UDC 004.8:616.89-07

H.D. Dziumenko
A.T. Terenchuk

NATURAL LANGUAGE PROCESSING ALGORITHMS FOR RECOGNIZING THE EMOTIONAL STATE OF THE USER

Vinnytsia National Technical University

Анотація

У тезах розглядаються сучасні алгоритми обробки природної мови (NLP) для автоматичного розпізнавання емоційного стану користувача в контексті веборієнтованих чат-систем психологічної підтримки. Проаналізовано методи аналізу тональності тексту, багатокласової класифікації емоцій, а також підходи на основі трансформерних архітектур. Описано особливості застосування цих методів у задачах первинної психологічної діагностики.

Ключові слова: обробка природної мови, розпізнавання емоцій, аналіз тональності, трансформерні моделі, психологічна діагностика, чат-системи.

Abstract

The paper considers modern natural language processing (NLP) algorithms for automatic recognition of the emotional state of users in web-oriented psychological support chat systems. Methods of sentiment analysis, multi-class emotion classification, and transformer-based approaches are analyzed. The specifics of applying these methods to the tasks of primary psychological diagnostics are described.

Keywords: natural language processing, emotion recognition, sentiment analysis, transformer models, psychological diagnostics, chat systems.

Introduction

The rapid growth of digital mental health platforms has created a pressing demand for intelligent systems capable of detecting users' emotional states from text-based interactions. Emotional recognition is a critical component in web-oriented chat systems designed for primary psychological diagnostics, as it enables adaptive responses and early identification of mental health concerns. Natural language processing (NLP) provides a set of powerful tools to automate this process, bridging the gap between human expression and machine understanding [1].

The relevance of this topic is reinforced by the increasing prevalence of anxiety, depression, and other psychological disorders globally, particularly following the COVID-19 pandemic [2]. Automated emotion detection can extend the reach of psychological support to populations with limited access to traditional mental healthcare services, making NLP-based diagnostic tools a timely and socially significant area of research.

Sentiment Analysis and Emotion Classification

The foundational task in emotional state recognition is sentiment analysis – the classification of text into positive, negative, or neutral polarity. Classical approaches rely on lexicon-based methods, such as VADER and SentiWordNet, which assign emotional valence scores to individual words and aggregate them at the sentence level [3]. While computationally efficient, these methods struggle with sarcasm, idiomatic expressions, and contextual nuance, which are frequently encountered in mental health conversations.

A more granular approach is multi-class emotion classification, where text is assigned to one of the basic emotional categories defined by psychological models such as Ekman's six basic emotions (joy, sadness, anger, fear, disgust, surprise) or Plutchik's wheel of emotions [4]. Machine learning classifiers – including Support Vector Machines (SVM), Naive Bayes, and gradient boosting – trained on labelled corpora such as ISEAR or EmoInt, have demonstrated acceptable performance on these tasks. However, they are sensitive to feature engineering quality and domain shift.

Transformer-Based Approaches

The introduction of the Transformer architecture [5] and the subsequent pre-trained language models – BERT, RoBERTa, and their domain-specific variants – fundamentally advanced the state of emotion

recognition in text. These models are pre-trained on massive corpora and fine-tuned on emotion-labelled datasets, capturing deep semantic, syntactic, and contextual relationships that lexicon-based and classical ML methods cannot.

For psychological chat applications, MentalBERT and Mental-RoBERTa [6] are particularly noteworthy, as they are pre-trained on clinical and mental health-related texts. Fine-tuning such models on datasets like GoEmotions (27 categories derived from Reddit) achieves macro-F1 scores exceeding 0.60 on multi-label emotion classification tasks, substantially outperforming traditional approaches. The multi-label formulation is especially important in psychological contexts, where users often express mixed or co-occurring emotions.

Application in Web-Oriented Psychological Chat Systems

The practical deployment of NLP-based emotion recognition in web chat systems introduces additional engineering challenges. Real-time inference requires model optimization via quantization or distillation – for instance, DistilBERT achieves approximately 97% of BERT's performance at 40% fewer parameters [7]. Dialogue context management is equally important: emotion recognition must account for the conversational history rather than processing each message in isolation. Architectures combining recurrent networks or attention mechanisms with transformer encoders have been proposed for conversational emotion recognition (CER) tasks [8].

Privacy and ethical considerations are paramount when processing sensitive psychological data. Data minimization, on-device inference, and differential privacy mechanisms should be integrated into system design to comply with GDPR and medical data protection standards. Furthermore, the risk of systematic bias in training data – particularly with respect to gender, age, and cultural background – must be audited and mitigated to ensure equitable diagnostic support [2].

Conclusions

Modern NLP algorithms provide an effective foundation for automated emotional state recognition in psychological chat systems. Transformer-based models, especially those pre-trained on domain-specific mental health corpora, significantly outperform classical lexicon-based and ML approaches. The integration of conversational context, real-time optimization, and robust ethical frameworks are essential for the responsible deployment of such systems. Future research should focus on low-resource language adaptation and multimodal fusion to further enhance diagnostic accuracy.

REFERENCES

1. Cambria E., Das D., Bandyopadhyay S., Feraco A. A Practical Guide to Sentiment Analysis. Springer, 2017. 285 p.
2. Shatte A. B. R., Hutchinson D. M., Teague S. J. Machine learning in mental health: a scoping review of methods and applications. Psychological Medicine. 2019. Vol. 49, No. 9. P. 1426–1448.
3. Hutto C. J., Gilbert E. VADER: A Parsimonious Rule-based Model for Sentiment Analysis of Social Media Text. Proceedings of the AAAI Conference on Web and Social Media. 2014. Vol. 8. P. 216–225.
4. Plutchik R. The Nature of Emotions. American Scientist. 2001. Vol. 89, No. 4. P. 344–350.
5. Vaswani A., Shazeer N., Parmar N. et al. Attention Is All You Need. Advances in Neural Information Processing Systems. 2017. Vol. 30. P. 5998–6008.
6. Ji S., Zhang T., Ansari L., Fu J., Tiwari P., Cambria E. MentalBERT: Publicly Available Pretrained Language Models for Mental Healthcare. Proceedings of LREC 2022. P. 7184–7190.
7. Sanh V., Debut L., Chaumond J., Wolf T. DistilBERT, a distilled version of BERT: smaller, faster, cheaper and lighter. arXiv:1910.01108. 2019.
8. Poria S., Hazarika D., Majumder N., Naik G., Cambria E., Mihalcea R. MELD: A Multimodal Multi-Relational Emotion Recognition Dataset for Conversational Sentiment Analysis. Proceedings of ACL 2019. P. 527–536.

Дзюменко Георгій Дмитрович – студент групи 1ПІ-22б, факультет інформаційних технологій та комп'ютерної інженерії, Вінницький національний технічний університет, м. Вінниця, e–mail: heorhiidziumenko@gmail.com.

Теренчук Анатолій Тимофійович – кандидат технічних наук, доцент, старший викладач кафедри програмного забезпечення, Вінницький національний технічний університет, м. Вінниця, e-mail: anateren59@gmail.com.

Dziumenko Heorhii D. – student of group 1PI-22b, Faculty of Information Technologies and Computer Engineering, Vinnytsia National Technical University, Vinnytsia, e-mail: heorhiidziumenko@gmail.com.

Terenchuk Anatolii T. – Candidate of Technical Sciences, Associate Professor, Senior Lecturer of the Software Engineering Department, Vinnytsia National Technical University, Vinnytsia, e-mail: anateren59@gmail.com.