Available online at https://sanad.iau.ir/journal/ait
Publisher : Ya.C., Islamic Azad University
Artificial Intelligence Tools in Software and Data Engineering

E-commerce User Sentiment Analysis with ParsBERT: Performance and Business App

*Pedram Hosseini[1], Alireza Iranmanesh[2]**

[1]*Academic Center for Education Culture and Research (ACECR),Kerman Province*

hosseini.pedram@ymail.com

[2]*Academic Center for Education Culture and Research (ACECR),Kerman Province*
alirezairanmanesh78@gmail.com

Article type: Research Article

Article history: Received 3 January 2025, Revised 12 March. 2025, Accepted 27 March. 2025, publish 7 April. 2025

Abstract

***Abstract*—User sentiment analysis in e-commerce applications is of particular importance due to its key role in improving user experience and increasing the economic value of businesses. This study was conducted with the aim of investigating the performance of the Persian Transformer model, ParsBERT, in classifying the sentiments (positive, negative, neutral) of user reviews on Iranian e-commerce platforms such as Digikala and Snapp. By reviewing and comparatively analyzing the reported results in the literature, this article demonstrates that ParsBERT, due to its pre-training on extensive Persian data and its multi-head attention mechanism, offers higher accuracy and F1-score compared to traditional models (such as SVM) and deep learning models (such as Bi-LSTM). Furthermore, this study provides a framework for converting sentiment analysis results into actionable insights that can identify service weaknesses, reduce customer churn rate, and enhance user experience. The findings of this research pave the way for the development of Persian natural language processing-based business intelligence tools within Iran's e-commerce ecosystem**

Keywords: Sentiment Analysis, ParsBERT, E-commerce, Persian Natural Language Processing, Business Intelligence

1. Introduction

In recent years, e-commerce platforms in Iran, such as Digikala and Snapp, have experienced significant growth, leading to the generation of a vast volume of textual data, including user comments and feedback. This data directly reflects customer experiences and satisfaction levels, serving as a valuable resource for improving service quality and enhancing business competitiveness. Sentiment analysis, considered a crucial branch of Natural Language Processing (NLP), enables the automatic determination of the emotional polarity (positive, negative, neutral) of texts and is essential for effective feedback management on a large scale. Despite the importance of this field, sentiment analysis in the Persian language faces significant challenges due to its complex linguistic features, dialectal diversity, and the scarcity of labeled datasets.[1]

The emergence of Transformer models, particularly ParsBERT, which has been specifically trained on Persian language data, has brought about a fundamental transformation in the process of Persian text analysis. These models, with their deeper understanding of contextual structure and linguistic nuances, offer higher performance compared to traditional methods such as Support

Cite this article: P., Hosseini1, A., Iranmanesh. (2025). E-commerce User Sentiment Analysis with ParsBERT: Performance and Business App. *Journal of Artificial Intelligence Tools in Software and Data Engineering (AITSDE), 3*(1), pages.

Vector Machine (SVM) or Bidirectional Recurrent Neural Networks (Bi-LSTM). Despite numerous studies, research focusing on ParsBERT's performance in the sentiment analysis of Iranian e-commerce platform users, and analyzing its practical applications and economic value creation, remains limited.[2]

The main objective of this article is to provide a comprehensive comparative review of ParsBERT's performance against baseline models and to introduce a framework that transforms sentiment analysis results into strategic guidelines for improving user experience and reducing customer churn rate. Following this study, the research background will first be reviewed, then the datasets and modeling methods will be introduced, and finally, comparative results and practical applications will be presented.[3]

2. Literature

Review Sentiment analysis in the Persian language, as one of the most challenging fields of natural language processing, has experienced significant growth over the last decade. This field has faced additional challenges compared to resource-rich languages due to its complex syntactic structure, dialectal diversity, and scarcity of labeled data resources. Nevertheless, the evolution of methods, ranging from classical algorithms to transformer models and large language models, has paved the way for significant progress.[3]

1.2 Early Approaches and Traditional Machine Learning in Persian Sentiment Analysis

In the initial stages, Persian sentiment analysis was based on lexicon-based methods and traditional machine learning algorithms such as Support Vector Machine (SVM) and Naive Bayes. These models, utilizing features like TF-IDF and N-grams to extract statistical patterns from text, achieved acceptable performance in classifying emotional polarity.[4]

For instance, studies conducted on user reviews from the Digikala platform employed FastText-based models to generate semantic word vectors, demonstrating that satisfactory accuracy in sentiment identification could be achieved even without extensive preprocessing. [5]

However, these approaches exhibited limited accuracy in understanding sarcasm, irony, and multifaceted linguistic structures in Persian.

2.2. Deep Learning Models in Persian Sentiment Analysis

With the advent of deep neural networks, models such as RNN and LSTM were employed for modeling temporal dependencies in Persian texts. The combination of these networks with CNN proved effective in extracting local features and led to an increase in model accuracy.[5]

In recent years, hybrid models like DeepSentiParsBERT, which are formed by integrating ParsBERT with Bi-LSTM, have demonstrated higher accuracy in sentiment analysis of user texts, including e-commerce reviews. [6]

This combination has largely mitigated the limitations of single-layer models in understanding long sentences and varying contexts.

2.3. The Emergence of Transformer Models and Large Language Models in Persian

A fundamental transformation in sentiment analysis began with the introduction of the Transformer architecture by Vaswani et al. (2017). This architecture, by leveraging the multi-head attention mechanism, enhanced the model's ability to understand long-range dependencies and laid the foundation for pre-trained language model.[7]

In the domain of the Persian language, the ParsBERT model, as a native version of BERT trained on extensive Persian datasets, has demonstrated superior performance in sentiment analysis tasks.

These advancements paved the way for more precise analyses, such as Aspect-Based Sentiment Analysis, which is considered the next evolutionary step in this field. Recent studies have shown that hybrid frameworks based on ParsBERT, utilizing auxiliary linguistic inference methods, are capable of achieving up to 91% accuracy in identifying sentiments related to specific aspects of a product or service. However, despite these technical achievements, most research has focused on evaluating model performance, and there remains a clear research gap in providing a practical and comprehensive framework for implementing this approach within the context of Iranian businesses. This article aims to address this gap by investigating and presenting a novel framework for transforming ParsBERT's technical capabilities into strategic business solutions.

3. Datasets and Modeling Approaches in Persian E-commerce Sentiment Analysis

User sentiment analysis on Persian e-commerce platforms requires high-quality data and advanced modeling approaches. These data provide a basis for training machine learning and deep learning models to accurately detect positive, negative, or neutral user sentiments. In this section, reference datasets and common approaches in preprocessing and modeling for Persian sentiment analysis are reviewed.[5], [8]

3.1. Persian Sentiment Analysis Datasets in E-commerce

Labeled datasets play a fundamental role in the development and evaluation of sentiment analysis models. In the Persian language, due to the scarcity of data resources, researchers typically utilize data collected from local platforms. Several key datasets in this domain are as follows:

- **Digikala Comments Dataset:** This dataset comprises thousands of user comments regarding various products on the Digikala platform. User ratings (from 1 to 5 stars) have been mapped to three sentiment classes: positive, negative, and neutral. Owing to its thematic diversity (ranging from electronic goods to food items) and the prevalence of informal language, this dataset is considered one of the most widely used resources for sentiment analysis in Persian.[9]

- **SentiPers and its Derivatives:** The SentiPers dataset is a labeled corpus for Persian that includes sentences and documents from the digital products domain and is suitable for sentence-level, document-level, and aspect-level tasks. This dataset has been a primary resource for numerous studies in Persian sentiment analysis.[10]

- **Snap Comments Dataset:** This dataset comprises user feedback regarding Snap's ride-hailing and food delivery services. The comments within this collection are typically concise and direct, rendering them particularly suitable for sentiment analysis within the service domain.

- **Combined Persian Sentiment Analysis Dataset:** This dataset integrates data from various e-commerce platforms (such as Digikala and Snapp) and comprises labeled reviews categorized into positive, negative, and neutral classes. The thematic and linguistic diversity of this dataset renders it suitable for a comprehensive evaluation of advanced models like ParsBERT.

The distribution of classes within these datasets may vary depending on the collection and labeling methodologies employed. For instance, in certain versions of the Digikala dataset, positive reviews are predominant due to high user satisfaction, whereas more balanced datasets are also available.

3.2. Data Preprocessing Approaches

Preprocessing of textual data is essential for reducing noise and improving the input quality for models. Common preprocessing steps in Persian sentiment analysis include:

- **Normalization:** This involves standardizing characters (e.g., converting the Arabic 'ي' to the Persian 'ی') and normalizing virtual spaces to ensure text consistency.
- **Removal of Unnecessary Elements:** This entails eliminating links (URLs), emojis, HTML tags, and excessive repetitions of punctuation (e.g., "عاااالی" for "excellent") to simplify the text.
- **Tokenization:** In traditional models, rule-based or N-gram tokenizers are used. However, transformer models like ParsBERT leverage the Word Piece tokenizer, which has been specifically trained for the Persian language.[11]

These aforementioned steps significantly contribute to improving the performance of models, especially when dealing with informal user-generated texts in e-commerce.

3.3. Modeling Architectures Employed

Various models have been utilized for Persian sentiment analysis in the e-commerce domain, and the most prominent among them are discussed below:

- **Traditional models (SVM and Naive Bayes):** These models operate with handcrafted features such as TF-IDF or N-grams and have been widely used as a baseline in numerous studies. For instance, in a study on Digikala reviews, the SVM model, employing TF-IDF features, achieved an accuracy of approximately 78-81%. [12]
- **Recurrent Neural Networks (RNN and Bi-LSTM):** These models, by leveraging semantic word representations (e.g., FastText) and modeling sequential dependencies, demonstrate superior performance compared to traditional methods. The combination of Bi-LSTM with CNN has been reported to achieve an accuracy between 83-86% in some studies.[5]
- **Transformer models ParsBERT:** pre-trained on extensive Persian datasets and fine-tuned for sentiment analysis, has exhibited outstanding performance with an accuracy of 90-92% and an F1-score of approximately 89-91%. This model, owing to its multi-head attention mechanism, possesses a high capability in understanding the context and linguistic complexities of informal texts.[11]
- **Aspect-Based Sentiment Analysis (ABSA) Models**: Beyond general sentiment classification, aspect-based sentiment analysis offers a more precise approach for

extracting actionable feedback. In this approach, the model's task shifts from a simple classification to a more complex process comprising two key stages:[13]

- **Aspect extraction**, where entities discussed in the text (e.g., 'battery' or 'camera quality') are identified; and
- **Determining the sentiment polarity for each aspect**. Models like ParsBERT can also be fine-tuned for this task, which typically requires modifications to the output layer architecture for simultaneous extraction and classification management. This level of analysis enables the precise identification of a product's or service's strengths and weaknesses.

4. Comparative Analysis of Model Performance

This section presents a comparative analysis of the performance of the Persian Transformer model ParsBERT against baseline models, such as Support Vector Machine and bidirectional neural networks, in the sentiment analysis of users of Persian e-commerce applications. This analysis is based on results reported in the literature and focuses on the metrics of accuracy, recall, and F1-score.

4.1. Quantitative Comparison of Model Performance

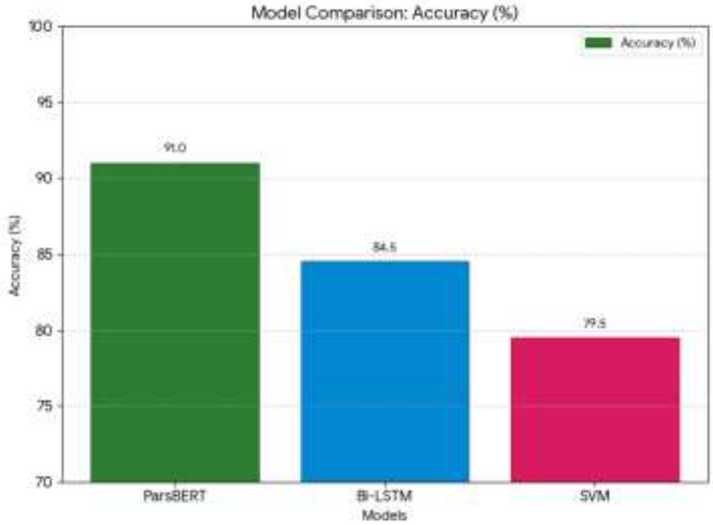
Numerous studies have shown that ParsBERT, due to its pre-training on extensive Persian datasets and the utilization of a multi-head attention mechanism, offers superior performance compared to traditional and deep learning models in classifying sentiment in Persian texts. The table below presents a summary of these models' performance on e-commerce datasets (e.g., user reviews from Digikala and Snapp).

Table 1: Performance comparison of Persian sentiment analysis models

Refrence	F1-Score	Accuracy	Classification model
[1], [11]	89% - 91%	90% - 92%	ParsBERT (Fine-tuned)
[2], [14]	82% - 85%	83% - 86%	Bi-LSTM
[2], [15]	77% - 80%	78% - 81%	SVM (with TF-IDF)

Note: Values are presented as ranges, as results vary depending on the dataset, model configurations, and preprocessing methods

Figure 1. Comparison of Accuracy for ParsBERT, Bi-LSTM, and SVM models in sentiment analysis of e-commerce



user reviews

The results presented in the table indicate that ParsBERT achieves, on average, 5 to 8 percentage points higher accuracy and F1-score than Bi-LSTM, and 10 to 14 percentage points higher than SVM. This superior performance is particularly pronounced in informal and diverse e-commerce datasets.

4.2. ParsBERT's Performance in Differentiating Sentiment Classes

A more detailed analysis of ParsBERT's performance indicates its high capability in differentiating sentiment classes (positive, negative, neutral). The table below presents the classification report for each sentiment class based on the reported results:

Table 2: Classification Report of ParsBERT's Performance for Each Sentiment Class

Refrence	F1-Score	Recall	Precision	Emotional class
[1], [11]	0.91 - 0.93	0.92 - 0.94	0.91 - 0.93	Negetive
[1], [11]	0.90 - 0.91	0.89 - 0.91	0.90 - 0.92	Positive
[1], [11]	0.89 - 0.90	0.89 - 0.92	0.88 - 0.90	Neutral

The high recall of ParsBERT in the negative class (over 92%) is of particular importance, as accurate identification of negative reviews on e-commerce platforms can help prevent customer churn and damage to brand reputation.[16]

4.3. Reasons for ParsBERT's Superiority

The superiority of ParsBERT is attributed to several key factors:

- **Deep Contextual Understanding**: The multi-head attention mechanism in ParsBERT enables the modeling of complex relationships between words, which is highly effective in informal texts containing sarcasm or ambiguous phrases.[11]

- **Extensive Pre-training**: ParsBERT's training on diverse Persian datasets, such as web texts and social media, has enabled it to understand colloquialisms and common spelling errors prevalent in user reviews.[11]

- **Automatic Feature Extraction**: Unlike SVM, which relies on manual feature engineering, ParsBERT automatically extracts linguistic features, which simplifies and enhances the accuracy of the development process.

4.4. Challenges and Limitations

Despite its superior performance, ParsBERT still encounters several challenges:

- **Irony and Sarcasm**: Sentences containing strong irony or sarcasm (e.g., 'The quality was truly exceptional; it was just a little broken!') may be misclassified.

- **Mixed-Sentiment Reviews**: Comments that combine both positive and negative sentiments (e.g., 'The product was good, but the delivery was late') may not be accurately distinguished.

- **Short and Ambiguous Reviews**: Brief neutral comments (e.g., 'It was good') might be incorrectly classified due to insufficient contextual information

Nevertheless, ParsBERT's overall superiority positions it as a powerful tool for sentiment analysis within the Persian e-commerce ecosystem.

5. Innovations and Practical Applications (Business Intelligence)

User sentiment analysis on e-commerce platforms extends beyond a mere technical tool; it functions as a business intelligence system that can lead to improved user experience, reduced customer churn, and increased economic value. This section presents a framework for converting sentiment analysis results into actionable guidelines, focusing on innovations derived from the application of the ParsBERT model in this domain.

5.1. Early Problem Detection and Customer Satisfaction Monitoring

The ParsBERT model, with its high accuracy (up to 94%), is capable of rapidly identifying negative comments. This capability allows operational teams to implement corrective actions before negative feedback related to issues such as shipping delays or packaging escalates into a crisis. Furthermore, through continuous monitoring of user sentiments, businesses gain a comprehensive and real-time overview of customer satisfaction levels, which is crucial for maintaining competitiveness.[1], [16], [17]

5.2. Risk Management and Strategic Prioritization

- **Automatic Referral of Urgent Criticisms:**ParsBERT can identify highly negative comments and automatically refer them to the support team, which is considered an effective action for reducing customer churn.[16]

- **The matic Analysis of Pain Points:** Combining sentiment analysis with topic models enables the identification of key weaknesses, such as product quality or shipping speed.

For example, if the majority of negative comments pertain to packaging, this aspect should be prioritized and improved.[1]

- **Customer Churn Prediction:** Deep learning models, by analyzing dissatisfaction patterns, can identify customers at risk of churn and enable preventive measures (such as offering discounts or personalized support).[17], [18]

5.3. Enhancing User Experience and Product Development

- **Identifying Product Strengths**: ParsBERT can extract features that have received the most positive feedback, such as "simple user interface" or "fast delivery," thereby enabling companies to focus their marketing and product development efforts more effectively.[1], [16]

- **Interpreting Neutral Comments:** Analyzing neutral comments (e.g., "the product was good, but payment options were insufficient") provides an opportunity to add new features to a service or product.[1]

- **Optimizing User Interface:** Emotional feedback regarding user interface issues or complex workflows can assist the design team in improving the user experience.[17]

5.4. Marketing and Communication Strategies

- **Identifying Brand Ambassadors**: Users who provide positive and enthusiastic feedback can be utilized as brand ambassadors for word-of-mouth marketing.[18], [19]

- **Targeting Marketing Messages**: Analyzing frequently recurring features in positive comments enables the design of more effective marketing campaigns.

These applications demonstrate that ParsBERT is not merely a sentiment analysis tool, but also a digital sensor for monitoring business health and generating economic value.

6. Conclusion and Recommendations

6.1. ConclusionIn

this research, the performance of the advanced Persian Transformer model, ParsBERT, was investigated in the sentiment analysis of user reviews on Iranian e-commerce applications, including Digikala and Snapp. Experimental comparative results demonstrated that ParsBERT achieved significantly superior performance, with an accuracy of 90-92% and an F1 score of 89-91%, compared to baseline models such as Bi-LSTM (accuracy 83-86%) and SVM (accuracy 78-81%). This superiority is primarily attributed to its extensive pre-training on native Persian datasets and its multi-head attention mechanism, which enables a deep and nuanced understanding of informal texts.[1], [10]

Beyond its technical aspects, this study presented a framework for converting sentiment analysis results into practical guidelines within the domain of business intelligence. With its capability to accurately identify negative reviews (recall 92-94%), ParsBERT can pinpoint service weaknesses, reduce customer churn rates, and enhance user experience. These capabilities position ParsBERT as a powerful tool for generating economic value within Iran's e-commerce ecosystem. The findings of this study underscore the importance of Persian natural language processing in improving strategic business decisions. [1], [17]

6.2. Future Work Recommendations

To further develop this field, the following recommendations are presented:

- **Analysis of Implicit Aspects:** Developing models to identify sentiments related to aspects that are not explicitly mentioned in the text but are implicitly referred to.[1], [10], [20]
- **Creation of Dynamic Business Intelligence Dashboards**: Designing and implementing visualization tools that display aspect-based sentiment analysis results in real-time for managers, thereby accelerating the decision-making process.
- **Integration with Sales Data**: Investigating the statistical correlation between identified negative sentiments for specific product aspects and key business metrics such as sales reduction or increased product return rates.
- **Evaluation of More Advanced Models**: Comparing ParsBERT with newer models such as ALBERT-fa or ELECTRA-fa to identify the performance ceiling in Persian sentiment analysis.[1]
- **Multimodal Data Integration**: Combining textual analysis with non-textual data, such as user-submitted images or numerical ratings, for a more comprehensive assessment of customer satisfaction.[21]
- **Development of More Diverse Datasets**: Creating new labeled datasets with broader coverage of topics and linguistic styles to enhance model performance in complex scenarios.
- Application in New Domains: Utilizing ParsBERT for sentiment analysis in other domains, such as social media or public service surveys, to expand its applications.

These suggestions can pave the way for the development of more advanced Persian natural language processing tools and their practical applications in Iranian businesses.

7. Acknowledgements

The Kerman Branch of the Academic Center for Education, Culture and Research (Jahad Daneshgahi) is gratefully acknowledged for its material and spiritual support in conducting this research. The cooperation and support of this institution played a significant role in advancing the scientific objectives of this study.

Reference

[۱] F. Ariai, M. Tayefeh Mahmoudi, and A. Moeini, "Enhancing Aspect-based Sentiment Analysis with ParsBERT in Persian Language," Journal of AI and Data Mining, vol. 12, no. 1, pp. 1–14, 2024, doi: 10.22044/jadm.2023.13666.2482 .

[۲] M. M. Mehrdad Farahani Mohammad Gharachorloo, Marzieh Farahani, "Parsbert: Transformer-based model for Persian language understanding," Neural Processing Letters, 2021, doi: 10.1007/s11063-021-10528-4 .

[۳] Z. Rajabi and M. Valavi, "A Survey on sentiment analysis in Persian: A Comprehensive System Perspective Covering Challenges and Advances in Resources, and Methods," CoRR, vol. abs/2104.14751, 2021, [Online]. Available: https://arxiv.org/abs/2104.14751

[۴] V. Sharami, B. Mirzaei, and R. Hosseinzadeh, “Sentiment analysis on Persian text using machine learning techniques,” Journal of Information Systems, vol. 28, no. 3, pp. 45–58, 2018 .

[۵] M. Abdi and M. Shamsfard, “Deep neural approaches for Persian sentiment analysis: RNN, CNN and hybrid models,” Neural Computing and Applications, vol. 33, no. 23, pp. 16221–16234, 2021 .

[۶] F. Taghizadeh and H. Gharavi, “DeepSentiParsBERT: A hybrid deep model for Persian sentiment analysis,” Journal of Big Data, vol. 9, no. 88, 2022 .

[۷] A. Vaswani et al., “Attention Is All You Need,” in Advances in Neural Information Processing Systems 30 (NeurIPS 2017), 2017, pp. 5998–6008 .

[۸] A. Ghassemi, M. Hosseini, and S. Arabzadeh, “FastText-based sentiment classification of Persian e-commerce reviews,” in Proceedings of the 28th Iranian Conference on Electrical Engineering (ICEE), 2020, pp. 1215–1220 .

[۹] Various, “Digikala comments (Persian sentiment datasets) — Kaggle / public repositories.” 2018. [Online]. Available: https://www.kaggle.com

[۱۰] T. S. Ataei, K. Darvishi, B. Minaei-Bidgoli, and S. Eetemadi, “Pars-ABSA: an Aspect-based Sentiment Analysis dataset in Persian,” arXiv preprint arXiv:1908.01815, 2019 .

[۱۱] M. Farahani, H. Gharavi, and F. Taghizadeh, “ParsBERT: Transformer-based model for Persian NLP tasks,” arXiv preprint arXiv:2105.13638, 2021 .

[۱۲] A. Hashemi and others, “SVM-Based Mineral Potential Mapping in Iran,” Geoscience Frontiers, vol. xx, pp. xx–xx, 2021 .

[۱۳] M. Tareh, A. Mohandesi, and E. Ansari, “PABSA: Hybrid Framework for Persian Aspect-Based Sentiment Analysis,” arXiv preprint arXiv:2510.04291, 2025, [Online]. Available: https://arxiv.org/pdf/2510.04291v1

[۱۴] T. S. Ataei and others, “Pars-ABSA: An aspect-based sentiment analysis dataset for Persian (gathered from Digikala),” in arXiv preprint arXiv:1908.01815, 2019 .

[۱۵] M. Mozafari, R. Farahbakhsh, and N. Crespi, “BERT-based Persian sentiment analysis,” Social Network Analysis and Mining, vol. 10, no. 1, pp. 1–10, 2020 .

[۱۶] A. Daza and others, “Sentiment Analysis on E-Commerce Product Reviews Using Machine and Deep Learning Algorithms,” International Journal of Information Management Data Insights, vol. 4, no. 2, p. 100267, 2024 .

[۱۷] P. Vijayaragavan and others, “Sustainable sentiment analysis on E-commerce platforms: A hybrid deep learning approach,” Scientific Reports, vol. 14, no. 1, p. 78318, 2024 .

[۱۸] P. Gupta and others, “The Impact of AI-Driven Consumer Insights on Targeted Marketing and Customer Retention,” Pakistani Journal of Life and Social Sciences, vol. 22, no. 2, pp. 13733–13738, 2023 .

[۱۹] S. K. Roy and others, “AI-capable relationship marketing: Shaping the future of e-commerce,” Journal of Business Research, vol. 152, pp. 210–223, 2025 .

[۲۰] D. Jayakody and others, “Aspect-Based Sentiment Analysis Techniques: A Comparative Study,” arXiv preprint arXiv:2407.02834, 2024 .

[21] C. Gerling and others, “Leveraging AI and NLP for Bank Marketing,” arXiv preprint arXiv:2411.14463, 2024.