

Особливої уваги потребує захист хмарних сервісів і віртуальних інфраструктур, які стають дедалі популярнішими. Забезпечення належної сегментації, шифрування, контролю доступу, а також використання сучасних систем моніторингу дозволяє мінімізувати ризики, пов'язані з використанням цих технологій.

Отже, забезпечення інформаційної безпеки [4] є багатогранним завданням, що вимагає постійного розвитку технологій, адаптації до нових викликів і тісної співпраці між усіма учасниками процесу. Тільки завдяки впровадженню комплексних заходів можна ефективно протистояти кіберзагрозам та забезпечувати стабільність і безпеку інформаційних систем у динамічному цифровому середовищі.

Список використаних джерел

1. Найкращі протоколи VPN та відмінності між типами VPN. URL: <https://nordvpn.com/uk/blog/protokoly/>. (дата звернення 2.04.2025).
2. Роль брандмауера у забезпеченні особистої кібербезпеки – захист від загроз і зламів в мережі. URL: <https://mediacom.com.ua/rol-brandmauera-v-osobistij-kiberbezpeti>. (дата звернення 2.04.2025).
3. Хмарна безпека. URL: <https://nordvpn.com/uk/cybersecurity/cloud-security/>. (дата звернення 2.04.2025).

МЕТОДИ І АЛГОРИТМИ АНАЛІЗУ КОРЕЛЯЦІЙ МІЖ ТЕКСТОВИМИ ДАНИМИ І ПОВЕДІНКОВИМИ ПОКАЗНИКАМИ КОРИСТУВАЧІВ

Ясінський Андрій Михайлович

здобувач другого (магістерського) рівня вищої освіти спеціальності Комп'ютерні науки,
Тернопільський національний педагогічний університет імені Володимира Гнатюка,
yasinskyj_am@fizmat.tnpu.edu.ua

Лень Андрій Володимирович

кандидат історичних наук, асистент кафедри інформатики та методики її навчання,
Тернопільський національний педагогічний університет імені Володимира Гнатюка,
lenandr@tnpu.edu.ua

У сучасному інформаційному середовищі значну роль відіграє аналіз великих обсягів даних, які активно генеруються користувачами у цифровому просторі. Текстові повідомлення, коментарі, пости та інші форми контенту, створені у межах соціальних мереж, інформаційних ресурсів і освітніх платформ, містять велику кількість прихованих закономірностей, які відображають поведінкові характеристики та інтереси користувачів. Тому актуальним є питання вивчення методів, які дозволяють встановити взаємозв'язки між текстовими даними та показниками взаємодії з контентом. Застосування відповідних методів аналізу відкриває можливості для глибшого розуміння цифрової поведінки, прогнозування тенденцій та формування ефективних стратегій залучення.

Одним із головних напрямів у цьому контексті є використання алгоритмічних методів аналізу текстових даних, які дозволяють не лише структурувати інформацію, а й виявити приховані закономірності у поведінці користувачів.

Проведення повноцінного аналізу вимагає комплексного підходу, що охоплює етапи збору, обробки, статистичного аналізу та візуалізації даних. На початковому етапі збираються як текстові дані, так і поведінкові показники користувачів, які можуть включати кількість переглядів, рівень залученості, взаємодії тощо. Подальша обробка даних включає очищення, структурування, нормалізацію та підготовку до аналізу, що забезпечує достовірність і репрезентативність результатів [1, с. 102].

Статистичний аналіз відіграє ключову роль у виявленні взаємозв'язків між різними параметрами. Зокрема, використовуються методи для визначення сили та характеру кореляцій між елементами тексту (ключові слова, семантичні групи тощо) та поведінковими реакціями аудиторії. Даний підхід дозволяє перейти від поверхневого спостереження до кількісного опису тенденцій і закономірностей, що властиві цифровому контенту. З технічного погляду, реалізація такого аналізу вимагає застосування широкого спектра інструментів та ресурсів [3].

Особливої уваги заслуговує питання практичного застосування отриманих результатів. Аналіз кореляцій між контентом і реакціями аудиторії може бути використаний для оптимізації структури та стилістики навчальних матеріалів, підвищення ефективності інформаційної комунікації в онлайн-навчанні, а також у сфері маркетингу, психології поведінки та соціологічних досліджень. Такий підхід відкриває нові можливості в освітній сфері, зокрема – у формуванні персоналізованих освітніх стратегій на основі даних про поведінку та інтереси користувачів [2].

Особливу роль у дослідженні відіграє візуалізація отриманих результатів. Побудова графіків і діаграм сприяє не лише кращому розумінню отриманих висновків, а й дозволяє виявити складні взаємозв'язки, які важко помітити під час табличного аналізу. Візуалізація перетворює абстрактні дані на наочну картину, яка може бути ефективно використана у практичній діяльності – від дослідження аудиторії до розробки цифрових продуктів [5].

Автоматизація процесів збору та обробки даних дозволяє значно підвищити ефективність і швидкість аналізу. Це створює можливості для застосування цих методів у різних галузях, таких як освіта, медіа, маркетинг і соціальні комунікації, забезпечуючи їх масштабованість і універсальність.

Для реалізації комплексного аналізу текстових та поведінкових даних широко застосовуються сучасні інструменти програмування та обробки інформації. Одним із найпопулярніших середовищ є мова програмування Python, що забезпечує широкі можливості, які забезпечують повний цикл роботи з даними – від збору до візуалізації результатів. Для збору інформації з вебресурсів та соціальних платформ широко використовуються бібліотеки Requests, BeautifulSoup і Playwright, що дозволяють автоматизувати процес отримання текстового контенту [4, с. 134].

На етапі обробки та аналізу даних ключову роль відіграють такі інструменти, як Pandas – для структурування й маніпулювання даними, NumPy – для чисельних обчислень, а також re для роботи з регулярними виразами, що дозволяє ефективно очищувати текстові дані. Для текстового аналізу активно

застосовуються бібліотеки spaCy та NLTK, які дають змогу проводити токенізацію, лематизацію, визначення частин мови та інші базові операції. Також важливу роль відіграє Scikit-learn – як інструмент для статистичного аналізу, зокрема для обчислення коефіцієнтів кореляції, кластеризації та зведення даних [4, с. 147].

Для візуалізації результатів найчастіше використовуються Matplotlib, Seaborn та Plotly, що дозволяють створювати інформативні графіки й діаграми, які наочно відображають виявлені зв'язки. Використання цих інструментів у комплексі забезпечує гнучкість і точність дослідження, дозволяючи адаптувати аналітичні моделі під потреби конкретного проєкту чи сфери застосування.

Аналіз кореляцій між текстовим контентом та поведінковими характеристиками користувачів є перспективним напрямом досліджень у контексті цифрової аналітики. Інтеграція методів збору, обробки, статистичного аналізу та візуалізації даних дозволяє створити цілісну модель вивчення взаємодії між інформаційними повідомленнями та реакцією аудиторії. Такий підхід сприяє не лише теоретичному осмисленню цифрової поведінки, але й формує підґрунтя для практичного застосування отриманих результатів у різноманітних галузях. У свою чергу, автоматизація процесів обробки даних дозволяє значно підвищити ефективність аналітичної діяльності та забезпечує можливість гнучкої адаптації до нових інформаційних викликів сучасності.

Список використаних джерел

1. Бахрушин В. Є. Методи аналізу даних : навчальний посібник для студентів. Запоріжжя : КПУ, 2011. 268 с.
2. Кореляційний аналіз даних: Що це таке і як його застосовувати. URL: <https://ua5.org/model/1728-korelyaczijnyj-analiz-danyh-shho-cze-take-i-yak-jogo-zastosovuvaty.html> (дата звернення: 08.02.2025).
3. Основне про аналіз даних: посібник для початківців. URL: <https://www.microsoft.com/uk-ua/microsoft-365/excel/data-analysis> (дата звернення: 04.02.2025).
4. Чичкарьов Є. А., Зінченко О. В., Єльченко С. В. Прикладне програмування на Python. Ч. 1. Основи програмування на Python : навч. посіб. Київ : ДУТ, 2022. 160 с.
5. Що таке аналіз даних? Дослідження, типи та приклад. URL: <https://www.guru99.com/uk/what-is-data-analysis.html> (дата звернення: 02.03.2025).