УДК 004.91
Ю. Громик, І. Бойко, канд. фіз.-мат. наук; доц.
(Тернопільський національний технічний університет імені Івана Пулюя, Україна)

РОЗРОБКА ІНФОРМАЦІЙНОЇ СИСТЕМИ ДЛЯ ЗДІЙСНЕННЯ АНАЛІЗУ ТОНАЛЬНОСТІ ТЕКСТУ ІЗ ЗАСТОСУВАННЯМ ТЕХНОЛОГІЇ ГЛИБОКОГО МАШИННОГО НАВЧАННЯ ТА МОВИ PYTHON

UDC 004.91
Y. Gromyk, I. Boyko, Ph.D., Assoc. Prof.

DEVELOPMENT OF AN INFORMATION SYSTEM FOR SENTIMENT ANALYSIS USING DEEP MACHINE LEARNING AND PYTHON

Аналіз тональности тексту – це найпоширеніший інструмент класифікації тексту, який аналізує вхідні дані та визначає, чи є забарвлення поданого тексту позитивним, негативним чи нейтральним. Система аналізу тональности для аналізу тексту поєднує методи обробки природної мови (NLP) і машинного навчання, щоб надати настроїв об’єктам, темам, категоріям у реченні або фразі.

Обробка природної мови належить до області інформатики, а конкретніше – машинного навчання, яка зосереджена на тому, аби надати комп’ютерам здатність розуміти текст таким самим чином, як це роблять люди.

Аналіз тональности тексту фокусує свою увагу на полярности (позитивний, негативний чи нейтральний текст), а також почуттях та емоціях (злість, щастя, сум і т.д.), терміновости чи зацікавленості.

Зараз існує багато компаній, що використовують аналіз тональности тексту задля полегшення обробки великих обсягів тексту, які вони збирають, як зворотний зв’язок. Аналіз тональности тексту допомагає аналітикам даних у великих підприємствах оцінювати громадську думку, проводити детальне дослідження ринку, стежити за репутацією бренду та продукту та розуміти досвід клієнтів. Крім того, компанії, що займаються аналітикою даних, часто інтегрують сторонні API аналізу сентиментів у власні системи збору та аналізу активности клієнтів, моніторинг соціальних мереж, аби надавати корисну інформацію власним користувачам та розуміти поведінку користувачів задля покращення продукту.

Своєю чергою, глибоке машинне навчання (DL) вважається еволюцією машинного навчання. Воно об’єднує алгоритми, які мають на меті імітувати роботу людського мозку, інакше відомі як штучна нейронна мережа, і дозволило багато практичних застосувань машинного навчання, включаючи автоматизацію підтримки клієнтів та автономні автомобілі. Мова програмування Python дозволяє скористатися великою кількістю різноманітних бібліотек для як обробки тексту, так і його аналізу, до таких бібліотек належать: tensorflow, keras, pytorch, spacy.