

КОЕФІЦІЄНТ КОРЕЛЯЦІЇ І НАПРЯМИ ЙОГО ЗАСТОСУВАННЯ

Донецький національний університет імені Василя Стуса, м. Вінниця

Вступ. У сучасних інформаційних технологіях важливо не лише збирати дані, а і вміти виявляти між ними приховані зв'язки. Одним із ключових інструментів для цього є кореляційний аналіз – метод статистичної оцінки сили і напрямку зв'язку між двома або більше змінними. Він широко використовується в машинному навчанні, обробці сигналів, розпізнаванні образів, криптографії та кібербезпеці.

У цій роботі розглянемо деякі застосування коефіцієнта кореляції в різних галузях інформаційних технологій.

Основна частина. Кореляційний аналіз відіграє ключову роль у виявленні статистичних зв'язків між змінними та знаходить широке застосування у різних галузях науки й техніки. Для характеристики кореляційного зв'язку використовується коефіцієнт кореляції. Коефіцієнтом кореляції r_{xy} між двома ознаками X і Y називається відношення їх коваріації до добутку середніх квадратичних відхилень цих величин [1]:

$$r_{xy} = \frac{\text{cov}_{xy}}{\sigma_x \sigma_y},$$

де $\text{cov}_{xy} = \overline{xy} - \bar{x} \cdot \bar{y}$ – кореляційний момент;

$\sigma_x = \sqrt{\overline{x^2} - \bar{x}^2}$ та $\sigma_y = \sqrt{\overline{y^2} - \bar{y}^2}$ – середні квадратичні відхилення;

$$|r_{xy}| \leq 1.$$

У машинному навчанні кореляційний аналіз використовується для визначення, які вхідні ознаки мають сильний лінійний зв'язок із цільовою змінною. Ознаки з високою кореляцією можуть бути більш інформативними для побудови прогностичних моделей. Наприклад, у задачі прогнозування ціни акцій можна оцінити кореляцію між різними економічними індикаторами та ціною акцій, щоб відібрати найбільш релевантні індикатори [2].

Для виявлення залежностей для багатофакторної моделі використовується кореляційна матриця, яка відображає парні коефіцієнти кореляції між усіма факторами і має вигляд:

$$\Delta = \begin{pmatrix} 1 & r_{yx_1} & r_{yx_2} & \dots & r_{yx_m} \\ r_{x_1y} & 1 & r_{x_1x_2} & \dots & r_{x_1x_m} \\ r_{x_2y} & r_{x_2x_1} & 1 & \dots & r_{x_2x_m} \\ \dots & \dots & \dots & \dots & \dots \\ r_{x_my} & r_{x_mx_1} & r_{x_mx_2} & \dots & 1 \end{pmatrix},$$

де r_{x_iy} – коефіцієнт кореляції між x_i , $i = 1, \dots, m$, та y ;

$x_1, \dots, x_m$ – незалежні змінні;
 y – залежна змінна [3].

Висока кореляція між двома або більше факторами у моделі регресії (мультиколінеарність) ускладнює інтерпретацію коефіцієнтів регресійної моделі та призводить до нестабільності моделі. Метод головних компонент (PCA) може бути застосований для зменшення впливу мультиколінеарності шляхом перетворення корельованих змінних на некорельовані головні компоненти [4].

До того ж кореляційний аналіз застосовується для обробки сигналів, зокрема допомагає виявляти приховані сигнали та аналізувати їх характеристики, а автокореляція і кроскореляція дають змогу виявити періодичні компоненти та ступінь подібності між двома сигналами зі зсувом відповідно [5]. В інформаційній безпеці використовується для мережевої активності, щоб виявити аномальні події (наприклад, різкі зміни між обсягом трафіка та кількістю з'єднань можуть свідчити про кібератаки, зокрема сканування мережі або DDoS-атаки), а також дає змогу виявляти підозрілу поведінку на ранньому етапі між подіями, що є критично важливим для захисту інформаційних систем [6]. У криптографічних системах важливо, щоб згенеровані послідовності мали низьку автокореляцію, що зменшує ймовірність їх прогнозування [7]. Інструменти кореляційного аналізу використовуються для розпізнавання образів та комп'ютерного зору, зокрема допомагає оцінювати симетрію об'єктів і виявляти структурні зміни. Зниження кореляції між симетричними ключовими точками обличчя може свідчити про деформації, що має практичне значення в біометричних системах, медичній діагностиці та задачах відстеження руху об'єктів [8].

Висновки. Кореляційний аналіз є потужним інструментом для виявлення, оцінювання та інтерпретації зв'язків між змінними в інформаційних системах. Його універсальність дають змогу застосовувати цей метод у задачах від вибору релевантних ознак у машинному навчанні до виявлення кіберзагроз та прихованої інформації. Завдяки своїм аналітичним можливостям, коефіцієнт кореляції сприяє побудові ефективних моделей, точнішій інтерпретації даних і забезпеченню безпеки систем.

Список використаних джерел

1. Зайцев Є. І. Теорія ймовірностей і математична статистика: навч. посіб. Київ: Алерта, 2017. 440 с.
2. Hastie T., Tibshirani R., Friedman J. The Elements of Statistical Learning: Data Mining, Inference, and Prediction. Springer, 2009.
3. Економетрика в електронних таблицях: навчальний посібник / Н. К. Васильєва, А. О. Мироненко, Н. М. Самарець, Н. О. Чорна. Дніпро: Видавець Біла К. О., 2017. 152 с. URL: <https://dspace.dsau.dp.ua/bitstream/123456789/713/1/Posibnyk.pdf>
4. Jolliffe I. T. Principal Component Analysis. Educational manual. Springer New York, NY. 2002. 488 p. URL: [http://cda.psych.uiuc.edu/statistical_learning_course/Jolliffe%20I.%20Principal%20Component%20Analysis%20\(2ed.,%20Springer,%202002\)\(518s\)_MVsa_.pdf](http://cda.psych.uiuc.edu/statistical_learning_course/Jolliffe%20I.%20Principal%20Component%20Analysis%20(2ed.,%20Springer,%202002)(518s)_MVsa_.pdf)
5. Johnson F., Duric Z., Jajodia S. Information Hiding: Steganography and Watermarking – Attacks and Countermeasures. Kluwer Academic Publishers, 2001. URL: <https://link.springer.com/book/10.1007/978-1-4615-4375-6>
6. Cover T. M., Thomas A. Elements of Information Theory. 2nd ed. Wiley-Interscience, 2006. URL: https://cs-114.org/wp-content/uploads/2015/01/Elements_of_Information_Theory_Elements.pdf

7. Manning D., Raghavan P., Schütze H. Introduction to Information Retrieval. Cambridge University Press, 2009. 581 c. URL: <https://nlp.stanford.edu/IR-book/pdf/irbookonlinereading.pdf>
8. Alan V. Oppenheim and Ronald W. Schafer. Discrete-Time Signal Processing. Prentice Hall, 1999. 997 p. URL: <http://103.203.175.90:81/fdScript/RootOfEBooks/E%20Book%20collection%20-%202023/EEE/DiscreteTimeSignalProcessing2n.pdf>

УДК 004.89

*Гуменчук П. С., здобувач вищої освіти 4 курсу спеціальності 122 Комп'ютерні науки,
Римар П. В., старший викладач кафедри інформаційних технологій*

РОЗРОБКА ВЕБІНТЕРФЕЙСУ СИСТЕМИ ДЛЯ РОБОТИ З ГЕНЕРАТИВНИМ ШТУЧНИМ ІНТЕЛЕКТОМ

Донецький національний університет імені Василя Стуса, м. Вінниця

Вступ. Генеративний штучний інтелект (ГШІ) демонструє значний прогрес у створенні різного типу контенту, зокрема і відеороликів. Це відкриває нові можливості для різних галузей, однак ефективне використання ГШІ часто обмежується складністю наявних інструментів та відсутністю зручних користувацьких інтерфейсів.

Актуальність роботи зумовлена необхідністю розробки доступних та інтегрованих програмних рішень, які б дали змогу широкому колу користувачів, включно з тими, хто не має глибоких технічних знань, ефективно взаємодіяти з моделями ГШІ для генерації відео. Існує потреба в системах, що забезпечують повний цикл роботи (від налаштування параметрів до управління результатами).

Аналіз наукових публікацій показує, що незважаючи на значні досягнення в розробці самих генеративних моделей, як-от GAN [1], та їх наступників StyleGAN [2], а також наявність загальних рекомендацій щодо взаємодії «людина – ШІ» [3], комплексні рішення, що поєднують зручний вебінтерфейс, релевантний API, який задовольняє своїми можливостями весь потенційний функціонал вебзастосунку та засоби управління для систем генерації відео, залишаються недостатньо дослідженими. Наявні інструменти часто або є вузькоспеціалізованими, або не забезпечують цілісного користувацького досвіду, зокрема в частині управління сесіями та історією запитів [4].

Мета роботи – розробити архітектуру та ключові компоненти вебінтерфейсу системи для роботи з генеративним штучним інтелектом, орієнтованої на створення відеоконтенту, що забезпечує інтуїтивну взаємодію, управління користувацькими даними та процесом генерації.

Постановка задачі. Для досягнення мети необхідно виконати таке:

- 1) спроектувати архітектуру клієнтської частини вебінтерфейсу на основі сучасного фреймворка;
- 2) розробити структуру бази даних для зберігання інформації про користувачів, сесії користувачів, запити та згенеровані відео;