Method of recognizing attitude components in online discussions based on neural network models[1]

Anfisa Chuganskaya[1] [0000-0001-9727-6406], Danil Kireev[1] [0000-0001-7498-4987], Natalia Chudova[1] [0000-0001-9306-1280], Dmitry Devyatkin[1] [0000-0002-0811-725X], Yulia Kuznetsova[1] [0000-0001-9380-4478],

[1] [a]Federal Research Center "Computer Science and Control" of Russian Academy of Sciences, 44/2 Vavilova str., Moscow, 119333 Moscow, Russia
anfisa.makh@gmail.com

Abstract. Modern research on large social groups has focused on the use of big data. One of the significant areas of study is the analysis of attitudes and their components. The goal of this work was to explore the possibility of creating an automated tool for recognizing attitude components in online discussions. The hypothesis was that there would be a difference in the effectiveness of recognizing attitude components using neural network models pre-trained on a Russian-language corpus of online discussions. The study was based on psycholinguistic methods and language model training in the study of a large social group of users of Internet discussions. Neural network models of the BERT and ELECTRA families were used. The accuracy and F1 are used to compare models, and it was maximized, when selecting the hyperparameters. The ruRoBERTa-large model achieved the best results with an F_1 of 75.97% and an accuracy of 72.69%. Applying machine learning methods based on pre-trained transformer language models to a corpus of texts in which social media users express their attitudes towards social objects allows us to simulate the work of a psychologist who evaluates the representation of attitude components in the statements with an acceptable level of accuracy.

Keywords: Internet discussion, Text corpus, Attitude, Large social group, Neural network, Language model, BERT, Affective component, Behavioral component, Cognitive component.

1 Introduction

In modern science, interdisciplinary research is becoming increasingly important. In the last decade, the field of big data analysis and large social groups has become a place where psychologists and data scientist can work together. The ability to study public sentiment through observing people's behavior on social media, which is of interest to

[1] The study was partially supported by the Russian Science Foundation Grant No. 25-28-02431 "The Cognitive Component of Attitudes in Online Discussions on Social and Historical Topics".

sociologists, political scientists, cultural experts, social psychologists, and management practitioners, has been significantly influenced by advancements in automatic text analysis and machine learning.

The reaction of network communities to socially significant events of various types usually manifest itself in the discussion of attitudes. In social psychology, an attitude (social setting) is understood as a specific state of readiness for a certain behavior or attitude, which is formed as a result of previous social experience and has a guiding and (or) dynamic influence on an individual's reactions to all objects or situations with which they are associated (Andreeva, Bogomolova, Petrovskaya, 2002). In 1942, M. Smith proposed the structure of attitudes in the form of three components (Andreeva, 2018):

- cognitive (opinions, judgments about the object or situation of the attitude, knowledge about the properties, purpose, and methods of action of subjects in a social situation);

- affective (emotional assessment of the object of the attitude or situation, expression of sympathy or antipathy);

- behavioral (consistent behavior and the degree of readiness to implement the attitude towards the object or situation in behavior).

At the same time, attitudes have a significant function in the form of cognitive economy and lower costs of adapting to different situations due to the availability of a ready-made attitude (Ross and Nisbett, 2000).

This paper presents the results of the first stage of work on creating an algorithm for recognizing attitude components in the texts of online discussions. It is assumed that the emergence of such a tool will allow for monitoring of social networks as part of solving the problem of increasing the efficiency of public communication and studying large social groups. In combination with modern methods of automatic text analysis, such as thematic modeling and emotional predicate analysis, knowledge about the cognitive, evaluative, and behavioral components of attitudes discussed on social media can provide insights into the perceptions of the subject of discussion, attitudes towards it, and behavioral scenarios considered successful or ineffective by participants in the discussions.

The hypothesis was that there is a difference in the effectiveness of recognizing attitude components by neural network models pre-trained on a Russian-language corpus of online discussions.
The aim of the work was to investigate the possibility of modeling the recognition of attitude components by specialists (psychologists) using machine learning tools based on language models. Accordingly, the following tasks were solved:

1. Creation of a labeled corpus of online discussion texts containing attitudes;
2. Experimental study of the effectiveness of machine learning methods for a corpus of texts containing attitudes;
3. Evaluation of the prospects for modeling expert text analysis using machine learning methods.

2 Method

A corpus of texts from online discussions on the Pikabu platform was collected to solve the problem of selecting discussions. The choice of this user communication platform was based on the criteria of having sufficiently detailed text posts and comments in communication, ethical permission to use anonymous social network data for scientific research (Rule 4.1.5 of the Pic URL: https://pikabu.ru/information/rules-sn (Date of access: 9.10.2023)].). For the collected corpus of texts, the sample is characterized by the following demographic composition: 58% men and 42% women; 77% aged 19-45 and 18% over 45 (Mediakit, 2022). To train the models was used a labeled corpus of 2620 statements - fragments of discussions in the social network Pikabu.ru (Social network Pikabu. URL: https://pikabu.ru (Accessed: 09.02.2026)). The markup corpus was pre-processed: the texts were automatically divided into sentences using the Razdel library (Razdel, 2024).
The markup of the automatically collected text corpus was carried out by specialists – expert psychodiagnostics, having a Ph.D. degree in Psychology. Three psychologists took part in the creation of the corpus; the texts of discussions are collected by a crawler from social network. The markup was carried out in a specially created text corpus labeling interface, developed at FRC CSC RAS. Before inclusion in the analyzed corpus, text fragments were preprocessed: service information, user names, dates, and non-printable characters were removed. For each sentence, the consistency of the expert ratings in one of three categories (the component of the attitude: affective, cognitive, and behavioral, or a statement that was not related to the subject of the conversation and did not explicitly express the author's attitude towards that subject) was evaluated, which was calculated as the ratio of the number of experts who assigned the same class label to the number of experts who assigned that class to the sentence, which was needed because sentences splitting sometimes divided marked text in half.
To solve the problem of creating a method for automatically identifying attitude components from the texts of online discussions, transformer language models (Vaswani A. et al., 2017) were used. We chose transformer language models as they are the state-of-the-art solution for natural language processing tasks and suitable for working with big data, which is relevant within the framework of social psychology of large social groups.
We used models pre-trained on Russian texts (Zmitrovich D. et al., 2023). The following families of language models were used: BERT - the original model with a transformer architecture described in (Devlin J. et al., 2018.); RoBERTa - an optimized version of the previous model described in (Liu Y. et al., 2019); ELECTRA - a modified version of the BERT model described in (Clark K. et al., 2020). To be precise, we used

the following models: ruBERT-base, ruBERT-large, ruRoBERTa-large, ruELECTRA-small, ruELECTRA-medium, and ruELECTRA-large.

We trained the models to take the context into account: a pair of 2 sentences that are located before and after the current sentence were added to the left and to the right, accordingly, of the current sentence. It is worth noting that the context of the sentences was selected based on the tree structure of the discussion, meaning that the parent comment being responded to was used for the left context, and the comments left in response to the current comment were used for the right context, allowing for a more accurate representation of the context that the comment was left and labeled in.

3 Results

Before training, the filtered dataset was preprocessed so that each example was a single structured XML-style string: the left context was concatenated and enclosed in a "left-context" tag, the text for labeling was enclosed in a "text" tag, and the right context was concatenated and enclosed in a "right-context" tag. This was done to combine the example text with its context and create a structure that can be recognized by language model. Moreover, all the tags were added as special tokens, which allows to reduce their impact on the limited size of the model's input sequence. Furthermore, this dataset was divided into training and test datasets in proportions of 90% and 10% with stratified distribution, i.e. with equal class ratios in the training and test datasets.

The transformers framework (Wolf T. et al., 2019) and the PyTorch library (Paszke A. et al., 2019) were used to train the neural network models. To evaluate the quality of the models' predictions, the F_1 and accuracy metrics were calculated, and an error matrix was constructed. Before testing, the hyperparameters of the neural network models were selected using the Optuna library (Akiba T. et al., 2019) to maximize F_1 on 100 iterations.

The results of training the language models revealed differences in the quality of identifying the components of attitudes in the corpus of online discussion texts.

All models based on BERT showed a sufficient and high level of identifying the components of attitudes, reaching the required 70% level of accuracy. The highest result of automatic identification was made for the cognitive component, more than 80%. The lowest values were obtained for the affective component of attitude, less than 60%. However, higher results for the affective component were obtained for the ruRoBERTa-large model.

Figure 1 shows the results of the model's error matrix on the test set of the described dataset of discussions on Pikabu, where the absolute number of sentences is represented in the cells. The recognition accuracy is represented on the scale as a color indicator.

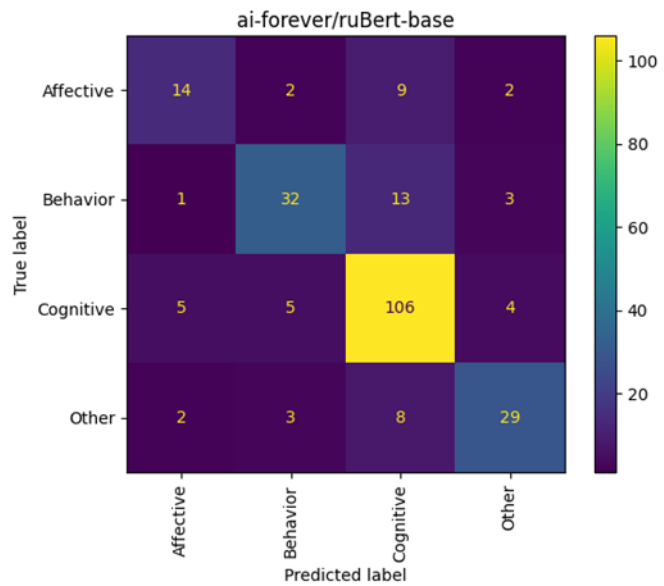

Fig. 1. The number of examples with the quality level of recognition by the ruBERT-base neural network model.

The resulting figure characterizes the high accuracy of recognizing the cognitive component of attitudes (106 examples, highlighted in yellow, which indicates high percentage). These results are largely due not only to the structure of the psycholinguistic representation of rational explanations within cognitive descriptions, but also to the large corpus of texts in this category. The category of behavioral types within attitudes was the second most accurate (32 examples, highlighted in turquoise, which indicates an approximation of 60% recognition). The results of the F_1 criterion and accuracy are presented in Table 1.

Table 1. The quality of attitude component recognition by the ruBERT-base, ruBERT-large, ruRoBERTa-large language model.

	ruBERT-base		ruBERT-large		ruRoBERTa-large	
Components of attitudes	F_1, %	accuracy, %	F1, %	accuracy, %	F1, %	accuracy, %
Affective component	57,14	51,85	56,6	55,56	74,51	70,37
Behavioral component	70,32	65,3	64,52	61,22	65,96	63,26
Cognitive component	82,81	88,33	81,89	86,67	84,92	89,17
No component	72,5	69,05	63,16	57,14	78,48	73,8
Total	**70,7**	**76,05**	**66,54**	**72,69**	**75,97**	**78,99**

It is worth noting the nature of errors in assigning categories during automatic recognition of attitude components. The largest error is associated with incorrect classification of the behavioral component into the cognitive group, which is generally expected when considering the speech components of behavior scenarios.

Another group of neural network models is represented by ruELECTRA. The summary of F_1 and accuracy is presented in Table 2.

Table 2. The quality of recognition of attitude components by ruELECTRA language models.

Model	**ruELECTRA-small**		**ruELECTRA-medium**		**ruELECTRA-large**	
Components of attitudes	**F1, %**	**Accuracy, %**	**F1, %**	**Accuracy, %**	**F1, %**	**Accuracy, %**
Affective component	52,17	44,44	51,06	44,44	64	59,26
Behavioral component	63,16	61,22	70,83	69,39	63,37	65,31
Cognitive component	80,78	85,83	83,95	85	79,67	80
No component	72,5	69,05	75,56	80,95	69,05	69,05
Total	**67,15**	**73,11**	**70,35**	**76,47**	**69,02**	**72,69**

The obtained data characterize a slightly lower accuracy of determining the components of attitudes by this model (within 70%) than by ruBERT models. However, the affective component is recognized at a critically low level and for two variations of the model does not even reach 50% of the recognition accuracy level. And the indicators for the cognitive component of attitudes are presented for the group of ruELECTRA models at the level of 80% and above, with the highest recognition accuracy of ruELECTRA-small (85%).

To compare the models, the F1 score is used, which was maximized when the hyperparameters were selected. If we look at the results, we can see the scores for other neural network models. In general, the ruRoBERTa-large model showed the best results with an F_1 of 75.97% and an accuracy of 72.69%.

In second place is the ruBERT-base model with an F_1 of 70.7% and an accuracy of 76.05%. In third place is the ruELECTRA-medium model with an F_1 of 70.35% and an accuracy of 76.47%. It is worth noting that the ruBERT-base and ruELECTRA-medium models showed similar results and may share the second place.

The cognitive component was best identified using the ruRoBERTa-large model with an F_1 of 84.92% and an accuracy of 89.17%, and was generally identified best by each model, but this result was expected as the cognitive component is the most frequent example in the dataset.

4 Conclusion

The data obtained allow us to consider the conducted pilot study to be successful – in none of the six cases does the overall recognition accuracy fall below 70%, which means that the possibility of automatic recognition of statements containing attitudes and their components in the texts of online discussions is confirmed. The result obtained is non-trivial, since the high variability of the speech design of social attitudes, combined with the significant linguistic non-normativity of online discussions, creates obvious difficulties in the automatic analysis of texts, which did not allow us to expect an easy success in the task of creating an algorithm (Nestik, 2025). In the next stage of the project, it is planned to carry out additional training on the extended corpus of the ruRoBERTa-large model, which showed the best results (the recognition accuracy for statements containing a cognitive component was more than 89%).

The ease with which the algorithm learned to recognize the speech patterns indicates that the speech manifestations of the attitude components themselves have a speech system (Kuznetsova, Mishlanov, Salimovskiy, Chudova, 2022). According to M.N. Kozhina's definition, "The speech systematism of a functional style is the interconnection and interdependence of the linguistic means used in a given field at different levels, both horizontally and vertically, based on the fulfillment of a single communicative task by these means, which is determined by the purpose of the extralinguistic basis of the corresponding speech variety, and which are related to each other by a specific functional meaning that expresses the specificity of the style" (Kozhina, 1972). Thus, the obtained result can be described as follows: in social networks, people regularly express their social attitudes, and their statements are made in three functional styles: evaluative judgment, instruction (giving advice), and information. Accordingly, we can also talk about the realism of the task of creating an "attitude profile" of a discussion. Automatic recognition of attitude components will allow us to calculate the proportion for each discussion in the future – in some discussions, people are more likely to share their emotions and assessments of the subjects discussed, in others, to offer their scenarios of behavior, and in others, to share knowledge and reflections on the subject of discussion.

As can be seen, in the discussions that we used to analyze the components of attitudes, the cognitive component prevailed, and therefore the focus on the information and cognitive aspirations of the discussants was more pronounced. This ability to quickly and automatically identify the nature of discussions taking place on the network opens up prospects for large-scale empirical research in the field of large group psychology (Nestik, Zhuravlev, 2019).

It is important to note the limitations associated with this exploratory study. The creation of a psychological data analysis tool requires the use of more data and additional validation procedures. The study's limitation lies in the use of data from a single online discussion platform, the need to increase the sample size for machine learning, and the additional iteration of quality assurance for the reliability of attitude component recognition by language models. In the future, it is advisable to use component-specific analysis to improve the quality and reliability of the assessment.

In conclusion, we would like to highlight an important procedural aspect of such developments. The success of machine learning in language models, such as "large language models", depends on the quality of the training corpus's annotation. The clearer the classification of statements for the experts, the more effective the resulting algorithm. It is evident that not all psychological typologies are equally valuable in this regard, so the choice of the subject of psychodiagnostics should be based on the "closeness" of the psychological concept to the realm of language.

The study showed the following:

- machine learning methods of so-called language models applied to a corpus of texts in which participants of discussions on social networks share their attitudes, allow to model with an acceptable level of accuracy the work of an expert psychologist, assessing the representation of the components of attitudes in statements;

- for the creation of the training corpus, the clarity of the theoretical concept used by the psychologists' annotators is important: the allocation of speech realizations of various authors' intentions should be based on the unambiguous understanding of the phenomena studied by experts;

- when creating a recognition algorithm, it is important both to have a more or less uniform representation of different types of statements in the training corpus, and to use language models that have been pre-trained on different principles.

References

1. Abakumova, I. V., Boguslavskaya, V. F., Grishina, A.V.: Ethnoreligious attitudes of contemporary Russian students toward labor migrants as a social group. Psychology in Russia: State of the Art, 9(1), 112-120. (2016).
2. Akiba, T. et al.: Optuna: A next-generation hyperparameter optimization framework //Proceedings of the 25th ACM SIGKDD international conference on knowledge discovery & data mining. – C. 2623-2631. (2019).
3. Allport, G. W.: The Nature of prejudice. MA. (1954).
4. Andreeva, G. M. Social Psychology. Moscow, Aspect Press. (2018).
5. Andreeva, G.M., Bogomolova, N.N., Petrovskaya, L.A.: Foreign Social Psychology of the 20th Century. Moscow: Aspect Press (2002).
6. Asmolov, A.G.: Activity and Attitude. Moscow: Moscow State University (1979).

7. Branitsky, A. A., Sharma, Ya. J., Kotenko, I. V., Fedorchenko, E. V., Krasov ,A. V., & Ushakov I. A.: determination of the mental state of users of the reddit social network based on machine learning methods // Information management systems, (1 (116)), 8-18. (2022).
8. Clark K. et al. Electra: Pre-training text encoders as discriminators rather than generators //arXiv preprint arXiv:2003.10555. (2020).
9. Devlin J. et al. Bert: Pre-training of deep bidirectional transformers for language understanding //arXiv preprint arXiv:1810.04805. (2018).
10. Devyatkin D. Extraction of Cognitive Operations from Scientific Texts // Communications in Computer and Information Science, 2019, 1093, pp. 189-200. (2019).
11. Devyatkin D., Chudova N., Chuganskaya A., Sharypina D. Methods for Recognition of Frustration-Derived Reactions on Social Media // Artificial Intelligence. RCAI 2021 Lecture Notes in Computer Science. — 2021 — Vol. 12948 — URL: https://doi.org/10.1007/978-3-030-86855-0_2 (accessed: 08.01.2023).
12. Farnadi, G., Sitaraman G., Sushmita, S. Celli, F. Kosinski, M. Stillwell D. (2016). Computational personality recognition in social media. User Modeling and User-Adapted Interaction, 26 (2–3), 109-142.
13. Fedotova V. A. (2017). Age-related differences in values and economic attitudes among Russians. Psychology in Russia: State of the Art, 10(1), 105-116.
14. Gulevich O., Osin E. N. Generalized trust and military attitudes in Russia: The role of national and global human identification // British Journal of Social Psychology. 2023. Vol. 62. No. 3. P. 1566-1579.
15. Houston M., Strebe V., and Stephenson J. Perspectives on Social Psychology / Translated from English. Moscow: EKSMO, 2001.
16. Kozhina M.N. On the Systematic Nature of the Scientific Style in Comparison with Some Other Styles / Perm University. Perm, 1972, pp. 115-116. (1972)/
17. Kurtukova, A.V., Romanov, A.S., Fedotova, A.M., Shelupanov, A.A. (2022). The use of machine learning methods and feature selection based on a genetic algorithm in solving the problem of determining the second Russian-language text for cybersecurity. Reports of TUSUR, 25, 1, 79-85. (2022).
18. Kuznetsova Yu. M., Mishlanov V. A., Salimovskiy V. A., Chudova N. V. The category of speech system in modeling affective processes (based on the material of network communication). Media Linguistics, 9 (3), 190–209. 2022. DOI: 10.21638/spbu22.2022.302 (2022).
19. La Piere R. Attitude versus action / (Eds.) Fishbein M., John N. Attitude Theory and Measurement. N.Y., 1967.
20. Leontiev, A.A. Psychology of communication. Moscow: Smysl Publishing House, 1999. 365 p.
21. Liu Y. et al. Roberta: A robustly optimized bert pretraining approach //arXiv preprint arXiv:1907.11692. – 2019.
22. Measures of Personality and Social Psychological Attitudes (Robinson, Shaver, & Wrightsman (Ed.), 1991.
23. Nestik, T. A. Collective Image of the Future: Socio-Psychological Analysis. – Moscow: Institute of Psychology of the Russian Academy of Sciences, 2025. – 657 p.
24. Nestik, T.A., Zhuravlev A.L. Analysis of Big Data in Psychology and Social Sciences: Promising Research Directions // Psychological Journal. – 2019. – Vol. 40. – No. 6. – Pp. 5-17. (2019)/
25. Oliseenko V.D., Abramov M.V. Prediction of the results of the 16-factor test R. Kettella based on the analysis of text posts of users of the social network // Scientific and Technical bulletin of information technologies, mechanics and optics. 2023b. Vol. 23, No. 2. pp. 279-288. doi: 10.17586/2226-1494-2023-23-2-279-288

26. Oliseenko V.D., Eirich M., Tulupyev A.L., Tulupyeva T.V. BERT and ELMo in task of classifying social media user's posts // Lecture Notes in Networks and Systems. 2023a. V. 566. P. 475-486. https://doi. org/10.1007/978-3-031-19620-1_45
27. Osipov, G.S.: Methods of Artificial Intelligence. Moscow, Fizmatlit. (2011).
28. Paszke A. et al. Pytorch: An imperative style, high-performance deep learning library //Advances in neural information processing systems. – 2019. – T. 32.
29. Picabu Media Kit. 2022. URL: https://specials.pikabu.ru/apps/docs/c/pikabu-media-kit.pdf (accessed on 04.04.2026).
30. Pikabu Social network. URL: https://pikabu.ru (Date of access: 9.10.2023).
31. Pikabu. Rules of Social Networks. URL: https://pikabu.ru/information/rules-sn (Date of access: 9.10.2025).
32. Razdel - segmentation of Russian-language text into tokens and sentences. URL: https://natasha.github.io/razdel / (Date of access: 02/12/2024).
33. Razina N.V. (2014) Attitudes to motherhood in different cultures. Psychology in Russia: State of the Art, 8(2), 93-104.
34. Rokeach M. Belief, Attitudes, and Values: A Theory of Organization and Change. San Francisco: Jossey- Bass, Inc. 1968.
35. Ross L., Nisbett R. (2000) Man and Situation. Lessons in Social Psychology / Translated from English by V. V. Rumynsky. — Moscow: Aspect Press. 429 p.
36. Smith, M. B., Bruner, J. S., & White, R. W. (1956). Opinions and Personality. New York: John Wiley and Sons.
37. Stefanenko T.G. Ethnopsychology: Practicum. - M.: "Aspect Press", 2006.
38. The World in 100 Years: A Collection of Articles. Moscow: Ves' Mir Publishing House. 2016 - 776 p. (2016)/
39. Vallat R. Pingouin: statistics in Python //J. Open Source Softw. – 2018. – T. 3. – №. 31. – C. 1026.
40. Vaswani A. et al. Attention is all you need //Advances in neural information processing systems. – 2017. – T. 30. Pp. 5998–6008. DOI: 10.48550/arxiv.1706.03762.
41. Vergunov E.G., E.I. Nikolaeva, Y.V. Bobrova To the question about psychometrical reliability of another psychological methods // Theoretical and Experimental Psychology - 2019 - Vol. 12 - № 1 - C. 61-68.
42. Wolf T. et al. Huggingface's transformers: State-of-the-art natural language processing //arXiv preprint arXiv:1910.03771. – 2019.
43. Zmitrovich D. et al. A Family of Pretrained Transformer Language Models for Russian //arXiv preprint arXiv:2309.10931. – 2023.