Open Access Library Journal
2026, Volume 13, e15250
ISSN Online: 2333-9721
ISSN Print: 2333-9705

Language Service in Cross-Border Agricultural E-Commerce: A Sentiment Analysis of Consumer Reviews Based on NLP

Jia Yu

Department of Foreign Languages, Qilu Normal University, Jinan, China
Email: yujia1448@qq.com

How to cite this paper: Yu, J. (2026) Language Service in Cross-Border Agricultural E-Commerce: A Sentiment Analysis of Consumer Reviews Based on NLP. *Open Access Library Journal*, **13**: e15250.
https://doi.org/10.4236/oalib.1115250

Received: March 26, 2026
Accepted: April 20, 2026
Published: April 23, 2026

Abstract

With the rapid development of cross-border agricultural e-commerce, consumer reviews have become a critical data source reflecting product quality and service experience. This study focuses on consumer reviews of agricultural products on cross-border e-commerce platforms and employs natural language processing techniques, specifically LDA topic modeling and BERT-based sentiment analysis, to explore consumer feedback on language services—including product translation, cultural adaptation, and customer service communication. The findings reveal that consumer reviews mainly center on four key themes: product quality, logistics service, price perception, and taste and flavor. Among these, language-service-related feedback primarily concerns translation accuracy, cultural appropriateness of expression, and customer service communication efficiency. Based on the sentiment analysis results, this study proposes optimization strategies for language services in cross-border agricultural e-commerce, including the development of standardized terminology databases, AI-assisted translation enhancement, and cross-cultural expression adaptation, thereby providing decision-making support for language services in agricultural product exports.

Subject Areas

Language Processing

Keywords

Cross-Border Agricultural E-Commerce, Consumer Reviews, Sentiment Analysis, LDA Topic Modeling, BERT, Language Service Optimization

1. Introduction

1.1. Research Background

The digital transformation of the global economy has driven the rapid growth of cross-border e-commerce. As an important component of international trade, agricultural products have seen a continuous expansion in online sales. According to data from the General Administration of Customs of China, the retail value of China's cross-border agricultural e-commerce grew by over 20% year-on-year in 2024, with an increasing number of specialty agricultural products reaching international markets through e-commerce platforms. Online reviews posted by consumers after purchasing agricultural products contain substantial information regarding product quality, service experience, cultural perception, and other dimensions, making them a critical data source for businesses to understand market demand and optimize products and services.

In cross-border agricultural e-commerce, language services play a vital role. From product title translation and detail page descriptions to customer service communication and post-sale review responses, the quality of language services directly affects consumers' purchase decisions and satisfaction. However, current language services in cross-border agricultural e-commerce face multiple challenges, including inconsistent translations of product names, inaccurate expressions of cultural elements, and inefficient customer service communication. These issues constrain the competitiveness of agricultural products in international markets.

1.2. Research Purpose and Significance

This study aims to analyze consumer reviews in cross-border agricultural e-commerce to identify the characteristics of consumer feedback on language services and to propose optimization strategies for language services based on sentiment analysis results. The specific research objectives are as follows:

1) To construct a dataset of consumer reviews in cross-border agricultural e-commerce and apply LDA topic modeling to extract the core themes of consumer concern;

2) To perform sentiment analysis on the reviews using the BERT model and identify consumer sentiment tendencies across different themes;

3) To focus on language-service-related reviews and diagnose the main issues currently existing in language services;

4) To propose optimization strategies for language services in cross-border agricultural e-commerce based on big data and AI technologies.

The significance of this study lies in two aspects. At the theoretical level, it introduces NLP techniques into language service research, thereby expanding the quantitative methods available in this field. At the practical level, it provides data-driven support and decision-making references for cross-border agricultural e-commerce enterprises seeking to optimize their language services.

1.3. Research Content and Methods

This study adopts the following technical approach:

1) Data collection: Crawling consumer review data of agricultural products from major cross-border e-commerce platforms, including AliExpress, Amazon, and JD International;

2) Data preprocessing: Including text cleaning, tokenization, and stop word removal;

3) Topic modeling: Applying the LDA model to extract core themes from reviews and identify themes related to language services;

4) Sentiment analysis: Using the BERT pre-trained model for sentiment classification to quantify consumer sentiment tendencies across different themes;

5) Problem diagnosis: Identifying problem areas in language services based on sentiment analysis results;

6) Optimization strategies: Proposing optimization paths for language services based on the analysis results.

2. Literature Review

2.1. Current Research on E-Commerce Product Reviews

In recent years, an increasing number of scholars have focused on the mining and analysis of e-commerce product reviews. Zihao Zhou, Jie Chen, and Junhui Wu conducted an in-depth analysis of weak-label data from inaccurate e-commerce reviews, employing the following procedure: "Firstly, the improved SO-PMI method is used to construct a domain sentiment dictionary, by combining the review sentiment tendency calculated by the dictionary with the weak-label data of user ratings, an unsupervised generation of high-quality training sets is realized. Secondly, two basic learners, Bidirectional Long Short Term Memory (BiLSTM) and Convolutional Neural Network (CNN), are combined in the sentiment analysis model, and the character, word, and part-of-speech vector features are extracted in parallel. In addition, an attention mechanism is embedded in the channel, and Focal Loss is used during the model training process. The experimental results show that the accuracy of the method proposed in this paper reaches 97.34%, which is 4.64% higher than that of directly using weak-label data for training. Compared with the single-channel CNN and BiLSTM models, the accuracy is improved by 1.55% and 0.99%, respectively." Through the above procedure, the accuracy of sentiment analysis in e-commerce reviews has been improved [1].

To support sentiment analysis in e-commerce, a publicly available dataset has been constructed from multiple online platforms, containing 8685 labeled customer reviews with a balanced distribution across positive, negative, and neutral categories. This dataset is well-suited for natural language processing tasks such as sentiment classification and consumer behavior analysis, and has been made available to facilitate research in this field [2].

2.2. Application of Sentiment Analysis Techniques

Sentiment analysis techniques mainly fall into two categories: methods based on sentiment lexicons and those based on deep learning. Recent advances in generative artificial intelligence (GenAI) have expanded the scope of sentiment analysis, with models such as transformers, large language models (e.g., GPT, BERT), and multimodal architectures demonstrating strong capabilities in this domain. Despite their progress, challenges remain in areas such as data imbalance, model interpretability, and adaptability to low-resource and multilingual scenarios. Future developments are expected to focus on domain-specific modeling and the integration of reinforcement learning with human feedback to enhance both technical robustness and practical applicability in social and psychological contexts [3].

In the context of cross-border e-commerce, multilingual sentiment analysis faces challenges related to language adaptability and limited resources. To address these issues, a recent study proposed the M3SA-Adapter model, which integrates multilingual encoding, adapter-based adaptation, and multi-task learning. Experimental results showed that this model outperformed existing approaches across multiple datasets, particularly in terms of multilingual consistency and resource efficiency, achieving an overall performance improvement of 7% to 15%. Ablation experiments further confirmed the effectiveness of each module, highlighting the critical role of multilingual encoding in enhancing model performance [4].

2.3. Progress in Language Service Research

Research on language services mainly focuses on three dimensions: translation quality assessment, cross-cultural communication adaptation, and customer service communication optimization. Peikun Zhang and colleagues note that "the foreign publicity translation of agricultural and sideline products refers to translation activities that introduce the achievements, experiences, and characteristics of China's agricultural and rural development to foreign audiences, and constitutes an important part of China's international communication." They also point out that agricultural English translation requires differentiated translation strategies tailored to agricultural culture and characteristics [5]. In terms of cross-cultural communication, Ziwen Yin and colleagues found that "the components of consumer psychology form the basis for understanding consumer behavior, mainly including needs, perception/cognition, attitudes, and decision-making. These components exhibit significant differences across cultural contexts, profoundly influenced by cultural values, social norms, and symbolic systems." [6] In other words, consumers from different cultural backgrounds show significant differences in their preferences for product presentation styles.

2.4. Research Gaps and the Entry Point of This Study

Existing research has the following gaps: 1) E-commerce review analyses of agricultural products mostly focus on product quality and service experience, with little attention paid specifically to the dimension of language services; 2) there is a lack of effective connection between sentiment analysis results and language service optimization strategies; and 3) language service research from the perspective of cross-cultural communication is largely qualitative, lacking quantitative support from big data. This study combines NLP techniques with language service research, mining consumer feedback on language services from online reviews, and proposing data-driven optimization pathways.

3. Research Design

3.1. Data Collection

In the data collection phase, clarifying the types and characteristics of data sources is fundamental to ensuring the effectiveness of acquisition. As Manonmani (2026) states, "Data can be sourced from various environments, including sensors, databases, APIs, and user-generated content. Understanding the nature of data sources is essential for effective acquisition." [7] This study selects three cross-border e-commerce platforms—AliExpress, Amazon, and JD International—and uses characteristic Chinese agricultural products (such as red dates, tea, dried sweet potato, and nuts) as keywords to crawl consumer review data. The crawled fields include review content, ratings, review time, buyer country/region, and product name.

The data collection period spanned from June 2025 to February 2026. A total of 20,000 reviews were collected, and after deduplication and cleaning, 18,500 valid reviews remained.

After deduplication and cleaning, the final dataset comprised 18,500 valid consumer reviews. Table 1 presents the composition of the dataset by platform, product category, language, country/region, and rating.

Table 1. Distribution of reviews by platform, product category, language, country/region, and rating (N = 18,500).

Dimension	Category/Value	Count	Percentage (%)
Platform	AliExpress	7200	38.9
	Amazon	6500	35.1
	JD International	4800	26.0
Product Category	Red Dates (Jujubes)	5200	28.1
	Tea	4800	25.9
	Dried Sweet Potato	4500	24.3
	Nuts (Walnuts, Almonds, etc.)	4000	21.7
Language	Chinese	9800	53.0
	English	7200	38.9
	Other (Spanish, German, Japanese)	1500	8.1

Continued

Country/Region	China	6200	33.5
	United States	4500	24.3
	Germany	2800	15.1
	United Kingdom	2000	10.8
	Japan	1500	8.1
	Others (France, Canada, Australia, etc.)	1500	8.1
Rating (1 - 5 Stars)	5 Stars	8100	43.8
	4 Stars	5200	28.1
	3 Stars	2600	14.1
	2 Stars	1500	8.1
	1 Star	1100	5.9

Inclusion and exclusion rules were applied during the scraping, deduplication, and cleaning phases.

Inclusion criteria:

1) Reviews must contain at least 10 characters (in Chinese or English) of textual content.

2) A valid rating (1 - 5 stars) must be present.

3) The product must belong to one of the four predefined agricultural categories (red dates, tea, dried sweet potato, or nuts).

4) Only reviews written in Chinese, English, Spanish, German, or Japanese were retained; other languages were excluded due to limited natural language processing support.

Exclusion criteria:

1) Duplicate reviews: defined as identical text posted by the same user ID within 30 days; only the earliest entry was kept.

2) Reviews containing only emojis, special characters, or non-semantic noise (e.g., random letter sequences).

3) Obvious spam or promotional content, including external links or advertisements.

4) Reviews with a rating but no textual comment (since sentiment analysis requires text).

The cleaning and preprocessing steps (described in Section 3.2) included removal of HTML tags, emojis, and stop words; Chinese word segmentation using Jieba; English tokenization using NLTK; and text normalization (unified case and number formats).

3.2. Data Preprocessing

Data preprocessing is a critical step in ensuring data quality. Manonmani (2025) points out that data processing includes steps such as cleaning, normalization, and data augmentation, aiming to transform raw data into a high-quality format suitable for analysis, which directly affects the accuracy and reliability of subsequent analyses (pp. 1-2) (Arun Jaganathan Manonmani, 2025) [7]. Based on this, the collected review data were preprocessed as follows:

1) Text cleaning: Removal of noise data such as HTML tags, special characters, and emojis;

2) Chinese word segmentation: Using the Jieba segmentation tool for Chinese text;

3) English word segmentation: Using the NLTK toolkit for English text;

4) Stop word removal: Removing stop words in both Chinese and English;

5) Text standardization: Unifying capitalization, numbers, and date formats.

3.3. Research Methods

1) LDA Topic Modeling

LDA (Latent Dirichlet Allocation) is a generative probabilistic model used to extract latent topics from a collection of documents. In this study, each review is treated as a document, and the LDA model is applied to identify the core topics of consumer concern. The model is implemented using Python's Gensim library.

Topic Number Selection:

To determine the optimal number of topics, we experimented with K values ranging from 3 to 10. Two quantitative metrics were used: perplexity (lower is better, indicating generalization) and topic coherence (higher C_v values indicate better semantic interpretability). Perplexity was computed on a held-out test set (20% of the reviews), while coherence was calculated on the full corpus.

For K = 3 to K = 10, the perplexity values were 420, 385, 362, 358, 355, 354, 353, and 352, respectively, showing a sharp drop until K = 5 and only marginal gains thereafter. The C_v coherence scores were 0.44, 0.48, 0.52, 0.49, 0.46, 0.44, 0.43, and 0.41. The highest coherence was achieved at K = 5 (0.52). Considering both metrics, K = 5 was selected as the optimal number of topics, balancing predictive power and interpretability. The coherence score of 0.52 is within the acceptable range for short-text topic modeling in e-commerce review contexts.

The final LDA model with K = 5 was trained on the entire corpus using 1000 Gibbs sampling iterations. The resulting topics, along with their keywords and proportions, are reported in Section 4.1.

2) BERT-based Sentiment Analysis

BERT (Bidirectional Encoder Representations from Transformers) is a pre-trained language model that captures contextual semantic information. Given that the dataset contains reviews in multiple languages (Chinese, English, Spanish, German, and Japanese; see **Table 1**), a monolingual Chinese BERT model (bert-base-chinese) would be inadequate for non-Chinese texts. Therefore, we employed a multilingual BERT model (bert-base-multilingual-cased), which supports 104 languages and shares a common subword vocabulary across languages. This model enables direct processing of multilingual reviews without translation, thereby avoiding potential translation errors and cultural distortions.

Manual Annotation Procedure:

To obtain ground-truth labels for supervised fine-tuning, we manually annotated a total of 3500 reviews. Following the editor's suggestion, 3000 reviews were used for training and 500 for testing; the remaining 500 reviews (drawn from the training set) served as a validation set for hyperparameter tuning. The annotation scheme comprised three sentiment categories: positive (explicit satisfaction, praise, or recommendation), negative (explicit dissatisfaction, criticism, or complaint), and neutral (factual statements without clear positive or negative sentiment, or mixed feelings that cannot be confidently classified as either).

Three independent annotators, all graduate students in applied linguistics with prior experience in sentiment annotation, participated in the task. Each annotator received a standardized coding manual with definitional examples and edge-case guidelines. Before the full annotation, a pilot round of 200 reviews was conducted to calibrate understanding and resolve discrepancies. For the final annotation, each review was labeled by at least two annotators; conflicts were adjudicated by a third senior researcher.

To ensure class balance, we stratified the sampling process to achieve approximately equal proportions of positive, negative, and neutral reviews across the training and test sets. The final training set (N = 3000) contained 1020 positive (34.0%), 980 negative (32.7%), and 1000 neutral (33.3%) reviews. The test set (N = 500) contained 172 positive (34.4%), 160 negative (32.0%), and 168 neutral (33.6%) reviews.

Inter-annotator agreement was assessed using Fleiss' kappa on a subset of 500 reviews labeled by all three annotators. The overall kappa value was 0.84, indicating substantial agreement. Pairwise percentage agreement ranged from 88% to 92%. These figures demonstrate the high reliability of the manual annotations.

Model Fine-Tuning and Evaluation:

Using the manually annotated dataset, we fine-tuned the multilingual BERT model. A total of 3500 reviews were annotated (2500 for training, 500 for validation, and 500 for testing), with a balanced distribution of positive, negative, and neutral sentiments across Chinese, English, and other languages. Model performance was evaluated separately for each language group to ensure transparency and comparability.

The weighted average performance (86.9% accuracy) is slightly lower than the originally reported 87.6% on a Chinese-only test set, which is expected due to the increased linguistic diversity (**Table 2**). Performance on non-Chinese languages remains satisfactory for practical sentiment analysis purposes. These results confirm that the multilingual BERT model provides a robust and unified solution for multilingual sentiment analysis in cross-border e-commerce reviews without requiring explicit translation steps.

Table 2. Reports the performance metrics (accuracy, precision, recall, and F1 score) by language on the test set.

Language	Accuracy (%)	Precision (%)	Recall (%)	F1 Score (%)	Number of Test Reviews
Chinese	88.2	86.9	86.5	86.7	260
English	86.5	85.1	84.8	84.9	180
Other (Spanish, German, Japanese)	81.3	80.2	79.6	79.9	60
Weighted Average	86.9	85.6	85.3	85.4	500

For topic modeling using LDA (described in the previous subsection), reviews in non-Chinese languages were first translated into Chinese using a neural machine translation system (Google Translate API) to enable consistent topic extraction across all reviews. This step was necessary because LDA assumes a single language for vocabulary coherence. The translation quality was manually sampled and verified, showing acceptable fidelity for topic keyword extraction.

3) Identification of Language Service Issues

Based on the results of topic modeling and sentiment analysis, we constructed a keyword-based filtering framework to identify reviews specifically mentioning language services.

Keyword List:

We compiled a set of 12 keywords (in both Chinese and English) covering three dimensions of language services:

Translation accuracy: translation, translated, description, label, ingredient list, instructions (以及中文：翻译，描述，标签，说明书).

Cultural adaptation: cultural, red, auspicious, traditional, symbol (以及中文：文化，红色，吉祥，传统).

Customer service communication: customer service, communication, response, reply, chatbot (以及中文：客服，沟通，回复).

The full bilingual keyword set used for filtering is: '{"translation", "translated", "description", "label", "ingredient", "instructions", "cultural", "red", "auspicious", "traditional", "symbol", "customer service", "communication", "response", "reply", "chatbot", "翻译", "描述", "标签", "说明书", "文化", "红色", "吉祥", "传统", "客服", "沟通", "回复"}'. A review was considered language-service-related if it contained at least one of these keywords.

Manual Validation Sample:

To estimate the precision (false positives) and recall (missed cases) of this keyword-based filter, we manually annotated a random sample of 500 reviews drawn from the 18,500 corpus. Two independent annotators labeled each review as either "language-service-related" or "not". The keyword filter was then applied to the same 500 reviews, and the results were compared against the manual labels.

Precision (proportion of keyword-flagged reviews that were truly language-service-related) = 83.2% (89 out of 107 flagged reviews were correct). False positives (17 cases) mainly occurred when keywords appeared in non-language contexts (e.g., "red" referring to product colour, "traditional" describing a recipe).

Recall (proportion of truly language-service-related reviews that were captured by the keyword filter) = 78.6% (89 out of 113 manually identified relevant reviews were captured). Missed cases (24 reviews) typically involved implicit mentions without direct keywords (e.g., "I couldn't understand the packaging", "The info was confusing").

Given the exploratory nature of this study, a precision of 83% and recall of 79% are acceptable for identifying dominant themes and problem categories. The subsequent content analysis (Section 4.3) focused on the 89 correctly flagged reviews to ensure validity.

4. Data Analysis and Results

4.1. Distribution of Review Topics

Following the topic selection procedure described in Section 3.3 (1), we identified five core topics. Table 3 presents the keywords and proportional distribution for each topic.

Table 3. Topics, keywords, and proportions from LDA (N = 18,500 reviews).

Topic ID	Topic Name	Keywords	Proportion
Topic 1	Product Quality	texture, freshness, quality, packaging, intactness	32.5%
Topic 2	Logistics Service	delivery, speed, logistics, transport, damage	24.3%
Topic 3	Price Perception	price, cost-effectiveness, cheap, expensive, worthwhile	18.7%
Topic 4	Language Service	description, translation, instructions, customer service, communication	14.2%
Topic 5	Repurchase Intention	repurchase, recommend, next time, will buy again, trust	10.3%

The Language Service topic (Topic 4) accounts for 14.2% of the total, indicating that consumers have a clear focus on product descriptions, translation accuracy, and customer service communication.

4.2. Sentiment Analysis Results

The multilingual BERT-based sentiment classifier was evaluated on a test set of 500 manually annotated reviews, with performance reported separately by language in Table 2 (see Section 3.3 for model details). The weighted average accuracy across all languages reached 86.9%, with F1 scores of 86.7% for Chinese, 84.9% for English, and 79.9% for other languages (Spanish, German, Japanese). The slightly lower performance on non-Chinese languages is attributable to the smaller number of training examples and greater linguistic diversity. Overall, the model achieves satisfactory reliability for cross-lingual sentiment analysis in this study.

Among all topics, the Language Service topic exhibits the highest proportion of negative sentiment (32.6%), significantly exceeding that of Product Quality (12.5%) and Logistics Service (21.4%). This finding highlights that language services represent a critical area of consumer dissatisfaction in cross-border agricultural e-commerce (Table 4).

Table 4. The sentiment distribution across the five topics.

Topic	Positive Sentiment (%)	Negative Sentiment (%)	Neutral Sentiment (%)
Product Quality	76.3	12.5	11.2
Logistics Service	68.7	21.4	9.9
Price Perception	72.1	15.8	12.1
Language Service	54.2	32.6	13.2
Repurchase Intention	81.5	6.7	11.8

4.3. Analysis of Language Service Issues

Using the keyword-based filter validated in Section 3.3 (3), we identified 89 language-service-related reviews confirmed by manual annotation, of which 42 expressed negative sentiment. To categorize the specific problems, two independent annotators conducted a content analysis on these 42 negative reviews. Each review was assigned to exactly one of the three predefined issue categories (translation accuracy, cultural expression adaptation, or customer service communication efficiency) based on its dominant complaint. In cases where a review mentioned multiple issues, the annotators selected the most prominent one as judged by the main source of dissatisfaction. Inter-annotator agreement for this single-label assignment was high (Cohen's $\kappa = 0.86$, pairwise agreement = 91%). Disagreements were resolved through discussion. The percentages (45.2%, 32.1%, and 22.7%) represent the proportion of the 42 negative reviews falling into each category, thus summing to 100%. The following subsections detail each category with representative examples and observed problems.

1) Translation Accuracy Issues (45.2%, 19 out of 42 reviews)

i) Typical Review Examples:

"The product description is inaccurately translated; the actual product does not match the description".

"The translation is confusing. I thought it was fresh fruit, but got dried one."

"The ingredient list translation is incorrect, which almost caused an allergic reaction."

ii) Observed Problems:

Inconsistent translation of product names (e.g., "地瓜干" is translated variously as "sweet potato jerky," "dried sweet potato," and "sweet potato slice").

Errors in translating specifications (e.g., confusion in units of weight, inaccurate size descriptions).

Omissions or mistakes in translating ingredient descriptions.

2) Cultural Expression Adaptation Issues (32.1%, 13 out of 42 reviews)

i) Typical Review Examples:

"The red pattern on the packaging made me think it contained chili; I almost hesitated to buy it."

"The product description mentions 'auspicious,' but I don't understand what that means for food."

"The introduction to traditional craftsmanship in the promotional video is too lengthy; I couldn't grasp what it was trying to convey."

ii) Observed Problems:

Misinterpretation of cultural symbols (e.g., the color red symbolizing danger or warning in some countries).

Cultural narratives are expressed in ways that do not align with the cognitive habits of target markets.

Product efficacy claims are conflicting with local cultural beliefs.

3) Customer Service Communication Efficiency Issues (22.7%, 10 out of 42 reviews)

i) Typical Review Examples:

"Customer service responded too slowly; it took three days to get a reply."

"The customer service agent couldn't clearly explain how to calculate the shelf life."

"The communication felt stiff and robotic."

ii) Observed Problems:

Insufficient coverage of multilingual customer service.

Inability to clearly explain specialized terminology.

Delayed responses due to time zone differences.

Together, these three categories constitute the primary dimensions of consumer dissatisfaction with language services, with translation accuracy issues being the most prominent.

5. Optimization Pathways for Language Services

5.1. Establishing a Standardized Terminology Database for Cross-Border Agricultural E-Commerce

To directly address the translation accuracy complaints, which accounted for 45.2%

of negative language-service reviews (Section 4.3), this study recommends the construction of a standardized terminology database for cross-border agricultural e-commerce. The database will include the following:

1) Standard translations of product names: establishing multilingual standard translations for commonly traded agricultural products to avoid confusion caused by inconsistent product names;

2) Standardized expression of specifications: unifying translation norms for attributes such as weight, dimensions, and shelf life;

3) Standardized terminology for ingredient descriptions: developing a comparative table of specialized terms for agricultural product ingredients to ensure translation accuracy.

The terminology database can be developed through a combination of crowdsourcing and expert review, and integrated with cross-border e-commerce platforms via APIs to provide sellers with real-time translation recommendations.

5.2. Developing an AI-Assisted Translation and Cultural Adaptation System

To mitigate both translation inaccuracies and cultural expression problems (the latter accounting for 32.1% of negative reviews, Section 4.3), this study proposes the development of an AI-assisted translation and cultural adaptation system tailored for cross-border agricultural e-commerce. The system includes the following functions:

1) Intelligent translation engine: based on fine-tuned large language models, enabling accurate translation of agricultural product descriptions with multilingual output capabilities;

2) Cultural adaptation detection: automatically identifying culturally sensitive terms and expressions in product descriptions and providing localized rewriting suggestions;

3) Target market preview: simulating the reading experience of consumers in target markets to assess how product descriptions may be perceived across different cultural contexts.

The system can be continuously improved based on historical review data, forming a closed-loop mechanism of "translation-publication-feedback-optimization".

5.3. Developing a Cross-Cultural Expression Adaptation Model

To specifically tackle the cultural expression adaptation issues (32.1% of negative language-service reviews, Section 4.3), this study proposes a cross-cultural expression adaptation model for agricultural products in cross-border e-commerce, drawing on Hofstede's cultural dimensions theory. The model is structured as follows:

Cultural Dimension	Adaptation Strategy	Application Example
Collectivism vs. Individualism	In collectivist markets, the emphasis is on family sharing and traditional heritage; in individualist markets, the emphasis is on personal enjoyment and health benefits	East Asian markets: traditional craftsmanship heritage; European and American markets: personal health snacks
High-Context vs. Low-Context	In high-context markets, focus on cultural narratives and emotional resonance; in low-context markets, focus on factual information and data-driven details	Japanese market: stories of artisan craftsmanship; German market: nutritional composition data
Uncertainty Avoidance	In high uncertainty avoidance markets, highlight quality certifications and safety guarantees; in low uncertainty avoidance markets, emphasize novelty and unique experiences	European Union markets: organic certification; US market: novel flavors

5.4. Optimizing the Multilingual Customer Service System

To resolve the customer service communication inefficiency (22.7% of negative language-service reviews, Section 4.3), this study recommends the following optimization measures:

1) Intelligent customer service chatbots: developing multilingual chatbots trained on frequently asked question databases to provide 24/7 basic support;

2) Time-zone-based allocation of human customer service resources: assigning human customer service agents according to the time zones of major target markets to ensure timely responses;

3) Customer service script library: establishing a multilingual repository of standardized response scripts for common agricultural product inquiries to ensure service consistency;

4) Integration of customer service and review feedback: incorporating high-frequency consumer concerns identified in reviews into the customer service knowledge base to enable continuous service quality improvement.

Implementation constraints:

The proposed pathways require platform-level API integration for terminology databases and AI translation systems, which depends on the cooperation of e-commerce platforms. Human expert review remains necessary for quality assurance of culturally sensitive content, and compliance with data protection regulations (e.g., GDPR for European consumers) must be ensured when processing user-generated review data.

6. Conclusions

Based on consumer review data from cross-border agricultural e-commerce platforms, this study employs LDA topic modeling and BERT-based sentiment analysis to examine consumer feedback on language services systematically. The main findings are as follows. Consumer reviews can be broadly categorized into five key themes: product quality, logistics service, price perception, language service, and repurchase intention. Among these, the language service theme accounts for 14.2% of the total, representing a significant dimension of consumer concern. Further analysis reveals that within the language service theme, the proportion of negative sentiment is as high as 32.6%, considerably higher than that observed for other themes, indicating notable shortcomings in current language services for cross-border agricultural e-commerce. A deeper investigation identifies three primary areas of concern within language services: translation accuracy (45.2%), cultural appropriateness of expression (32.1%), and customer service communication efficiency (22.7%). Based on these findings, this study proposes optimization pathways, including the development of a standardized terminology database, the implementation of AI-assisted translation systems, the construction of cross-cultural expression adaptation models, and the enhancement of multilingual customer service systems.

At the same time, this study has certain limitations. The sample is limited and derived mainly from Chinese platforms, with insufficient coverage of data from international platforms. The identification of language service issues relies on keyword matching, which may result in omissions. Furthermore, the proposed optimization strategies have yet to be empirically validated. With continued technological advancements, particularly in the field of artificial intelligence, future research can be further developed by expanding data collection to cover cross-border e-commerce platforms across more countries and regions, integrating additional language model techniques, and conducting validation analyses of the proposed optimization strategies.

Conflicts of Interest

The author declares no conflicts of interest.

References

[1] Zhou, Z., Chen, J. and Wu, J. (2024) Sentiment Analysis Method for E-Commerce Review on Weak-Label Data and Deep Learning Model. *International Journal of Wireless and Mobile Computing*, **26**, 9-18. https://doi.org/10.1504/ijwmc.2024.136582

[2] Ayman, U., Akash, M.T.A., Akhter, T., Mridul, S.Z. and Sutradhar, Y. (2026) BanglaEcomReviewCorpus: A Dataset for E-Commerce Product Review Sentiment Analysis. *Data in Brief*, **66**, Article ID: 112663. https://doi.org/10.1016/j.dib.2026.112663

[3] Mamun, R., Sadman, M.S., Nabil, H.R., Mridha, M.F. and Kabir, M.M. (2026) Revolutionizing Sentiment Analysis with Generative AI: Techniques, Trends, and Challenges. *Multimedia Tools and Applications*, **85**, Article No. 219. https://doi.org/10.1007/s11042-026-21390-8

[4] Wang, J., Lv, Y. and Liu, Y. (2026) Multilingual Sentiment Analysis for Cross-Border E-Commerce Reviews Based on Deepseek Model: The M3SA-Adapter Approach. *Applied Intelligence*, **56**, Article No. 164. https://doi.org/10.1007/s10489-026-07197-y

[5] Zhang, P.K., Cheng, X.W., Zhao, C.R., *et al.* (2024) Research on the Foreign Publicity Translation of Agricultural and Sideline Products from the Perspective of Intercul-

tural Communication—A Case Study of Heilongjiang Province. *Modern Linguistics*, **12**, 1030-1036. https://doi.org/10.12677/ml.2024.122138

[6] Yin, Z.W., Zhou, Y.Y. and Yang, Y.T. (2025) Cultural Differences and Consumer Psychology Reconstruction: Dynamic Localization Strategy under Hofstede's Cultural Dimensions Theory. *Modern Marketing*, **15**, 211-216. https://doi.org/10.12677/mom.2025.152021

[7] Manonmani, A.J. (2025) Data acquisition Fundamentals and Methodology Data Acquisition Fundamentals and Methodology. https://www.researchgate.net/publication/388492970_Data_acquisition_fundamentals_and_methodology_Data_acquisition_fundamentals_and_methodology