

USING NATURAL LANGUAGE PROCESSING TO AUTOMATE USER FEEDBACK ANALYSIS

SKORIN YURI (yuriy.skorin@hneu.net)

PETRENKO BOHDAN (pebogdan1988@gmail.com)

Kharkiv National Economic University named after Semen Kuznets

The purpose of the study is to develop a system for automated analysis of user feedback based on natural language processing methods to detect tone, key aspects and topics. The object of the study is the processes of processing and analyzing user feedback in a digital environment. The subject of the study is natural language processing methods and models for automating the analysis of text feedback. The research methods include machine learning, deep neural networks, statistical methods of text analysis and methods for assessing the quality of models. Classical algorithms, neural network models and transformer architectures are used. The statistical significance of the results was experimentally confirmed and recommendations for choosing models for various scenarios were developed. The research results can be used in e-commerce, service companies, software development, marketing and analytics to automate the analysis of feedback and identify trends in large arrays of text data.

Statement of the problem.

In today's digital world, user review analysis has become a critical element for any business looking to improve the quality of its products and services. However, the scale and unstructured nature of this data pose significant challenges for traditional analysis. The main problem addressed in this study is the inefficiency and limitations of manual analysis of large volumes of text-based user reviews. This problem has several important aspects:

1. Exponential growth in data. Today's businesses receive thousands of new reviews every day across multiple channels. According to the Customer Experience Trends Report 2024, online reviews are growing at an estimated 18% annually, nearly double the 10% growth rate seen in 2020. The rate is even higher for the largest e-commerce platforms, with Amazon seeing an annual growth rate of 22% and services marketplaces seeing an annual growth rate of 25%.

2. Subjectivity and inconsistency of analysis. Human interpretation of text is subjective, which can lead to different conclusions about the same review. Independent studies conducted in 2023-2024 in the field of cloud technologies and financial services show that the level of consistency between different analysts when assessing the tone of reviews is only 60-75%, and when identifying specific problems mentioned in reviews, this figure drops to 55% [1].

3. Critical time costs. Manual analysis requires significant time resources. According to McKinsey experts (2024), an analyst spends an average of 2-3 minutes on a thorough analysis of one review of medium length, which for a company with 1000 reviews daily means 30-50 hours of work every day, or 5-8 analysts working full-time exclusively on review processing. This approach makes it impossible to respond promptly to critical issues.

4. Limited ability to detect hidden patterns. Human analysts are often unable to detect hidden trends and relationships in large data sets, especially when these patterns manifest over a long period of time or across different user segments. Forrester Research (2024) found that about 68% of valuable insights from user feedback remain undiscovered by manual analysis, especially when it comes to correlations between different aspects of the product and user reactions.

5. The increasing complexity of multilingual analysis. The globalization of business has led to the need to analyze reviews in many languages. By 2025, this figure is predicted to be 55-60%.

6. Lack of standardization and subjectivity of assessments. Different analysts may use different assessment methodologies and criteria, which makes it difficult to compare results and develop a unified strategy for responding to feedback. Such variability makes it impossible to compare data over time and makes it difficult to track progress in improving products and services.

7. Rising costs for review analytics. Economic studies show a rapid increase in the cost of manual review analysis. According to the Customer Experience Association, the average annual cost of review

analytics for the average company has increased to \$280,000 in 2024, and is projected to be \$350,000-400,000 by the end of 2025.

Presentation of the main material.

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 part of the modern business ecosystem. The ability to effectively analyze these reviews has become a critical success factor for companies seeking to remain competitive. Traditional manual review analysis methods face many limitations. Analysts can only effectively process a limited amount of text, and the subjectivity of human interpretation leads to inconsistent results. Moreover, human resources often focus on the most recent reviews, completely ignoring historical data that could reveal long-term trends. From a business perspective, user reviews contain invaluable information about: the strengths and weaknesses of a product or service; customer expectations and needs; comparisons with competitors; ideas for improvement and innovation; problems that need immediate solutions. However, extracting this information from large amounts of unstructured text data remains a challenge. This is where natural language processing (NLP) technologies come to the rescue. NLP is an interdisciplinary field that combines computer science, artificial intelligence, and linguistics to enable interaction between computers and human language. Applying NLP to analyze user feedback is a complex task that includes a number of subtasks: text preprocessing (tokenization, lemmatization, stop word removal); sentiment analysis; key aspects and themes extraction; classification of feedback into different categories; emotional coloring detection; generalization and report generation. These tasks require the use of various models and algorithms - from classical statistical methods to modern deep learning approaches using transformers and large language models (LLM). To solve the tasks set, a complex methodological approach will be used that combines machine learning, deep learning, natural language processing, and data analysis.

The research will be conducted in several stages:

1. Analytical stage. At this stage, a detailed analysis of existing methods and models, their advantages and limitations will be conducted. Modern research in the field of NLP and feedback analysis will be studied, the most promising approaches and technologies will be identified;
2. Design and development stage. At this stage, specific methods and models will be developed to solve the tasks set. The development will include both the adaptation of existing approaches and the creation of new methods and architectures optimized for feedback analysis;
3. Experimental stage. This stage includes conducting experiments to evaluate the effectiveness of the developed methods and models. The experiments will be conducted on real sets of reviews from different domains (e-commerce, hotel business, software, services, etc.) and in different languages. Standard metrics such as accuracy, completeness, F1-measure, as well as specialized metrics to assess the quality of aspect and topic detection will be used for evaluation;
4. Analytical and generalization stage. At this stage, the results of the experiments will be analyzed, patterns will be identified, and conclusions will be formulated regarding the effectiveness of the developed methods and models. The main advantages and limitations of the proposed approaches, as well as directions for further research, will be identified.

Conclusions.

The study conducted on the application of natural language processing to automate the analysis of user feedback allowed us to obtain a comprehensive understanding of the effectiveness, advantages and limitations of various NLP methods and models in solving this problem. As a result of the study, all the goals and tasks were achieved, which allowed us to formulate the following conclusions:

Analysis of the current state of the problem confirmed the critical importance of automating the analysis of user feedback in the context of the constant growth of text data volumes. Traditional manual analysis faces fundamental limitations associated with time costs, subjectivity of assessments and the inability to effectively process large data sets. These limitations are especially acute in the context of business globalization, when companies receive feedback in different languages and from different sources.

Theoretical research on natural language processing methods has demonstrated the evolution of approaches to text analysis - from simple dictionary methods to complex transformer architectures. Each of these approaches has its own advantages and limitations, and the choice of a particular method depends on the specifics of the task, available computing resources, and requirements for accuracy and speed of

analysis. Based on experimental research, it was found that transformer models provide the highest accuracy for analyzing the tone of responses, significantly surpassing classical methods, such as. It was found that the effectiveness of all models significantly depends on the specificity of the domain and the quality of data preprocessing. Domain adaptation and additional training on specific data corpora significantly increase the accuracy of analysis, especially for specialized industries with their own terminology.

The cost-effectiveness of implementing automated feedback analysis systems is confirmed by a significant reduction in the time and human resources required to process feedback, as well as by improving the quality of the insights obtained, which allows companies to make more informed decisions regarding the development of products and services. The practical significance of the results obtained lies in the possibility of using them to create effective automated user feedback analysis systems that can be integrated into the business processes of companies of various scales and industries. Such systems allow not only to reduce the costs of feedback analysis, but also to obtain deeper and more objective insights into the perception of products and services by users, to identify hidden trends and patterns, and to respond promptly to critical problems.

Thus, the conducted research makes a significant contribution to the development of methods for automating user feedback analysis based on natural language processing and creates a 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.

UDC 004.93'1:336.74

CONCEPT OF CONSTRUCTION OF AN INFORMATION AND MEASUREMENT SYSTEM BASED ON VIRTUAL COMPUTER SIMULATORS

SKORIN YURI (yuriy.skorin@hneu.net)

Kharkiv National Economic University named after Semen Kuznets

An analysis of traditional methods and measuring instruments and proposals for an alternative solution to the problem were made. It was noted that in addition to their intended use, i.e. as virtual measuring instruments, the use of virtual devices for building virtual simulators based on them is quite promising, which will ensure increased visibility and quality of training, which, in turn, creates the prerequisites for including them in existing ones or creating new virtual devices based on them. The study involves highlighting virtual devices as basic for building virtual simulators based on them, which ensure increased efficiency and visibility of the educational process and create the prerequisites for creating and improving distance learning systems.

Statement of the problem.

The gradual development of computer technology, computerization of all sectors of the economy suggests the use of such a powerful technological potential as computerization in improving the measurement process in information and measuring systems.

The search for a solution led to the need to create computer simulators based on virtual devices, analogues of which already exist and demonstrate huge advantages over the so-called traditional devices, which provides an incentive and opportunity to create, based on the virtualization of the measurement