

Висновки. Залучення в сферу інтернет-маркетингу штучного інтелекту представляє собою широке поле для проведення досліджень та розробок. Застосування штучного інтелекту в цифровому маркетингу має потенціал суттєво підвищити ефективність кампаній у взаємодії з клієнтами, яка досягається за допомогою створення високоякісних дописів, яскравих зображень, анімацій та відеоматеріалів, спрямованих на привертання уваги користувачів до конкретного товару та підтримання їх на різних етапах процесу, що в свою чергу створює високий рівень привабливості для різноманітної аудиторії, сприяючи залученню нових клієнтів та підвищенню популярності пропонованого товару.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Blazeska D. The Impact Of Digital Trends On Marketing. / D. Blazeska // UTMS Journal of Economics. – 2020. – Vol. 11.– P. 48-58. Режим доступу: <https://www.proquest.com/docview/2444992500?sourcetype=Scholarly%20Journals> .
2. Проскурніна Н.В. Штучний інтелект у маркетинговій діяльності. Зовнішня торгівля: економіка, фінанси, право. 2020. № 4. С. 129-140. Режим доступу: http://nbuv.gov.ua/UJRN/uazt_2020_4_11 .
3. Чубукова О. Ю. Складові інноваційної економіки – освіта, технологічні уклади, когнітивні технології. Науковий вісник Полісся. 2016. №3 (7). С. 130-133. Режим доступу: http://nbuv.gov.ua/UJRN/nvp_2016_3_20 .
4. Івлєва А. О. Дослідження технологій створення дизайн-концепцій UI з використанням нейромереж : пояснювальна записка до атестаційної роботи здобувача вищої освіти на другому (магістерському) рівні, спеціальність 186 Видавництво та поліграфія / А. О. Івлєва; М-во освіти і науки України, Харків. нац. ун-т радіоелектроніки. – Харків, 2022. – 63 с. Режим доступу: <https://openarchive.nure.ua/handle/document/21334> .

УДК: 004.8

ДОСЛІДЖЕННЯ ЕФЕКТИВНОСТІ МЕТОДУ ВИЯВЛЕННЯ ЦІЛЮВИХ ОБ'ЄКТІВ ПРЕДМЕТНОЇ ОБЛАСТІ

Мазурець О.В., Віт Р.В.
(exe.chong@gmail.com, vit.roman.vit@gmail.com)
Хмельницький національний університет

Досліджено ефективність методу виявлення цільових об'єктів предметної області, який відрізняється від аналогів урахуванням ключових слів та іменникових сутностей предметної області, що дало змогу підвищити точність виявлення цільових об'єктів предметної області внаслідок врахування іменникових сутностей.

Методи виявлення цільових об'єктів у предметній області є критично важливими для ефективного аналізу та обробки великих обсягів інформації [1]. В умовах зростаючої складності даних, які охоплюють різноманітні предметні області, необхідність розробки та вдосконалення методів автоматизованого виявлення цільових об'єктів стає все більш актуальною [2, 3]. Це особливо важливо в таких сферах, як штучний інтелект, а саме системи обробки природної мови та інформаційний пошук. Відсутність надійних та ефективних методів виявлення цільових об'єктів може призвести до втрати важливої інформації, зниження точності прийняття рішень та збільшення витрат на аналіз даних. Враховуючи швидкий розвиток технологій та постійне зростання обсягів інформації, дослідження методів виявлення цільових об'єктів набуває особливої ваги [4].

Виявлення цільових об'єктів у заданій предметній області передбачає застосування спеціальних алгоритмів та методів, спрямованих на ідентифікацію та класифікацію елементів, які мають ключове значення для аналізу конкретної задачі [5, 6]. У роботі цільові об'єкти будуть

шукатись у текстових даних, а під терміном «цільові об’єкти» буде матись на увазі сукупність множини ключових слів та множини NER з групуванням шляхом лематизації.

Виявлення цільових об’єктів у системах NLP, зокрема розпізнавання іменованих сутностей, відіграє важливу роль у багатьох завданнях аналізу тексту та обробки інформації. Основна мета NER полягає в ідентифікації і класифікації значущих елементів тексту, таких як імена людей, назви організацій, географічні назви, дати та інші сутності, які мають специфічне значення для конкретного контексту. Це завдання є ключовим для ряду практичних задач, таких як інформаційний пошук, машинний переклад, обробка юридичних документів та аналіз даних у соціальних медіа [7].

Метод виявлення цільових об’єктів предметної області, який призначений для автоматизації процесу ідентифікації ключових елементів у великих масивах текстових даних, спрямований на підвищення точності та ефективності аналізу текстової інформації. Цей метод використовує алгоритми машинного навчання для адаптивного розпізнавання об’єктів, враховуючи специфіку предметної області, що дозволяє значно скоротити час обробки даних і знизити ризик упущення важливої інформації.

Вхідними даними методу є досліджуваній текст та попередньо оброблений збалансований корпус текстів досліджуваної предметної області. Першим етапом є підготовка досліджуваного тексту для аналізу, який включає в себе токенізацію, лематизацію та видалення стоп-слів.

Наступним етапом є пошук ключових слів різними методами, такими як TF-IDF, TF, YAKE! та методом дисперсної оцінки. Кожним перерахованим методом відбувається формування множини ключових слів.

На третьому етапі здійснюється виявлення цільових об’єктів, яке включає в себе декілька кроків. Цільові об’єкти є об’єднаною множиною ключових слів знайденими різними методами без повторів та множиною NER що згруповані шляхом лематизації.

Таким чином описано метод виявлення цільових об’єктів предметної області, що призначений для автоматизації процесу ідентифікації ключових елементів у великих масивах текстових даних.

Для валідації запропонованого методу для пошуку цільових об’єктів предметної області було розроблено програмний застосунок мовою C# для перетворення текстового контенту файлів із тестової вибірки у множину цільових об’єктів предметної області.

Оскільки українська мова повсякденного спілкування значно відрізняється від літературної через велику кількість діалектів, слів-запозичень та слів-покручів, наявні частотні словники не здатні охопити всю множину української мови. Для створення вектора значущих слів українською мовою було вирішено об’єднати кілька частотних словників, з відсіканням стоп-слів. Після об’єднання та фільтрації довжина вектора значущих слів склала 1500 елементів.

Таблиця 1 – Евклідові відстані між текстами одного спрямування

	Текст 1	Текст 2	Текст 3	Текст 4	Текст 5	Текст 6	Текст 7	Текст 8	Текст 9	Текст 10
Текст 1	0	10.3	11.2	9.75	14.7	25.7	23.4	28.6	29.6	24.7
Текст 2	10.3	0	15.7	17.1	16.4	30.21	24.5	26.3	23.34	26.5
Текст 3	11.2	15.7	0	9.4	8.89	27.6	24.9	23.8	25.7	27.1
Текст 4	9.75	17.1	9.4	0	5.47	32.4	30.7	26.1	27.6	23.6
Текст 5	14.7	16.4	8.89	5.47	0	19.4	23.45	26.12	28.4	24.7
Текст 6	25.7	30.21	27.6	32.4	19.4	0	9.78	6.99	9.1	14.3
Текст 7	23.4	24.5	24.9	30.7	23.45	9.78	0	11.9	12.45	7.98
Текст 8	28.6	26.3	23.8	26.1	26.12	6.99	11.9	0	6.33	8.91
Текст 9	29.6	23.34	25.7	27.6	28.4	9.1	12.45	6.33	0	13.5
Текст 10	24.7	26.5	27.1	23.6	24.7	14.3	7.98	8.91	13.5	0

Для дослідження ефективності запропонованого підходу було створено окреме консольне програмне забезпечення мовою Python, яке передбачає використання отриманого списку цільових об’єктів для досліджуваних текстів, та словників для окреслених тем «Карпати буд каркас» та «Блог садівника». Відповідно, знайдені цільові об’єкти були переведені у векторне представлення розміром 1500 (як розмір словника) методом One-Hot Encoding. Надалі було перевірено Евклідові

відстань між текстами одного спрямування (5 текстів категорії «Карпати буд каркас» та 5 текстів «Блог садівника»), а також були обраховані Евклідові відстані між векторами протилежних категорій. Дані експерименту наведено в таблиці 1.

Результати отримані з таблиці 1 проілюстровані на рисунку 1. Матриця відстаней таблиці 1 та рисунку 1 демонструють чітке розділення текстів на дві основні групи з різним змістом. Перша група текстів (1–5), що належать категорії «Карпати буд каркас» має тісніші зв'язки між собою, аналогічно як друга група (6–10) також має менші внутрішні відстані (категорія «Блог садівника»), але водночас має великі відстані до текстів з першої групи, що свідчить про те, що ці групи належать до різних тематик. Тексти всередині кожної групи мають невеликі відстані, що свідчить про їхню тематичну схожість.

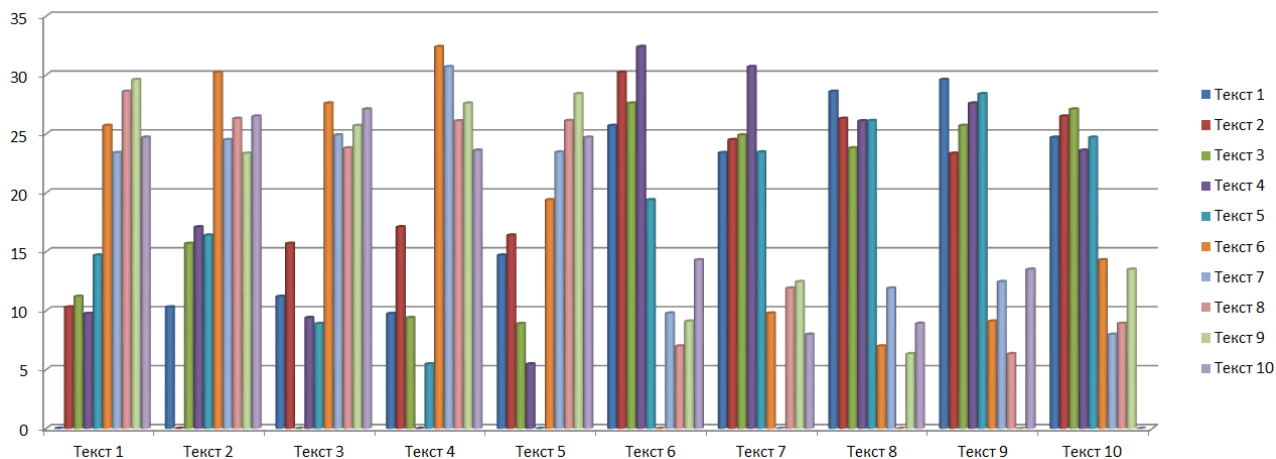


Рисунок 1 – Евклідові відстані між тестовими текстами двох категорій

Отже, проведене дослідження ефективності методу виявлення цільових об'єктів предметної області показало спроможність запропонованого методу ефективно виявляти цільові об'єкти. Вдалось підвищити точність виявлення цільових об'єктів предметної області внаслідок врахування не лише ключових слів, а й урахуванням іменникових сутностей.

Список використаних джерел

[1] Мазурець О., Віт Р. Інтелектуальний метод виявлення цільових об'єктів предметної області для класифікації текстової інформації. Матеріали XII Міжнародної науково-практичної конференції «Інформаційні управляючі системи та технології ІУСТ-ОДЕСА-2024». 23-25.09.2024. Одеса. 2024. С.205-208.

[2] Krak I., Zalutcka O., Molchanova M., Mazurets O., Manziuk E., Barmak O. Method for Neural Network Detecting Propaganda Techniques by Markers With Visual Analytic. CEUR Workshop Proceedings, 2024, vol. 3790, pp. 158-170.

[3] Молчанова М.О., Мазурець О.В., Собко О.В., Віт Р.В., Назаров В.В. Алгоритм виявлення аб'юзивного вмісту в україномовному аудіоконтенті для імплементації в об'єктно-орієнтовану інформаційну систему. Науковий журнал «Вісник Хмельницького національного університету» серія: Технічні науки. Хмельницький, 2024. №1 (331). С. 101-106.

[4] Залуцька О.О., Молчанова М.О., Віт Р.В., Мазурець О.В. Конфігурування нейронної мережі для класифікації емоційної тональності текстової інформації за показниками семантичної зв'язності. Збірник наукових праць за матеріалами XV Всеукраїнської науково-практичної конференції «Актуальні проблеми комп'ютерних наук АПКН-2023». Хмельницький, 2023. с. 102-107.

[5] Krak I., Zalutcka O., Molchanova M., Mazurets O., Bahrii R., Sobko O., Barmak O. Abusive Speech Detection Method for Ukrainian Language Used Recurrent Neural Network. CEUR Workshop Proceedings, 2024, vol. 3688, pp. 16-28.

[6] Krak I., Zalutcka O., Molchanova M., Mazurets O., Manziuk E., Barmak O. Method for Neural Network Detecting Propaganda Techniques by Markers With Visual Analytic. CEUR Workshop Proceedings, 2024, vol. 3790, pp. 158-170.

[7] Slobodzian V., Kovalchuk O., Molchanova M., Sobko O., Mazurets O., Barmak O., Krak I. Text Data Vectorization Model of Ukrainian-Language Internet Communication Content. CEUR Workshop Proceedings, 2022, vol. 3171, pp. 561–571.

УДК: 004.8

ДОСЛІДЖЕННЯ ЕФЕКТИВНОСТІ МЕТОДУ АНАЛІЗУ ПРОЯВУ ПОСТТРАВМАТИЧНОГО СТРЕСОВОГО РОЗЛАДУ В КОРИСТУВАЦЬКОМУ КОНТЕНТІ

Мазурець О.В., Овчарук О.М.
(exe.chong@gmail.com, off4aruk@gmail.com)
Хмельницький національний університет

Досліджено ефективність методу аналізу прояву посттравматичного стресового розладу в користувацькому контенті, який дозволяє сфокусувати увагу на ПТСР та зменшити його плутанину із іншими розладами психіки за рахунок навчання на даних з іншими розладами у протилежній категорії. Було отримано такі результати за метриками: Accuracy 0.934, Precision 0.948, F1 0.841, AUC 0.872. Досліджуваний метод дозволяє не лише визначити наявність посттравматичного стресового розладу у текстовому контенті, а також інтерпретувати отримані результати за допомогою моделей «Transformers Interpret» та «BertViz».

Посттравматичний стресовий розлад є виснажливою психічною хворобою, з якою страждають приблизно 10% населення. З розвитком цифрових технологій і соціальних медіа з'явилася нова можливість для вивчення проявів ПТСР через аналіз користувацького контенту. Ця тема набуває особливої актуальності в контексті збройних конфліктів, природних катастроф і інших травматичних подій, які часто залишають глибокий відбиток на психіці людей.

Аналіз користувацького контенту, такого як повідомлення в соціальних мережах, блоги, форуми і коментарі, дозволяє отримати унікальні інсайти щодо емоційного і психічного стану постраждалих [1, 2]. Застосування методів машинного навчання та обробки природної мови для виявлення ознак ПТСР у текстах користувачів є сучасним і перспективним напрямом наукових досліджень, адже можуть не тільки ідентифікувати окремі випадки ПТСР, але й допомогти в розробці превентивних заходів і програм підтримки для осіб, які перебувають у зоні ризику.

Підхід до аналізу прояву посттравматичного стресового розладу в користувацькому контенті показано на рисунку 1. Підхід демонструє перетворення вхідної інформації у вигляді текстового контенту для аналізу та навченої контекстно-орієнтованої нейромережевої моделі з токенизатором у вихідну інформацію у вигляді відсотку прояву посттравматичного стресового розладу в користувацькому контенті.

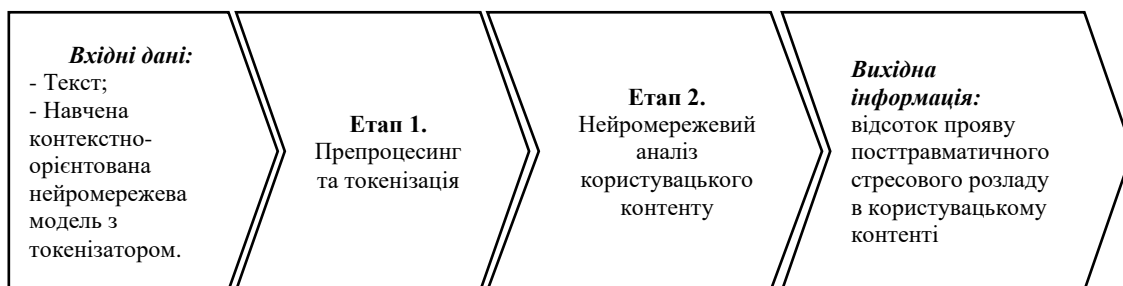


Рисунок 1 – Підхід до аналізу прояву посттравматичного стресового розладу в користувацькому контенті

На першому етапі здійснюється препроцесинг та токенизація текстового контенту. Контент перевіряється на довжину та на непустоту. Однак, розділові знаки, смайли та решта онзнак тексту