

**Міністерство освіти і науки України
Національний технічний університет України «КПІ» імені Ігоря Сікорського
Кафедра обчислювальної техніки ФІОТ**

**ЗВІТ
з лабораторної роботи №4
з навчальної дисципліни «Вступ до технології Data Science»**

Тема:

**РЕАЛІЗАЦІЯ ПРОЦЕСІВ ІНТЕЛЕКТУАЛЬНОГО АНАЛІЗУ ДАНИХ: МІНІ
ПРОЕКТИ В ГАЛУЗЯХ OLAP, Data Mining, Text Mining, Voice Recognition**

Виконав:

Студент 3 курсу кафедри ІІІ ФІОТ,
Навчальної групи ІІІ-11
Лошак В.І.

Перевірив:

Професор кафедри ОТ ФІОТ
Писарчук О.О.

Київ 2023

I. Мета роботи:

Виявити дослідити та узагальнити особливості інтелектуального аналізу даних та технологій OLAP, Data Mining, Text Mining, Voice Recognition

II. Завдання:

1. Лабораторія провідної IT-компанії реалізує проект із розробки ERP системи для автоматизації процесів управління у сфері завдань електронної комерції. Вам, як Data Science Engineer поставлене наступне завдання. Розробити програмний скрипт мовою Python що реалізує обчислювальний алгоритм ERP-системи із впровадженням технології інтелектуального аналізу даних відповідно до технічних умов:

II рівень складності 9 балів, викладених у табл.2.

Розробити програмний скрипт, що реалізує виявлення факту появи обраного товару в Інтернет магазині.

III. Результати виконання лабораторної роботи.

Для виконання лабораторної були використані бібліотеки python такі як: selenium—збереження вмісту сторінки html. Також BeautifulSoup для обробки сторінки та пошуку даних на ній.

Для парсингу обрано сайт rozetka, а точніше його сторінку за посиланням: https://rozetka.com.ua/ua/mobile-phones/c80003/producer=apple/#search_text=iphone що містить дані про останні моделі айфона. Мета лабораторної — зреагувати на вихід нового iphone і зберегти дані про нього в файл.

Оскільки rozetka використовує механізм блокування парсингу, використано headless мод доступу до браузеру бібліотеки selenium щоб отримати дані.

```
1 chrome_options = Options()
2 chrome_options.add_argument("--headless")
3 driver = webdriver.Chrome()
4
5 url = "https://rozetka.com.ua/ua/mobile-phones/c80003/producer=apple/#search_text=iphone"
6 driver.get(url)
7 page_content = driver.page_source
8
9 driver.quit()
10
```

Рис. 1 — завантаження даних в програму для подальшої обробки.

```
title="Мобільний телефон Apple iPhone 15 128GB Pink (MTP13RX/A)" href="https://rozetka.com.ua/ua/apple-iphone-15-128gb-pink/p395460636/"><span class="goods-tile__title">
```

Рис. 2 — Дані з назвою

```
<p class="ng-star-inserted"><span
class="goods-tile__price-
value">25&nbsp;499<span
class="goods-tile__pr
ice-currency
currency">₴</span></
span>
</p><!--><app-buy-button
```

Рис. 3 — Дані з ціною

```
<div
class="goods-tile__availability goods-tile__availability--available ng-star-inserted">
Готовий до відправлення <svg
aria-hidden="true">
```

Рис. 4 — Дані про наявність

Для видобутку необхідних даних зі коду сторінки використано наступний скрипт.

```

1 def extract_product_details(soup):
2     product_details = []
3     product_containers = soup.find_all('div', class_='goods-tile__inner')
4
5     for product in product_containers:
6
7         name = product.find('span', class_='goods-tile__title').get_text(strip=True)
8         price = product.find('span', class_='goods-tile__price-value').get_text(strip=True)
9         availability = product.find('div', class_='goods-tile__availability').get_text(strip=True)
10
11         product_details.append({'name': name, 'price': price, 'availability': availability})
12
13     return product_details
14

```

Рис. 5 – Вивантаження необхідних даних про товари внутрішньої структури даних

```

1 # run this cell to create a timestamped csv file with the results
2 now = datetime.datetime.now()
3 soup = BeautifulSoup(page_content, 'html.parser')
4 product_details = pd.DataFrame(extract_product_details(soup))
5 product_details.to_csv(f'rozetka_{now.date()}.csv', index=False)
✓ 1.0s

```

Рис. 6 — код використаний для збереження необхідних даних в файл з часовою відміткою для подальшого аналізу.

Після збереження вивантажимо файл в програму щоб пересвідчитись в його цілісності і достовірності.

✓ 0.0s

	name	price	availability
0	Мобільний телефон Apple iPhone 15 128GB Pink (...)	38 499₴	Готовий до відправлення
1	Мобільний телефон Apple iPhone 15 Pro Max 256G...	63 499₴	Закінчується
2	Мобільний телефон Apple iPhone 12 128GB Purple...	25 499₴	Готовий до відправлення
3	Мобільний телефон Apple iPhone 15 Pro 256GB Bl...	58 499₴	Готовий до відправлення
4	Мобільний телефон Apple iPhone 15 Plus 128GB P...	43 499₴	Готовий до відправлення
5	Мобільний телефон Apple iPhone 13 128GB Midnig...	31 999₴	Готовий до відправлення
6	Мобільний телефон Apple iPhone 14 128GB Purple...	36 999₴	Готовий до відправлення
7	Мобільний телефон Apple iPhone 14 Pro Max 256G...	57 499₴	Є в наявності
8	Мобільний телефон Apple iPhone 11 128GB Black ...	23 999₴	Готовий до відправлення
9	Мобільний телефон Apple iPhone 14 Plus 128GB M...	37 999₴	Готовий до відправлення
10	Мобільний телефон Apple iPhone 15 Pro 128GB Bl...	53 499₴	Готовий до відправлення
11	Мобільний телефон Apple iPhone 14 Pro Max 128G...	52 499₴	Готовий до відправлення
12	Мобільний телефон Apple iPhone 14 128GB Yellow...	36 999₴	Готовий до відправлення
13	Мобільний телефон Apple iPhone 14 128GB PRODUC...	36 999₴	Готовий до відправлення
14	Мобільний телефон Apple iPhone 14 128GB Starli...	36 999₴	Готовий до відправлення
15	Мобільний телефон Apple iPhone 14 128GB Blue (...)	36 999₴	Готовий до відправлення
16	Мобільний телефон Apple iPhone 14 Pro 256GB Go...	52 499₴	Готовий до відправлення
17	Мобільний телефон Apple iPhone 15 256GB Pink (...)	43 499₴	Готовий до відправлення
18	Мобільний телефон Apple iPhone 13 256GB Green ...	36 999₴	Готовий до відправлення
19	Мобільний телефон Apple iPhone 11 64GB White (...)	20 999₴	Готовий до відправлення

Рис.7 — Дані про результати пошуку iphone 14.12.2023

```

name,price,availability
Мобільний телефон Apple iPhone 15 128GB Pink (MTP13RX/A),38[4992,Готовий до відправлення
Мобільний телефон Apple iPhone 15 Pro Max 256GB Blue Titanium (MU7A3RX/A),63[4992,Закінчується
Мобільний телефон Apple iPhone 12 128GB Purple (MJNP3FS/A),25[4992,Готовий до відправлення
Мобільний телефон Apple iPhone 15 Pro 256GB Blue Titanium (MTV63RX/A),58[4992,Готовий до відправлення
Мобільний телефон Apple iPhone 15 Plus 128GB Pink (MU103RX/A),43[4992,Готовий до відправлення
Мобільний телефон Apple iPhone 13 128GB Midnight (MLPF3HU/A),31[9992,Готовий до відправлення
Мобільний телефон Apple iPhone 14 128GB Purple (MPV03RX/A),36[9992,Готовий до відправлення
Мобільний телефон Apple iPhone 14 Pro Max 256GB Deep Purple (MQ9X3RX/A),57[4992,Є в наявності
Мобільний телефон Apple iPhone 11 128GB Black (MMDH3),23[9992,Готовий до відправлення
Мобільний телефон Apple iPhone 14 Plus 128GB Midnight (MQ4X3RX/A),37[9992,Готовий до відправлення
Мобільний телефон Apple iPhone 15 Pro 128GB Blue Titanium (MTV03RX/A),53[4992,Готовий до відправлення
Мобільний телефон Apple iPhone 14 Pro Max 128GB Gold (MQ9R3RX/A),52[4992,Готовий до відправлення

```

Рис.7 — Дані про результати пошуку iphone 14.12.2023(вигляд всередині csv файлу)

IV. Висновки.

У ході цієї роботи, я практикував парсинг сайтів для отримання потрібних даних, форматування даних та їх підготовку до статистичного навчання та оцінки. Варто зазначити, що функціонал такої програми підлягає вдосконаленню в подальшому, існує багато опцій. Наприклад, використати email розсилку яка буде сповіщати користувача, що товар з'явився у продажу. Можлива розробка систем e-commerce на основі даного функціоналу.

Виконав: студент ФІОТ Лошак В.І. ПІ-11