**PYTHON ДЛЯ АНАЛІЗУ ДАНИХ**

**Викладач:** *кандидат технічних наук, ст. викладач Добровольський Геннадій Анатолійович*

**Кафедра:** *комп’ютерних наук, І корпус, ауд. 39*

**E-mail:** *gen.dobr@gmail.com*

**Телефон:** *(061) 289-12-57*

**Інші засоби зв’язку:** *Moodle (форум курсу, приватні повідомлення), Telegram:* [*https://t.me/gen\_dobr*](https://t.me/gen_dobr)

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Освітня програма, рівень вищої освіти:** | | Інформаційні системи та технології, комп’ютерні науки  Бакалавр | | | | | |
| **Статус дисципліни:** | | Вибіркова | | | | | |
| **Кредити ECTS** | 5 | **Навч. рік:** | 2022-23 | **Рік навчання** | 3 | **Тижні** | 14 |
| **Кількість годин** | 150 | **Кількість змістових модулів[[1]](#footnote-2)** | 8 | **Лекційні заняття** – 28  **Лабораторні заняття** – 28  **Самостійна робота** – 94 | | | |
| **Вид контролю:** | | Залік | |  | | | |
| **Посилання на курс в Moodle** | | | https://moodle.znu.edu.ua/course/view.php?id=13793 | | | | |
| **Консультації:** | | | | | | | |

**ОПИС КУРСУ**

*В межах дисципліни «Python для аналізу даних» вивчаються бібліотеки мови Python для візуалізації, маніпуляції табличними даними, аналізу графів.*

***Метою*** *викладання навчальної дисципліни «Python для аналізу даних» є формування у студентів та слухачів навичок презентації та обробки табличних даних, аналізу графів.*

***Основними завданнями*** *вивчення дисципліни є вивчення принципів та способів візуалізації числових даних; читання, перетворення, об’єднання, групування, збереження табличних даних, операції з графами.*

**ОЧІКУВАНІ РЕЗУЛЬТАТИ НАВЧАННЯ**

*У результаті вивчення дисципліни студенти повинні:*

***знати:***

* *ознаки якісної візуалізації даних;*
* *методи візуального представлення числових даних;*
* *методи читання табличних даних із різних джерел;*
* *методи перетворення табличних даних;*
* *методи збереження табличних даних;*
* *прості методи аналізу текстів;*
* *методи аналізу графів.*

***вміти:***

* *використовувати бібліотеку matplotlib для візуалізації даних;*
* *використовувати бібліотеку pandas для обробки табличних даних;*
* *використовувати бібліотеку NLTK для обробки текстових даних;*
* *використовувати бібліотеку networkX для обробки та аналізу графів.*

*Згідно з вимогами освітньо-професійної програми студенти повинні досягти таких* ***компетенцій****:*

* *Здатність до абстрактного мислення, аналізу та синтезу;*
* *Здатність застосовувати знання у практичних ситуаціях;*
* *Здатність до пошуку, оброблення та аналізу інформації з різних джерел;*
* *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;*
* *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;*
* *Здатність реалізовувати високопродуктивні обчислення на основі хмарних сервісів і технологій, паралельних і розподілених обчислень при розробці й експлуатації розподілених систем паралельної обробки інформації*

*та наступних* ***програмних результатів навчання****:*

* *Використовувати методи обчислювального інтелекту, машинного навчання, нейро-мережевої та нечіткої обробки даних, генетичного та еволюційного програмування для розв’язання задач розпізнавання, прогнозування, класифікації, ідентифікації об’єктів керування тощо;*
* *Застосовувати методи та алгоритми обчислювального інтелекту та інтелектуального аналізу даних в задачах класифікації, прогнозування, кластерного аналізу, пошуку асоціативних правил з використанням програмних інструментів підтримки багатовимірного аналізу даних на основі технологій DataMining, TextMining, WebMining.*
* *Розробляти, аналізувати та використовувати математичні моделі предметних областей, що використовуються в інформаційних системах.*

**ОСНОВНІ НАВЧАЛЬНІ РЕСУРСИ**

*Презентації лекцій, методичні рекомендації до виконання лабораторних робіт та індивідуального завдання, тести у системі Moodle*

**КОНТРОЛЬНІ ЗАХОДИ**

***Поточні контрольні заходи***

Поточний контроль передбачає такі ***теоретичні*** завдання:

* *усне опитування з теоретичного матеріалу за темою на початку кожного лабораторного заняття. Перелік питань з кожної лабораторної роботи розміщено у файлі з завданням до лабораторної роботи у системі Moodle.*
* *поточний тест за пройденим матеріалом.*

Поточний контроль передбачає таке ***практичне*** завдання:

* *виконання лабораторних робіт.*

***Підсумкові контрольні заходи:***

***Індивідуальне завдання*** *– візуальне дослідження даних.*

***Підсумковий тест*** *– підсумкове тестування з курсу за обмежений час у системі Moodle.*

|  |  |  |  |
| --- | --- | --- | --- |
| **Контрольний захід** | | **Термін виконання** | **% від загальної оцінки** |
| **Поточний контроль (max 60%)** | |  |  |
| *Змістовий модуль 1 (розділ 1)* | *Опитування з теоретичного матеріалу* | *Тиждень 1* | **2** |
| *Лабораторна робота 1* | *Тиждень 1* | **4** |
| *Змістовий модуль 2 (розділ 1)* | *Опитування з теоретичного матеріалу* | *Тиждень 2,3* | **2** |
| *Лабораторна робота 2* | *Тиждень 2,3* | **4** |
| *Змістовий модуль 3 (розділ 1)* | *Опитування з теоретичного матеріалу* | *Тиждень 4,5* | **2** |
| *Лабораторна робота 3* | *Тиждень 4,5* | **4** |
| *Змістовий модуль 4 (розділ 1)* | *Опитування з теоретичного матеріалу* | *Тиждень 6, 7* | **2** |
| *Лабораторна робота 4* | *Тиждень 6, 7* | **4** |
|  | *Поточний тест* | *Тиждень 7* | **6** |
| *Змістовий модуль 5 (розділ 2)* | *Опитування з теоретичного матеріалу* | *Тиждень 8,9* | **2** |
| *Лабораторна робота 5* | *Тиждень 8,9* | **4** |
| *Змістовий модуль 6 (розділ 2)* | *Опитування з теоретичного матеріалу* | *Тиждень 10,11* | **2** |
| *Лабораторна робота 6* | *Тиждень 10,11* | **4** |
| *Змістовий модуль 7 (розділ 2)* | *Опитування з теоретичного матеріалу* | *Тиждень 12, 13* | **2** |
| *Лабораторна робота 7* | *Тиждень 12, 13* | **4** |
| *Змістовий модуль 8 (розділ 2)* | *Опитування з теоретичного матеріалу* | *Тиждень 14* | **2** |
| *Лабораторна робота 8* | *Тиждень 14* | **4** |
|  | *Поточний тест* | *Тиждень 14* | **6** |
| **Підсумковий контроль (max 40%)** | |  |  |
| *Підсумковий тест* | |  | **20** |
| *Індивідуальне завдання* | |  | **20** |
| **Разом** | |  | **100%** |

**Шкала оцінювання: національна та ECTS**

|  |  |  |  |
| --- | --- | --- | --- |
| За шкалоюECTS | За шкалою університету | За національною шкалою | |
| Екзамен | Залік |
| A | 90 – 100 (відмінно) | 5 (відмінно) | Зараховано |
| B | 85 – 89 (дуже добре) | 4 (добре) |
| C | 75 – 84 (добре) |
| D | 70 – 74 (задовільно) | 3 (задовільно) |
| E | 60 – 69 (достатньо) |
| FX | 35 – 59 (незадовільно – з можливістю повторного складання) | 2 (незадовільно) | Не зараховано |
| F | 1 – 34 (незадовільно – з обов’язковим повторним курсом) |

**РОЗКЛАД КУРСУ ЗА ТЕМАМИ І КОНТРОЛЬНІ ЗАВДАННЯ**

|  |  |  |  |
| --- | --- | --- | --- |
| **Тиждень**  **і вид заняття** | **Тема заняття** | **Контрольне завдання** | **Кількість балів** |
| Змістовий модуль 1. | | | |
| Тиждень 1, 2  Лекція 1 | Можливості пакета Pandas. Основні структури даних, методи їх створення та зміни. |  |  |
| Тиждень 1, 2 Лабораторна робота 1 | Знайомство з пакетом Python Pandas. | Усне опитування з теми.  Виконання завдань лабораторної роботи | 6 |
| Змістовий модуль 2. | | | |
| Тиждень 3  Лекція 2 | Перевірка статистичних гіпотез |  |  |
| Тиждень 3  Лабораторна робота 2 | Перевірка статистичних гіпотез | Усне опитування з теми.  Виконання завдань лабораторної роботи | 6 |
| Змістовий модуль 3. | | | |
| Тиждень 4, 5  Лекція 3 | Очищення даних засобами пакета Pandas |  |  |
| Тиждень 4, 5  Лабораторна робота 3 | Очищення даних | Усне опитування з теми.  Виконання завдань лабораторної роботи | 6 |
| Змістовий модуль 4. | | | |
| Тиждень 6, 7  Лекція 4 | Візуалізація табличних даних |  |  |
| Тиждень 6, 7  Лабораторна робота 4 | Візуалізація табличних даних | Усне опитування з теми.  Виконання завдань лабораторної роботи | 6 |
| Тиждень 7  Поточний тест |  | Тестові завдання в системі Moodle | 6 |
| Змістовий модуль 5. | | | |
| Тиждень 8, 9  Лекція 5 | Обробка текстових даних засобами пакету NLTK |  |  |
| Тиждень 8, 9  Лабораторна робота 5 | Класифікація тексту | Усне опитування з теми.  Виконання завдань лабораторної роботи | 6 |
| Змістовий модуль 6. | | | |
| Тиждень 10  Лекція 10 | Тематичне моделювання |  |  |
| Тиждень 10  Лабораторна робота 6 | Тематичне моделювання | Усне опитування з теми.  Виконання завдань лабораторної роботи | 6 |
| Змістовий модуль 7. | | | |
| Тиждень 11, 12  Лекція 7 | Збір даних в Інтернеті |  |  |
| Тиждень 11, 12  Лабораторна робота 7 | Збір даних в Інтернеті | Усне опитування з теми.  Виконання завдань лабораторної роботи | 6 |
| Змістовий модуль 8. | | | |
| Тиждень 13, 14  Лекція 8 | Аналіз графів, пакет networkX |  |  |
| Тиждень 13, 14  Лабораторна робота 8 | Дослідження графа | Усне опитування з теми.  Виконання завдань лабораторної роботи | 6 |
| Тиждень 14  Поточний тест |  | Тестові завдання в системі Moodle | 6 |

**ОСНОВНІ ДЖЕРЕЛА**

1. pandas - Python Data Analysis Library <https://pandas.pydata.org/>
2. Data Manipulation with pandas <https://www.datacamp.com/courses/data-manipulation-with-pandas> .
3. Steven Bird, Ewan Klein, and Edward Loper, Natural Language Processing with Python, <https://www.nltk.org/book/>
4. Kyran Dale, Data Visualization with Python and JavaScript: Scrape, Clean, Explore & Transform Your Data, O'Reilly Media: 2016. 592 pp
5. A Tutorial on NetworkX: Network Analysis in Python, https://medium.com/swlh/a-tutorial-on-networkx-network-analysis-in-python-part-i-43c1d35830b6
6. A Practical Introduction to Web Scraping in Python. https://realpython.com/python-web-scraping-practical-introduction/

**РЕГУЛЯЦІЇ І ПОЛІТИКИ КУРСУ[[2]](#footnote-3)**

**Відвідування занять. Регуляція пропусків.**

*Відвідування лекційних і лабораторних занять є обов’язковим. Студенти, які за певних обставин не можуть відвідувати лабораторні заняття регулярно, мусять впродовж тижня узгодити із викладачем графік індивідуального відпрацювання пропущених занять.*

*.*

**Політика академічної доброчесності**

*Усі письмові роботи, що виконуються слухачами під час проходження курсу, перевіряються на наявність плагіату за допомогою спеціалізованого програмного забезпечення UniCheck. Відповідно до чинних правових норм, плагіатом вважатиметься: копіювання чужої наукової роботи чи декількох робіт та оприлюднення результату під своїм іменем; створення суміші власного та запозиченого тексту без належного цитування джерел; рерайт (перефразування чужої праці без згадування оригінального автора). Будь-яка ідея, думка чи речення, ілюстрація чи фото, яке ви запозичуєте, має супроводжуватися посиланням на першоджерело.*

*Роботи, у яких виявлено ознаки плагіату, до розгляду не приймаються і відхиляються без права перескладання. Якщо ви не впевнені, чи підпадають зроблені вами запозичення під визначення плагіату, будь ласка, проконсультуйтеся з викладачем.*

*Висока академічна культура та європейські стандарти якості освіти, яких дотримуються у ЗНУ, вимагають від дослідників відповідального ставлення до вибору джерел. Посилання на такі ресурси, як Wikipedia, бази даних рефератів та письмових робіт (Studopedia.org та подібні) є неприпустимим. Рекомендовані бази даних для пошуку джерел:*

*Електронні ресурси Національної бібліотеки ім. Вернадського:* [http://www.nbuv.gov.ua](http://www.nbuv.gov.ua/)

*Цифрова повнотекстова база даних англомовної наукової періодики JSTOR:* <https://www.jstor.org/>

**Використання комп’ютерів/телефонів на занятті**

*Використання мобільних телефонів, планшетів та інших гаджетів під час лекційних та лабораторних занять дозволяється виключно у навчальних цілях. Будь ласка, не забувайте активувати режим «без звуку» до початку заняття.*

*Під час виконання заходів контролю (поточних та підсумкового тестів) використання гаджетів заборонено. У разі порушення цієї заборони роботу буде анульовано без права перескладання.*

**Комунікація**

*Базовою платформою для комунікації викладача зі студентами є Moodle.*

*Важливі повідомлення загального характеру розміщуються викладачем на форумі курсу. Для персональних запитів використовується сервіс приватних повідомлень. Відповіді на запити студентів подаються викладачем впродовж трьох робочих днів. Для оперативного отримання повідомлень про оцінки та нову інформацію, розміщену на сторінці курсу у Moodle, будь ласка, переконайтеся, що адреса електронної пошти, зазначена у вашому профайлі на Moodle, є актуальною, та регулярно перевіряйте папку «Спам».*

*Якщо за технічних причин доступ до Moodle є неможливим, або ваше питання потребує термінового розгляду, направте електронного листа з позначкою «Важливо» на адресу gen.dobr@gmail.com. У листі обов’язково вкажіть ваше прізвище та ім’я, курс та шифр академічної групи.*

**ДОДАТОК ДО СИЛАБУСУ ЗНУ – 2020-2021**

***ГРАФІК НАВЧАЛЬНОГО ПРОЦЕСУ 2020-2021 н. р. (гіперпосилання на сторінку сайта)***

***АКАДЕМІЧНА ДОБРОЧЕСНІСТЬ.*** Студенти і викладачі Запорізького національного університету несуть персональну відповідальність за дотримання принципів академічної доброчесності, затверджених ***Кодексом академічної доброчесності ЗНУ*:** <https://tinyurl.com/ya6yk4ad>. *Декларація академічної доброчесності здобувача вищої освіти* (додається в обов’язковому порядку до письмових кваліфікаційних робіт, виконаних здобувачем, та засвідчується особистим підписом): <https://tinyurl.com/y6wzzlu3>.

***ОСВІТНІЙ ПРОЦЕС ТА ЗАБЕЗПЕЧЕННЯ ЯКОСТІ ОСВІТИ.*** Перевірка набутих студентами знань, навичок та вмінь (атестації, заліки, іспити та інші форми контролю) є невід’ємною складовою системи забезпечення якості освіти і проводиться відповідно до *Положення про організацію та методику проведення поточного та підсумкового семестрового контролю навчання студентів ЗНУ*: <https://tinyurl.com/y9tve4lk>**.**

***ПОВТОРНЕ ВИВЧЕННЯ ДИСЦИПЛІН, ВІДРАХУВАННЯ.*** Наявність академічної заборгованості до 6 навчальних дисциплін (в тому числі проходження практики чи виконання курсової роботи) за результатами однієї екзаменаційної сесії є підставою для надання студенту права на повторне вивчення зазначених навчальних дисциплін. Порядок повторного вивчення визначається *Положенням про порядок повторного вивчення навчальних дисциплін та повторного навчання у ЗНУ*: <https://tinyurl.com/y9pkmmp5>. Підстави та процедури відрахування студентів, у тому числі за невиконання навчального плану, регламентуються *Положенням про порядок переведення, відрахування та поновлення студентів у ЗНУ*: <https://tinyurl.com/ycds57la>.

***НЕФОРМАЛЬНА ОСВІТА.*** Порядок зарахування результатів навчання, підтверджених сертифікатами, свідоцтвами, іншими документами, здобутими поза основним місцем навчання, регулюється *Положенням про порядок визнання результатів навчання, отриманих у неформальній освіті*: <https://tinyurl.com/y8gbt4xs>.

***ВИРІШЕННЯ КОНФЛІКТІВ.*** Порядок і процедури врегулювання конфліктів, пов’язаних із корупційними діями, зіткненням інтересів, різними формами дискримінації, сексуальними домаганнями, міжособистісними стосунками та іншими ситуаціями, що можуть виникнути під час навчання, регламентуються *Положенням про порядок і процедури вирішення конфліктних ситуацій у ЗНУ*: <https://tinyurl.com/ycyfws9v>. Конфліктні ситуації, що виникають у сфері стипендіального забезпечення здобувачів вищої освіти, вирішуються стипендіальними комісіями факультетів, коледжів та університету в межах їх повноважень, відповідно до: *Положення про порядок призначення і виплати академічних стипендій у ЗНУ*: <https://tinyurl.com/yd6bq6p9>; *Положення про призначення та виплату соціальних стипендій у ЗНУ*: <https://tinyurl.com/y9r5dpwh>.

***ЗАПОБІГАННЯ КОРУПЦІЇ.*** Уповноважена особа з питань запобігання та виявлення корупції (Воронков В. В., 1 корп., 29 каб., тел. +38 (061) 289-14-18).

***ПСИХОЛОГІЧНА ДОПОМОГА.*** Телефон довіри практичного психолога (061)228-15-84 (щоденно з 9 до 21).

***РІВНІ МОЖЛИВОСТІ ТА ІНКЛЮЗИВНЕ ОСВІТНЄ СЕРЕДОВИЩЕ.*** Центральні входи усіх навчальних корпусів ЗНУ обладнані пандусами для забезпечення доступу осіб з інвалідністю та інших маломобільних груп населення. Допомога для здійснення входу у разі потреби надається черговими охоронцями навчальних корпусів. Якщо вам потрібна спеціалізована допомога, будь-ласка, зателефонуйте (061) 228-75-11 (начальник охорони). Порядок супроводу (надання допомоги) осіб з інвалідністю та інших маломобільних груп населення у ЗНУ: <https://tinyurl.com/ydhcsagx>.

***РЕСУРСИ ДЛЯ НАВЧАННЯ. Наукова бібліотека***: [http://library.znu.edu.ua](http://library.znu.edu.ua/). Графік роботи абонементів: понеділок – п`ятниця з 08.00 до 17.00; субота з 09.00 до 15.00.

***ЕЛЕКТРОННЕ ЗАБЕЗПЕЧЕННЯ НАВЧАННЯ (MOODLE): HTTPS://MOODLE.ZNU.EDU.UA***

Якщо забули пароль/логін, направте листа з темою «Забув пароль/логін» за адресами:

· для студентів ЗНУ - moodle.znu@gmail.com, Савченко Тетяна Володимирівна

· для студентів Інженерного інституту ЗНУ - alexvask54@gmail.com, Василенко Олексій Володимирович

У листі вкажіть: прізвище, ім'я, по-батькові українською мовою; шифр групи; електронну адресу.

Якщо ви вказували електронну адресу в профілі системи Moodle ЗНУ, то використовуйте посилання для відновлення паролю https://moodle.znu.edu.ua/mod/page/view.php?id=133015.

***Центр інтенсивного вивчення іноземних мов***: http://sites.znu.edu.ua/child-advance/

***Центр німецької мови, партнер Гете-інституту***: https://www.znu.edu.ua/ukr/edu/ocznu/nim

***Школа Конфуція (вивчення китайської мови)***: http://sites.znu.edu.ua/confucius.

1. **1 змістовий модуль = 15 годин (0,5 кредита EСTS)** [↑](#footnote-ref-2)
2. **Тут зазначається все, що важливо для курсу: наприклад, умови допуску до лабораторій, реактивів і т.д. Викладач сам вирішує, що треба знати студенту для успішного проходження курсу!** [↑](#footnote-ref-3)