Petrenko Bogdan
student of higher education
Simon Kuznets Kharkiv National University of Economics
Kharkiv, Ukraine
Supervisor: Candidate of Technical Sciences, Associate Professor
Skorin Yuriy

NATURAL LANGUAGE PROCESSING TO ANALYZE USER FEEDBACK TO AUTOMATE FEEDBACK ANALYSIS AND IDENTIFY TRENDS IN LARGE AMOUNTS OF TEXT BUSINESS DATA

The purpose of the study is to develop a system of automated analysis of user feedback based on natural speech processing methods to identify sentiment, key aspects and topics.

The subject of the study is methods and models of natural speech processing to automate the analysis of text responses.

Research methods include machine learning, deep neural networks, statistical text analysis methods, and model quality assessment techniques.

Classical algorithms, neural network models, and transformer architectures are used. The results of the study can be used in e-commerce, service companies, software development, marketing and analytics to automate the analysis of reviews and identify trends in large amounts of text data.

In today's digital landscape, the amount of text data generated by users of products and services is growing rapidly.

In particular, user reviews have become an integral element of the modern business ecosystem.

The ability to effectively analyze these reviews has become a critical success factor for businesses looking to remain competitive.

Traditional methods of analyzing reviews based on manual processing face many limitations.

Analysts can only efficiently process a limited number of texts, and the subjectivity of human interpretation leads to inconsistencies in results.

Moreover, human resources often focus on the most recent feedback, completely leaving out historical data that could reveal long-term trends.

From a business perspective, user reviews provide invaluable information about:

- the strengths and weaknesses of a product or service;
- expectations and needs of customers;
- comparison with competitors;
- ideas for improvement and innovation;
- problems that need to be solved immediately.

However, extracting this information from a large array of unstructured text data remains a challenging task.

This is where natural speech processing (NLP) technologies come in. NLP is an interdisciplinary field that integrates computer science, artificial intelligence, and linguistics to enable interaction between computers and human language.

To solve the tasks, a comprehensive methodological approach will be used, which combines the methods of machine learning, deep learning, natural language processing and data analysis.

The practical significance of the results obtained lies in the possibility of using them to create effective systems for automated analysis of user feedback, which can be integrated into the business processes of companies of various sizes and industries.

Such systems allow not only to reduce the cost of analyzing reviews, but also to get deeper

and more objective insights into the perception of products and services by users, identify hidden trends and patterns, as well as promptly respond to critical problems.

As a result of the study, it was found that transformer models, in particular domain-adapted BERT, demonstrate the highest accuracy for sentiment analysis (F1-score 0.93), and combined RACL models are the most effective for aspect-oriented analysis (F1-score 0.85). The statistical significance of the results has been experimentally confirmed and recommendations for the selection of models for different scenarios have been developed.

The results of the study can be used in e-commerce, service companies, software development, marketing, and analytics to automate the analysis of reviews and identify trends in large amounts of text data.

Thus, the conducted research makes a significant contribution to the development of methods for automating the analysis of user reviews based on natural language processing and creates the basis for further improvement of such systems.

References:

1. Stanford Sentiment Treebank [Electronic resource] // Stanford NLP Group. 2023. – Access mode: https://nlp.stanford.edu/sentiment/.

2. State of AI Report 2023 [Electronic resource] // State of AI. 2023. Access mode: https://www.stateof.ai/.

3. TensorFlow Documentation [Electronic resource] // Google. 2024. Access mode: https://www.tensorflow.org/api_docs.

4. Beecham S. Software Process Improvement Problems in Twelve Software Companies: An Empirical Analysis / S. Beecham, T. Hall, A. Rainer, et al. // Empirical Software Engineering 2003 8(1): 7–42. [Electronic resource]. – Access mode: https://doi.org/10.1023/A:1021764731148.