

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
Сумський державний університет
Навчально-науковий інститут бізнесу, економіки та
менеджменту

**Проблеми та перспективи розвитку
фінансово-кредитної системи
України**
**Problems and Prospects for the Development of the Financial
and Credit System**

Матеріали
Міжнародної науково-практичної конференції
(Україна, Суми, 4-5 грудня 2023 р.)

Суми
Сумський державний університет
2023

ВИКОРИСТАННЯ МАШИННОГО НАВЧАННЯ ДЛЯ СТВОРЕННЯ ТЕКСТОВОЇ ІНФОРМАЦІЇ НА ОСНОВІ АНАЛІЗУ ФІНАНСОВИХ ЗВІТІВ КОМПАНІЙ

Півень Артем Вадимович
студент кафедри економічна кібернетика,
Сумський державний університет, м. Суми
Кушнерьов Олександр Сергійови,
асистент кафедри економічна кібернетика,
Сумський державний університет, м. Суми
Койбічук Віталія Василівна
доцентка, кандидатка економічних наук
кафедри економічна кібернетика,
Сумський державний університет, м. Суми

Генерація тексту за допомогою нейронних мереж стала однією з найбільш вражаючих досягнень у галузі штучного інтелекту та обробки природної мови (NLP). Використання глибинних нейронних мереж дозволяє створювати складні та змістовні текстові послідовності, що знаходять застосування у різних областях, від автоматичного створення контенту до чат-ботів.

Машинне навчання (ML), як підгалузь штучного інтелекту (AI) [1], спрямоване на створення алгоритмів та статистичних моделей, що дозволяють комп'ютерним системам автоматично покращувати свою ефективність шляхом набуття досвіду та використання даних.

Таким чином машинне навчання відкриває широкі перспективи для генерації текстових даних на підставі аналізу фінансових звітів компаній, активно впливаючи на процеси автоматизації та підвищення ефективності та якості аналізу. Декілька ключових аспектів використання машинного навчання в цій сфері можна визначити так:

- автоматизація первинного аналізу фінансових даних за допомогою алгоритмів машинного навчання;
- покращення аналізу фінансових звітів за допомогою глибинного навчання;
- використання моделей машинного навчання та обробки природної мови для генерації звітів і узагальнень;
- прогнозування фінансових тенденцій та розробка прогнозів за допомогою алгоритмів машинного навчання;
- індивідуалізований підхід до аналізу на основі методів машинного навчання;

- виявлення аномалій та ризиків за допомогою моделей машинного навчання;
- вдосконалення стратегій інвестування з використанням алгоритмів машинного навчання.

В процесі розробки модуля для автоматичного аналізу тексту було використано ряд ключових бібліотек. Однак, перш ніж перейти до основної функції модуля, слід розглянути допоміжну функцію `main`, яка відіграє важливу роль в підготовці та передпрацюванні текстових даних (рис. 1).

```
def main():
    st.title("Додаток для формування короткого фінансового звіту компанії PDF-тексту")
    uploaded_file = st.file_uploader("Завантажити PDF-файл", type="pdf")
    if uploaded_file:
        st.success("Файл успішно завантажено!")
        with st.spinner("Аналіз та формування короткого фінансового звіту..."):
            filename = uploaded_file.name
            base_name = os.path.basename(filename)
            name_without_extension = os.path.splitext(base_name)[0]
            pdfFileObject = open(filename, 'rb')
            pdfReader = PdfReader(pdfFileObject)
            text = []
            for i in range(0, len(pdfReader.pages)):
                pageObj = pdfReader.pages[i].extract_text()
                pageObj = pageObj.replace('\n', '')
                pageObj = pageObj.replace('\xa0', '')
                text.append(pageObj)
            new_text = join_elements(text, 3)
            new_text_len = len(new_text)
            summary = summarize_text(new_text)
            st.subheader("Результат")
            st.write(summary)
            with open(f'./{name_without_extension}_summary_ui.txt', 'w') as out:
                out.write(summary)
```

Рисунок 1 — Лістинг функції `main`

Функція `summarize_text` (рис. 2) призначена для здійснення автоматичного аналізу тексту. Вхідним параметром функції є `text`, який містить список абзаців тексту. Створюємо змінну `prompt`, яка містить текстову інструкцію для моделі, що пояснює її завдання. Це завдання полягає у аналізі фінансового звіту, тобто тексту. Запитуємо модель, використовуючи функцію `get_completion`, передаючи їй `prompt`. Отримуємо відповідь моделі, яка містить підсумок тексту. Додаємо підсумок до змінної `summary_chunks`.

На фінальному етапі роботи модуля, створюємо змінну `main_prompt`, яка містить загальну інструкцію для моделі щодо підсумовування тексту. Знову запитуємо модель, використовуючи функцію `get_completion`, передаючи їй `main_prompt`. Отримуємо відповідь моделі, яка містить результат підсумків тексту зі всіх абзаців.

```
def summarize_text(text):
    summary_chunks = ''
    for i in range(len(text)):
        prompt = f'''
        Your task is to act as a Text Summarizer.
        I'll give you text from pages of a book from beginning to end.
        And your job is to summarize text from these pages.
        Do not be conversational. I need a plain answer.
        Text is shared below, delimited with triple backticks:
        """{text[i]}"""
        '''
        try:
            response = get_completion(prompt)
        except:
            response = get_completion(prompt)
            summary_chunks = summary_chunks + ' ' + response + '\n\n'
            time.sleep(5)
    main_prompt = f'''
    Your task is to act as a text summarizer.
    I will provide you with text from pages of a book from beginning to end,
    and your job is to summarize and provide an assessment of their financial position the text from these pages.
    Do not be conversational. I need a plain answer.
    Begin with "Короткий фінансовий звіт компанії {company_name}
    (Replace with the name of the company from the text is shared below, delimited with triple backticks):
    The text is shared below, delimited with triple backticks:
    """{summary_chunks}"""
    The result will be translated into the Ukrainian language.
    '''
    try:
        summary = get_completion(main_prompt)
    except:
        summary = get_completion(main_prompt)
    st.success('Програма успішно завершила свою роботу!')
    return summary
```

Рисунок 2 — Лістинг функції summarize_text

Виконана програмна реалізація дозволяє модулю виконувати аналіз будь якого фінансового звіту за будь яким шаблоном та формуванням звітності. Результат аналізу також записується в текстовий документ для зручності. Враховуючи написання модулю на мові Python його з легкістю можливо інтегрувати у вже існуючу систему, що робить його адаптивним та гнучким у використанні.

У розроблений модуль інтегрований інтерфейс (рис. 3), який дозволяє відобразити реакцію коду на дії користувача та дає користувачу необхідну інформацію про процес аналізу. Для перевірки роботи модулю та його ефективності було обрано фінансові звіти компанії Apple Inc. CONDENSED CONSOLIDATED STATEMENTS OF OPERATIONS (Unaudited) [2].

Коли модуль запущено користувач може обрати звіт та завантажити для аналізу (рис. 3) модуль надасть йому повідомлення про це.

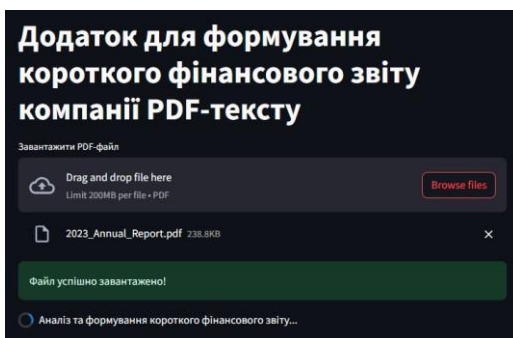


Рисунок 3 — Інтерфейс модулю та повідомлення

Проаналізувавши фінансові звіти компанії Apple було отримано хороший результат у вигляді тексту який структуровано та чітко описує ключові моменти фінансового звіту. Не дивлячись на різноманіття отриманої інформації її тип та об'єм модуль виконує свої функції швидко та ефективно.

Результати аналізу та демонстрація сторінок звіту компаній наведені на рисунках 4 та 5.

Apple Inc.
CONDENSED CONSOLIDATED STATEMENTS OF OPERATIONS (Unaudited)
(In millions, except number of shares which are reflected in thousands and per share amounts)

	Three Months Ended		Six Months Ended	
	April 1, 2023	March 26, 2022	April 1, 2023	March 26, 2022
Net sales:				
Products	\$ 73,629	\$ 77,457	\$ 170,317	\$ 181,886
Services	20,907	19,821	41,673	39,337
Total net sales ⁽¹⁾	94,836	97,278	211,990	221,223
Cost of sales:				
Products	46,795	49,290	107,560	113,599
Services	5,065	5,429	12,122	10,822
Total cost of sales	52,860	54,719	119,682	124,421
Gross margin	41,976	42,559	92,308	96,802
Operating expenses:				
Research and development	7,457	6,387	15,166	12,693
Selling, general and administrative	6,201	6,193	12,808	12,642

Рисунок 4 — Фрагмент звіту компанії Apple

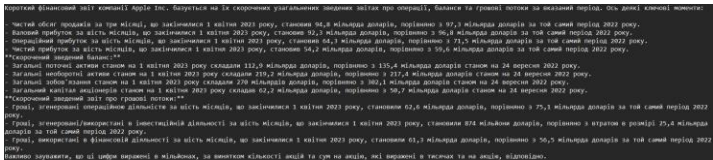


Рисунок 5 — Результат роботи розробленого модуля

Отже, застосування машинного навчання в аналізі фінансових звітів компаній відкриває нові горизонти для більш глибокого та ефективного використання даних, забезпечуючи компаніям конкурентні переваги та підтримуючи прийняття обґрунтованих рішень.

Список використаних джерел

1. Штучний інтелект, машинне навчання - різниця, теорія і практика. (б. д.). Evergreen - web розробка і діджиталізація бізнесу за допомогою AI продуктів. <https://evergreens.com.ua/ua/articles/machine-learning-overview.html>
2. Apple Inc. CONDENSED CONSOLIDATED STATEMENTS URL: apple.com/newsroom/pdfs/FY23_Q2_Consolidated_Financial_Statements.pdf