https://doi.org/10.31891/csit-2025-4-13
UDC 004.93'1:336.74

SKORIN Yuriy
Semen Kuznets Kharkiv National University of Economics

IMPROVEMENT OF THE AUTOMATED NLP SYSTEM AS A FACTOR IN IMPROVING THE QUALITY OF MARKETING STRATEGY FORMATION

Natural language processing in company marketing is transforming data analytics, offering new opportunities to understand customers and optimize strategies. Natural language processing simplifies processes such as sentiment analysis, segmentation, and ad targeting. It is important to consider data accuracy, security, and query management skills training for effective use of technology. One of the main challenges in marketing analytics is the transformation of initial numerical data into understandable and useful conclusions for humans. The way to solve the problem are natural language processing technologies and generative artificial intelligence, which allow you to turn complex data into accessible and useful information for work. Traditional manual analysis of reviews in marketing analytics has long ceased to meet modern business requirements, because it requires huge human resources, which makes the process extremely costly. Natural language processing offers a solution to this problem through the use of algorithms capable of automatically analyzing the semantics of the text, determining the tone of statements, and isolating key topics from large data sets. The purpose of this study is to develop a system of automated analysis of user reviews based on the developed effective methods and models for automated analysis of user reviews in the field of marketing of companies using natural speech processing technologies. The paper describes the problem to be solved and formulates a scientific task; analyzes approaches, methods and models for solving research problems; sets research tasks, analyzes theoretical approaches to solving research problems; considers theoretical aspects of natural language processing; investigates various models and algorithms for analyzing feedback, and also conducts an experimental assessment of their effectiveness on real data; models, algorithms and analysis of their adequacy in solving research problems; methodological support for the organization of research is being improved. The results of the study can be used to develop software solutions that will allow companies to better understand the needs of their customers, quickly respond to problems and improve the quality of their products and services.

Keywords: natural language processing, marketing, automation of user feedback analysis, method, analysis, evaluation, business process.

СКОРІН Юрій
Харківський національний економічний університет імені Семена Кузнеця

УДОСКОНАЛЕННЯ АВТОМАТИЗОВАНОЇ СИСТЕМИ NLP, ЯК ФАКТОР ПІДВИЩЕННЯ ЯКОСТІ ФОРМУВАННЯ МАРКЕТИНГОВОЇ СТРАТЕГІЇ

Оброблення природної мови в корпоративному маркетингу трансформує аналітику даних, відкриваючи нові можливості для розуміння клієнтів і оптимізації стратегій, спрощує такі процеси, як аналіз настроїв, сегментація реклами. Важливо враховувати точність даних, безпеку та навчання навичкам управління запитами для ефективного використання технологій. Одним із головних викликів у маркетинговій аналітиці є перетворення початкових числових даних на зрозумілі та корисні для людини висновки. Шлях до вирішення проблеми – це технології оброблення природної мови та генеративний штучний інтелект, які дозволяють перетворювати складні дані на доступну та корисну інформацію для роботи. Традиційний аналіз оглядів у маркетинговій аналітиці давно перестав відповідати сучасним вимогам бізнесу, оскільки він вимагає величезних людських ресурсів, що робить процес надзвичайно дорогим. Оброблення природної мови пропонує рішення цієї проблеми завдяки використанню алгоритмів, здатних автоматично аналізувати семантику тексту, визначати тон тверджень і виділяти ключові теми з великих наборів даних. Метою статті є розроблення системи автоматизованого аналізу відгуків користувачів на основі розроблених ефективних методів і моделей автоматизованого аналізу відгуків користувачів у сфері маркетингу компаній із використанням технологій обробки природної мови. У статті описується задача, яку потрібно розв'язати, ставиться наукове завдання; аналізує підходи, методи та моделі для розв'язання дослідницьких задач; встановлює дослідницькі завдання, аналізуються теоретичні підходи до розв'язання дослідницьких задач; розглядаються теоретичні аспекти оброблення природної мови; досліджуються різні моделі та алгоритми для аналізу зворотного зв'язку, а також проводиться експериментальне оцінювання ефективності на реальних даних; досліджуються моделі, алгоритми та аналізується їх адекватність для розв'язання дослідницьких задач; удосконалюється методологічна підтримка організації досліджень. Результати дослідження можна використати для розроблення програмних рішень, які дозволять компаніям краще розуміти потреби своїх клієнтів, швидко реагувати на проблеми та підвищувати якість своїх продуктів і послуг.

Ключові слова: обробка природної мови, маркетинг, автоматизація аналізу зворотного зв'язку користувачами, метод, аналіз, оцінка, бізнес-процес.

Received / Стаття надійшла до редакції 01.09.2025
Accepted / Прийнята до друку 24.09.2025

Introduction

In today's digital landscape, the amount of text data generated by users of products and services is skyrocketing. 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. According to research, more than 90% of consumers read reviews online before buying, and 84% trust them as much as personal recommendations. Natural language processing in company marketing has the potential to transform data analytics by offering new opportunities to understand customers and optimize strategies.

One of the main challenges of marketing analytics is referencing initial numerical data and converting it into understandable and useful conclusions. In this case, natural language processing technologies and generative artificial intelligence, which turn complex numerical data into simple and useful information for work, will help. NLP models are trained on a wealth of information from the internet, including social media posts and news. The main goal of using NLP and generative AI is to gain specific insights from marketing data based on the knowledge collected.

Benefits of NLP for marketing data analysis:
– improving the decision-making process with the help of available information; improving the efficiency of data analysis and improving the accuracy of trend forecasting and understanding of customer needs.

Integrating NLP into marketing analytics involves several stages:
– definition of purpose: a clear understanding of the tasks or problems you want to solve with NLP;
– use of ready-made models: huge models that have already been studied on large amounts of data, which allows you to preserve NLP capabilities without high training costs;
– model adaptation: ready-made models often know general information from the Internet, but not from your data, two approaches can be taken to avoid this: fine-tuning – the model is trained from your company's data or through intermediary solutions that link your data to the finished model;
– testing and implementation: after setup and integration, it is necessary to test the system to ensure the accuracy and usefulness of the results obtained, and then implement them in daily work;
– analysis and adjustments: Regular analysis of the results will help identify areas for improvement, adjust settings or methods for using the model.

Application of NLP in marketing, which helps improve brand strategy, personalize customer experience, and optimize content for better engagement and visibility: sentiment analysis for brand perception; customer segmentation and personalization; content optimization and SEO. Working with NLP in marketing is a series of unique tasks, one of the main of which is operational engineering. Traditional methods of analyzing reviews based on manual processing have many limitations.

Analysts can only efficiently process a limited number of texts, and the subjectivity of human interpretation leads to difficulties in establishing and maintaining contacts with other people, which manifests itself in isolation and problems with mutual understanding (inconsistency). From a business point of view, 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.

Using NLP to analyze user feedback is a complex task that involves a number of subtasks:
– pre-processing of text;
– determination of tonality;
– coverage of key aspects and topics;
– classification of reviews into different categories;
– identification of emotional coloring;
– summarizing and generating reports.

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 (LLMs). The market for NLP-based feedback analysis solutions is growing rapidly. According to analysts, by 2026, the global text analysis software market will reach $16 billion, demonstrating an annual growth rate of more than 17%. This shows the growing interest of businesses in technologies capable of efficiently handling user feedback. However, despite significant advances in the field of NLP, automated analysis of user feedback still faces many problems. Among them are the polysemy of words, idiomatic expressions, sarcasm, variety of languages, spelling mistakes, jargon, and many other linguistic nuances that require a deep understanding of the context. Solving these problems requires a comprehensive approach that combines the most modern NLP techniques with an understanding of the business context and specifics of a particular industry.

Analysis of NLP models. Classical machine learning algorithms, despite the emergence of more complex neural network approaches, remain relevant for certain scenarios due to their ease of implementation, interpretation, and efficiency with limited data. Their work is based on the manual construction of features and the application of proven statistical principles.

1. The naïve Bayesian classifier (Naive Bayes) represents one of the simplest but most effective approaches to text classification. The method is based on Bayes' theorem with the assumption of conditional independence of features (words) in a given class. To analyze the sentiment of responses, the model calculates the probability of text belonging to a class (positive/negative) based on word frequencies. The advantages of Naive Bayes for review analysis include the speed of training and classification (processing hundreds of thousands of reviews in seconds), minimal computational resource requirements, efficient handling of small training samples, and natural processing of multiclass classification. The disadvantages of the method are critical for complex feedback: the assumption of the independence of words ignores the context (the phrase "not good" is classified incorrectly due to independent analysis of the words "not" and "good"), the inability to take into account the order of words leads to a loss of

semantics, sensitivity to the distribution of classes in the study sample, and problems with new words that were not encountered during teaching.

2. The Support Vector Machine (SVM) support vector method represents a more powerful approach looking for the optimal hyperplane for separating classes in a high-dimensional feature space. For the response classification problem, SVM maximizes the distance (margin) between the hyperplane and the nearest points of each class. Advantages of SVM for feedback analysis: high accuracy with the correct selection of parameters, efficiency in high-dimensional spaces, robustness to overtraining due to margin maximization, the ability to work with nonlinear dependencies through cores. Disadvantages: sensitivity to the choice of parameters, high computational complexity with large data sets, difficulty in interpreting results, the need to normalize features.

3. Ensemble methods represent the next level of complexity by combining the results of a set of base models to improve accuracy and stability. Random Forest (RF) and Gradient Boosting (GB) turned out to be the most effective for analyzing reviews. RF builds a set of decision trees on different subsamples of data and traits, aggregating their predictions through voting or averaging. The RF learning algorithm includes: for each tree, sampling bootstrap samples from the training set, randomly selecting a subset of features in each node, building the tree to the maximum depth without cropping. GB, in particular XGBoost and LightGBM, builds trees sequentially, where each successive tree learns from the mistakes of the previous ones.

4. Deep learning models for sequence processing. Recurrent neural networks (RNNs) were the first significant breakthrough in the application of deep learning to NLP, allowing models to account for the sequential nature of text. Unlike classical methods, RNN processes text as a sequence, maintaining an internal state (memory) that is updated at each step. Convolutional neural networks for text (TextCNN) have proposed an alternative approach to word processing, adapting a successful architecture from computer vision. The advantages of TextCNN for review analysis include high speed due to the possibility of full parallelization (unlike RNN), efficient detection of persistent phrases and expressions ("absolutely excellent", "complete disappointment"), fewer parameters compared to LSTM, and the ability to use pre-trained embeddings (Word2Vec, GloVe). Disadvantages include limited consideration of long-term dependencies (filters usually only cover 3-7 words), difficulty modeling word order outside the convolution window, and lower efficiency on very long reviews compared to BiLSTM.

5. Transformer models and their application for analyzing reviews. The Transformer architecture revolutionized NLP, completely abandoning recurring relationships in favor of the self-attention mechanism. This allowed for better parallelization, more efficient handling of long sequences, and modeling of complex dependencies. BERT applied the transformer encoder to create contextual representations of words through bidirectional pre-learning. Key innovations include Masked Language Modeling (MLM), where 15% of tokens are randomly masked for prediction, and Next Sentence Prediction (NSP) for understanding relationships between sentences.

Fine-tuning BERT for the sentiment classification problem involves adding a classification layer and training the entire end-to-end model on the marked-up feedback:

– RoBERTa improved BERT by removing the NSP task, which turned out to be ineffective;
– dynamic masking (different masks at each era); larger batch sizes (8K instead of 256);
– more training data (160GB of text);
– longer training.

These modifications resulted in a steady improvement of 1-2% on most NLP tasks. For analyzing reviews, RoBERTa is especially effective on complex domains with technical vocabulary. The choice of the optimal model for a particular review analysis scenario depends on many factors. Classical methods (XGBoost) or API LLM are suitable for prototyping and rapid hypothesis testing. For medium-scale production systems (10K-100K reviews/day), DistilBERT or domain-adapted BERT is optimal. For critical applications with maximum accuracy requirements, RoBERTa or domain-adapted BERT with careful fine-tuning should be used. Thus, the modern arsenal of NLP models for analyzing feedback is represented by a wide range from simple classical methods to complex transformer architectures. No model is universally optimal – the choice is determined by specific requirements for accuracy, speed, resources and interpretation for a particular business scenario.

Literature review

The article [1] focuses on how to manage innovative marketing in today's competitive market. The research shows that many companies in Ukraine don't use new marketing methods well. The article talks about new marketing tools that use parts of artificial intelligence and machine learning. It also lists the main skills needed for digital marketing, like video editing, graphic design, coding, making programs, writing, knowing software, experience with SEO and SEM, using software, managing projects, analyzing data, using marketing automation, and other related skills. The article also highlights the importance of protecting customers' personal information, as it's a key part of digital marketing. It looks at different technologies and methods for finding and securing users' personal data.

The article [2] says that digital marketing helps bring customers' attention to a company, its brand, products, and services. Using artificial intelligence along with digital marketing can improve customer service and help promote the brand more effectively. The article looks at how artificial intelligence is changing modern marketing strategies. It's important to note that AI has many benefits, but it also brings up problems and ethical concerns. The article points out how AI can be used to create personalized messages that make customers more interested and loyal. The article also says that AI has a big role in the future of marketing, but it's important to use it carefully, considering both its strengths and the ethical and practical issues involved. The wide use of AI in digital marketing shows how important it is in today's world. The paper [3] explains that making a marketing plan is a key way for a company to grow and do well. Using SWOT analysis, the paper looks at the company's strengths and weaknesses, as well as chances and challenges in the market. From this, the main areas for the marketing plan are determined, which help improve management, boost sales, and make the company more competitive.

The article [4] talks about how emotions can be detected from texts and explains the main methods researchers use when building NLP text systems. It also mentions some recent ideas that are being explored in this area. Finally, it lists some unanswered questions and future possibilities for NLP research.

Publication [5] shows how NLP is changing data analysis in marketing, giving companies better ways to understand their customers and improve their strategies. This makes tasks like sentiment analysis, customer grouping, and ad targeting easier.

The article [6] says the main goal of NLP is to build systems that can work well with people, making technology easier to use and better at understanding everyday communication. As NLP grows, it changes how we use machines, offering new ways to communicate effectively and naturally. It also outlines some key approaches to NLP, including controlled, uncontrolled, natural language understanding, and natural language generation.

The author [7] explains that unlike artificial languages, natural languages change over time and are hard to define with clear rules. They suggest looking at NLP in a wide way, covering any computer work with natural language. The author points out that NLP can range from simple tasks like counting word use to understand writing styles, to more complex tasks where systems truly "understand" human language and can give useful answers.

The textbook [8] gives a detailed look at the basics, models, and methods of NLP, providing both theory and practical knowledge for using modern language tools to solve engineering and analysis problems.

Publications [9] explore what NLP is, how it connects to artificial intelligence, how it helps people work more efficiently, how it improves user experience, how it creates new data for analysis, and gives examples of NLP tasks.

The article [10] examines the significant impact of artificial intelligence (AI) on modern company strategies. Automating routine tasks with the help of AI reduces the cost of time and resources, allowing specialists to focus on the strategic aspects of their work. This helps businesses not only handle inquiries efficiently but also increase customer loyalty through timely responses. The introduction of AI in campaigns can significantly increase the conversion rate thanks to targeted advertising messages and personalized recommendations. At the same time, the introduction of AI poses new challenges for companies, especially in the field of ethics and data protection. Ensuring the privacy and security of customer data becomes a critical task, as misuse or improper storage of information can lead to reputational and legal consequences.

Purpose, objectives and methods of research

In today's digital world, user feedback 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:

– exponential growth in data volumes;
– subjectivism and inconsistency of analysis;
– critical time costs;
– limited ability to detect hidden patterns;
– the increasing complexity of multilingual analysis;
– lack of standardization and subjectivity of assessment;
– rising costs for review analytics.

However, applying NLP to analyze user feedback also faces a number of challenges that need to be addressed:

– understanding context and ambiguity; revealing irony, sarcasm and hidden meanings;
– processing of unstructured and "noisy" text; multilingualism and intercultural aspects;
– adaptation to the specifics of the domain; assessment of the reliability of reviews and detection of false reviews;
– ethical issues and biases of models. Based on the above and taking into account current trends and challenges in the field of NLP, it is possible to formulate the purpose of the study.

The purpose of the study is, firstly, to develop and evaluate the effectiveness of a comprehensive system of automated analysis of user feedback based on modern methods of natural language processing, which provides high accuracy in identifying sentiment, key aspects and topics, identifying critical problems and opportunities for improvement in feedback in different languages, and secondly, forming practical recommendations for the implementation of work results in the business environment.

Table 1

Average time spent on manual analysis of reviews of varying complexity

Type of response	Average length (words)	Time to analyze (min)	Number of reviews per working day (8 hours)	Number of analysts per 1000 reviews/day
Short	10-30	1-2	240-480	2-4
Medium	31-100	2-4	120-240	4-8
Long	101-300	4-8	60-120	8-16
Detailed	>300	8-15	30-60	16-33

Unlike existing studies that focus on specific aspects of review analysis, this work offers an integrated approach that combines state-of-the-art NLP techniques to solve the entire complex of problems related to automated analysis of user reviews. Particular attention is paid to the adaptability of the system to different domains and languages, as well as its ability to provide high accuracy of analysis in the face of limited training data.

Solving this scientific problem is of great practical importance for business, as it will allow companies to:

– quickly identify problematic aspects of products and services, even with large volumes of feedback;
– make informed decisions on improving products based on objective analysis; to effectively allocate resources to the most critical areas of development;
– increase customer satisfaction through prompt response to their needs;
– gain a competitive advantage through a deeper understanding of the market and consumers;
– significantly reduce the cost of analyzing reviews while improving the quality of results;
– proactively identify new trends and opportunities based on the analysis of large data sets.

Table 2 presents a comparison of the effectiveness of different approaches for key feedback analysis tasks based on recent research. Successfully integrating a feedback analysis system into business processes requires not only choosing the right technologies, but also understanding the business context and needs of a particular organization. According to McKinsey, companies that effectively use user feedback analytics demonstrate 15-20% higher customer retention rates and 10-15% higher average revenue per customer [7].

The most successful implementations of review analysis are characterized by several key features:

– first, they provide integration with existing systems such as CRMs, customer support systems, and analytics platforms, this allows you to combine data from different sources and gain a comprehensive understanding of customer needs and problems;
– second, effective feedback analysis systems provide results in an understandable and actionable format, they not only generate statistical reports, but also highlight specific problems that need attention and provide recommendations for solving them, this is especially important for users who do not have technical training, but make strategic decisions based on the results of the analysis.

Table 2

Comparison of the effectiveness of different approaches of key review tasks

Method	Sentiment Analysis (F1)	Topic Detection (NMI)	Cross-Lingualism	Speed (Feedback/sec)	Need for Data for Learning	Explainability
Vocabulary Methods	0.65-0.70	N/A	Low	500-700	Minimum	High
ML with manual signs	0.75-0.85	0.50-0.65	Low	100-300		Medium
CNN/RNN	0.82-0.88	0.60-0.70	Medium	50-100	High	Low
BERT/RoBERTa	0.88-0.92	0.65-0.75	Medium-High	20-50	High	Low-Medium
LLM + Few-shot	0.90-0.95	0.70-0.80	High	5-15	Minimum	Average
Specialized models	0.92-0.96	0.75-0.85	Medium-high	20-40	Medium-high	Medium-high

Thirdly, successful review analysis systems provide full-cycle automation – from collecting and analyzing reviews to generating insights and recommendations. This allows you to significantly reduce the time from receiving feedback to responding to it, which is critically important in today's fast-changing business environment. Summarizing the analysis of approaches, methods and models for automating the analysis of user feedback, several key conclusions can be drawn. The modern landscape of feedback analysis technologies is represented by a wide range of methods – from simple dictionary approaches to complex neural network architectures and large language models.

The choice of a specific method depends on many factors: the volume and quality of the available data, the specifics of the domain, the requirements for processing speed, the required level of accuracy and interpretation of the results. Transformer models, especially the specialized domain versions of BERT and RoBERTa, demonstrate the best ratio of accuracy, speed, and resource intensity for most practical feedback analysis tasks. They provide high accuracy in understanding the context and nuances of language, which is critical for correctly interpreting user feedback. Large language models with few-shot training open up new opportunities for the rapid implementation of feedback analysis systems in new domains and for new products. They are especially valuable in situations with limited training data or when you need to quickly adapt to new types of feedback. Aspect-oriented sentiment analysis is the most informative approach for businesses, as it allows you to gain a detailed understanding of the strengths and weaknesses of products or services. The combination of ABSA with topic and aspect detection techniques provides the most complete picture of user feedback. Multilingual models provide better results for multilingual analysis compared to the translation + analysis approach, especially when understanding cultural and linguistic nuances is important.

These models are becoming increasingly important in a globalized business environment. The optimal solution for most business problems is a combination of different methods, which allows you to ensure high accuracy, efficiency and adaptability of the feedback analysis system. This approach allows you to use the advantages of each method and minimize their limitations. Based on the analysis of the problems and existing approaches to automating the analysis of user feedback, key challenges and promising areas of research have been identified. A clear statement of research tasks will allow you to structure further work and focus on the most relevant aspects of the problem. The main goal of this study is to develop and evaluate the effectiveness of a comprehensive system of automated analysis of user feedback based on modern methods of natural language processing, which will provide high accuracy in identifying sentiment, key aspects and topics, the ability to work with feedback in different languages and effectively integrate into the business processes of organizations of various sizes.

To achieve this goal, it is necessary to solve the following specific tasks: development of effective methods for pre-processing user feedback; research and comparative analysis of modern models of classification of reaction tones, which involves adaptation and evaluation of the effectiveness of various NLP models for analyzing the mood of reactions – from classical approaches to machine learning to modern transformer architectures and large language models; development of a method of aspect-oriented sentiment analysis for different domains, which is aimed at developing an approach that will allow not only to determine the general tone of response, but also to identify specific aspects of products or services and the corresponding tone in relation to each of them; creation of effective methods for identifying topics and key aspects in the feedback corpus, which includes the development and evaluation of methods for automatically identifying the main topics discussed in detailed reviews, and highlighting key aspects of products or services; development of an approach to multilingual feedback analysis, which focuses on creating methods that will allow you to effectively analyze reviews in different languages without significant loss of accuracy and while preserving linguistic and cultural characteristics; creation of methods for interpreting and visualizing the results of the analysis of feedback, which includes the development of approaches to ensure the clarity of the analysis results and their effective visualization for end users; development and evaluation of a comprehensive system of automated analysis of feedback, which provides for the combination of developed methods and models into a single system that will provide a full cycle of feedback analysis – from collection and pre-processing to the generation of insights and recommendations; experimental assessment of the effectiveness of the developed methods on real data, which includes conducting complex experiments to assess the effectiveness of the developed methods and the system as a whole on real sets of reviews from different domains and in different languages. To solve the tasks, a comprehensive methodological approach will be used, which combines methods of machine learning, deep learning, natural language processing and data analysis.

The study will be carried out in several stages:

- analytical stage, at which a detailed analysis of existing methods and models, their advantages and limitations will be carried out;
- the design and development stage, at which specific methods and models will be developed to solve the tasks;
- experimental stage, which includes conducting experiments in order to assess the effectiveness of the developed methods and models;
- analytical and generalizing stage, at which the results of experiments will be analyzed, patterns will be identified and conclusions will be formulated regarding the effectiveness of the developed methods and models.

Results

Automated analysis of user feedback relies on a solid theoretical foundation of natural language processing, machine learning, and artificial intelligence techniques. Pre-processing of text data is a fundamental stage in the process of analyzing user feedback and largely determines the effectiveness of post-processing. At the heart of text preprocessing is the theory of tokenization, which defines the principles of splitting text into smaller units – tokens. In the context of analyzing user feedback, tokenization faces a number of specific challenges. Various theoretical approaches have been developed to address these problems, including statistical methods, dictionary-based approaches, and, most recently, subword tokenization techniques.

Based on the analysis, the following conclusions can be drawn about the adequacy of various models and algorithms for the tasks of automated analysis of user feedback:

- for the analysis of moods, the most adequate are transformer models, additionally trained on data from a specific subject area, which provide an optimal balance between accuracy and computational efficiency;

- for aspect-oriented analysis of moods, it is recommended to use combined models that simultaneously identify aspects and determine their tonality;
- to identify topics and aspects, it is advisable to combine classical algorithms of thematic modeling with neural network approaches, using the advantages of each method;
- for multilingual analysis, it is optimal to use specialized multilingual models with additional adaptation to the specifics of domains and languages;
- to ensure scalability and high performance of the automated feedback analysis system, it is recommended to implement a multi-level architecture using models of varying complexity at different levels.

Overall, hybrid approaches that combine the strengths of different models and algorithms, providing an optimal balance between accuracy, computational efficiency, scalability, and interpretation, are the most adequate for comprehensively solving the problems of automated analysis of user feedback. In Table 3 summarizes the key methods of sentiment analysis and their theoretical foundations.

Table 3

Methods of sentiment analysis and identification of aspects

Method	Theoretical background	Key principle
Dictionary methods	Lexicography and affect theory	The tonality of a text is calculated as the sum of the tonalities of individual words
Machine Learning with	Teacher Theory of Learning from Labeled Data	Optimization of the Classification Function on Labeled Data
Deep Learning	Neural Network Theory	Automatic Learning of Hierarchical Representations of Text
ABSA	Multitasking Learning	Simultaneous Identification of Aspects and Sentiment Analysis of Them
Thematic	modeling Probabilistic modeling	Presentation of documents as a mixture of topics, and topics as distributions of words

The considered theoretical foundations provide an understanding of the principles of operation of various methods and models, their capabilities and limitations, which is a prerequisite for the effective implementation of the system of automated feedback analysis and the development of innovative approaches in this area.

Discussion

Effective research in the field of automation of user feedback analysis requires careful planning, a structured approach and the use of appropriate techniques. Methodological aspects of the organization of scientific research, starting from the collection and preparation of data and ending with the evaluation of results and their interpretation. The methodical approach to data collection and preparation involves the implementation of a number of successive stages, each of which has its own characteristics and requirements.

The methodological approach provides for the sequential implementation of the following stages:

- selection of the basic model architecture in accordance with a specific task;
- division of data into training and validation and test samples;
- the process of training and optimization of models includes;
- selection of the loss function in accordance with the set one;
- selection of an optimizer and determination of training parameters;
- regularization to prevent requalification;
- monitoring of the learning process and early stopping when reaching a plateau on a validation sample;
- search for optimal hyperparameters using grid search, random search, or Bayesian optimization.

The methodological approach to assessing the effectiveness of feedback analysis models involves the use of a set of metrics adapted to specific tasks:

- for mood classification tasks, the main metrics are: F1-score – harmonious average accuracy and memorization;
- for aspect-oriented sentiment analysis, the assessment is carried out in two stages: assessment of the quality of identifying aspects;
- assessment of the accuracy of the classification of tonality in relation to correctly identified aspects;
- for thematic modeling problems, specific metrics are used, such as: consistency of the topic;
- surprise; variety of topics;
- for multilingual model evaluation tasks, it is important to test in different languages and check the stability of the model;
- to interpret the results, it should be carried out taking into account the context of the study and the characteristics of the data;
- to ensure the reliability of the results obtained, it is recommended to conduct statistical analysis;
- analysis of the sensitivity of models to changes in data and parameters; checking the stability of the results on different subsamples of data.

The methodical approach to the organization of experiments involves adherence to a structured protocol.

Each experiment must have a clearly defined goal and hypotheses that are tested.

The documentation of the experiment should include:

– a description of the input data (source, scope, characteristics, method of pre-processing);
– detailed information about the model architecture and its parameters; description of the learning process (loss function, optimizer, learning speed, number of epochs);
– results of evaluation of validation and test samples against all relevant metrics; error analysis and interpretation of results;
– comparison with basic models and previous experiments.

Methodological support for research in the field of automation of user feedback analysis is a comprehensive process that covers all stages from data collection to implementation of results. The developed methodological recommendations provide a structured approach to research, which allows you to obtain reliable and reproducible results. For a comprehensive assessment of various methods of automated analysis of user feedback and verification of theoretical results, a comprehensive experimental design was developed, covering four key tasks: general sentiment analysis, aspect-oriented sentiment analysis, identification of themes and aspects in reviews, and multilingual analysis. For each task, appropriate datasets, estimation techniques and comparison models were selected.

The following data sets were used in the experimental study:

1. For general sentiment analysis: Amazon Product Reviews (version 2024) – a subset of 500,000 reviews for different product categories (electronics, books, clothing, household goods); Yelp Open Dataset (2023-2025) – 300,000 restaurant and service industry reviews [3, 30]; IMDb Movie Reviews – 50,000 movie reviews.

2. For aspect-oriented sentiment analysis: SemEval-2024 ABSA – dataset with marked aspects and corresponding tonality; Restaurant Reviews Dataset – a specialized set of 25,000 restaurant reviews with aspect and tone markup; Tech Products Reviews – 35,000 tech device reviews with detailed aspect markup.

3. To identify themes and aspects: Hotel Reviews Corpus – 200,000 hotel reviews for thematic modeling; AppStore Reviews – 150,000 mobile app reviews; Product Discussion Forums – 100,000 texts from discussions of various products.

4. For multilingual analysis: MultiLing Sentiment Dataset – reviews in ten different languages (English, German, French, Spanish, Italian, Portuguese, Chinese, Japanese, Arabic); Cross-lingual E-commerce Reviews – a set of parallel product reviews in different languages.

All datasets have been pre-processed according to the methodology described earlier. This included:

– cleaning up HTML tags and special characters;
– normalization of text (lowercase casting, removal of duplicate spaces); tokenization; lemmatization using spaCy;
– removing stop words for thematic modeling tasks;
– saving emoticons and emojis as special tokens for sentiment analysis; processing abbreviations and slang using specialized dictionaries.

To ensure the objectivity and statistical significance of the results, all experiments were conducted using cross-validation (5-fold), and the assessment was carried out according to a set of relevant metrics for each task. To model the problem of general sentiment analysis, several approaches have been implemented and evaluated, ranging from classical machine learning methods to modern transformer models. The task was to classify feedback into three classes: positive, negative and neutral.

The results of experiments on the classification of the tonality of responses demonstrate a clear pattern: transformer models are significantly superior to classical approaches in terms of accuracy, but require more computing resources. The domain-adapted BERT performed the highest with an average F1-score of 0.93 across all datasets, which is 17 percentage points higher than the baseline Naive Bayesian (0.76).

Additional experiments with the analysis of the impact of the volume of training data showed that transformer models are especially sensitive to the number of training examples. With limited datasets (less than 5000 samples), the difference between classical ensemble methods and transformer models is significantly reduced, which confirms the importance of having large marked enclosures to fully unleash the potential of modern approaches. Aspect-oriented sentiment analysis (ABSA) presents a more complex task compared to general classification, since it requires not only the definition of tonality, but also the identification of specific aspects to which it relates. Within the framework of the experiments, two main approaches were implemented and compared: pipeline models and joint models.

The experiments were conducted on three domains: restaurant reviews, technical product reviews, and hotel reviews. Specific aspects have been identified for each domain. Additional experiments were aimed at studying the ability of models to detect and classify tonality for different types of aspects. For each domain, 5 to 8 categories of aspects were identified, and the results were evaluated separately for each category.

The results of experiments on aspect-oriented sentiment analysis demonstrate several important patterns: combined models that simultaneously solve the problem of identifying aspects and classifying their tonality show better results compared to sequential approaches. The RACL model showed the highest performance with an average F1-score ABSA of 0.85, which is 9 percentage points higher than the baseline BiLSTM-CRF + ATAE-

LSTM sequential model (0.76); the performance of all models varies depending on the data domain, the highest scores are seen for restaurant reviews, where models reach an F1-score ABSA of up to 0.86, while for tech product reviews, the score is lower (up to 0.84).

This can be attributed to the greater complexity and diversity of aspects in technical products; analysis of the effectiveness of models for different categories of aspects revealed significant differences. All models demonstrate the highest accuracy for aspects that are often found in reviews (for example, "Food" in restaurant reviews with an F1-score of up to 0.88), instead, rare aspects or those with a high difficulty of detection, similar to the problem of general sentiment analysis, more complex models demonstrate higher accuracy due to higher computational requirements, the RACL model, which has the highest precision, also has the largest size (750 MB) and the longest inference time (35 ms per sample); impact on different stages of ABSA. It is interesting to note that the improvement in the transition from sequential to combined models is most noticeable in the combined F1 ABSA metric, this suggests that the combined models make more effective use of the relationship between the tasks of identifying aspects and determining their tone. Pooled models, especially RACLs, demonstrate a better ability to cope with these challenges through collaborative learning for interrelated tasks and the use of contextual attention mechanisms.

To assess the adequacy of the studies carried out, an integrated approach was applied, including quantitative and qualitative evaluation methods. To ensure the statistical significance of the results and avoid random patterns, a number of statistical validation methods were used. A five-fold cross-validation procedure made it possible to evaluate the performance of models on different subsets of data and minimize the impact of random distribution on training and test samples.

The standard deviation of the results between the folds for all models is in the permissible range of 1-3%, which indicates the stability of the models and their ability to generalize. To confirm the statistical significance of the difference between the results of different models, a paired Bonferroni-corrected t-test was applied for multiple comparisons [1]. The results obtained confirm that the advantage of the domain-adapted BERT and the RACL model over other approaches is statistically significant ($p < 0.01$). An additional method of checking reliability was bootstrap oversampling, which confirmed the stability of the results obtained when changing the data distribution. 95% confidence intervals for the F1 metric of transformer models are within ±1.5%, which indicates a high reliability of estimates.

An important aspect of assessing the adequacy of an experimental study is the analysis of the representativeness of the data sets used. To verify this aspect, a detailed analysis of the characteristics of the data and their correspondence to real use cases was carried out.

The analysis of the distribution of sentiment classes revealed a certain imbalance in the data, which is typical for real user reviews – positive reviews dominate (55-60%), negative reviews make up 25-30%, and neutral reviews make up only 10-15%. To compensate for this imbalance, weighted loss functions and class balancing techniques were used during model training. Analysis of the distribution of the length of reviews showed that the datasets cover a wide range – from short (5-10 words) to extended (more than 200 words) reviews, which corresponds to a variety of real-world scenarios.

The models were evaluated separately on subsamples of different lengths of feedback to ensure their effectiveness for all types of inputs. Lexical diversity and specific features of responses (slang, emoticons, technical terms) are also adequately represented in the datasets, which is confirmed by frequency response analysis and lexical diversity assessment methods (TTR–Type–Token Ratio). To assess the representativeness of the data in terms of the diversity of topics and domains, the Topic Modeling (LDA) method was used, which found that the datasets cover all key topics and aspects of products in their respective domains. To ensure the validity of measurements and the objectivity of model evaluation, an analysis of the compliance of the selected metrics with the research tasks was carried out. For aspect-oriented tonality analysis, a combined metric was used, taking into account both the accuracy of identifying aspects and the correctness of determining their tonality.

This metric better reflects the practical value of the models than individual metrics for each stage. Additionally, for a comprehensive evaluation of the models, metrics of the efficiency of the use of computing resources (training time, inference time, memory usage) were used, which are critical for practical application. An important aspect of the adequacy of experimental research is the consistency of the results obtained with theoretical predictions and data available in the literature.

The results of the experiments confirm the theoretical provisions presented earlier, in particular:

– the superiority of transformer models over classical approaches in natural language comprehension problems is confirmed experimentally (F1-score 0.93 versus 0.76);
– the effectiveness of domain adaptation to improve the results of the analysis of specific texts is confirmed by higher indicators of domain-adapted BERT compared to the standard one;
– the theoretical assumption about the superiority of combined models over sequential models for aspect-oriented sentiment analysis is confirmed experimentally (F1-score ABSA 0.85 vs. 0.76);
– the dependence of model efficiency on the volume of training data is consistent with theoretical predictions – transformer models demonstrate greater sensitivity to the volume of data compared to classical approaches.

To objectively assess the adequacy of an experimental study, it is necessary to recognize its limitations and analyze their impact on the reliability and generalizability of the results.

Table 4

Comparison of dataset characteristics with real user reviews

Characteristics	Experiment datasets	Real user reviews	Representativeness score
Class Distribution	60% / 25% / 15% (Positive/Negative/Neutral)	58% / 27% / 15%	High
Average Review Length	75 Words	68 Words	High
Lexical Diversity (TTR)	0.32	0.29	Medium-High
Share of reviews with emoticons	22%	25%	High
Share of reviews with spelling errors	35%	42%	Average
Domain Coverage	5 Major Domains	Multiple Domains	Medium-High

The main limitations of the study include:

– domain limitations – although the study covers several important domains (electronics, restaurants, movies, hotels), a number of specific areas (medical services, education, financial products) remained uncovered;
– language limitations – despite the inclusion of 10 languages in the multilingual experiment, the features of some rare languages and dialects were not taken into account;
– limitations of computing resources – for the largest transformer models (for example, RoBERTa-large), it was impossible to conduct a full cycle of hyperparameter optimization due to the limitations of available computing resources;
– focus on textual data – the study does not look at multimodal feedback that includes images or videos, which are becoming increasingly common in today's digital platforms.

To evaluate the effect of these limitations on the reliability of the results, a series of additional experiments were conducted on smaller samples with different data configurations. The results show that the identified patterns persist when the composition of the data changes, which indicates the stability and generalizability of the conclusions obtained. The adequacy of the experimental study was also evaluated in terms of the practical applicability of the results obtained to solving real business problems. To assess the practical applicability, a pilot feedback analysis system based on domain-adapted BERT and the RACL model was developed, which was tested on real data from one of the e-commerce platforms.

The system has demonstrated the ability to effectively analyze user feedback, identify problematic aspects of products, and track changes in user sentiment over time. Comparison of the automated analysis with the manual analysis carried out by a team of experts on a sample of 500 reviews showed high consistency of results (Cohen's kappa = 0.82 for sentiment classification and 0.79 for aspect detection). At the same time, the automated system processed reviews 120–150 times faster, which confirms its practical value. Based on a comprehensive analysis of statistical reliability, representativeness of data, validity of metrics and compliance with theoretical predictions, it can be concluded that the experimental study conducted is highly adequacy to solve the tasks.

The main factors confirming the adequacy of the study:

– the use of a complex set of statistical methods to ensure the reliability of the results obtained;
– the use of representative data sets reflecting the diversity of real user reviews;
– selection of appropriate evaluation metrics that take into account all aspects of the model's effectiveness;
– consistency of experimental results with theoretical forecasts and existing studies;
– successful pilot implementation, confirming the practical applicability of the developed models.

Thus, the conducted experimental study is sufficient to assess the effectiveness of various methods of automating the analysis of user feedback and develop practical recommendations for their implementation.

Conclusions

As a result of the study, all the goals and objectives were achieved, which made it possible to formulate the following conclusions. A theoretical study of natural language processing methods in the field of marketing 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 specific method depends on the specifics of the task, available computing resources, requirements for accuracy and speed of analysis.

Based on the experimental study, it was found that transformer models, in particular domain-adapted BERT, provide the highest accuracy of response sentiment analysis (F1-score up to 0.93), which is significantly superior to classical methods such as Naïve Bayes (F1-score 0.76) and XGBoost (F1-score 0.85). In the field of aspect-oriented tonal analysis, combined models have shown a significant advantage over sequential approaches, which simultaneously solve the problem of identifying aspects and classifying their tonality. The RACL model demonstrated the highest performance with an average F1 ABSA score of 0.85, which is 9 percentage points higher than the baseline BiLSTM-CRF + ATAE-LSTM sequential model (0.76). It has been established that the efficiency of all models significantly depends on the specifics of the subject area and the quality of preliminary data processing. Domain adaptation and additional training on specific data corpora greatly increase the accuracy of analysis, especially for specialized industries with their own terminology.

The study confirmed that multilingual models such as XLM-RoBERTa provide efficient analysis of responses in different languages without significant loss of accuracy compared to models trained for specific languages. This is especially important for global companies that operate in different markets and receive feedback in many languages. Statistical validation of the results using cross-validation, bootstrapping and t-tests confirmed the reliability and statistical significance of the results obtained. 95% confidence intervals for the F1 metric of transformer models are within ±1.5%, which indicates high stability of the results.

The developed pilot system of automated feedback analysis in the field of marketing of companies demonstrated high consistency with the manual analysis of experts (Cohen kappa = 0.82 for sentiment classification and 0.79 for identifying aspects) at a much higher processing speed (120-150 times), which confirms the practical applicability of the proposed solutions.

The cost-effectiveness of the implementation of automated feedback analysis systems is confirmed by a significant reduction in the time and human resources required to process feedback, as well as an increase in the quality of the insights received, which allows companies to make more informed decisions about the development of products and services.

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 different sizes and industries.

The novelty lies in the development of an original pilot system for automated analysis of user feedback, improvement and adaptation of the methodological approach to a specific task. Thus, the conducted research makes a significant contribution to the development of methods for automating the analysis of user reviews in the field of marketing companies based on natural language processing and creates the basis for further improvement of such systems.

Declaration on the use of generative artificial intelligence tools

In preparing this paper, the author(s) used Grammarly to: check grammar and spelling, paraphrasing and rewording. After using this tool/service, the author has reviewed and edited the content and is fully responsible for the content of the publication.

Author Contributions

Conceptualization, methodology, validation, formal analysis, investigation, resources, data curation, writing - original draft preparation, writing and editing, visualization, supervision, project administration – Skorin Yuriy, as a author, read and agree with the published version of the manuscript.

References

1. Ковпака Анастасія, Мосійчук Ірина, Клімова Інна. Інструменти інноваційного маркетингу в системі управління підприємством [Електронний ресурс]. Економіка. Управління. Інновації. Випуск 2(29), 2021. URL:https://www.researchgate.net/publication/356951598_INNOVATION_MARKETING_TOOLS_IN_THE_ENTERPRISE_MANAGEMENT_SYSTEM. DOI: 10.35433/ISSN2410-3748-2021-2(29)-4.

2. Артур Струнгар. Вплив штучного інтелекту на стратегії цифрового маркетингу: поточні можливості та перспективи розвитку [Електронний ресурс]. Економіка та суспільство. Випуск 62, 2024. URL:https://www.researchgate.net/publication/381795905_VPLIV_STUCNOGO_INTELEKTU_NA_STRATEGII_CIFROVOGO_MARKETINGU_POTOCNI_MOZLIVOSTI_TA_PERSPEKTIVI_ROZVITKU. DOI: 10.32782/25240072/-2024-62-160.

3. Бутенко В.М., Тоюнда А.І.. Формування маркетингової стратегії підприємства [Електронний ресурс]. Підприємництво та інновації. (24), 61-67. 2022. URL: https://doi.org/10.32782/2415-3583/24.10. DOI: https://doi.org/10.32782/2415-3583/24.10.

4. Ірина Вікторівна Іванова, Тетяна Михайлівна Боровик, Таміла Григорівна Залозна, Альона Юріївна Руденко. Використання штучного інтелекту в маркетингу. Випуск 7, №2, 32-42, 2023. URL: https://web.stanford.edu/~jurafsky/slp3/. DOI: https://doi.org/10.15276/mdt.7.2.2023.3.

5. Joseph T. Natural Language Processing (NLP) for Sentiment Analysis in social media. International Journal of Computing and Engineering. 2024. Vol. 6, № 2, 35–48. URL: https://carijournals.org/journals/index.php/IJCE/article/view/2135. DOI: https://doi.org/10.47941-/ijce.2135.

6. Ranjan R., Ahmad H. Comprehensive Study on Sentiment Analysis: From Rule-based to modern LLM based system [Електронний ресурс]. arXiv preprint. 2024. URL: https://arxiv.org/abs/2409.09989. DOI: https://doi.org/10.7717/peerj-cs.2592.

7. Zhang W., Deng Y., Liu B., Pan S., Bing L. Sentiment Analysis in the Era of Large Language Models: A Reality Check. Findings of the Association for Computational Linguistics: NAACL 2024, 881–906. URL:https://www.researchgate.net/publication/371009625_Sentiment_Analysis_in_the_Era_of_Large_Language_Models_A_Reality_Check. DOI: https://doi.org/-10.48550/arXiv.2305.15005.

8. Viktoria Adamyk, Oleksandr Ivanovskyi. Use of artificial intelligence in the marketing system: current trends and challenges. Herald of Economics, 2025, №1, 230-43, URL: https://visnykj.wunu.edu.ua/index.php/htneu/article/view/1738. DOI: https://doi.org/10.35774/-visnyk2025.01.230.

9. Hnoievyi V. H., Koren O. M. (2021). Suchasni tendentsii tsyfrovoho marketynhu ta yikh vplyv na formuvannia marketynhovoi stratehii [Modern trends in digital marketing and their influence on the formation of a marketing strategy]. Akademichnyi ohliad, 2025, № 1(54), 49–55. URL: https://www.researchgate.net/publication/353471733_MODERN_DIGITAL_MARKETING_TRENDS_AND-_THEIR_INFLUENCE_ON_THE_MARKETING_STRATEGY_FORMATION. DOI: https://doi.org/10.32342/2074-5354-2021-1-54-5.

10. Kalinichenko Andrii. The impact of artificial intelligence on company promotion strategies. Економіка та суспільство. Випуск 69, 2024. URL:https://www.researchgate.net/publication/353471733_MODERN_DIGITAL_MARKETING_TRENDS_AND_THEIR_INFLUENCE_ON_THE_MARKETING_STRATEGY_FORMATION. DOI: https://doi.org/10.32782/2524-0072/2024-69-67.

Yurii Skorin **Юрій Скорін**	Candidate of Technical Sciences, Associate Professor, Associate Professor of the Department of Information Systems Semen Kuznets Kharkiv National University of Economics https://orcid.org/0009-0004-5218-6369 e-mail: Yuriy.Skorin@hneu.net	Кандидат технічних наук, доцент, доцент кафедри інформаційних систем Харківський національний економічний університет імені Семена Кузнеця