіями, що використовуються в галузях AI та Data Science, зокрема Python та його бібліотеки (Pandas, NumPy, Matplotlib, Seaborn, Scikit-learn, Tensorflow тощо)

Формат занять

Лекції, лабораторні роботи, самостійна робота, консультації. Підсумковий контроль – залік.

Компетентності

ЗК2. Здатність застосовувати знання у практичних ситуаціях.

ЗК7. Здатність до пошуку, оброблення та аналізу інформації з різних джерел.

СК2. Здатність до виявлення статистичних закономірностей недетермінованих явищ, застосування методів обчислювального інтелекту, зокрема статистичної, нейромережевої та нечіткої обробки даних, методів машинного навчання та генетичного програмування тощо.

СК11. Здатність до інтелектуального аналізу даних на основі методів обчислювального інтелекту включно з великими та погано структурованими даними, їхньої оперативної обробки та візуалізації результатів аналізу в процесі розв'язування прикладних задач.

СК17. Здатність проектувати та розробляти системи штучного інтелекту на основі використання штучних нейронних мереж різноманітної архітектури.

Результати навчання

РН4. Використовувати методи обчислювального інтелекту, машинного навчання, нейромережевої та нечіткої обробки даних, генетичного та еволюційного програмування для розв'язання задач розпізнавання, прогнозування, класифікації, ідентифікації об'єктів керування тощо.

РН17. Застосовувати штучні нейронні мережі, що розроблені на основі використання бібліотек, розширень та фреймворків мови Python, для вирішення практичних задач інтелектуального аналізу даних.

Обсяг дисципліни

Загальний обсяг дисципліни 120 год. (4 кредити ECTS): лекції – 32 год., лабораторні роботи – 32 год., самостійна робота – 56 год.

Передумови вивчення дисципліни (пререквізити)

Для успішного проходження курсу необхідно мати знання та практичні навички з наступних дисциплін: "Вступ до комп'ютерних наук", "Основи програмування".

Особливості дисципліни, методи та технології навчання

Лекції проводяться інтерактивно з використанням мультимедійних технологій. На лабораторних заняттях використовується проєктний підхід до навчання, акцентується увага на застосуванні інформаційних технологій. Навчальні матеріали доступні студентам.

Програма навчальної дисципліни

Теми лекційних занять

Тема 1. Основи AI та Data Science. Основи синтаксису Python. Середовища розробки.

Поняття, основні завдання, взаємозв'язок штучного інтелекту та науки про дані. Класифікація та загальний огляд етапів та методів. Встановлення Python. Середовища розробки.

Тема 2. Основи мови програмування Python.

Введення та виведення даних. Змінні та типи даних: огляд типів даних в Python, операції та функції, що виконуються над різними типами даних. Умовні конструкції та цикли. Функції: визначення та виклик функцій, параметри та аргументи функцій, повернення значень з функцій. Робота зі списками та рядками. Обробка виключень. Класи. Використання вбудованих та сторонніх модулів.

Тема 3. Jupyter notebook та бібліотека Numpy.

Встановлення та основи роботи з Jupyter notebook. Google Colaboratory. Історія створення Numpy. Масиви ndarray. Порівняння масивів NumPy та списків Python. Типи даних NumPy. Створення

масивів. Операції над масивами. Універсальні функції. Конкатенація та розділення масивів NumPy. Broadcasting. Порівняння, маски та булева логіка. Fancy Indexing.

Тема 4. Візуалізація даних.

Бібліотеки для візуалізації даних: Matplotlib, Seaborn, Plotly. Інтерфейси Matplotlib. Встановлення стилів. Прості лінійні графіки. Налаштування графіків. Типи графіків. Візуалізація похибок. Створення декількох графіків на одному рисунку. Створення теплових карт та інтерактивних графіків.

Тема 5. Обробка даних за допомогою Pandas.

Об'єкти Pandas. Властивості Series та DataFrame. Побудова об'єктів Pandas. Об'єкт Index. Індексція та вибір даних. Операції з даними. Збереження індексу. Універсальні функції та вирівнювання індексу. Операції між DataFrame та Series. Маніпуляції з даними. Групування даних. Обробка пропущених даних. Ієрархічна індексція. Поєднання наборів даних: concat(), append(), merge(), join(). Зведені таблиці. Візуалізація в Pandas.

Тема 6. Робота з текстом

Обробка тексту методами стандартної бібліотеки Python. Бібліотека re та регулярні вирази пошуку, заміни та аналізу тексту. Шаблони і класи символів. Квантифікатори. Групи. Обробка рядків в Pandas. Векторизація рядкових операторів. Методи Pandas, що використовують регулярні вирази. Токенізація, стемінг, лемматизація тексту з бібліотекою NLTK. Аналіз настрою, сентименту та інші завдання NLP. Набори україномовних текстів для задач машинного навчання. Основні можливості бібліотек spaCy, TextBlob, gensim.

Тема 7. Робота з графікою. PIL та opencv-python

Встановлення бібліотек. Представлення зображень в бібліотеках. Колірні моделі. Формати графічних даних. Основні маніпуляції з графічними даними: зміна розмірів, обрізання, транспонування, поворот, вставка одного зображення в інше, створення «водяних знаків». Визначення об'єктів на зображенні. Pytesseract та розпізнавання тексту.

Тема 8. Бібліотека інструментів машинного навчання Scikit-learn

Огляд можливостей, основні поняття: навчальні та тестові набори даних, моделі та їх гіперпараметри. Навчання з учителем та без учителя. Оцінка моделей: метрики ефективності, кросвалідація, матриця помилок. Експорт та використання моделей.

Тема 9. Бібліотека інструментів глибокого навчання Tensorflow.

Створення моделей за допомогою Tensorflow Keras API. Визначення архітектури нейронних мереж: повністю з'єднані мережі, згорткові мережі, рекурентні мережі. Компіляція моделі: вибір функції втрат та оптимізатора. Навчання моделі на тренувальних даних. Оцінка ефективності моделі за допомогою метрик. Регуляризація та уникнення перенавчання. Перенос навчання. Збереження моделей.

Тема 10. Веб-розробка для Data Science

Огляд та встановлення мікрофреймворка Flask. Створення, налаштування та запуск першого застосунка. Використання шаблонів, передача змінних у шаблони та використання їх для відображення даних. Реагування на різні типи HTTP-запитів (GET, POST, PUT, DELETE). Отримання та обробка даних, що надходять з клієнта. Використання Flask-WTF. Інтеграція моделей машинного навчання у веб-додатки для представлення результатів. Розгортання додатку.

Теми практичних занять

Практичні роботи в рамках дисципліни не передбачені.

Теми лабораторних робіт

Тема 1. Основи роботи з Python та Jupyter notebook.

Тема 2. Аналіз та візуалізація даних засобами Numpy та Matplotlib.

Тема 3. Аналіз даних засобами Pandas.

Тема 4. Об'єднання наборів даних у Pandas та робота з пропущеними даними.

Тема 5. Обробка цифрових зображень.

Тема 6. Основи роботи з моделями машинного навчання в бібліотеці Scikit-learn.

Тема 7. Основи роботи з моделями глибокого навчання Tensorflow.

Тема 8. Створення веб-інтерфейсів для моделей машинного навчання та аналізу даних

Самостійна робота

Опрацювання лекційного матеріалу.

Підготовка до лабораторних занять та контрольних робіт (тестів).

Самостійне вивчення тем та питань, які не викладаються на лекційних заняттях.

Література та навчальні матеріали

Основна література

1. Wes McKinney. Python for Data Analysis: Data Wrangling with pandas, NumPy, and Jupyter, 3rd edition, 2022, 540 p. URL: <https://wesmckinney.com/book/>
2. Jake VanderPlas. Python Data Science Handbook, 2d edition, O'Reilly Media, 2022, 550 p. URL: <https://jakevdp.github.io/PythonDataScienceHandbook/>
3. Robert Johansson. Numerical Python. Scientific Computing and Data. Science Applications with Numpy, SciPy and Matplotlib. Second Edition, 2019, 685 p. URL: <https://rjohansson.github.io/numericalpython.html>
4. 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. URL: <https://ir.lib.vntu.edu.ua/bitstream/handle/123456789/36454/120997.pdf?sequence=2&isAllowed=y>
5. Копей В. Б. Мова програмування Python для інженерів і науковців : навчальний посібник / В. Б. Копей – ІваноФранківськ : ІФНТУНГ, 2019. – 272 с. URL: https://shron1.chtyvo.org.ua/Kopei_Volodymyr/Mova_prohramuvannia_Python_dlia_inzheneriv_i_naukovtsiv.pdf
6. Костюченко А. О. Основи програмування мовою Python: навчальний посібник. Ч.: ФОП Балакіна С.М., 2020. 180 с. URL: <https://epub.chnpu.edu.ua/jspui/bitstream/123456789/5584/1/%D0%9E%D1%81%D0%BD%D0%BE%D0%B2%D0%B8%20%D0%BF%D1%80%D0%BE%D0%B3%D1%80%D0%B0%D0%BC%D1%83%D0%B2%D0%B0%D0%BD%D0%BD%D1%8F%20%D0%BC%D0%BE%D0%B2%D0%BE%D1%8E%20Python.pdf>

Додаткова література

7. Методичні вказівки до лабораторної роботи "Основи роботи з бібліотекою Pandas" з курсу "Обробка даних Python" [Електронний ресурс] : для студентів спец. 121 Інженерія програмного забезпечення, 122 Комп'ютерні науки, 124 Системний аналіз / уклад.: С. М. Коваленко, С. В. Коваленко ; Нац. техн. ун-т "Харків. політехн. ін-т". – Електрон. текст. дані. – Харків, 2022. – 44 с. – URL: <http://repository.kpi.kharkov.ua/handle/KhPI-Press/61382>
8. Методичні вказівки до лабораторної роботи "Основи роботи з бібліотекою Matplotlib" з курсу "Обробка даних Python" [Електронний ресурс] : для студентів спец. 121 Інженерія програмного забезпечення, 122 Комп'ютерні науки, 124 Системний аналіз, 126 Інформаційні системи і технології / уклад.: С. М. Коваленко, С. В. Коваленко, О. В. Шматко ; Нац. техн. ун-т "Харків. політехн. ін-т". – Електрон. текст. дані. – Харків, 2021. – 28 с. – URL: <http://repository.kpi.kharkov.ua/handle/KhPI-Press/55469>.
9. Методичні вказівки до лабораторної роботи "Основи роботи з бібліотекою NumPy" з курсу "Обробка даних Python" [Електронний ресурс] : для студентів спец. 121 Інженерія програмного забезпечення, 122 Комп'ютерні науки, 124 Системний аналіз, 126 Інформаційні системи і технології / уклад.: С. М. Коваленко, С. В. Коваленко, О. В. Шматко ; Нац. техн. ун-т "Харків. політехн. ін-т". – Електрон. текст. дані. – Харків, 2021. – 48 с. – URL: <http://repository.kpi.kharkov.ua/handle/KhPI-Press/55468>
10. Методичні вказівки до лабораторної роботи "Основи роботи в середовищі Jupyter Notebook" з курсу "Обробка даних Python" [Електронний ресурс] : для студентів спец. 121 Інженерія програмного забезпечення, 122 Комп'ютерні науки, 124 Системний аналіз, 126 Інформаційні системи і технології / уклад.: С. М. Коваленко, С. В. Коваленко, О. В. Шматко ; Нац. техн. ун-т "Харків. політехн. ін-т". – Електрон. текст. дані. – Харків, 2021. – 28 с. – URL: <http://repository.kpi.kharkov.ua/handle/KhPI-Press/55466>
11. Путівник мовою програмування Python [Електронний ресурс]. – Режим доступу : <https://pythonguide.rozh2sch.org.ua>

Система оцінювання

Критерії оцінювання успішності студента та розподіл балів

100% підсумкової оцінки складаються з результатів оцінювання у вигляді заліку (40%) та поточного оцінювання (60%):

- 8 лабораторних робіт (по 5%);
- 2 контрольні роботи (по 10%).

Шкала оцінювання

Сума балів	Національна оцінка	ECTS
90–100	Відмінно	A
82–89	Добре	B
75–81	Добре	C
64–74	Задовільно	D
60–63	Задовільно	E
35–59	Незадовільно (потрібне додаткове вивчення)	FX
1–34	Незадовільно (потрібне повторне вивчення)	F

Норми академічної етики і політика курсу

Студент повинен дотримуватися «Кодексу етики академічних взаємовідносин та доброчесності НТУ «ХПІ»: виявляти дисциплінованість, вихованість, доброзичливість, чесність, відповідальність. Конфліктні ситуації повинні відкрито обговорюватися в навчальних групах з викладачем, а при неможливості вирішення конфлікту – доводитися до відома співробітників дирекції інституту. Нормативно-правове забезпечення впровадження принципів академічної доброчесності НТУ «ХПІ» розміщено на сайті: <http://blogs.kpi.kharkov.ua/v2/nv/akademichna-dobrochesnist/>

Погодження

Силабус погоджено

25.08.2023

Завідувач кафедри
Юрій ДОРОФЄЄВ

25.08.2023

Гарант ОП
Марина ГРИНЧЕНКО