

УДК 004.8+004.02:[004.91+004.93]

**В. Кохан, Є. Тиш, канд. техн. наук**

(Тернопільський національний технічний університет імені Івана Пулюя, Україна)

## **МЕТОДИ ОЦІНЮВАННЯ ЕМОЦІЙНОГО НАХИЛУ ТЕКСТІВ ЗАСОБАМИ ШТУЧНОГО ІНТЕЛЕКТУ**

UDC 004.8+004.02:[004.91+004.93]

**V. Kokhan, Ie. Tysh, Ph.D.**

## **METHODS OF EVALUATION OF SENTIMENT ANALYSIS OF TEXTS BY MEANS OF ARTIFICIAL INTELLIGENCE**

Штучний інтелект – популярний на сьогодні напрям проведення багатьох досліджень, який дозволяє робити неймовірні речі з використанням можливостей сучасних комп'ютерних технологій. До досліджень в цій галузі залучають спеціалістів усіх галузей знань у залежності від очікуваних результатів.

Виданнями, що спеціалізуються на штучному інтелекті [1], було визначено ключові завдання штучного інтелекту, такі як: комп'ютерний зір, машинне навчання, обробка природної мови та інші. Кожне таке завдання, або комбінація завдань, дозволило окреслити великий перелік простіших, конкретніших цілей, рішення яких вже почали впливати на наше життя.

Для прикладу оцінка емоційного нахилу (sentiment analysis) використовує напрацювання машинного навчання та обробки природної мови для класифікації текстових даних за їх емоційним нахилом та об'єктивністю. Опрацьовані дані потім застосовуються компаніями та брендами для того, щоб зрозуміти як їхні користувачі та фанати сприймають рішення компанії з того чи іншого питання, або навіть для виявлення проблем, про які компанія могла не здогадуватись. Така інформація є свого роду компасом суспільної думки, який показує реакцію суспільства на діяльність компанії, при цьому не потребуючи десятків чи сотень працівників, які будуть перечитувати усі відгуки від користувачів.

Існує декілька методів автоматизованої оцінки емоційного нахилу. Методи засновані на правилах і словниках, статистичні та комбіновані методи. Методи засновані на правилах і словниках – дозволяють проаналізувати текст за допомогою попередньо складених словників та правил лінгвістичного аналізу [2]. Суть цього методу полягає у присвоєнні кожному слову значення зі словника, якщо воно є, а за загальну оцінку тексту приймають суму оцінок усіх слів. Хоча безпосередньо застосування цього методу є доволі простим – основна частина роботи припадає на складання словника з правильними вагами слів для досліджуваної галузі. Для прикладу, слово «великий» буде мати позитивне значення, якщо мова буде йти про обсяг пам'яті жорсткого диска і негативне, якщо мова буде йти про розміри телефону.

Статистичні методи через хороші результати в інших завданнях штучного інтелекту набирають дедалі більшої популярності. Сюди належать машинне навчання без вчителя, та метод заснований на теоретико-графових моделях. У першому випадку текст розбивають на ключові терміни, які людина позначає як позитивні чи негативні, маючи таку інформацію – система робить висновок про емоційний нахил всього тексту. Точність такого методу вища, але потребує великого обсягу даних для тренування моделі. Другий метод припускає що різні слова мають різний вплив на емоційний нахил тексту, тому потребує створення спеціальних графів досліджуваного тексту, які потім проходять процес ранжування вершин, класифікації знайдених слів і лише після того дозволяють отримати результат.

### **Література.**

1. Russell S. Artificial Intelligence: A Modern Approach, 4th US ed. [Електронний ресурс] / S. Russell, P. Norvig – Режим доступу до ресурсу: <http://aima.cs.berkeley.edu/>.
2. Cambria E. Affective Computing and Sentiment Analysis [Електронний ресурс] / Erik Cambria // IEEE Intelligent Systems. – 2016. – Режим доступу до ресурсу: <https://ieeexplore.ieee.org/document/7435182>.