УДК 004.8:004.67
Hennadii Kochev[1], Nataliia Zhukova[2]
[1] PhD student NU "Zaporizhzhia Polytechnic"
[2] PhD (Philology), Associate Professor NU "Zaporizhzhia Polytechnic"

NEURAL NETWORK METHODS FOR PSYCHOLOGICAL TRAITS IDENTIFICATION BY SOCIAL NETWORK TEXT CONTENT

H.H. Kochev, N.M. Zhukova

The rapid growth of social networks brought huge amounts of user-generated text data. Every day, millions of users post comments and messages that reflect their feelings, attitudes, and behaviors. This creates new opportunities to automatically apply neural network methods to identify psychological traits from text. Such approaches are increasingly relevant across areas such as digital marketing, behavioral analytics, human-computer interaction, and online safety.

A key part of this field is the use of deep learning techniques for natural language processing (NLP). In particular, transformer-based models have greatly improved the ability to understand complex language patterns and relationships within text. Models such as BERT [1], RoBERTa [2], XLNet [3] can extract hidden psychological signals that traditional statistical methods cannot easily detect.

Social media text presents unique challenges and features that must be considered during model development. These features are:

1. informal and rapidly changing language (slang, abbreviations, hashtags);
2. short messages that depend on context;

3. multilingual usage and code-switching;

4. high levels of noise and variability in writing styles.

To address these difficulties, modern neural network workflows often involve multiple preprocessing and feature-extraction steps. Common components of workflows are:

1. text normalization and tokenization suited for social media formats;

2. generating embeddings with contextualized language models;

3. fine-tuning on specific datasets;

4. integrating metadata (timestamps and interaction patterns) when available.

One of the most valuable features of transformer-based models is the self-attention mechanism [4, p. 2], which enables them to capture long-range relationships between words and phrases. This enables the identification of subtle psychological cues, such as emotional tone, cognitive complexity, and interpersonal orientation, within a text sequence.

Identifying psychological traits often relies on established theoretical frameworks. One of the most common is the Five-Factor Model (Big Five) [5], which describes personality traits like openness, conscientiousness, extraversion, agreeableness, and neuroticism. Neural networks can be trained to predict these traits by learning from datasets where text samples are labeled with corresponding psychological scores.

Building these models typically involves several key steps:

1. pre-training on large general datasets to learn language representations;

2. fine-tuning on labeled datasets with psychological annotations;

3. evaluating performance using metrics: accuracy, F1 score, and correlations with human assessments;

4. iterative optimization to reduce bias and improve general performance.

Despite this, the quality of the training data remains a crucial factor affecting model performance. Psychological traits often show up indirectly in language, making annotation complex and subjective. There are multiple well-known problems in creating datasets for such a task, like limited availability of high-quality labeled data, subjectivity and inconsistency in human annotations, imbalanced distributions of personality traits, and others.

Beyond technical issues, ethical and legal aspects are vital when deploying these systems. Analyzing personal text data raises important concerns about privacy, consent, and potential misuse. To implement a system properly and responsibly, scientists need to ensure that user data is anonymized, obtain explicit consent for data collection and analysis, and provide transparency and clarity in model decisions.

Additionally, there is a growing need to develop models that fit specific languages and cultural contexts. Many existing NLP models are mostly trained on English data, limiting their effectiveness for other languages. For example,

adapting models to the Ukrainian language needs the creation of special datasets and consideration of local communication styles.

To sum up, neural network methods, especially those that use transformer architectures, provide powerful tools for identifying psychological traits from social media text. Future efforts should focus on improving data quality, addressing ethical issues, and creating models that adapt to different cultures to ensure fair and effective use across diverse groups.

REFERENCES

1. Devlin J., Chang M., Lee K., Toutanova K. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. 2019. URL : https://arxiv.org/pdf/1810.04805/ (Accessed : 05.04.2026).

2. Liu Y. RoBERTa: A Robustly Optimized BERT Pretraining Approach. 2019. URL: https://arxiv.org/pdf/1907.11692/ (Accessed : 05.04.2026).

3. Yang Z. XLNet: Generalized Autoregressive Pretraining for Language Understanding. 2019. URL : https://arxiv.org/pdf/1906.08237/ (Accessed : 05.04.2026).

4. Vaswani A. Attention Is All You Need : conference proceedings et al. *31st Conference on Neural Information Processing Systems* (NIPS 2017). Long Beach, CA, USA, 2017. P. 1-15. URL : https://arxiv.org/pdf/1706.03762/ (Accessed : 05.04.2026).

5. Najm N. Big Five Traits: A Critical Review. *Gadjah Mada International Journal of Business*. 2019. №21(2). P. 159-186. URL : https://www.researchgate.net/publication/335834113_Big_Five_Traits_A_Critical_Review/ (Accessed : 05.04.2026).