

УДК 004.8:[004.934:159.942]

РОЗПІЗНАВАННЯ ЕМОЦІЙ У МОВІ ЛЮДИНИ З ВИКОРИСТАННЯМ ШІ ТА GPTs

Шабанова А. А.

E-mail anastasiia.shabanova@nure.ua

Харківський національний університет радіоелектроніки, каф. КІТС
м. Харків, Україна

The focus of this research is to examine the speech emotion recognition technologies and their integration with Generative Pre-Trained Transformers models. Speech emotion recognition systems can be used in education, medicine, marketing and other fields. By capturing subtle emotional nuances, GPT-based systems improve applications in customer service, mental health, and marketing, enabling more empathetic chatbots, psychological analysis, and data-driven strategies. This can facilitate the creation of more personalized systems and enable data-driven strategy adjustments.

Емоції є ключовим елементом людської комунікації, і їх правильне розпізнавання відіграє важливу роль у створенні природної взаємодії між людиною та машиною. Традиційні методи аналізу тексту часто не можуть точно визначити контекстуальні та приховані емоції, тому використання сучасних підходів на основі штучного інтелекту стає необхідністю.

Speech emotion recognition (SER) – процес ідентифікації афективних і емоційних станів з мовних сигналів за допомогою таких методів, як виділення ознак, вибір і класифікація, часто з використанням алгоритмів машинного навчання для точного визначення різних категорій емоцій у голосовому введенні [1-2]. Технологія використовує різноманітні характеристики голосу, такі як тон і висота, що часто передають приховані емоції.

Розпізнавання мовлення можна практично використовувати в багатьох сферах, наприклад, але не виключаючи виявлення розчарування або задоволеності клієнтів у режимі реального часу в службах обслуговування клієнтів і кол-центрах; моніторинг самопочуття пацієнта за допомогою його мовлення або дистанційних консультацій під час сеансів терапії; виявлення ознак тривоги та депресії для перевірки психічного здоров'я на підприємствах; спостереження за захопленням або неприязню клієнта до маркетингу; підвищення якості навчання через адаптацію до настрою учнів; персоналізація ігрового досвіду, посилення заходів безпеки. Цінна інформація про психічний та емоційний стан людини, яка містить потенційні ключі до її емоційного благополуччя, прихована в нашому мовленні.

Generative Pre-Trained Transformers (GPT) та інші трансформерні моделі значно покращили точність розпізнавання емоцій, оскільки вони здатні аналізувати текст у глибокому контексті. GPT-моделі навчаються на

великих масивах текстових даних і використовують контекстні векторні представлення, що дозволяє їм не лише визначати зміст повідомлення, але й аналізувати його емоційний тон. Це особливо корисно у таких сферах, як аналіз настроїв у соціальних мережах, автоматизоване модерацію контенту, розробка чат-ботів та голосових асистентів.

Методи розпізнавання емоцій із використанням GPT включають model fine-tuning на спеціалізованих наборах даних, що містять мітки емоцій. Також широко застосовуються комбінації GPT із традиційними моделями Sentiment Analysis, такими як BERT, RoBERTa та T5. Моделі дозволяють глибше аналізувати контекст і враховувати не лише тон, висоту голосу та окремі слова, а й їхній зв'язок з емоційним станом людини.

Практичні застосування цієї технології охоплюють широкий спектр сфер. Наприклад, у сфері обслуговування клієнтів чат-боти на основі GPT можуть не тільки відповідати на запити, а й визначати емоційний стан користувача, пропонуючи більш персоналізовані відповіді. В психологічному консультуванні ШІ може допомагати аналізувати текстові повідомлення пацієнтів, виявляючи ознаки депресії чи стресу. Крім того, емоційний аналіз широко застосовується у маркетингу для оцінки реакції аудиторії на бренди, продукти чи рекламні кампанії.

Однак попри значні досягнення існує низка викликів та обмежень. Основна проблема – це здатність моделей правильно інтерпретувати емоції у різних культурних контекстах, оскільки способи вираження емоцій можуть значно відрізнятися. Також є ризики маніпуляції емоціями користувачів, що викликає етичні питання щодо використання подібних технологій. GPT-моделі можуть містити вбудовану упередженість, що впливає на точність їхніх прогнозів.

Перспективи розвитку включають інтеграцію мультимодальних моделей, які об'єднують аналіз тексту, голосу та відео для більш точного визначення емоцій. Також ведуться дослідження щодо зменшення упередженості в моделях та створення прозорих алгоритмів ухвалення рішень. У майбутньому розширене використання GPT у сфері емоційного аналізу може призвести до створення більш емоційно коректних та “людяних” взаємодій між людиною та штучним інтелектом, що відкриває нові можливості для комунікації та автоматизації.

Список використаних джерел:

1. A Comprehensive Review of Speech Emotion Recognition Systems / T. M. Wani et al. IEEE Access. 2021. Vol. 9. P. 47795–47814. URL: <https://doi.org/10.1109/access.2021.3068045>.
2. Аксак Н.Г., Безсонов О.О., Новосельцев І.В., Руденко О.Г. Розпізнавання зміни розміру і кольору зображення на основі згорткової нейронної мережі. *Науково-технічний журнал "Біоніка інтелекту"*. 2018. Т. 2. № 91. С. 114-119. DOI: [https://doi.org/10.30837/bi.2018.2\(91\).17](https://doi.org/10.30837/bi.2018.2(91).17)