

## АНАЛІЗ ЕМОЦІЙ У ТЕКСТІ

Чуб І.М., к.т.н., доц.  
Колокольніков В.С., здобувач РВО бакалавр  
Державний біотехнологічний університет  
м. Харків, Україна

**Анотація:** Аналізується технологія штучного інтелекту для виявлення та розуміння людських емоцій у тексті. Дослідження покликане поширити інформацію пов'язану з проблемами та методами аналізу емоцій в тексті.

**Ключові слова:** розпізнавання емоцій, аналіз настроїв, штучний інтелект, природна мова

Аналіз тональності – це застосування технологій обробки природної мови (NLP), які навчають комп'ютерне програмне забезпечення розуміти текст як і, як і люди.

За останні роки штучний інтелект стає все більш просунутим і затребуваним для швидкого вирішення різних завдань, виконання яких іншими способами вимагало б набагато більше часу, коштів і людей.

Вже зараз штучний інтелект може аналізувати текст, фотографії, відео та аудіозаписи, витягаючи з цих вихідних даних будь-яку інформацію. Також він може перетворювати проаналізовані дані різними способами. Наприклад, зробити детальний опис картинки у текстовому вигляді; згенерувати нову подібну картинку, використовуючи цей опис; перетворити текст на голос та багато іншого.

Зараз створюються різні методи штучного інтелекту для аналізу емоцій у тексті. Це допомагає прогнозувати майбутні події, відгуки людей про продукт або послуги, ідентифікувати групи людей за інтересами та розроблювати машини, що здатні імітувати поведінку людських емоцій. Основні дані для аналізу емоцій вилучаються з сайтів соціальних мереж, блогів, груп, що використовуються людьми для вираження своєї думки.

Система загального опитування (General Survey System) була першою системою, створеною в 1966 році для аналізу емоцій [1]. В 1990 році Джейнс Вібе дає назву терміну «суб'єктивність» для вивчення інформаційного пошуку [2]. У 1997 році Васілеос визначив смислову спрямованість прикметників у тексті [3]. Денік у 2009 році провів дослідження щодо корисності попередніх оцінок полярності з емоційної бази даних (SentiWordNet) [4].

Проблеми автоматичного розпізнавання емоцій в текстах є досить складними, а з іншого боку — актуальними і цікавими. Складності проблема набуває через наступні причини: ситуативність фрази (для одного тексту позитивна, а для іншого негативна); розташування слів (позитивний текст та негативний висновок); наявність іронії тощо.

Зараз ми використовуємо автоматичне розпізнавання емоцій в тексті у багатьох сферах. Це і маркетингові дослідження, і пошукові системи, рекомендаційні системи, оцінка настроїв новин, людино-машинний інтерфейс тощо. Проблема розпізнавання емоцій з'явилась в центрі наукових досліджень

порівняно недавно, десь на початку 2000-х років. Дуже детально оглянули цю проблему Бо Панг і Ліліан Лі [5].

Існують також і інші постановки завдань, наприклад: роздільне розпізнавання емоцій з точки зору письменника та читача; спочатку аналіз суб'єктивності, а потім тональності, з рештою класифікація за емоціями.

Метою роботи є оптимізація коду, що наведений у [6]. Це підвищить швидкість та якість виконання програми, та забезпечить більшу обробку проаналізованих даних. Було досліджено декілька текстів та продемонстровано переваги аналізу емоцій в тексті із застосуванням розробленого методу.

У статті було проведено аналіз роботи бібліотеки Vader, спеціальної програми для аналізу емоційного відтінку тексту. Виявлено, що ця бібліотека має свої недоліки, зокрема, вона не може адекватно аналізувати тексти, що містять багато речень цілком, що може спотворювати результати. Для вирішення цієї проблеми було запропоновано методику, яка базується на введенні понять парсінгу, нормалізації та розрахунку максимальних позитивних та негативних оцінок. За допомогою цієї методики було розроблено програму на мові програмування Python для аналізу складних текстів. Результати аналізу демонстраційного тексту, отримані за допомогою цієї програми, показали більш точний результат у порівнянні з результатом роботи бібліотеки Vader без застосування методики. Таким чином, удосконалений спосіб аналізу емоцій дозволяє проводити більш детальний аналіз текстів та виявляти всі емоційні відтінки, що допомагає приймати більш обґрунтовані рішення щодо тексту в цілому.

#### Список літератури

1. Stone P.J., Dunphy D.C., Smith M.S., Ogilvie D.M. (1966). The general inquirer: A computer approach to content analysis. – Cambridge: The MIT Press, p. 651. DOI:10.2307/1161774
2. Janyce Wiebe M. (1990). Identifying Subjectivity characters in Narrative // Proc.13th International Conference on Computational Linguistics. Helsinki, pp. 401-406. DOI:10.3115/997939.998008
3. Vasileios H., Kathleen R. M. (1997). Predicting the Semantic Orientation of Adjectives // Proc. 8th Conference on European chapter of the Association for Computational Linguistics. Spain, pp 174-181. DOI:10.3115/976909.979640
4. Denecke K. (2009). Are SentiWordNet scores suited for multi-domain sentiment classification? // Proc. 4th International Conference on Digital Information Management. USA, pp. 33-38. DOI: 10.1109/ICDIM.2009.5356764
5. Pang, B. and Lee, L. (2008) Opinion Mining and Sentiment Analysis. Foundations and Trends® in Information Retrieval, 2, 1-135. <https://doi.org/10.1561/1500000011>
6. Обробка природної мови (NLP) у Python з кодом (Частина 1. Аналіз тональності) [Електронний ресурс] // Medium. – 2021. – Режим доступу до ресурсу: <https://oleg-dubetchky.medium.com/обробка-природної-мови-nlp-у-python-частина-1-83d588b3ad71>