https://doi.org/10.31449/inf.v49i28.8232 Informatica **49** (2025) 211–226 **211**

BERT-based Consumer Sentiment Analysis for Personalized Marketing Strategies

Min Li[1], Shujuan Guo[2*], Runchen Liu[3]
[1]Department of International Business, Wanbo Institute of Science and Technology, Hefei 230031, China
[2]School of Business, Anhui Xinhua University, Hefei 230088, China
[3]School of Finance and Accounting, Anhui Xinhua University, Hefei 230000, China
E-mail: gsj202305@163.com
[*]Corresponding Author

Keywords: consumer sentiment analysis, personalized marketing, natural language processing, deep learning, bidirectional encoder representations from transformers

This study aims to construct an efficient consumer sentiment analysis model and apply it to optimize personalized marketing strategies for enhancing user conversion rates and satisfaction. The yf_amazon dataset and Chinese sentiment classification corpus are selected as foundational data by integrating natural language processing and deep learning technologies. Preprocessing procedures including tokenization, stop word removal, noise filtering, and sentiment annotation, are employed to clean and structure consumer reviews. A sentiment classification model based on Bidirectional Encoder Representations from Transformers (BERT) and Transformer architectures is developed and fine-tuned with parameters including a learning rate of 2e-5 and 4 training epochs. Evaluation metrics encompass accuracy, precision, recall, and F1 score. Experimental results demonstrate that the proposed model achieves 92.5% accuracy and a 0.90 F1 score in sentiment classification tasks. This model outperforms conventional Convolutional Neural Networks (83.6% accuracy, 0.83 F1 score) and Long Short-Term Memory models (86.2% accuracy, 0.85 F1 score). Further segmentation of user sentiment clusters based on analysis results enables the implementation of differentiated recommendation strategies and time-limited marketing campaigns. Ultimately, it increases user conversion rates from 25.0% to 32.7% (a 7.7 percentage point improvement) while significantly enhancing consumer satisfaction with personalized recommendations. These findings indicate that Transformer-based sentiment analysis models hold substantial practical value for advancing personalized marketing applications.

Povzetek: Študija združi BERT/Transformer za analizo sentimenta ter oblikuje personalizirane marketinške strategije. Uvaja dinamično povezovanje segmentacije čustev s priporočili in povratnimi zankami, omogoča bolj kvalitetno ciljanje kot klasični SVM/CNN pristopi.

1 Introduction

The patterns of interaction between consumers and enterprises increasingly rely on online platforms, making a wealth of feedback, reviews, and social media data vital resources for studying consumer behavior [1-3]. Sentiment analysis, as a significant branch of Natural Language Processing (NLP), has become an essential tool for enterprises to understand customer sentiment, predict customer behavior, and formulate marketing strategies [4]. Concurrently, the rise of deep learning (DL) technology has brought unprecedented accuracy and efficiency to sentiment analysis [5, 6]. Traditional sentiment analysis methods, such as support vector machine (SVM) and naive Bayes (NB), have achieved some success in some tasks. However, they face significant limitations in dealing with large-scale and unstructured data. These methods rely on the characteristics of manual design, and their ability to capture complex emotional patterns and contextual relationships in text data is limited, resulting in unsatisfactory results in multi-domain and complex datasets. DL methods, represented by Deep Neural Networks (DNNs), Convolutional Neural Networks (CNNs), Recurrent Neural Networks (RNNs), and the latest Transformer models, can automatically extract features from vast amounts of textual data. These greatly enhance the precision and generalization ability of sentiment analysis [7-9].

The core of personalized marketing lies in tailoring marketing content and interaction methods suitable for different consumers based on their behavior and emotional characteristics, to maximize consumer satisfaction and enterprise benefits [10]. By integrating the sentiment analysis results, enterprises can segment markets and optimize marketing decisions, making them more precise and efficient. Despite the remarkable progress made by NLP and DL in sentiment analysis and personalized marketing, there are still challenges to be addressed [11, 12]. Firstly, large-scale sentiment data analysis poses higher demands on algorithms' computational efficiency and accuracy. Secondly, organically combining consumer sentiment with personalized marketing strategies remains a significant challenge in practice [13-15]. In response to these issues, this study fuses NLP with DL technologies to construct an efficient consumer sentiment analysis framework and propose a design scheme for personalized marketing strategies.

This study explores the application effect of sentiment analysis models based on Bidirectional Encoder Representations from Transformers (BERT) in improving customer conversion rate. Especially under the strategy of consumer sentiment analysis, optimizing personalized recommendations and marketing activities is critical. To better answer this question, the following hypotheses are put forward. Hypothesis 1: The BERT-based sentiment analysis model demonstrates significant potential for enhancing customer conversion rates. This hypothesis stems from BERT's superior capability in capturing emotional cues and comprehending lengthy texts, enabling more precise identification of consumer sentiment to drive higher conversion performance. Hypothesis 2: Comparative analysis reveals that the BERT-based sentiment analysis model achieves substantially better performance metrics—including accuracy, recall, and F1 score—than conventional models (e.g., LSTM and SVM) in sentiment classification tasks. This advantage primarily originates from BERT's architectural strengths in processing long texts and complex emotional expressions, yielding superior precision and effectiveness in sentiment analysis applications.

The innovations are mainly reflected in the following three aspects:

1) Integration of advanced technologies: A new consumer sentiment analysis model is designed using NLP and DL, leveraging BERT's semantic understanding capabilities and improving the accuracy and efficiency of sentiment classification through custom optimization strategies.

2) Personalized marketing strategies: A systematic set of personalized marketing strategies is proposed based on multidimensional sentiment analysis results, which are empirically validated to effectively increase customer conversion rates and satisfaction.

3) Dynamic feedback mechanism: A consumer feedback-based dynamic adjustment mechanism is designed to optimize marketing strategies in real-time. This enables timely adjustments to personalized recommendations based on changes in consumer behavior, thereby enhancing customer experience.

2 Literature review

As one of the pivotal applications of NLP, sentiment analysis has become a hot spot in consumer behavior research [16]. Hung & Alias (2023) [17] posited that sentiment analysis involved uncovering emotional tendencies within the text to comprehend consumers' attitudes, emotions, or opinions. Liu (2024) [18] explored the role of personalization in modern digital marketing and its impact on consumer engagement. This approach was particularly applicable to unstructured data such as social media comments and product reviews, assisting enterprises in identifying consumer satisfaction and needs. However, traditional sentiment analysis models, which depended on manual feature extraction, failed to fully harness the potential of large-scale data and exhibited significant limitations when handling complex linguistic structures.

With the rapid development of NLP technology, researchers have gradually introduced more complex models to process consumer sentiment data. As social media and other data sources increase, massive text data has increased sharply. Meanwhile, some traditional models show problems such as low computational efficiency and insufficient model generalization ability when dealing with large-scale unstructured data. In recent years, word vector models such as Word2Vec and BERT have provided more powerful textual representation capabilities for sentiment analysis, effectively capturing emotional nuances in text. For example, Yan et al. (2022) [19] remarkably improved the accuracy and generalization ability of sentiment classification based on DL technology.

Moreover, CNNs and RNNs have been widely used in text classification tasks, and RNNS are particularly suitable for processing time series data, such as consumer feedback and product reviews. For instance, Priya & Deepalakshmi (2023) [20] proposed fine-tuning the BERT and NLP techniques to analyze hotel review information for customer sentiment analysis. However, these models still faced challenges in maintaining global information when processing long texts and complex dependencies. The emergence of Transformer models provided a solution to this issue. As a typical Transformer model, BERT could capture global semantic information in text through bidirectional attention mechanisms, demonstrating exceptional performance in sentiment analysis tasks. Tan et al. (2022) [21] utilized data augmentation techniques with GloVe word embeddings, synthesizing a more lexically diverse set of samples through replacements with similar word vectors.

The personalized marketing strategy relies on a deep comprehension of consumer behavior and emotions to increase purchase intention and satisfaction. Tran et al. (2020) [22] demonstrated that personalized content significantly enhanced consumer brand loyalty, thus increasing enterprise revenue. Karabila et al. (2024) [23] proposed a novel recommendation system, which combined collaborative filtering with sentiment analysis to provide accurate and personalized suggestions. A recommendation model based on hybrid collaborative filtering is created by developing the BERT fine-tuning model for accurate sentiment classification. At the same time, BERT is used to gain insight into the product selection process in the improved recommendation system, thus enhancing the recommendation accuracy in e-commerce. Karbauskaitė et al. (2020) [24] formulated and addressed classification problems related to facial emotion recognition based on dimensional emotion models. Professional psychologists estimated the relationships between different emotions by representing them as points on a plane. Kaminskas & Vidugirienė (2016) [25] compared nonlinear input-output models that described relationships between signals of human emotions (excitement, frustration, and engagement/boredom) and virtual 3D facial features (interpupillary distance). To better illustrate the differences between existing models and the method employed in this study, Table 1 compares the characteristics and performance of various sentiment analysis models.

Table 1: Comparison of existing sentiment analysis models

Research	Method	Dataset	Evaluation indicator	Key findings
Hung & Alias (2023) [17]	Traditional sentiment analysis method	Social media reviews, product evaluations	Accuracy and sentiment classification	It relied on manual feature extraction, making it difficult to process large-scale data.
Liu (2024) [18]	Personalized sentiment analysis method	Digital marketing platform data	Customer satisfaction, emotional trends	It combined sentiment analysis with personalized recommendations, but its generalization ability was poor.
Priya & Deepalakshmi (2023) [20]	Fine-tuned the combination of BERT and NLP	Hotel evaluation data	Accuracy and sentiment classification	BERT improved sentiment analysis, but did not handle long text well.
Tan et al. (2022) [21]	Data augmentation +BERT and GloVe embeddings	Film review data	Accuracy and generalization ability	Data augmentation improved classification capabilities, but long text processing was still problematic.
Karabila et al. (2024) [23]	Combination of collaborative filtering with sentiment analysis	E-commerce data	Recommendation accuracy, user engagement	It was a recommendation system based on BERT. However, the integration of sentiment analysis and recommendation was not perfect.

Current sentiment analysis methodologies employing traditional models like CNN and RNN exhibit notable limitations despite their demonstrated utility. These models predominantly rely on manual feature engineering, failing to adequately extract deep semantic information—particularly when handling unstructured data. While CNNs specialize in local feature extraction, they disregard global textual patterns; RNNs, though capable of capturing certain temporal dependencies, still suffer from information loss and gradient vanishing issues in long-text processing scenarios.

The advent of Transformer architectures, particularly BERT, has driven substantial progress in sentiment analysis technology. BERT's bidirectional attention mechanism effectively captures contextual relationships, demonstrating exceptional performance in lengthy text analysis. Nevertheless, persistent challenges remain. BERT shows limited effectiveness in handling nuanced emotional hierarchies, cross-domain data adaptation, and multi-task learning contexts—with particular room for improvement in fine-grained analysis capability and cross-domain generalization performance.

This study's innovation lies in its proposed enhanced model architecture that integrates BERT with Transformer frameworks to address conventional models' limitations in processing complex texts and analyzing nuanced sentiment layers. This architecture optimizes BERT's pretraining procedure and significantly improves multi-task learning performance by incorporating fine-grained sentiment analysis mechanisms. This demonstrates superior generalization ability, particularly for cross-domain and complex sentiment analysis tasks. Consequently, the study bridges existing gaps in sentiment analysis depth and precision. These innovations substantially enhance the model's processing capacity while improving its practical applicability and adaptability. Especially in complex commercial and marketing contexts, it enables more accurate emotion trend analysis and consumer behavior prediction.

3 Research methodology

3.1 Data source and collection

The data sources for this study employ two publicly available Chinese sentiment analysis datasets: the yf_amazon and the Chinese Sentiment Classification Corpus (ChnSentiCorp). The yf_amazon dataset contains user reviews from the Amazon platform, complete with user ratings and review texts, making it suitable for sentiment analysis tasks. Due to its coverage of multiple product categories and the large number of reviews, the data is well-representative and can assist in identifying consumer emotional trends across different product domains. The ChnSentiCorp dataset is a typical Chinese product review dataset, compiling a vast number of product evaluations from e-commerce platforms, which is particularly suitable for studying consumer emotional tendencies towards products. yf_amazon dataset comprises approximately 50,000 reviews with sentiment label distribution as follows: positive (60%), negative (25%), and neutral (15%). The ChnSentiCorp dataset contains around 30,000 reviews distributed as positive (58%), negative (27%), and neutral (15%).

To further strengthen model generalization and verify its applicability across diverse social media and e-commerce environments, additional datasets from Weibo and JD.com are incorporated. The Weibo dataset encompasses extensive user sentiment expressions from social media platforms, capturing emotional attitudes and their temporal evolution. The JD.com dataset focuses on consumer product sentiment in e-commerce contexts, providing complementary perspectives to yf_amazon and ChnSentiCorp. Together, these datasets offer more diversified emotional data for validating model effectiveness and reliability across different data environments.

During data preprocessing, noise removal eliminates invalid content including advertisements, duplicate reviews, and irrelevant text through automated tools and rule-based filtering to ensure dataset validity. Subsequent automated annotation assigns sentiment labels to texts, extending beyond basic positive/negative/neutral classifications to include specific emotional states such as joy, anger, and disappointment. The setting of these emotional labels improves emotional recognition accuracy. Also, it enables more precise sentiment-guided personalization in recommendation systems and marketing strategy design.

The text tokenization process utilizes the Jieba tool to process Chinese texts by decomposing sentences into individual words. Jieba demonstrates superior accuracy in Chinese text tokenization, effectively resolving linguistic ambiguities while enhancing tokenization quality through customizable user dictionaries. Comparative analysis reveals that Jieba achieves higher tokenization efficiency than alternative Chinese tokenization tools, along with better adaptability for sentiment analysis tasks. The Jieba tool supports multiple tokenization modes (precision, search engine, and full mode), enabling optimal mode selection based on specific application requirements. Precision mode is selected for sentiment analysis tasks in this study to ensure maximal tokenization accuracy and emotional expression clarity.

Each tokenized word is an input feature vector, enabling the sentiment analysis model to effectively process and comprehend textual data. The tokenization process incorporates word frequency statistics to identify high-frequency terms and salient textual features, thus optimizing model training performance. Stop word removal constitutes a critical preprocessing step. Functionally insignificant particles (e.g., "of", "the", "in") are systematically eliminated using predefined stop lists to reduce computational overhead while improving model efficiency and prediction accuracy. The vectorization phase employs Word2Vec technology to transform textual data into numerically processable vector representations. By converting lexical items into fixed-dimensional vectors that preserve semantic relationships, Word2Vec facilitates enhanced contextual understanding by the model.

These methodological steps collectively generate high-quality training and testing datasets that support sentiment analysis model refinement and personalized marketing strategy development. This ensures statistically reliable and operationally valid analytical outcomes. The specific process is presented in Figure 1:

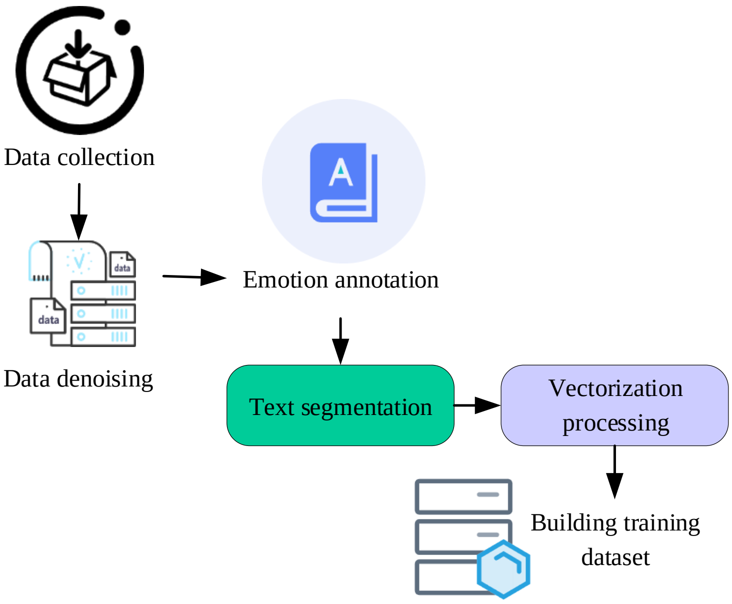

Figure 1: Process of data preprocessing

3.2 Design of the consumer sentiment analysis model

The proposed consumer sentiment analysis model is based on BERT and DL technologies to efficiently analyze and understand the emotions expressed by consumers in social media, product reviews, and customer feedback. The overall framework of the model is displayed in Figure 2:

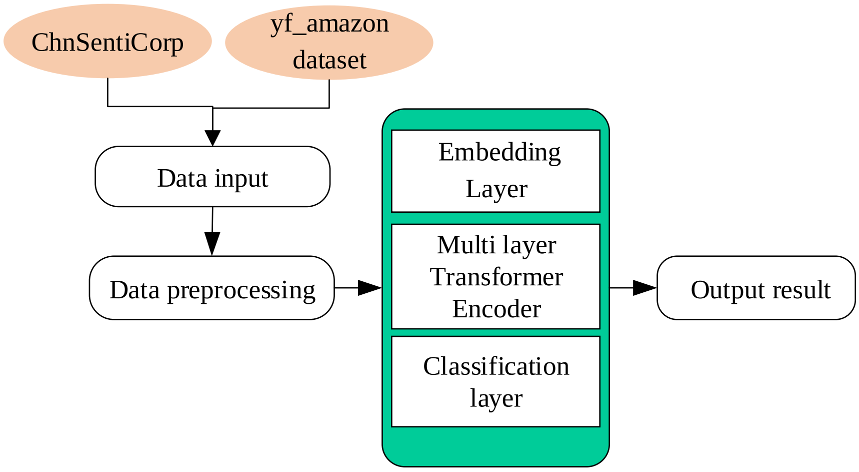

Figure 2: The overall framework

Figure 2 illustrates the design of the consumer sentiment analysis model, which is divided into three main parts: data input, model processing, and output results. During the data input stage, consumer sentiment data is derived from the two publicly available datasets mentioned previously: the ChnSentiCorp and the yf_amazon. These data have undergone preliminary preprocessing, encompassing denoising, sentiment annotation, tokenization, and vectorization, ensuring the quality of the data input into the model.

In the model processing stage, the core of this study lies in the DL architecture based on the BERT model. The uniqueness of the BERT model lies in its bidirectional encoder design, which can simultaneously consider contextual information, thereby providing a more accurate sentiment understanding. In the sentiment analysis task, the input format is simplified to "[CLS] + Sentence + [SEP]". "[CLS]" is a classification marker indicating the beginning of the entire sentence, and "[SEP]" is used to separate sentences. In this task, the input data is a single review text rather than sentence pair data, ensuring that the sentiment features of each comment can be fully learned and extracted. The model processing mainly includes an embedding layer, multiple Transformer encoders, and a classification layer. The embedding layer converts the input text into a vector representation, the Transformer encoder models the contextual relationships in the text, and the classification layer outputs sentiment labels based on the encoded text representation. Through this structure, BERT can effectively capture the deep semantic information required for sentiment analysis, thus improving the sentiment classification performance of the model.

Firstly, the input vocabulary sequence is transformed into a vector representation in the embedding layer. BERT maps each vocabulary to a high-dimensional space through word embedding techniques to obtain semantic information. In this way, the model can capture the contextual relationships of vocabulary. In the subsequent multi-layer Transformer encoder, the core of BERT consists of multiple stacked Transformer encoder layers, each of which processes information through self-attention mechanisms and feedforward neural networks. The calculation process of the attention mechanism reads:

$$\mathrm{Attention}(Q, K, V) = \mathrm{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V \tag{1}$$

Q, K, and V represent the query, key, and value vector of each word in the input text, respectively. These vectors calculate the correlation between each word and other words through the self-attention mechanism, to assign a weight to each word, and then extract the key information in the text. d_k refers to the dimension of the key vector, which plays a role in standardization and avoids numerical instability in the calculation process. After calculation, the softmax function normalizes the weighted value to obtain the importance of each word. Finally, it is combined with the value matrix V to output the weighted representation. Through the self-attention mechanism, the model can dynamically calculate the relevance of each word in the text, thereby extracting important features.

When processing long text sequences, BERT models and their variants frequently encounter the challenge of text length exceeding the model's maximum input limitation. The standard BERT architecture typically accepts a maximum of 512 tokens, though this study employs a reduced sequence length of 128 tokens. For input sequences exceeding this predetermined length threshold, a truncation strategy is implemented. Specifically, when text sequences surpass the 128-token limit, truncation occurs from the sequence's end. This is a methodological choice based on empirical observations that sentiment expressions often manifest early in textual content, making tail truncation preferable for preserving emotional information. Conversely, text sequences shorter than the maximum length require padding to maintain uniform input dimensions across all samples. The zero-padding technique is applied, whereby empty token positions are filled with zeros until reaching the specified 128-token length. This standardization ensures consistent input sizes for efficient batch processing while maintaining the model's architectural requirements.

In each layer, after the self-attention calculation, the input is added to the current layer's output through a residual connection, which helps with information propagation and model stability. Subsequently, layer normalization is applied to each layer's output to reduce the issue of vanishing gradients during training. After processing all the Transformer encoders, the final output is the representation of each input text. To obtain the sentiment classification results, the vector corresponding to the "[CLS]" token is extracted as the final text feature and input into a fully connected layer. The classification layer employs a softmax function to map the output to different sentiment categories, as shown in Equation (2):

$$p(y=k|x)=\frac{e^{z_k}}{\sum_{j=1}^{C}e^{z_j}} \quad (2)$$

z_k represents the logits value output by the model and the predicted score of category k. Equation (2) converts the model's output value z_k into a probability distribution through the softmax function. $p(y=k|x)$ refers to the probability that the input text belongs to category k; C is the total number of categories; $\sum_{j=1}^{C} e^{z_j}$ denotes the normalized term of the scores of all categories, ensuring that the probability sum of all categories is 1. Through this process, the model maps the emotional tendency of the input text to the corresponding sentiment category. To this end, the cross-entropy loss function is used for model optimization, defined as Equation (3):

$$L=-\frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{C} y_{ij}\log(p_{ij}) \quad (3)$$

N stands for the total number of samples; C denotes the number of sentiment categories; y_{ij} represents the true label; p_{ij} means the probability predicted by the model. During the model training process, the Adam optimization algorithm is employed to dynamically adjust the learning rate, aiming to achieve faster convergence and superior model performance. In the specific training steps, key hyperparameters are set, encompassing the learning rate, batch size, and number of training epochs. In addition, the model's generalization ability is evaluated by the layered k-fold cross-validation method. In this process, the dataset is randomly divided into k subsets, in which each time the k-1 subset is used as the training set, and the remaining subset is the verification set. Different from the traditional k-fold cross-validation, the layered k-fold cross-validation ensures that the category distribution in each subset is consistent with the overall dataset. Especially in the sentient analysis task, this is particularly important for dealing with data imbalance.

3.3 Dynamic feedback mechanism and personalized marketing strategy

The dynamic feedback mechanism is also introduced to improve the personalized marketing recommendation's accuracy and responsiveness. This mechanism adjusts recommendation strategies in real-time, adapting to changes in consumer behaviors and emotions. By monitoring and analyzing consumers' emotional fluctuations and behavioral changes in real-time, the mechanism ensures that the recommended content aligns with the user's current needs and emotional state.

In practical applications, the sentiment analysis model first performs real-time sentiment classification on user feedback (including comments, ratings, click behaviors, etc.) to determine current emotional tendencies. When significant emotional shifts are detected (e.g., a transition from positive to negative sentiment), the recommendation system dynamically adjusts its strategy accordingly. For users exhibiting negative sentiments, the system prioritizes recommendations of products or services likely to improve their emotional state while filtering out items that might exacerbate this sentiment. Concurrently, the system emphasizes positively reviewed offerings to facilitate emotional state transition.

Behavioral data plays an equally critical role, with the system continuously monitoring patterns in user interactions - particularly changes in clicking, browsing, and purchasing behaviors. Behavioral pattern variations trigger adaptive recommendation adjustments. For instance, frequent clicks on specific product categories or demonstrated emotional preferences for particular items result in increased recommendation frequency for those categories. The system further integrates reinforcement learning algorithms to continuously optimize the recommendation model based on behavioral and emotional feedback. By implementing a state-action-reward framework, the system learns optimal recommendation strategies for various emotional and behavioral contexts, iteratively refining content to maximize user satisfaction and conversion rates. This dynamic feedback mechanism, combining real-time sentiment analysis with behavioral tracking, enables precise, adaptive content recommendations. The integrated approach allows the recommendation system to responsively accommodate fluctuating consumer emotions and evolving needs, enhancing user experience and strengthening platform engagement.

Based on the results of consumer sentiment analysis, this study designs a set of personalized marketing strategies to increase customer engagement and satisfaction. Details are suggested in Figure 3:

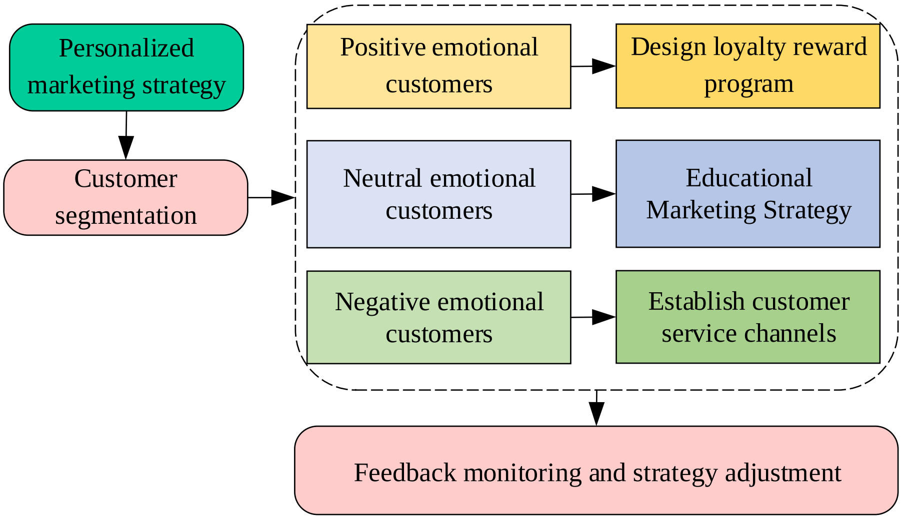

Figure 3: Personalized marketing strategy

Figure 3 demonstrates that consumers are initially classified into multiple segments through in-depth analysis of emotional data. This segmentation process relies on emotional tendencies identified by the sentiment analysis model, including positive, negative, and neutral emotions. At the same time, it incorporates more refined emotional labels such as joy, anger, and disappointment. These granular sentiment categories capture consumer sentiment diversity more precisely, thus enhancing emotion-based customer segmentation accuracy and facilitating more targeted marketing strategies. After the segmentation is completed, personalized marketing strategies are formulated according to the characteristics and needs of each subgroup to achieve more effective market promotion.

In the design process of personalized marketing strategy, priority should be given to the needs of customers with positive emotions, such as customers who show joy. These customers typically have a high satisfaction level with products and services and are inclined to share their positive experiences. For this group, enterprises can design loyalty reward programs, offer customized discounts and promotional activities, and encourage customers to share product experiences, facilitating brand loyalty and customer retention. By promoting these activities through social media platforms, enterprises can effectively increase brand exposure and customer engagement.

For customers with neutral sentiments, their feedback is often ambiguous, and they may not have high expectations for products or services. In this case, enterprises should adopt educational marketing strategies to increase customer trust and satisfaction by providing more information and support. Moreover, enterprises may employ email marketing campaigns or targeted advertisements to distribute product-related educational content, including detailed specifications and application guidelines. Such informational outreach enhances consumer product comprehension and purchase motivation while strategically incorporating promotional materials.

For customers exhibiting negative sentiments - particularly those labeled as angry or disappointed - prompt response and feedback management prove critical. Enterprises should implement effective customer care measures that extend beyond conventional service channels to improve satisfaction. Customized interventions like personalized product recommendations, immediate feedback collection, and improvement suggestions enable emotional guidance tailored to individual needs, achieving more focused marketing objectives. These measures directly address emotional requirements rather than merely resolving complaints, ultimately strengthening brand loyalty and conversion rates. Through this approach, customers with negative sentiments receive timely support and perceived brand value, thereby enhancing satisfaction and long-term engagement intentions.

When implementing personalized marketing strategies, enterprises fully utilize modern technological means, especially big data analysis and artificial intelligence technologies, to continuously monitor and evaluate the effectiveness of marketing activities. By conducting real-time analysis of marketing results, enterprises can adjust strategies promptly to adapt to the ever-changing market demands and customer preferences.

3.4 Data analysis and model evaluation

During the training and testing stages of the consumer sentiment analysis model, data analysis and model evaluation are crucial components. To ensure the validity and reliability of the model, this study employs a series of standards and metrics to assess model performance, including accuracy, recall, F1 score, and the confusion matrix, among others. The calculation can be written as Equations (4)-(7):

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \tag{4}$$

$$\text{Recall} = \frac{TP}{TP+FN} \tag{5}$$

$$F1 = 2 \cdot \frac{\text{Precision} \cdot \text{Recall}}{\text{Precision}+\text{Recall}} \tag{6}$$

$$\text{Precision} = \frac{TP}{TP+FP} \quad (7)$$

True Positive (TP) and True Negative (TN) represent the number of samples correctly predicted as positive and negative class; False Positive (FP) and False Negative (FN) refer to the number of samples incorrectly predicted as positive and negative class.

The experimental design allocates 80% of data for training and 20% for model validation, ensuring generalization ability during training. Hyperparameter optimization employs both grid search and random search methods. For learning rates, multiple values (0.0001, 0.001, 0.01) are evaluated, with 0.001 selected as optimal due to its balanced convergence speed and avoidance of vanishing and exploding gradients. Batch size comparisons (16, 32, 64) identify 32 as delivering optimal accuracy-computation efficiency trade-offs. The Adam optimizer with a 0.001 learning rate demonstrates superior convergence and final performance.

Further refinements include dynamic learning rate adjustment - reducing rates by 0.5 after epoch 10 for fine-tuning while preventing overfitting. Moreover, an early stopping mechanism is introduced, which automatically stops training when the validation set's performance does not improve significantly within 5 consecutive epochs to ensure the model's generalization ability. Additional safeguards against overfitting incorporate L2 regularization and a 0.1 dropout rate, collectively enhancing model stability. The main parameters used for the consumer sentiment analysis model and their settings are exhibited in Table 2:

Table 2: Parameter setting

Parameter	Setting value
Model type	BERT, Transformer
Learning rate	0.001
Batch size	32
Maximum sequence length	128
Training epochs	10
Optimizer	AdamW
Loss function	Cross-entropy loss
Activation function	ReLU
Dropout rate	0.1
Random seed	42
Division of training and verification sets	80% for the training set; 20% for the verification set
Overfitting prevention technology	Early stopping, L2 regularization
Cross-validation method	Layered k-fold cross-validation
Padding and truncation strategies	Padding to the maximum sequence length of 128, exceeding partial truncation

In model training and evaluation, K-fold cross-validation is employed to improve the model's generalization ability. In K-fold cross-validation, the training set is divided into K equal-sized subsets. The model is trained K times, with each iteration selecting a different subset as the validation set and the remainder as the training set. The final model performance evaluation result is obtained by averaging the K validation results. This method effectively prevents the model from overfitting to a specific training set, thus improving its performance on new data. A confusion matrix visualizes the model's classification results, analyzing its performance across different sentiment categories. The confusion matrix analysis reveals the model's recognition accuracy across different sentiment categories during actual evaluation. This approach also identifies potential inter-category confusion patterns, providing targeted guidance for model refinement. Text that exceeds the maximum sequence length 128 is truncated. For short text, zero-padding is employed to ensure that all input sequences have the same length, which can effectively improve the BERT model's processing efficiency.

4 Results and discussion

4.1 Results of consumer sentiment analysis

In consumer sentiment analysis, an in-depth examination of sentiment classification results is conducted to identify the key factors influencing changes in consumer emotions. Through sentiment classification of the yf_amazon and ChnSentiCorp datasets, significant differences in the distribution of positive, negative, and neutral sentiments are observed. Specifically, the proportion of positive sentiment is notably high in reviews, particularly regarding service experience and product quality. Negative sentiments are primarily concentrated on negative experiences related to dissatisfaction with product quality and service delays. In addition, fine-grained sentiment labels, such as joy, anger, and disappointment, are introduced to refine consumers' emotional expression and can more accurately capture subtle changes in emotional fluctuations. The distribution of refined sentiment labels is suggested in Figure 4:

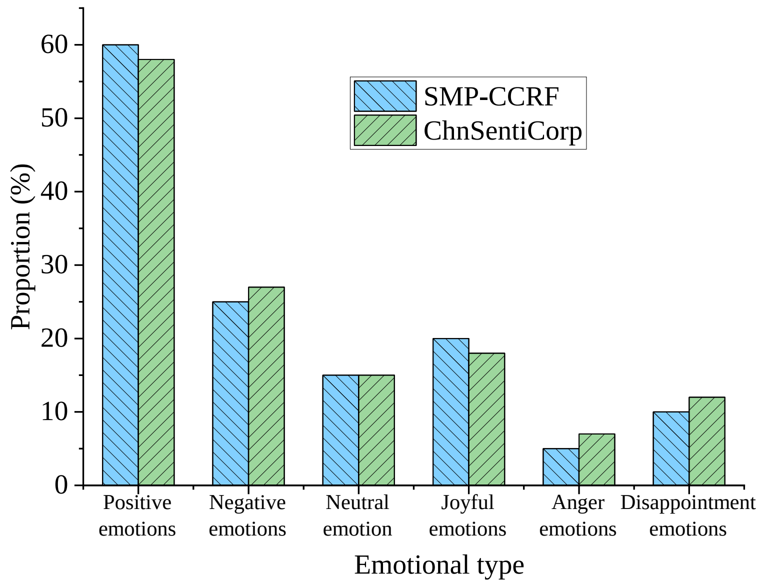

Figure 4: Results of consumer sentiment analysis

Figure 4 reveals that positive sentiments constitute the predominant proportion across both datasets, particularly in evaluations of product quality and service experience. This indicates that superior products and services significantly enhance consumer satisfaction. Negative sentiments primarily emerge from critical feedback regarding product defects and service delays, typically triggered by unsatisfactory shopping experiences. The relatively low proportion of neutral sentiments suggests most consumers tend to express more definitive emotional attitudes in their evaluations.

Within the fine-grained emotional labels, joy predominantly appears in positive assessments of product functionality and aesthetics, reflecting strong approval of need fulfillment. Anger mainly stems from strong reactions to substandard product quality or poor service attitudes. Disappointment closely correlates with discrepancies between consumer expectations and actual experiences, becoming particularly pronounced when products fail to meet anticipated standards. This granular sentiment classification enables precise identification of specific causes behind consumer sentiment fluctuations, thereby providing more accurate data support for personalized marketing strategy development. Brand image, product quality, and customer service demonstrate significant influence on consumer sentiments. Positive brand perception and high product quality consistently foster favorable sentiments, whereas inadequate service quality and negative reviews predominantly trigger adverse emotional responses.

4.2 Performance of the model

A 5-fold cross-validation is performed on the whole dataset to comprehensively evaluate the model's generalization ability. The dataset is divided into 5 subsets, with each taking turns as the validation set and the other 4 subsets as the training set. Based on this process, the performance results of 5 training and verification are obtained. Table 3 shows the results of K-fold cross-validation:

Table 3: The results of K-fold cross-validation

Fold	Accuracy	F1 score	Recall
1	91.5%	90.8%	89.3%
2	92.0%	91.2%	90.1%
3	91.8%	90.6%	89.8%
4	92.3%	91.5%	90.2%
5	91.9%	91.0%	90.0%

The average accuracy, F1 score, and recall of 5-fold cross-validation in Table 3 are 91.9%, 91.0%, and 89.9%. These results show that the model is stable under different data divisions. It has strong generalization ability, can identify positive samples well, and has a high recall. To further evaluate the model's performance in different populations, the dataset is divided into multiple subsets according to the characteristics of users such as age, gender, and region, evaluating the performance of each subset. Table 4 presents the model's accuracy, F1 score, recall, and precision in various groups.

Table 4: Performance evaluation results based on different population groups

Group	Accuracy	F1 score	Recall	Precision
Age (18-25years)	90.5%	89.3%	88.0%	91.0%
Age (26-35years)	92.0%	91.5%	90.3%	93.2%
Age (36-50 years)	91.2%	90.4%	89.6%	91.9%
Age (over 50)	89.8%	88.5%	87.2%	90.4%

Gender (Male)	91.7%	91.0%	90.1%	92.5%
Gender (Female)	90.6%	89.8%	88.5%	91.3%
Region (Urban)	92.5%	91.8%	90.7%	93.1%
Region (Rural)	89.6%	88.2%	87.0%	90.2%

Table 4 demonstrates notable variations in sentiment analysis performance across different age groups. Users aged 26-35 exhibit superior accuracy, F1 scores, and recall, indicating more consistent and explicit emotional expressions within this demographic. The slightly lower accuracy observed among users over 50 may stem from the more complex emotional language patterns characteristic of this age group. While performance metrics remain comparable between genders, male users show marginally higher accuracy and recall, potentially attributable to their tendency toward more direct emotional expressions in evaluations. Urban users outperform their rural counterparts in sentiment analysis tasks. This is likely due to more frequent social network engagement and standardized emotional expressions resulting from greater information exposure.

Figures 5 and 6 present comparative performance analyses across multiple models on both ChnSentiCorp and yf_amazon datasets. These models encompass the proposed model, LSTM, SVM, CNN, and models from references [26] and [27]. The evaluation incorporates standard metrics: F1 score, accuracy, and recall.

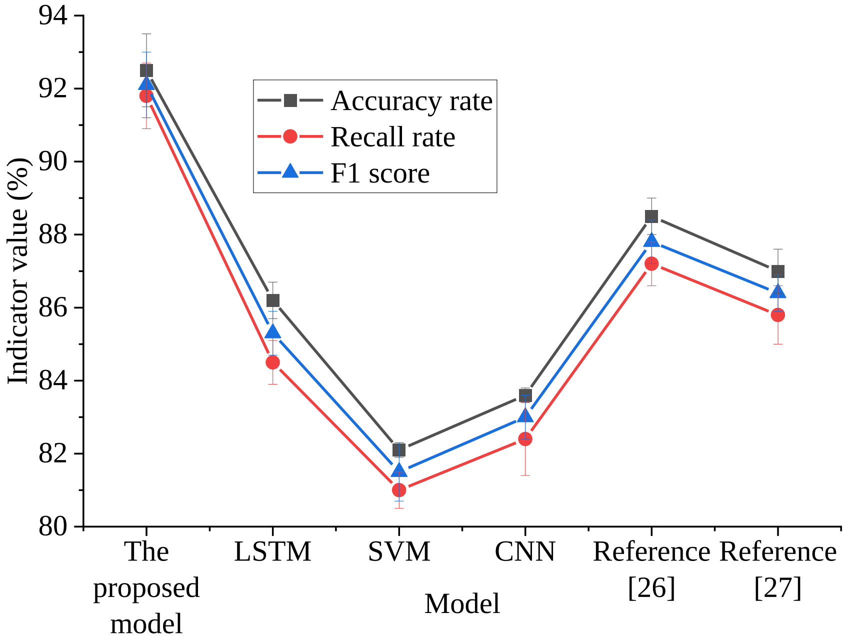

Figure 5: Performance comparison of different models (The ChnSentiCorp dataset)

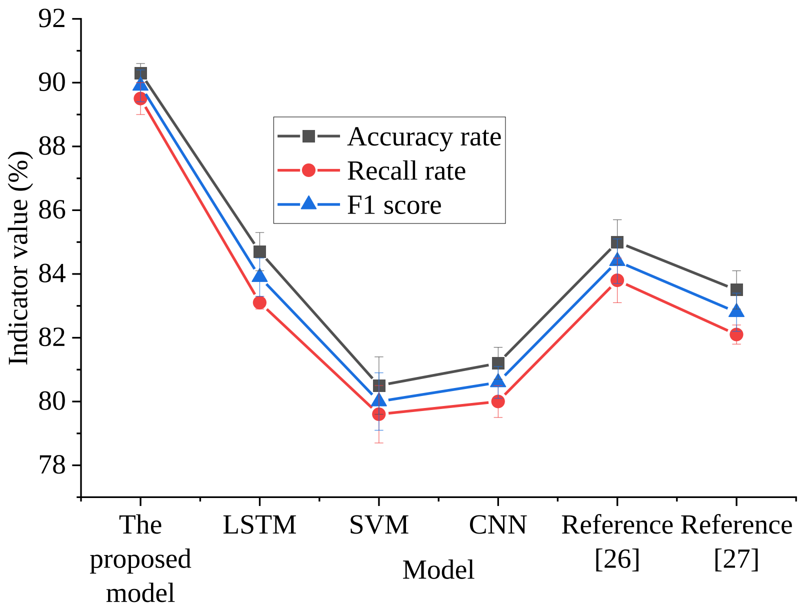

Figure 6: Performance comparison of diverse models (The yf_amazon dataset)

The comparative results in Figures 5 and 6 demonstrate that the proposed model consistently outperforms conventional models (LSTM, SVM, and CNN) across all evaluation metrics (accuracy, recall, and F1 score). This performance advantage holds consistently across both datasets (ChnSentiCorp and yf_amazon). Especially in the ChnSentiCorp dataset, the proposed model achieves 92.5% accuracy, significantly higher than LSTM (86.2%), SVM (82.1%), and CNN (83.6%). Similarly, on the yf_amazon dataset, the proposed model maintains superior performance with 90.3% accuracy compared to alternative models.

This performance enhancement primarily stems from the proposed model's BERT-based bidirectional self-attention mechanism. Traditional LSTM, SVM, and CNN models exhibit inherent limitations when processing textual data, particularly for sentiment analysis tasks involving long texts and complex dependencies. LSTMs, while effective at capturing sequential information, tend to lose critical semantic context in long-range dependencies. SVMs demonstrate competent classification performance but show constrained effectiveness for nonlinear problems. CNNs, despite their local feature extraction capabilities, struggle to capture global semantic patterns in text. In contrast, BERT's bidirectional attention mechanism simultaneously processes contextual information from both directions, enabling more precise sentiment classification judgments.

Furthermore, while references [26] and [27] incorporate data augmentation and other optimization techniques, they fail to overcome the fundamental limitations of LSTM, CNN, and SVM architectures in sentiment analysis. Reference [26] primarily employs word embedding methods to improve accuracy, but its reliance on static word vectors prevents adequate capture of contextual emotional nuances. Reference [27]'s hybrid approach combines traditional sentiment analysis models but lacks BERT's contextual comprehension capability, resulting in weaker performance for complex sentiment tasks. Overall, the proposed model's superior performance across both datasets derives from BERT's robust semantic understanding. Meanwhile, it effectively captures deep emotional relationships in text, excelling in long-text analysis and complex dependency tasks where conventional models show limitations.

To statistically validate performance differences, t-tests assess the significance of variations in accuracy, recall, and F1 scores between model pairs, with detailed results presented in Table 5.

Table 5: Statistical significance test results

Comparison model	p-value of accuracy	p-value of recall	P-value of F1 score
The proposed model vs LSTM	0.0012*	0.0013*	0.0011*
The proposed model vs SVM	0.0008*	0.0009*	0.0006*
The proposed model vs CNN	0.0024*	0.0021*	0.0020*
LSTM vs SVM	0.1459	0.1293	0.1385
LSTM vs CNN	0.2731	0.2837	0.2652
SVM vs CNN	0.4876	0.4325	0.4703

The statistical significance test results in Table 5 demonstrate that the performance differences between the proposed model and conventional models (LSTM, SVM, CNN) are statistically significant (p-values < 0.05). This shows that the proposed mode has superior performance across all evaluation metrics. Particularly in comparisons with LSTM, SVM, and CNN, the proposed model exhibits marked advantages in accuracy, recall, and F1 score. Furthermore, the performance differences between LSTM and SVM, LSTM and CNN, and SVM and CNN lack statistical significance (p-values > 0.05), providing additional evidence of the proposed model's relative superiority.

The confusion matrix can better show the model's performance in each emotion category. Table 6 lists the predictions of the different models on the positive, negative, and neutral categories.

Table 6: The confusion matrix (The proposed Model)

	Positive	Negative	Neutral	Total
Positive	680	30	20	730
Negative	25	750	15	790
Neutral	15	10	280	305
Total	720	790	315	1825

Table 6 reveals that the proposed model achieves optimal performance in sentiment classification, particularly for positive and negative categories, where correctly classified samples significantly outnumber other models. LSTM, SVM, and CNN models demonstrate considerable confusion between negative and neutral classifications. However, the proposed model maintains superior differentiation across all categories, albeit with relatively lower precision for neutral sentiment classification compared to positive and negative categories.

This performance gap likely reflects the inherent ambiguity of neutral emotional expressions. Comparative analysis of confusion matrices confirms the proposed model's enhanced discriminative capability across all three sentiment categories, demonstrating clear advantages in sentiment classification tasks.

Figure 7 displays the operational efficiency of each model, including the training and prediction time for each model:

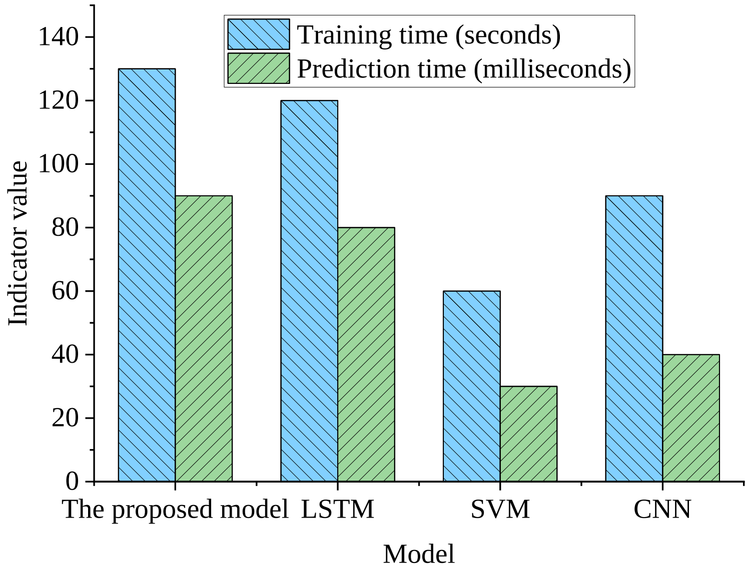

Figure 7: Comparison of operational efficiency of various models

Figure 7 suggests that the proposed model has a training time of 130 seconds, slightly longer than that of LSTM, SVM, and CNN. However, its prediction time reaches 90 milliseconds, indicating high real-time performance. The proposed model achieves efficient operation while maintaining good performance by optimizing data loading and batch processing strategies, as well as employing mixed-precision training. Compared to other models, the proposed model can provide accurate sentiment analysis results within a shorter prediction time, making it more suitable for real-time demands in practical applications. Table 7 summarizes the complexity of each model, encompassing the number of parameters and computational resource requirements:

Table 7: Comparison of model complexity

Model	The number of parameters	Computational resource requirements
The proposed model	30 M	Medium (can run on high-performance Central Processing Unit (CPU))
LSTM	20 M	Medium (can run on regular CPU)
SVM	1 M	Low (usable on regular computers)
CNN	15 M	Medium (requiring moderate optimization)

In Table 7, the number of parameters in the proposed model is 30M, with moderate computational resource requirements, making it suitable for operation on high-performance CPUs. Through optimization of the model structure and reduction of parameters, this study effectively lowers the model's complexity, allowing it to operate efficiently even in resource-constrained environments. Although LSTM, SVM, and CNN have lower parameter numbers and computational resource requirements, their performance in sentiment analysis does not match that of the proposed model. Consequently, in resource-limited scenarios, while other models may serve as alternative options, the proposed model achieves a good balance between accuracy and efficiency, making it suitable for a wider range of practical applications.

The computational cost is further explored to evaluate the feasibility and efficiency of the model in practical application. It includes the performance of training time, memory utilization, and inference time under different hardware settings. Experiments are conducted on the Graphics Processing Unit (GPU) and CPU, respectively. Table 8 shows the experimental results:

Table 8: Computational cost analysis under different hardware configurations

Item	GPU(Training/Inference)	CPU (Training/Inference)
Training time (per epoch)	12 minutes	45minutes
Maximum memory utilization	4 GigaByte (GB)	12GB
Inference time	200 milliseconds (ms)	800 ms

Table 8 reveals the substantial advantages of GPU over CPU in both model training and inference. Regarding training efficiency, the GPU completes each epoch in merely 12 minutes compared to the CPU's 45 minutes, demonstrating significantly enhanced processing capability, especially for large-scale datasets. Memory utilization analysis shows the GPU's maximum memory consumption at 4GB, substantially lower than the CPU's 12GB requirement, indicating the GPU's superior suitability for memory-intensive DL tasks. Inference time comparisons highlight hardware impact on operational efficiency. GPU achieves 200ms inference latency versus CPU's 800ms, making GPU preferable for real-time applications. While the CPU remains functionally capable, the GPU demonstrates clear superiority in both computational efficiency and memory management for large-scale data processing and time-sensitive inference tasks. These findings strongly recommend GPU as the preferred hardware platform for accelerating both training and inference processes in practical implementations.

4.3 The application effects of personalized marketing strategies

This section analyzes the practical effects of personalized marketing strategies based on sentiment analysis, targeting consumers who have purchased products from a well-known e-commerce platform. By deeply mining and analyzing the sentiment data of these consumers, the marketing team can better understand their needs and preferences, thus formulating more precise marketing strategies. After implementing the personalized marketing strategies, feedback and behavioral changes from consumers are tracked and surveyed. The statistical results are revealed in Figure 8:

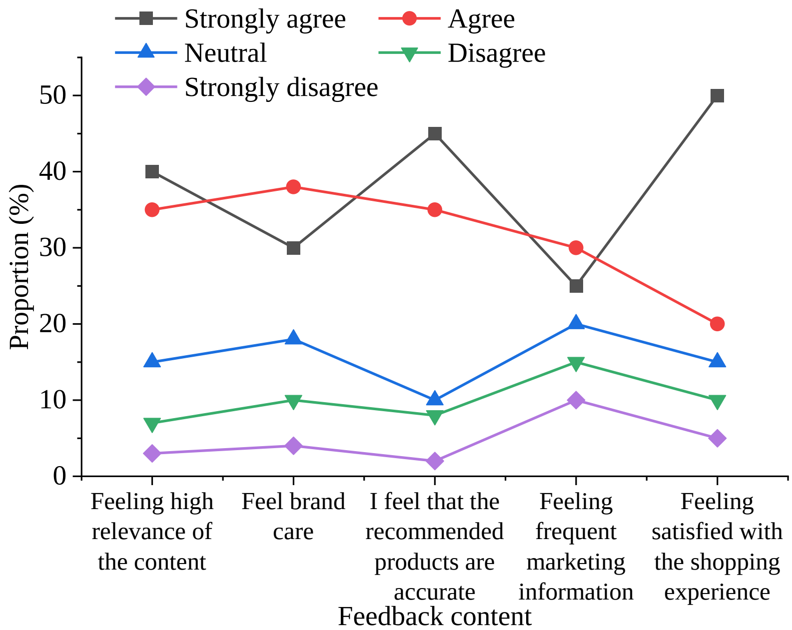

Figure 8: Statistics on consumer feedback level

Figure 8 provides a visual analysis of consumer feedback on personalized marketing strategies. In the feedback category “feeling high content relevance,” 40% of consumers indicate “strongly agree,” while 35% show “agree,” demonstrating a general recognition of the relevance of the recommended products. Regarding “feeling brand care,” 68% of consumers offer positive feedback, illustrating the effectiveness of the personalized care strategy. In the category “feeling accurate recommended products,” 45% of consumers state “strongly agree,” suggesting that the effectiveness of sentiment analysis is well validated in the recommendation system. Regarding feedback on "frequent marketing information," 25% of consumers select "strongly agree," indicating the need for careful adjustment of marketing information frequency to avoid negative consumer experiences caused by information overload. Consequently, personalized adjustment strategies for marketing information delivery frequency become essential to ensure information dissemination meets consumer needs without causing annoyance or resistance. The changes in customer conversion rates are depicted in Figure 9:

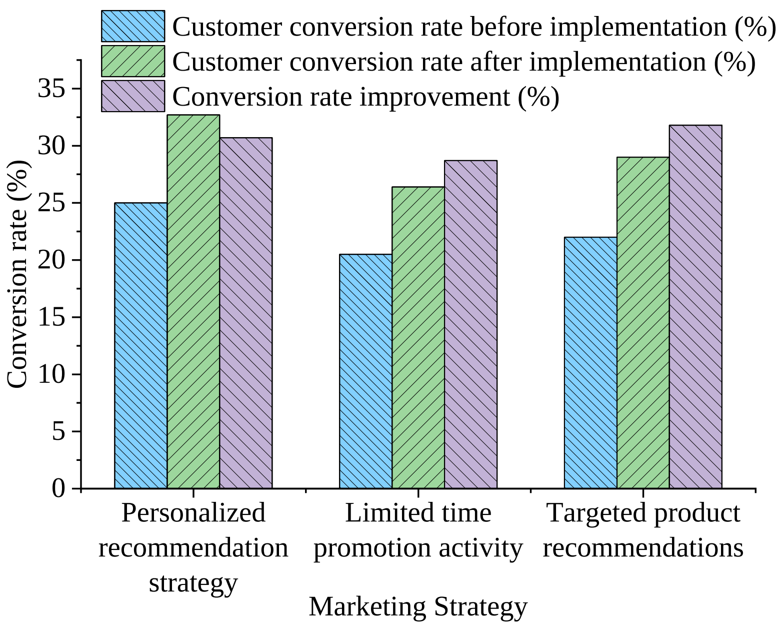

Figure 9: The changes in customer conversion rate

Figure 9 demonstrates that after implementing personalized recommendation strategies, customer conversion rates increase from 25.0% to 32.7%, representing a 7.7 percentage point improvement equivalent to approximately 30.8% relative growth. These results confirm that consumer sentiment analysis-based strategies effectively attract more potential customers. Other marketing strategies, such as limited-time promotional activities and targeted product recommendations, also show good conversion effects, enhancing the rates by 28.7% and 31.8%, respectively. However, these improvements may also be influenced by external factors. Market competition, seasonal variations, or fluctuations in consumer behavior could potentially impact conversion rates. Additionally, possible biases in data collection, such as uneven distribution of consumer preferences or regional market differences, may also affect the results. Hence, future research should investigate these potential influencing factors to enable a more accurate evaluation of personalized recommendation strategy effectiveness.

Overall, consumer feedback on personalized marketing strategies exhibits a positive trend, with these strategies effectively promoting consumer purchasing decisions and enhancing brand loyalty. However, some consumers still express concerns about the frequency of marketing messages. Enterprises should continuously optimize marketing content and strategies to maintain a positive consumer experience when implementing personalized marketing. By segmenting the degree of feedback, it is possible to more accurately identify consumer needs and preferences, offering data support for subsequent marketing strategy optimization.

4.4 Discussion

The research findings demonstrate that sentiment analysis-based personalized recommendation strategies yield significant improvements in customer conversion rates. Despite these encouraging performance gains, several potential external factors and biases may influence the results, warranting further investigation.

Market competition emerges as a critical factor affecting conversion rates. In highly competitive environments, rival companies' strategies, advertising campaigns, or promotional activities may temporarily alter customer purchasing decisions, thereby impacting conversion rates. While the current personalized recommendation strategy performs well in controlled settings, its effectiveness may diminish when confronted with real-world competitive pressures. Seasonal variations substantially influence purchasing behaviors. For instance, holiday promotions or seasonal discounts often increase customers' purchase likelihood during specific periods. Such cyclical fluctuations necessitate consideration since they may cause temporal variations in the strategy's effectiveness. A comprehensive evaluation of long-term performance requires repeated experiments across different timeframes to isolate seasonal effects on conversion rates.

Consumer behavior diversity represents another influential factor. Individual characteristics, including age, gender, purchasing habits, and cultural background, generate varied responses to personalized recommendations. These behavioral differences may produce uneven conversion rate improvements across demographic segments. Consequently, a uniform recommendation approach may not prove universally effective across all consumer groups, suggesting future research should explore segment-specific strategy optimization based on distinct user characteristics.

Moreover, potential biases in data collection processes warrant careful consideration. For instance, issues regarding the representativeness of consumer feedback may arise if sample data fail to accurately reflect the diversity of target populations, potentially leading to skewed analytical outcomes. Subsequent research should employ more representative datasets and implement controls for potential biases during data acquisition to ensure the generalizability and reliability of findings. Although personalized recommendation strategies demonstrate significant efficacy in improving conversion rates, external factors and underlying biases may still substantially influence the results. Future investigations should incorporate more sophisticated data analysis and experimental designs to systematically examine these factors, thus validating and optimizing the effectiveness of personalized recommendation strategies.

The impact of domain-specific language on sentiment analysis models constitutes another critical consideration. Distinct industry sectors (e.g., electronics versus fashion products) exhibit unique terminologies, expression patterns, and consumer sentiment tendencies, potentially affecting model generalizability. This study utilizes multiple datasets (yf_amazon and ChnSentiCorp) encompassing different product categories (such as electronics and consumer goods), enabling a preliminary assessment of cross-domain performance. However, given inherent variations in linguistic characteristics and emotional expression patterns across industries, model performance may vary significantly. It potentially excels in certain domains (e.g., fashion reviews) while underperforming in others (e.g., electronics evaluations). These discrepancies primarily manifest in differential usage frequencies of sentiment-bearing words and contextual comprehension challenges. For example, fashion reviews frequently employ affective descriptors (e.g., "stylish," "elegant"), whereas electronics reviews emphasize technical specifications (e.g., "performance," "battery life"). To address these limitations, future research could implement domain adaptation techniques to enhance cross-industry robustness. Additionally, for domain-specific sentiment analysis tasks, researchers may refine training data to strengthen domain-aware linguistic representations, thus improving models' cross-domain adaptability.

5 Conclusion

This study combines NLP with DL technologies to explore the effective integration of consumer sentiment analysis and personalized marketing strategies, achieving a series of important research outcomes. By employing the yf_amazon and ChnSentiCorp datasets, an efficient sentiment analysis model is constructed, and personalized marketing strategies targeting consumer sentiment are designed. First, the study innovates in designing the sentiment analysis model, proposing a model based on BERT and Transformer architecture. Meanwhile, it optimizes parameter settings and training processes, thus markedly improving the accuracy of sentiment classification and the model's operational efficiency. Second, the study devises a systematic set of personalized marketing strategies based on consumer sentiment analysis results, capable of more precisely meeting consumer demands. Lastly, this study lays the foundation for future research on personalized marketing strategies based on sentiment data. It provides data support and theoretical guidance for enterprises' marketing decisions by analyzing the relationship between sentiment changes and consumer behavior. Although this study has achieved certain results, some limitations still exist. For example, the datasets used are mainly from specific domains, which may affect the model's generalization abilities. Therefore, future research could consider expanding data sources to cover more industries and different types of consumer feedback to improve the model's adaptability. Future research directions include enhancing the sentiment analysis model's adaptability and robustness, particularly for processing more complex and diverse emotional expressions. To better capture nuanced emotional variations in lengthy narrative-style reviews, multimodal data and richer emotional labels are incorporated to improve the model's recognition capability across different emotional dimensions. Furthermore, advanced explainable AI techniques, including Local Interpretable Model-agnostic Explanations and Shapley values, are explored to address model interpretability challenges in sentiment classification. These techniques clarify the model's decision-making logic while generating actionable feedback for business applications. Additional testing across diversified datasets and real-world business scenarios could further validate the model's generalization ability and cross-domain applicability, thus improving its reliability and effectiveness in practical implementations.

References

[1] Alantari, H. J., Currim, I. S., Deng, Y., & Singh, S. (2022). An empirical comparison of machine learning methods for text-based sentiment analysis of online consumer reviews. International Journal of Research in Marketing, 39(1), 1-19.

[2] Rocklage, M. D., He, S., Rucker, D. D., & Nordgren, L. F. (2023). Beyond sentiment: the value and measurement of consumer certainty in language. Journal of Marketing Research, 60(5), 870-888.

[3] Iqbal, A., Amin, R., Iqbal, J., Alroobaea, R., Binmahfoudh, A., & Hussain, M. (2022). Sentiment analysis of consumer reviews using deep learning. Sustainability, 14(17), 10844.

[4] Singh, A., Jenamani, M., Thakkar, J. J., & Rana, N. P. (2022). Quantifying the effect of eWOM embedded consumer perceptions on sales: An integrated aspect-level sentiment analysis and panel data modeling approach. Journal of Business Research, 138, 52-64.

[5] Yuan, Z. (2022). Big data recommendation research based on travel consumer sentiment analysis. Frontiers in Psychology, 13, 857292.

[6] Kaur, G., & Sharma, A. (2023). A deep learning-based model using hybrid feature extraction approach for consumer sentiment analysis. Journal of big data, 10(1), 5.

[7] Hossain, M. S., & Rahman, M. F. (2023). Customer sentiment analysis and prediction of insurance products' reviews using machine learning approaches. FIIB Business Review, 12(4), 386-402.

[8] Taherdoost, H., & Madanchian, M. (2023). Artificial intelligence and sentiment analysis: A review in competitive research. Computers, 12(2), 37.

[9] Adak, A., Pradhan, B., & Shukla, N. (2022). Sentiment analysis of customer reviews of food delivery services using deep learning and explainable artificial intelligence: Systematic review. Foods, 11(10), 1500.

[10] Shaik Vadla, M. K., Suresh, M. A., & Viswanathan, V. K. (2024). Enhancing product design through AI-driven sentiment analysis of Amazon reviews using BERT. Algorithms, 17(2), 59.

[11] Noorian, A., Harounabadi, A., & Hazratifard, M. (2024). A sequential neural recommendation system exploiting BERT and LSTM on social media posts. Complex & Intelligent Systems, 10(1), 721-744.

[12] Ghobakhloo, M., & Ghobakhloo, M. (2022). Design of a personalized recommender system using sentiment analysis in social media (case study: banking system). Social Network Analysis and Mining, 12(1), 84.

[13] El-Ansari, A., & Beni-Hssane, A. (2023). Sentiment analysis for personalized chatbots in e-commerce applications. Wireless Personal Communications, 129(3), 1623-1644.

[14] Gooljar, V., Issa, T., Hardin-Ramanan, S., & Abu-Salih, B. (2024). Sentiment-based predictive models for online purchases in the era of marketing 5.0: a systematic review. Journal of Big Data, 11(1), 107.

[15] Roumeliotis, K. I., Tselikas, N. D., & Nasiopoulos, D. K. (2024). Leveraging Large Language Models in Tourism: A Comparative Study of the Latest GPT Omni Models and BERT NLP for Customer Review Classification and Sentiment Analysis. Information, 15(12), 792.

[16] Çetinkaya, Y. M., Külah, E., Toroslu, İ. H., & Davulcu, H. (2024). Targeted marketing on social media: utilizing text analysis to create personalized landing pages. Social Network Analysis and Mining, 14(1), 77.

[17] Hung, L. P., & Alias, S. (2023). Beyond sentiment analysis: a review of recent trends in text based sentiment analysis and emotion detection. Journal of Advanced Computational Intelligence and Intelligent Informatics, 27(1), 84-95.

[18] Liu, H. (2024). The role of personalization in modern digital marketing: how tailored experiences drive consumer engagement. Strategic Management Insights, 1(8), 34-40.

[19] Yan, C., Liu, J., Liu, W., & Liu, X. (2022). Research on public opinion sentiment classification based on attention parallel dual-channel deep learning hybrid model. Engineering Applications of Artificial Intelligence, 116, 105448.

[20] Priya, C. S. R., & Deepalakshmi, P. (2023). Sentiment analysis from unstructured hotel reviews data in social network using deep learning techniques. International Journal of Information Technology, 15(7), 3563-3574.

[21] Tan, K. L., Lee, C. P., Anbananthen, K. S. M., & Lim, K. M. (2022). RoBERTa-LSTM: a hybrid model for sentiment analysis with transformer and recurrent neural network. IEEE Access, 10, 21517-21525.

[22] Tran, T. P., van Solt, M., & Zemanek Jr, J. E. (2020). How does personalization affect brand relationship in social commerce? A mediation perspective. Journal of Consumer Marketing, 37(5), 473-486.

[23] Karabila, I., Darraz, N., EL-Ansari, A., Alami, N., & EL Mallahi, M. (2024). BERT-enhanced sentiment analysis for personalized e-commerce recommendations. Multimedia Tools and Applications, 83(19), 56463-56488.

[24] Karbauskaitė, R., Sakalauskas, L., & Dzemyda, G. (2020). Kriging predictor for facial emotion recognition using numerical proximities of human emotions. Informatica, 31(2), 249-275.

[25] Kaminskas, V., & Vidugirienė, A. (2016). A comparison of Hammerstein-type nonlinear models for identification of human response to virtual 3D face stimuli. Informatica, 27(2), 283-297.

[26] Märtin, C., Bissinger, B. C., & Asta, P. (2023). Optimizing the digital customer journey—Improving user experience by exploiting emotions, personas and situations for individualized user interface adaptations. Journal of Consumer Behaviour, 22(5), 1050-1061.

[27] Gao, Y., & Liu, H. (2023). Artificial intelligence-enabled personalization in interactive marketing: a customer journey perspective. Journal of Research in Interactive Marketing, 17(5), 663-680.