Journal of Organizational and End User Computing
Volume 37 • Issue 1 • January-December 2025

Sentiment Analysis and Consumer Purchase Intention Prediction Based on BERT and DBLSTM

Changzheng Yang
Ocean University of China, China

Mengmeng Guo
https://orcid.org/0009-0009-0067-0681
Renmin University of China, China

Muhammad Asif
https://orcid.org/0000-0003-1839-2527
National Textile University, Pakistan

ABSTRACT

With the rapid growth of e-commerce and social media, analyzing consumer sentiment and predicting purchase intentions are vital for understanding behavior and optimizing marketing strategies. This study proposes an integrated model combining Bidirectional Encoder Representation from Transformers (BERT) and Deep Bidirectional Long Short-Term Memory Network (DBLSTM). BERT efficiently extracts semantic features from consumer reviews using self-attention mechanisms, while DBLSTM captures sequential dynamics, leveraging temporal dependencies for predicting purchase intentions. The model was evaluated on real-world datasets and compared against other deep learning models. Results showed that the BERT-DBLSTM model outperformed others in sentiment analysis accuracy and purchase intention prediction, demonstrating higher generalization and prediction accuracy. This approach provides enterprises with precise market insights, enabling improved marketing strategies, enhanced user satisfaction, and increased conversion rates.

KEYWORDS

Purchase Intention Prediction, Sentiment Analysis, Self-Attention Mechanism, DBLSTM, BERT, Accuracy Improvement, Optimization of Marketing Strategies

INTRODUCTION

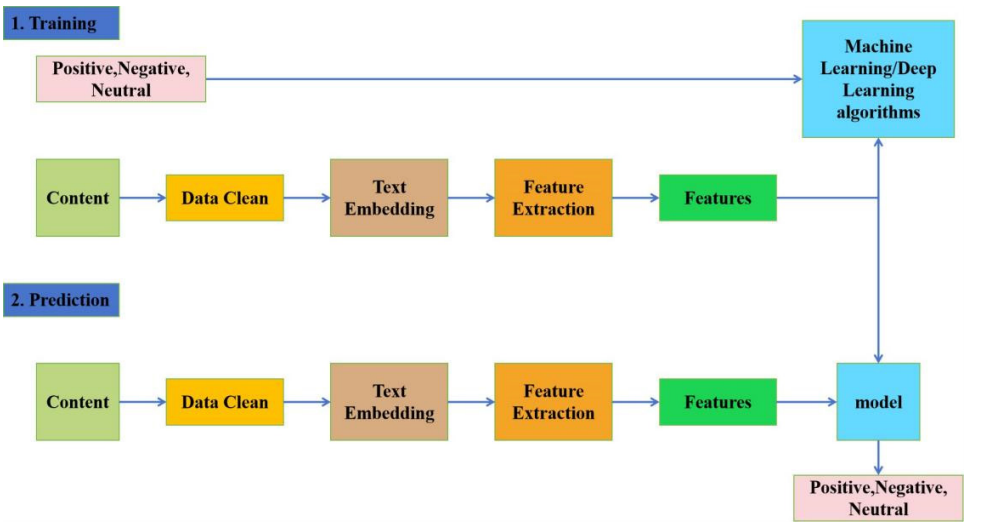
In today's global business environment, e-commerce has become essential, continually advancing and expanding. With a growing preference for online shopping, consumers are increasingly turning to e-commerce to meet their needs, making it a fundamental aspect of contemporary life. Yet, even with its rapid growth, the e-commerce sector faces significant challenges, with one of the foremost being the intricate relationship between consumer trust and purchase intention (Yu et al., 2021). Simultaneously, social media have emerged as key platforms where people share their opinions and emotions. Effectively extracting consumer sentiment and gaining a deeper insight into the decision-making processes behind their purchases have become essential goals. Sentiment analysis, a technique in natural language processing (NLP), takes on an important role in achieving these

DOI: 10.4018/JOEUC.372206

objectives (Grover, 2022); it is used to extract sentiment information from text to assist companies in understanding consumers' emotional tendencies and attitudes. Concurrently, consumer purchase intent prediction represents a valuable tool for gaining insight into the consumer decision-making process. The integration of these two approaches can facilitate the acquisition of more comprehensive market insights and the optimization of marketing strategies.

Sentiment analysis, a significant NLP technique, is becoming a crucial tool for comprehending consumer behavior and market dynamics. Sentiment analysis is a technique used to extract sentiment information from text, identify the people's affective tendencies (e.g., positive, negative, or neutral), and categorize them. Moreover, sentiment analysis can be expanded to incorporate more advanced methods capable of detecting specific emotions, intentions, or nuanced emotional expressions (e.g., anger, happiness, or sarcasm), as well as identifying context-specific sentiments often found in areas like product reviews (Cambria et al., 2017). Sentiment analysis system architecture (Devika et al., 2016), including data collection, data cleansing, text coding, feature extraction, training models, and evaluating models, is shown in Figure 1. Basing their analysis on this overall system architecture, researchers can extract valuable information from textual content to personalize the sentiment expressed in the textual data.

Figure 1. Working process of sentiment analysis



In recent years, deep learning has advanced considerably in both social media and e-commerce fields. It is particularly effective in analyzing consumer behavior, conducting sentiment analysis, and predicting purchase intentions (Alzahrani et al., 2022). All these methods largely rely on the powerful self-learning capabilities of deep learning networks, including feature extraction, contextual modeling, and big data processing. These capabilities facilitate more effective extraction of emotional information from text and enhance the comprehension and prediction of consumer purchase decisions. Meanwhile, these research advancements come with a range of challenges and issues. For instance, the exponential increase in data volume, along with the diversity and complexity arising from multiple data sources, will require more advanced computational resources and greater processing power to support deep learning model. Furthermore, e-commerce product reviews comprise a multitude of emotional and semantic elements, necessitating the development of sophisticated deep learning models to effectively grasp and comprehend their nuances. Additionally, e-commerce is a

multidisciplinary field that requires the integration of expertise from various disciplines, providing enterprises with a broader range of solutions.

Previous studies have proposed many methods to solve the problem of sentiment analysis, especially in the field of deep learning algorithms, and common deep learning methods include long short-term memory (LSTM; Murthy et al., 2020), bidirectional gated recurrent unit (BiGRU; Xu et al., 2024), a combination of convolutional neural network (CNN) and LSTM (CNN-LSTM; Rehman et al., 2019), and a combination of bidirectional encoder representation from transformers (BERT; Kenton et al., 2019) and LSTM (BERT-LSTNet; Ma et al., 2024). These models have demonstrated effectiveness in text feature extraction and time sequence modeling, enabling them to accurately classify and predict the sentiments expressed in the text. For instance, LSTM models can process sequence data by capturing contextual dependencies in text and analyzing the impact of context on sentiment. Bi-GRU, a variant of the LSTM algorithm, captures both preceding and succeeding sequence information, thereby enabling a more comprehensive understanding of context and facilitating more accurate interpretation of contextual sentiments. In contrast, CNN-LSTM integrates a CNN to extract local text features that are then used to analyze sentiment information and model temporal sequences. Combining the strengths of CNNs for feature extraction with LSTMs for sequential processing enables users of this approach to benefit from enhanced performance in sentiment analysis tasks. In an alternative approach, the pretrained language model BERT is leveraged to extract text features that are then processed by the LSTNet model used for predicting purchase intentions and sentiment analysis. This method effectively captures fine-grained contextual sentiment features, leading to a more precise sentiment analysis. Although each of these methods demonstrates unique strengths in sentiment analysis, each one also shares a reliance on the LSTM structure to model temporal sequences, aiming to classify text features for sentiment classification and purchase prediction. However, users of existing sentiment analysis methods often struggle to accurately capture fine-grained emotional shifts in text, especially when expressing complex or multilayered emotions (e.g., anger, happiness, or sarcasm). Meanwhile, these methods generally lack the ability to adequately model dynamic changes in user preferences, making it difficult to reflect shifts in user needs in a timely manner. This limitation causes purchase intention predictions to lag behind actual behavior. These methods have limitations in processing multiscale temporal features and capturing detailed features of signals.

In this paper we aim to overcome the limitations of the aforementioned algorithms and enhance the accuracy of purchase prediction and sentiment analysis, while enabling real-time monitoring of the evolution of consumer purchasing behavior. To achieve this goal, we designed an integrated model that combines SoftMax, BERT, and deep bidirectional LSTM (DBLSTM; Nguyen et al., 2016). Initially, we used BERT to extract rich semantic and contextual information from textual comments. BERT has become a common method for text feature extraction and is particularly effective in text data analysis. The application of BERT embedding to the comment text enables the more effective capture of sentiment and key information, the identification of fine-grained model features, and an enhancement in the accuracy of sentiment analysis.

Second, we used a DBLSTM network to perform sentiment analysis and purchase intention prediction on the features output by the BERT model. The DBLSTM network is composed of stacked bidirectional LSTM (Bi-LSTM) units, providing robust time-series modeling capabilities and enabling the processing of multiple sequences simultaneously. The analysis of temporal features allows for the monitoring of consumer decision-making processes in real time. This monitoring is accomplished by analyzing the trend of purchase intention over time using multiscale temporal features. This approach also enables the exploration of the causal relationship between consumers' emotions and their purchase intentions. Finally, the SoftMax module analyzes the sentiment category of each review and outputs the probability of purchase intention, thereby enhancing model interpretability and providing a clearer understanding of its predictions. Our proposed BERT-DBLSTM-SoftMax model leverages the BERT model's advanced text analysis capabilities to capture fine-grained

sentiment signals across texts, employs the DBLSTM model to process multiscale temporal features, and uses SoftMax's probability output to improve interpretability. The BERT model, with its powerful pretrained language representation capabilities, can deeply understand the underlying semantics of text, and its bidirectional transformer structure effectively captures contextual information. This capability overcomes the limitations of traditional methods in handling fine-grained sentiment changes and complex emotional patterns. Meanwhile, through the application of bidirectional LSTM networks (specifically a DBLSTM) for time-series analysis, the model effectively captures the evolving patterns in user behavior and long-term dependencies, making it crucial for accurately forecasting user purchase intent. By combining BERT's deep semantic understanding with DBLSTM's time-series modeling advantages, our model overcomes the shortcomings of traditional sentiment analysis methods in handling complex emotional patterns and dynamic changes, significantly improving the precision of purchase intention prediction and sentiment analysis. By examining the temporal trends of purchase intent, the model provides a more profound insight into the progression of consumer purchasing decisions, offering valuable insights to support e-commerce decision-making.

In this paper we made several contributions. First, we introduced a novel integrated model, termed BERT-DBLSTM-SoftMax, which uniquely combines BERT with DBLSTM to more effectively capture sentiment information in text and accurately predict consumer purchase intentions.

We also proposed a DBLSTM model for sentiment analysis and purchase prediction. By taking advantage of its excellent bidirectional context capture capability and the ability to handle long-term dependencies, we were able to extract subtle sentiment signals from text and improve the algorithm performance of the model.

We extensively tested the integrated model across multiple datasets, with results showing outstanding performance in both sentiment analysis and purchase prediction tasks. The integrated model provides novel insights into the fields of text sentiment analysis and purchase intention prediction, establishes a new paradigm for modeling approaches to solve complex problems in other fields, and highlights the importance of model integration.

RELATED WORKS

Sentiment Analysis Research

Sentiment analysis is an important part of NLP (Guo, 2022). It needs to identify the sentiment polarity in the text content; that is, whether the sentiment expressed by the text content is neutral, negative, or positive (Farimani et al., 2022). Sentiment analysis is extensively applied in fields such as social media, e-commerce, and market research. In e-commerce, it takes on particular importance because companies can measure consumer sentiment and attitudes to inform product improvements and marketing strategies, and ultimately increase customer satisfaction.

In the past decade, the field of sentiment analysis has experienced a period of rapid development. The research methods in this field mainly include the following aspects. The initial methods were predominantly lexicon based. Chang & Lin (2011) put forth the use of a sentiment lexicon to annotate words in a text with their corresponding sentiment. This approach is straightforward and readily implementable; however, it is constrained by the scope of the sentiment dictionary and the polysemous nature of word meanings. The second method is a machine learning method that treats sentiment analysis as a classification task and uses common classification methods. The majority of previous research on sentiment analysis can be classified under these two categories. The third approach is primarily based on CNN (Liao et al., 2017), recurrent neural networks (Basiri et al., 2021), and other models (Jain et al., 2024), all of which utilize deep learning techniques to extract textual features and subsequently classify the features for sentiment classification. With advancements in large-scale model technology, recent research trends have focused on employing pretrained language models (e.g., BERT [Pota et al., 2021]) and generative pretrained transformers (Zhan et al., 2024) for sentiment analysis. These models are more adept at capturing textual semantic information. Additionally,

multimodal sentiment analysis, which employs multimodal data (e.g., text, images, and audio), offers the potential for more comprehensive sentiment recognition. Despite these advancements, however, sentiment analysis still faces many challenges, such as imbalanced sentiment polarity, ambiguity, lengthy text and noisy data, and multilingual sentiment. In addition, the performance of generalized sentiment analysis models is not suitable for all fields, so transfer learning has become a research focus in this field.

Purchase Intention Prediction Research

Research on purchase intent prediction holds significant value in the e-commerce sector. By understanding consumers' purchase intentions as they browse products or services, companies can refine their marketing strategies, deliver personalized recommendations, and improve sales conversion rates. Predicting purchase intent allows e-commerce platforms to gain deeper insights into consumer needs, thereby enabling a more tailored shopping experience (Yuan et al., 2022). This approach is critical to driving sales growth and improving user satisfaction.

Purchase prediction is an important research field that focuses on predicting users' future purchasing behavior by analyzing user behavior, demographic information, historical purchase records, and other relevant features. With the development of e-commerce and digital marketing, purchase prediction has become a popular research area. In current research, techniques and methods for purchase prediction can be categorized into several broad groups. For example, early purchase prediction relied heavily on traditional statistical methods (e.g., decision trees, logistic regression, and linear regression). These methods are usually modeled by some human-designed key features (e.g., users' historical purchase frequency, average purchase amount, etc.). These methods are characterized by strong interpretability, which can help researchers intuitively understand the factors affecting purchase behavior, but they have certain limitations in dealing with complex nonlinear relationships and large-scale data.

Another method is based on machine learning algorithms, such as gradient boosted decision tree (Wu & Li, 2022), random forest (Joshi et al., 2018), K-nearest neighbor (Huang et al., 2018), and support vector machine (Saranya et al., 2020). These methods can handle multidimensional features and significantly outperform traditional methods in terms of prediction accuracy. However, they usually require extensive feature engineering work. In addition, these methods are challenging for real-time prediction of large-scale dynamic data. With the increase of data volume and the progress of neural networks, deep learning models are gradually applied in purchase prediction. Deep learning models can automatically learn complex features and achieve better results in large-scale data, especially in dealing with complex nonlinear relationships and large-scale data with significant advantages. Commonly used deep learning models include CNN (Sun, 2022), recurrent neural network (Sheil et al., 2018), LSTM (Sakar et al., 2019) and its variants, multilayer perceptron, and self-attention model (Huang et al., 2021).

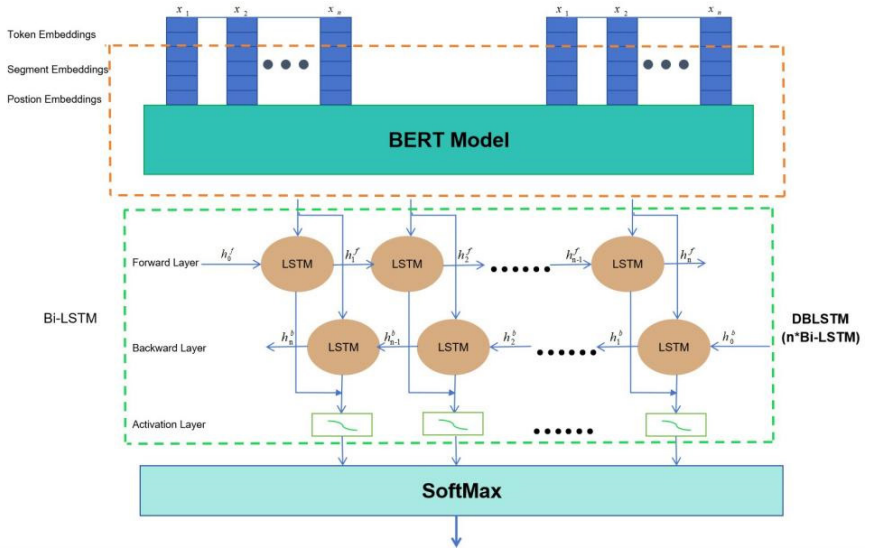
In recent years, purchase prediction based on pretrained language models has progressively become a major research topic. Especially in the field of e-commerce and social media. Pretrained language models can better capture contextual and semantic information, enabling the models to obtain valuable information from text data, such as user reviews and product descriptions. Such models further extend on the basis of sentiment analysis to recognize users' potential purchase intentions and provide more accurate input features for purchase prediction. With the development of smart devices, purchase prediction gradually adopted multimodal information fusion technology, which improves the prediction effect by combining multiple information sources (e.g., text, image, and audio) from users. Meanwhile, purchase prediction has a close relationship with recommender systems, and many studies have combined the two to achieve personalized recommendations and purchase tendency prediction for users. Recommendation methods based on collaborative filtering and matrix decomposition, when combined with deep learning, are able to better understand user preferences and apply them to purchase prediction. In addition, the

application of graph neural networks (Zhou & Hudin, 2024) in social networks also brings a new perspective to purchase prediction, which is able to predict potential purchase behaviors through the relationship graph between users.

METHODOLOGY

In this paper we propose an integrated model that includes BERT, DBLSTM, and SoftMax aimed at leveraging the strengths of each component to enhance accuracy and real-time performance of purchase intention prediction and sentiment analysis. We also examined our model's ability to monitor the relationship between sentiment features and purchase intention in product reviews in real time. The BERT module extracts sentiment features from text using its powerful semantic understanding and context capture capabilities. It can decode intricate contextual data and deliver more accurate and refined features for purchase intention and sentiment analysis. Moreover, the DBLSTM module forecasts time series data and examines the progression of purchase intentions as they shift over time. The model also can track consumers' decision-making process in real time, helping to clarify how consumer sentiment influences purchase intent. The SoftMax module processes the output of the DBLSTM module and predicts the sentiment category and purchase intention probability of each comment. This capability enhances the interpretability of the model and facilitates a more comprehensive understanding of its predictions. The integrated model structure is shown in Figure 2.

Figure 2. The integrated model structure



Note. BERT = bidirectional encoder representation from transformers, BiLSTM = bidirectional long short-term memory, DBLSTM = deep bidirectional long short-term memory, and LSTM = long short-term memory.

Overall, our integrated model innovatively combines the powerful text representation capability of BERT and the time series processing capability of DBLSTM, which is able to effectively grasp the emotional nuances in text and reliably forecast consumers' purchase intent. The introduced DBLSTM model can capture the subtle features of sentiment analysis and can also make in-depth predictions and track the temporal trends in purchase intentions. It uncovers the shifting patterns in consumers'

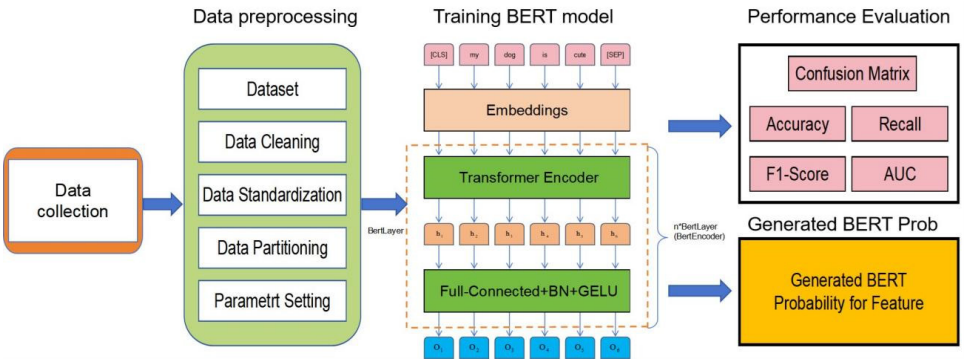
purchasing decisions, offering more comprehensive insights for e-commerce decision-making. This integrated model holds great potential for application in e-commerce review analysis, offering businesses deeper insights into consumer preferences, enhancing product and service quality, and strengthening competitive advantage. Furthermore, through the integration of multiple components, our model is versatile across different e-commerce contexts, delivering valuable support for informed decision-making.

BERT Model

BERT is a pretrained language model proposed by Google AI in 2018. It uses deep learning technology to improve text comprehension capabilities, with a particular emphasis on capturing context and semantic nuances, and it has been widely applied across NLP domains. BERT uses the encoder part of the transformer architecture to focus on the contextual understanding of text input. Unlike traditional unidirectional models, BERT uses a bidirectional training method, allowing it to take into account the context on both sides of a word, thus leading to a more comprehensive semantic representation. BERT's bidirectional nature means that, when analyzing a word, the model considers not only the preceding words but also those that follow. This design enhances the model's ability to interpret shifts in word meaning. BERT employs a self-attention mechanism to assess relationships between each word and other words within the input sequence. This mechanism enables the model to dynamically focus on context relevant to each word, thereby strengthening its feature extraction capabilities.

The BERT model is mainly composed of input representation and encoder and pretraining tasks. Its work flow chart is shown in Figure 3. Input representation means word embedding, positional embedding, and participle embedding of the input sentences. The encoder is composed using multiple stacked transformer encoder layers, each of which includes a self-attention mechanism and a feed-forward neural network capable of extracting text features at multiple levels.

Figure 3. Workflow chart of the BERT model



Note. AUC = area under the curve, BERT = bidirectional encoder representation from transformers, BN = batch normalization, and GELU = Gaussian error linear unit.

Before introducing the model, we will first explain the principle of the self-attention mechanism, which is composed of three components: queries (Q), keys (K), and values (V). The dimensions of both queries and keys are d_k, and the dimension of value is d_v. Here, d_k and d_v are the dimensions of keys and values by the linear projection, respectively.

The self-attention mechanism (Pan et al., 2024) can be represented as shown in equation (1).

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

In equation (1), Softmax(·) denotes the function normalized by Softmax. Furthermore, the multi-head (o heads) attention can be calculated as shown in equation (2).

$$MutliHead(Q,K,V) = W_{out}Concat\left(hea\,d_1,\ldots,hea\,d_n\right), \tag{2}$$

In this equation, $\text{head}_i = \text{Attention}\left(QW_i^Q, KW_i^K, VW_i^V\right), i \in \{1,...,o\}$, and $W_{\text{out}} \in R^{od_v \times d_{\text{model}}}$ (d_{model} is the dimension of model input) is the projection of final linear layer. Here $W_i^Q \in R^{d_{\text{model}} \times d_k}$, $W_i^K \in R^{d_{\text{model}} \times d_k}$ and $W_i^V \in R^{d_{\text{model}} \times d_v}$ are the projections by the linear layer, respectively. This self-attention mechanism can focus on every position within the text and capture dependencies over long distances. Furthermore, multi-head attention is calculated in parallel mode to reduce time consumption.

The feed-forward neural network employs two fully connected layers that act on the output vector from the self-attention mechanism. The formula is shown in equation (3).

$$\text{FFN}(x) = \text{ReLU}\left(xW_1 + b_1\right)W_2 + b_2, \tag{3}$$

In equation (3), x represents the input vector, while W_1 and b_1 denote the weight matrix and bias vector for the first fully connected layer in the feed-forward neural network. Similarly, W_2 and b_2 correspond to the weight matrix and bias vector for the second fully connected layer. ReLU is the activation function.

The pretraining tasks for the BERT model primarily include two components: the masked language model and next sentence prediction. These tasks can be represented by equations (4) and (5).

Equation (4) shows the formula for the masked language model.

$$P(x) = \prod_{i=1}^{n} P\left(x_i | x_{<i,}\theta\right)^{m_i}, \tag{4}$$

In this equation, $P(x)$ represents the target probability distribution, m_i indicates whether the i-th token is masked, and $x_{<i}$ denotes all tokens preceding the i-th token.

Equation (5) shows the formula for the next sentence prediction.

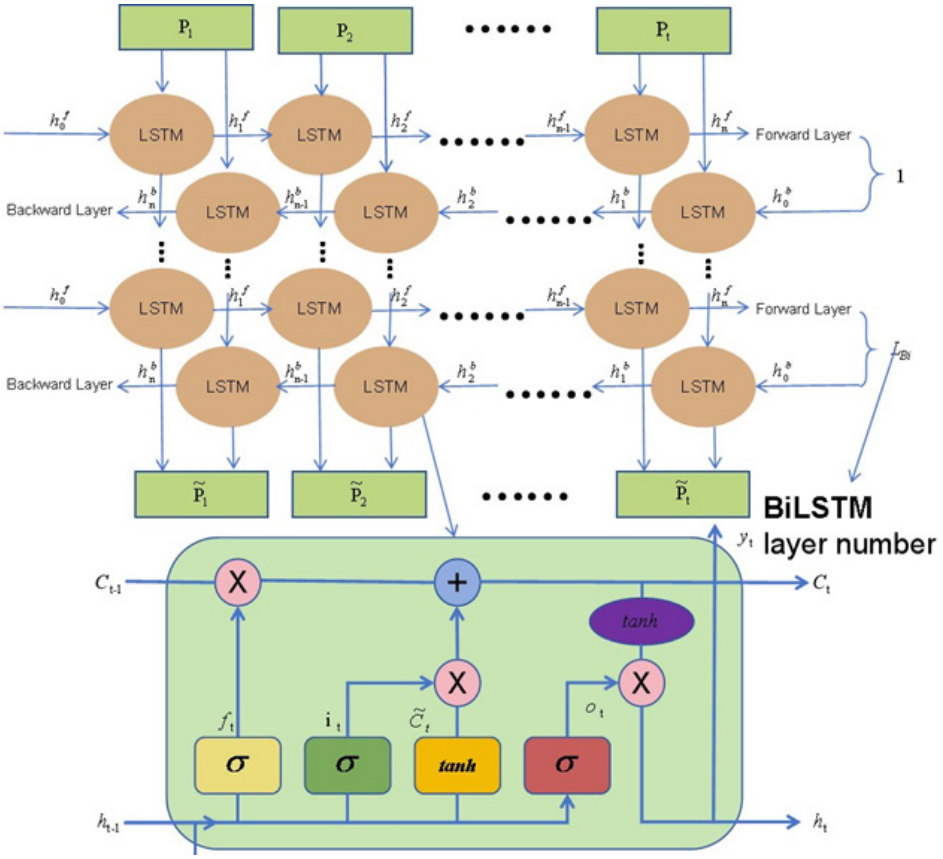
$$P_{\text{NSP}}(s,t) = \frac{exp(\varphi(s,t))}{exp(\varphi(s,t)) + exp\left(\varphi(s,t_{\text{rand}})\right)}, \tag{5}$$

In equation (5), s and t represent two individual sentences, $P_{\text{NSP}}(s,t)$ denotes the vector representation of the sentence pair, and t_{rand} indicates randomly selected sentences from other documents.

DBLSTM Module

DBLSTM stacks multiple BiLSTM layers (the structure is shown in Figure 4). Each BiLSTM layer contains multiple LSTMs connected by bidirectional rules. Unlike vanilla LSTM, BiLSTM combines forward and backward propagation. Consider the DBLSTM module of nth encoder layer as an example. For a single LSTM, which consists of a forgetting gate f_t, an input gate i_t, and an output gate o_t, the f_t consists of previous layer out f_{t-1} and current input p_t with weights W_f, $\widetilde{W}_f$ and bias b_f, which can be calculated as shown in equation (6). This formula is also in line with the study by Yu et al. (2022).

Figure 4. Overall network structure of the DBLSTM module



Note. BiLSTM = bidirectional long short-term memory, LSTM = long short-term memory.

$$f_t = \sigma\left(W_f p_t + \widetilde{W}_f h_{t-1} + b_f\right) \tag{6}$$

In equation (6), σ is the sigmoid activation function. Then, the input vector fed into i_t can be calculated as shown in equation (7).

$$i_t = \sigma\left(W_i p_t + \widetilde{W}_i h_{t-1} + b_i\right), \tag{7}$$

In equation (7), W_i, $\widetilde{W}_i$ and b_i are the weights and biases of the input gates. Meanwhile, the LSTM has a status update unit, which can be calculated as shown in equation (8).

$$\tilde{c}_t = \lambda\left(W_c p_t + \widetilde{W}_c h_{t-1} + b_c\right) \tag{8}$$

In equation (8), λ is the tanh active function. W_c, $\widetilde{W}_c$, and b_c are the weights and biases of the status update unit. The data update can be calculated as shown in equation (9).

$$c_t = f_t \otimes c_{t-1} + i_t \otimes \tilde{c}_t \tag{9}$$

In equation (9), ⊗ stands for the operation of element-wise multiplication. Then, the updated data enters o_t as shown in equations (10) and (11).

$$o_t = \sigma\left(W_o p_t + \widetilde{W}_o h_{t-1} + b_o\right) \tag{10}$$

$$h_t = o_t \otimes \lambda(c_t) \tag{11}$$

In these equations, h_t is the final output. W_o, $\widetilde{W}_o$, and b_o are the weights and biases of the output gate, respectively.

During the training of a single BiLSTM layer, the forward output $h_t^{cell,f}$ and the backward output $h_t^{cell,b}$ are initially computed using equation (11). Subsequently, the output of the BiLSTM layer is obtained as shown in equation (12).

$$h_{t,L_{Bi}}^{cell,Bi} = \sigma_{f,b}\left(h_{t,L_{Bi}}^{cell,f}, h_{t,L_{Bi}}^{cell,b}\right) \tag{12}$$

In equation (12), the function $\sigma_{f,b}$ represents the concatenate merge mode. Additionally, the output of the L_{Bi}th BiLSTM layer is expressed as shown in equation (13).

$$h_{t,L_{Bi}}^{cell,Bi} = \sigma_{f,b}\left(h_{t,L_{Bi}}^{cell,f}, h_{t,L_{Bi}}^{cell,b}\right) \tag{13}$$

The DBLSTM module can extract features of the input data by stacking BiLSTM. Unlike the fully connected layers in the vanilla transformer, this design can further improve the prediction performance of the proposed algorithm.

SoftMax Module

The SoftMax module is a widely used classification technique. The core idea is to perform multiclass classification by transforming input values into likelihood scores for each category. In this approach, each element of the input vector undergoes an exponential transformation through the SoftMax function, followed by a normalization step. This process ensures that each element indicates the likelihood of associating with a particular category. The SoftMax module produces a probability distribution over the classes for the classification task. In this study, we applied SoftMax to transform the features from the DBLSTM layer via the fully connected layer and then converted the output layer for sentiment analysis and purchase intention prediction into a probability distribution. Specifically, SoftMax converts the output layer of sentiment analysis into a probability distribution, determining the likelihood that the sentiment polarity expressed in the comment is positive or negative. Similarly, it converts the output layer of purchase intention prediction into probability to observe the user's satisfaction with the product more intuitively. The score matrix is converted into specific weight values through the Softmax function, allocating unique weights to each element as shown in equation (14).

$$\eta_n = \frac{\exp(e_n)}{\sum_{n=1}^{N_n} \exp(e_n)} \tag{14}$$

EXPERIMENT

Datasets

For this paper we used four datasets: the Twitter Sentiment Analysis Dataset (Wagh & Punde, 2018), the Yelp Review Dataset (Pondel et al., 2021), the Amazon Product Review Dataset (Alsubari et al., 2021), and the IMDB Movie Reviews Dataset (Sarker et al., 2022).

Twitter Sentiment Analysis Dataset

The Twitter Sentiment Analysis Dataset focuses on sentiment expression by users on the social media platform Twitter (now X). This dataset is valuable for sentiment analysis, public opinion monitoring, and social media research, enabling researchers to examine user sentiment across various topics. The Twitter Sentiment Analysis Dataset contains thousands of tweets, sentiment-related hashtags, and associated topic information, providing valuable perspectives on user emotions, buying intentions, and social media trends. This information makes it especially well suited for extensive sentiment analysis research.

Yelp Review Dataset

The Yelp Review Dataset is a well-known collection focused on the restaurant and entertainment sectors, containing customer feedback on locations like restaurants, bars, and cafes. This dataset generally includes millions of reviews, offering valuable user feedback for sentiment analysis and market research. Collected over several years, the Yelp Review Dataset encompasses reviews of various types of venues across diverse regional and cultural settings, allowing researchers to examine user behavior and sentiment across different environments.

Amazon Product Review Dataset

The Amazon Product Review Dataset is a large and varied compilation of authentic user feedback from Amazon.com, establishing it as a crucial asset for machine learning, natural language understanding, and sentiment evaluation. This extensive dataset comprises millions of reviews gathered over multiple years, spanning several decades and covering a broad range of product categories, such as electronics, books, apparel, and food. It also contains important details, such as product categories, user ratings, and product identifiers, thus offering valuable context for more in-depth research.

IMDB Movie Reviews Dataset

The IMDB Movie Reviews Dataset is dedicated to movie reviews, containing user feedback and ratings across a range of films. This dataset is crucial for examining consumer sentiment and purchase intent within the film industry. It contains an extensive collection of reviews, user ratings, and extra details, such as movie titles and genres. Analyzing the sentiment within these reviews offers meaningful insights for industry decision-making. Although its data volume is smaller compared with certain other datasets, it offers a rich source of review content well suited for sentiment analysis research. Furthermore, the dataset spans an extensive time period, including reviews from early films to recent releases, which enables researchers to study shifts in user sentiment over time.

Data Preprocessing

In the data preprocessing during our research, we completed three phases: data cleaning, data standardization, and data partitioning. We discuss each of these phases in this section.

Data Cleaning

In the data cleaning phase, we addressed potential inconsistencies and outliers to ensure high data quality. The procedure involves eliminating duplicate entries, addressing missing data, and removing irrelevant or distorted text. Additionally, we standardized text formatting by applying normalization methods, including spelling correction, root word extraction, and removing non-alphanumeric characters. The specific steps we completed were removing noisy text, removing duplicate comments, and normalizing text.

Removing noisy text is a crucial step in data preprocessing that is aimed at improving data quality and enhancing model performance. In this process, we first used regular expressions or other preprocessing tools to identify and remove garbled characters, ensuring that the text did not contain any

unrecognized or meaningless symbols. Next, for nonstandard language, we employed spaCy and Natural Language Toolkit to clean the text by removing spelling errors, grammatically incorrect words, and slang or nonstandard abbreviations that do not conform to standard language usage. Additionally, to further eliminate irrelevant information, we used regular expressions or HTML parsing tools (e.g., BeautifulSoup) to remove ads, links, email addresses, and other non-text content, ensuring that the data used contain only information relevant to the task.

To remove duplicate comments, we employed two methods: exact matching and approximate matching. For identical comments, we directly identified and removed duplicates through hash algorithms or string comparison, ensuring that each comment was unique and preventing repeated comments from adversely affecting model training. For comments that were similar, but not exactly the same, we used text similarity algorithms (such as Jaccard similarity, cosine similarity, etc.) to calculate the similarity between comments. If the similarity exceeded a predefined threshold, the comment was considered a duplicate and removed. This method is particularly useful for comments that express similar ideas, but with different wording, helping to further reduce redundant data, ensuring diversity and representativeness in the dataset, and ultimately improving the model's accuracy and generalization ability.

In the data cleaning phase, we also performed text normalization to ensure consistency in the text data. The specific operations included the following:

- Spelling correction: Automatically detect and correct spelling errors in the text using the spaCy tool to reduce the impact of spelling issues on model analysis.
- Stemming: Reducing words to their base form so that different word forms (e.g., speaking and spoke) are treated uniformly, enhancing the model's understanding of the same concept.
- Removing special characters: Removing special characters from text is a key step in text preprocessing. Common methods include using regular expressions to remove nonalphabetic, nonnumerical characters and punctuation; using BeautifulSoup to remove HTML tags; and leveraging spaCy and Natural Language Toolkit to remove punctuation and stop words. Additionally, regular expressions can be used to remove URLs, email addresses, and emojis, ensuring the purity of the text data to improve the accuracy of the analysis.

Data Standardization

Data standardization is a critical step to ensure consistency in the data. During the data standardization phase, we processed both text data and numerical features to ensure they had a consistent format and encoding. This processing involved two operations: text data standardization and numerical feature standardization.

In the text standardization phase, we ensured that all comment texts had a consistent format and encoding. Specifically, this phase included the following steps:

- Converting all text to lowercase to ensure the model treats each word consistently, regardless of letter case
- Standardizing text encoding to avoid errors or loss of information owing to inconsistent encoding formats

In this study, during the numerical feature standardization operations, for datasets containing numerical features, we applied the Z-score normalization method to standardize the numerical features. Z-score normalization transforms the values of each feature to a distribution with a mean of 0 and a standard deviation of 1, ensuring that all numerical features have the same scale. This process eliminates the impact of differences in value ranges on model training.

Data Partitioning

To assess the algorithm's performance more accurately, we split the dataset into three parts: a training set, a validation set, and a test set. The preprocessed data were divided in an 8:1:1 ratio for training, validation, and testing, respectively.

In the experiments, data preprocessing was essential to ensure that deep learning models could learn efficiently and generalize well. By performing data cleansing, standardization, and partitioning, we transformed raw data into a format optimized for model training and evaluation. This approach reduced noise, improved model effectiveness, and enhanced the reliability of the experimental results.

Experimental Environment

Hardware Environment

The system was equipped with an Intel Core i7-6700K central processing unit, 256 GB of RAM, 8 terabytes of solid-state drive storage, and eight NVIDIA Tesla V100S graphic processing units.

Software Environment

The system runs on Ubuntu 20.04 LTS, and the algorithms were developed in Python using the PyTorch 1.13.1 deep learning framework, with graphic processing unit acceleration provided by CUDA 11.2.

Model Training and Evaluation

Before model training, we needed to initialize and set the network structure and its parameters in the integrated model. Our parameter settings were as follows: data enhancement strategies, BERT model parameter settings, DBLSTM module parameter settings, optimizer and learning rate strategy, and regulation and loss function.

For BERT model parameter settings, we selected BERT-base-uncased as the pretrained model, with the following parameter settings: vocabulary size of 30,522, 12 heads for the multi-head self-attention mechanism, 12 transformer layers, a maximum sequence length of 256, and an embedding dimension of 768.

For DBLSTM module parameter settings, we set the number of BiLSTM layers to five, with each BiLSTM layer containing five LSTM layers, and a hidden layer dimension of 128.

For optimizer and learning rate strategy, we employed the Adam optimizer for network optimization, with the learning rate adjusted using cosine annealing. We set the initial learning rate to 0.001, with a decay coefficient of 0.000001. During training, we configured the batch size to 64, and epochs were 100.

For regularization and loss function, we applied L2 regularization as the chosen regularization technique, with cross-entropy serving as the loss function.

The model training process consisted of several key stages. First, we divided the dataset into a training set, validation set, and test set. The training set was used to train the model, the validation set assessed the model's performance during training, and the test set verified the model's generalization ability. Next, we conducted model training and parameter tuning. First, we needed to select the loss function. For this study, we selected the cross-entropy loss function owing to its superior performance in multiclass classification tasks. Cross-entropy effectively measures the discrepancy between the predicted probability distribution and the true labels, making it particularly suitable for tasks such as sentiment analysis and purchase intention prediction. These tasks often involve multiclass classification, where each review or user behavior can be classified into multiple sentiment or purchase intention categories. Thus, the cross-entropy loss function ensured that the model accurately assessed the predicted probability for each category, which in turn enhanced the precision of sentiment analysis and the accuracy of purchase intention prediction.

Second, we performed hyperparameter tuning. For the learning rate, we selected an initial value of 0.001 through experimental tuning and applied a cosine annealing strategy to adjust it, aiming to balance the convergence speed and stability during training. The learning rate choice was primarily based on empirical values and best practices from similar tasks. We set the batch size to 64, which was chosen experimentally to balance memory use and training efficiency. For the hidden layer dimension of the DBLSTM, we set it to 128, a value selected after multiple trials, in an effort to balance the model's representational capacity and computational complexity. In the process of selecting these hyperparameters, we also considered methods such as grid search and random search to further optimize the model's performance.

For model evaluation, we used common metrics such as accuracy, recall, and F1 score to comprehensively assess the model's performance. The definitions and formulas for these metrics are provided in equations (15)–(18).

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{15}$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \tag{16}$$

$$\text{Pr}ecision = \frac{TP}{TP + FP} \tag{17}$$

$$\text{F1 Score} = \frac{2 \times \text{Recall} \times \text{Precision}}{\text{Recall} + \text{Precision}} \tag{18}$$

In these equations, TN refers to the count of true negative sentiment instances, TP (true positives) refers to the count of true positive sentiment instances, FN refers to the count of false negative sentiment instances, and FP (false positives) refers to the count of false positive sentiment instances. To better reflect the model's performance in different sentiment classification and purchase intention prediction tasks, we conducted experiments on multiple datasets, validating the model's stability and robustness.

In sentiment analysis tasks, the accuracy and discriminative ability of the model are crucial, especially when dealing with imbalanced datasets. Traditional evaluation metrics such as accuracy may not fully reflect the true performance of the model. Therefore, we also used area under the curve (AUC) as a key metric to evaluate the sentiment analysis model. AUC is the area beneath the receiver operating characteristic curve, which shows the relationship between the true positive rate and false positive rate at various threshold settings. The formulas are shown in equations (19) and (20).

$$TPR = \frac{TP}{TP + FN} \tag{19}$$

$$\text{FPR} = \frac{\text{FP}}{\text{FP} + \text{TN}} \tag{20}$$

In these equations, FN = false negative, FP = false positive, FPR = false positive rate, TN = true negative, TP = true positive, and TPR = true positive rate.

Comparative Models

To highlight the superior performance of the proposed model in sentiment analysis prediction, we conducted comparisons against several existing models: BiLSTM (Dong et al., 2020), BiGRU (Kavatagi & Adimule, 2021), CNN-BiLSTM (Wankhade et al., 2024), CNN-BiGRU (Shan, 2023), BERT-BiLSTM (Li et al., 2022), BERT-BiGRU (Liu et al., 2020). All these models have hyperparameters appropriately tuned based on previous experience. To evaluate performance, we trained and tested these models on the same dataset. The performance comparison involved several

metrics, including F1 score, AUC, recall, and accuracy, to assess whether the BERT-DBLSTM-Softmax model offers a notable advantage.

Experimental Results and Analysis

Before evaluating the performance metrics of the model algorithms, we first provide a comparison of some fundamental indicators between our proposed model and the baseline models. These indicators include model parameters, floating-point operations per second, training time, and inference time. This comparison offers a clearer understanding of each model's hardware requirements and computational demands. The specific metrics are shown in Table 1. By examining these basic indicators, we can gain insight into each model's hardware adaptability, computational complexity, and real-time processing capabilities, laying the foundation for the subsequent performance evaluation.

Table 1. Model parameters of different methods

Method	Parameters (M)	Inference time (ms)	Training time (s/iteration)	FLOPS (G)
BiLSTM	51.48	3.92	0.051	16.51
BiGRU	44.65	3.65	0.049	15.63
CNN-BiLSTM	93.54	12.48	0.159	91.48
CNN-BiGRU	87.37	9.47	0.123	65.42
BERT-BiLSTM	119.32	4.43	0.063	29.03
BERT-BiGRU	116.53	4.21	0.059	27.54
Ours	123.67	5.38	0.077	38.47

Note. BERT-BiGRU = bidirectional encoder representation from transformers-bidirectional gated recurrent unit, BERT-BiLSTM = bidirectional encoder representation from transformers-bidirectional long short-term memory, BiLSTM = bidirectional long short-term memory, BiGRU = bidirectional gated recurrent unit, CNN-BiLSTM = convolutional neural network-bidirectional long short-term memory, CNN-BiGRU = convolutional neural network-bidirectional gated recurrent unit, FLOPS = floating-point operations per second.

To verify the effectiveness of the algorithm model, we conducted a large number of experiments. We also compared the competitive model with the evaluation metrics in multiple dimensions. The model experimental results are shown in Table 2. In comparing the performance of the BiLSTM and BiGRU algorithms across four datasets, we observed that BiLSTM consistently achieved slightly better results than BiGRU. Specifically, on the Amazon Product Review Dataset, BiLSTM outperformed BiGRU with improvements of 0.22% in accuracy, 1.74% in recall, 1.8% in F1 score, and 2.31% in AUC. On the Yelp Review Dataset, BiLSTM also showed higher accuracy, F1 score, and AUC, with increases of 0.96%, 0.45%, and 0.41%, respectively, although its recall rate was marginally lower. We observed similar trends on the IMDB Movie Reviews and Twitter Sentiment Analysis datasets, where BiLSTM demonstrated slightly better overall performance across all evaluation metrics compared with BiGRU. To demonstrate these performance comparisons more clearly, we further visualized these data, as shown in Figure 5.

Table 2. Comparison of different methods in different indicators from various datasets

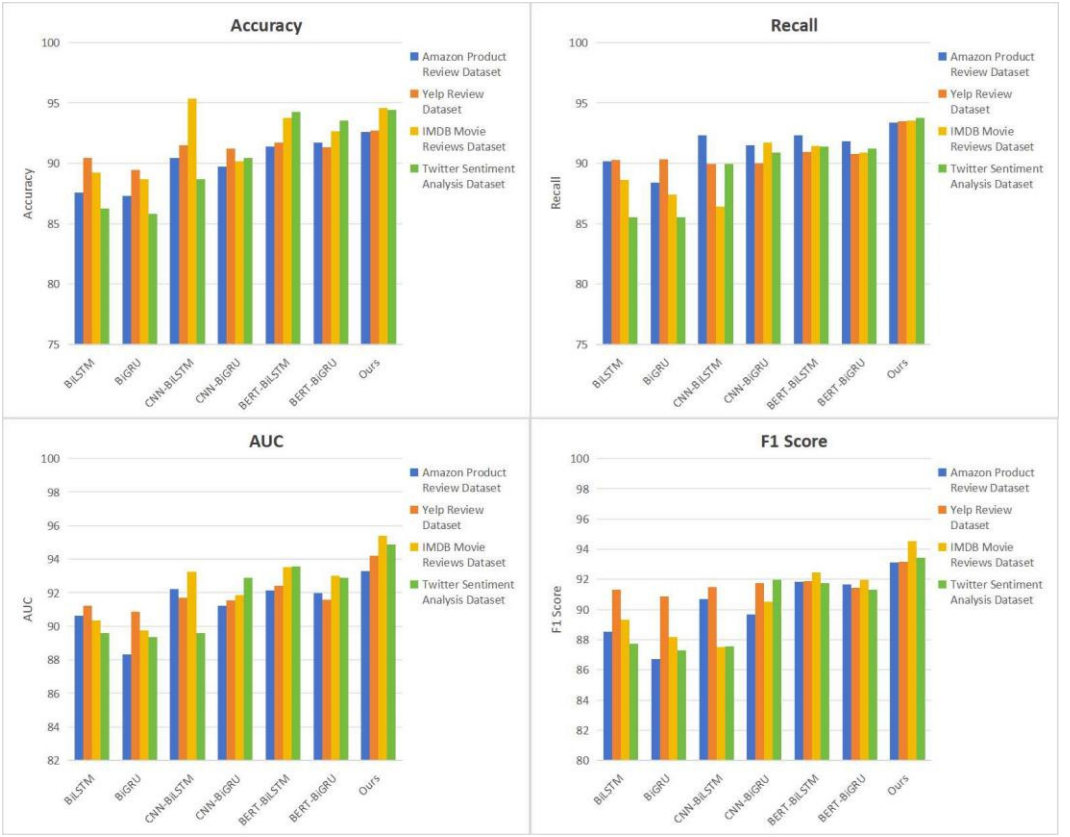
(a) Amazon Product Review Dataset and Yelp Review Dataset								
Method	**Amazon Product Review Dataset**				**Yelp Review Dataset**			
	Accu-racy	**Recall**	**F1 score**	**AUC**	**Accuracy**	**Recall**	**F1 score**	**AUC**
BiLSTM	87.55	90.17	88.52	90.63	90.43	90.27	91.32	91.21
BiGRU	87.31	88.43	86.72	88.32	89.47	90.31	90.87	90.87
CNN-BiLSTM	90.47	92.31	90.71	92.21	91.47	89.96	91.47	91.71
CNN-BiGRU	89.73	91.49	89.68	91.23	91.21	90.03	91.73	91.54
BERT-BiLSTM	91.41	92.32	91.85	92.13	91.74	90.94	91.87	92.41
BERT-BiGRU	91.74	91.84	91.64	91.99	91.32	90.76	91.43	91.58
Ours	92.58	93.38	93.12	93.27	92.73	93.47	93.14	94.21
(b) IMDB Movie Reviews Dataset and Twitter Sentiment Analysis Dataset								
Method	**IMDB Movie Reviews Dataset**				**Twitter Sentiment Analysis Dataset**			
	Accuracy	**Recall**	**F1 score**	**AUC**	**Accuracy**	**Recall**	**F1 score**	**AUC**
BiLSTM	89.24	88.62	89.33	90.33	86.23	85.51	87.72	89.58
BiGRU	88.67	87.43	88.19	89.76	85.81	85.54	87.31	89.37
CNN-BiLSTM	95.35	86.44	87.53	93.24	88.68	89.96	87.56	89.58
CNN-BiGRU	90.17	91.73	90.51	91.85	90.47	90.87	91.96	92.89
BERT-BiLSTM	93.76	91.45	92.47	93.51	94.24	91.41	91.76	93.56
BERT-BiGRU	92.65	90.87	91.96	93.03	93.54	91.21	91.32	92.89
Ours	94.58	93.54	94.52	95.41	94.41	93.78	93.43	94.87

Note. AUC = area under the curve, BERT-BiLSTM = bidirectional encoder representation from transformers-bidirectional long short-term memory, BiLSTM = bidirectional long short-term memory, BiGRU = bidirectional gated recurrent unit, CNN-BiLSTM = convolutional neural network-bidirectional long short-term memory, CNN-BiGRU = convolutional neural network-bidirectional gated recurrent unit.

We extended this analysis to other model pairs, comparing CNN-BiLSTM with CNN-BiGRU and BERT-BiLSTM with BERT-BiGRU. In both cases, the BiLSTM-based models showed a slight performance advantage over their BiGRU counterparts. This difference is likely due to the more complex structure of LSTM, which includes three gating mechanisms—input gate, forget gate, and output gate—and maintains both cell state and hidden state variables in its hidden layer. This architecture enables LSTM to better capture and retain long-distance dependencies within the data, especially important in tasks that require understanding subtle contextual semantics. By leveraging this gating mechanism, BiLSTM is able to more effectively maintain semantic coherence and capture long-range dependencies, leading to improved performance in sentiment analysis tasks where nuanced emotional understanding is crucial. At the same time, we compared CNN-BiLSTM and BiLSTM algorithms and found that CNN-BiLSTM has higher accuracy, recall, F1 score, and AUC than BiLSTM on four datasets. Similarly, compared with CNN-BiGRU and BiGRU algorithms, CNN-BiLSTM has slightly higher overall performance than BiGRU. This is due to the complementary advantages of CNN and BiLSTM, which enables CNN-BiLSTM to more effectively extract local features of text and capture global semantic dependencies, allowing CNN-BiLSTM to reduce redundant features while retaining important emotional information, thereby more efficiently understanding the emotional expression of text. In comparing the BERT-BiLSTM and CNN-BiLSTM algorithms, we found that BERT-BiLSTM consistently achieved higher accuracy, recall, F1 score, and AUC on the dataset. Similarly, when comparing BERT-BiGRU with CNN-BiGRU, we found that BERT-BiGRU

demonstrated slightly better overall performance. This performance difference is primarily due to BERT's architecture and pretraining approach. BERT's multilayer bidirectional transformer structure encodes each word in both directions, allowing it to capture more comprehensive contextual information. Additionally, BERT is pretrained on a large-scale corpus, equipping it with extensive vocabulary and deep semantic knowledge. This pretrained knowledge significantly enhances BERT's ability to recognize subtle emotional expressions, making it particularly effective for fine-grained sentiment analysis tasks.

Figure 5. Comparison of the evaluation indicators of the methods on different datasets



Note. BERT-BiGRU = bidirectional encoder representation from transformers-bidirectional gated recurrent unit, BERT-BiLSTM = bidirectional encoder representation from transformers-bidirectional long short-term memory, BiGRU = bidirectional gated recurrent unit, BiLSTM = bidirectional long short-term memory, CNN-BiGRU = convolutional neural network-bidirectional gated recurrent unit, CNN-BiLSTM = convolutional neural network-bidirectional long short-term memory.

Tables 1 and 2 clearly show that, despite having more parameters and longer inference times than BiLSTM and BiGRU, our model achieves significantly higher performance metrics. Even when scaled to a comparable model size, our model's performance still surpasses that of the best existing models. In summary, our proposed model demonstrates clear performance advantages in sentiment analysis tasks, consistently improving accuracy, recall, F1 score, and AUC across various datasets. Supported by a more complex network structure and enhanced feature extraction

capabilities, the model exhibits robust emotion recognition and classification abilities. It significantly outperforms BiLSTM, BiGRU, CNN-BiLSTM, and CNN-BiGRU and shows even stronger results when compared with similarly sized BERT-based models. These findings indicate that, despite the higher parameter count and increased inference time, the proposed model effectively enhances sentiment analysis performance, particularly excelling in tasks that require a nuanced understanding of emotional expressions. This capability makes the model a promising solution for fine-grained sentiment analysis.

CONCLUSION AND DISCUSSION

The BERT-DBLSTM-SoftMax model combines BERT's pretrained language representation capabilities with DBLSTM's memory network, Successfully addressing the shortcomings of traditional approaches in capturing contextual nuances and long-term relationships. The bidirectional transformer design of BERT allows for an in-depth understanding of the contextual relationships within the text, while DBLSTM, through bidirectional learning and long-term memory retention, precisely captures emotional and semantic nuances, enhancing the model's sentiment understanding and prediction capabilities. Compared with traditional single-model sentiment analysis methods, the BERT-DBLSTM-SoftMax model significantly improves prediction accuracy by combining BERT's powerful contextual understanding with DBLSTM's fine-grained sequence modeling. According to experimental findings, the model outperforms conventional approaches on multiple key metrics. The model not only breaks through the limitations of traditional sentiment analysis methods, improving prediction accuracy and robustness, but also demonstrates its potential for multi-domain sentiment analysis tasks through a more complex network architecture and deeper emotional feature extraction. In practical applications, this advantage provides businesses with more precise insights into user behavior, helping e-commerce platforms and social media analyze consumer emotional dynamics and purchase intentions, leading to more accurate marketing decisions and recommendations. By improving sentiment analysis accuracy, the model can better identify users' emotional tendencies, whether in product reviews, social media discussions, or online customer service interactions, helping businesses deliver a personalized user experience. In terms of purchase intention prediction, the model can more accurately capture users' purchase intentions, optimizing inventory management, promotional strategies, and advertising campaigns, ultimately enhancing overall marketing effectiveness.

Despite the significant performance gains, the BERT-DBLSTM model has some limitations. The model's large number of parameters and complex network structure result in high computational demands and longer inference times. Furthermore, the model currently focuses on text-only sentiment analysis and does not incorporate multimodal information like images or audio. In modern social media and e-commerce contexts, where users often express emotions through multimodal formats, such as text combined with images or videos, single-text analysis may be insufficient.

In the future, the BERT-DBLSTM model's practicality and application scope can be enhanced by model optimization and multimodal extensions. Techniques like parameter compression, pruning, or knowledge distillation can reduce computational costs and inference time, making the model more suitable for real-time applications and mobile devices. Additionally, expanding the model to handle multimodal sentiment analysis by integrating text, image, and audio features would improve its ability to interpret complex emotional expressions, meeting the diverse needs of social media and e-commerce sentiment analysis. These enhancements would further increase the model's value, enabling it to play a greater role in various sentiment analysis and user behavior prediction tasks.

CONFLICT OF INTEREST

We have no conflict of interest.

FUNDING STATEMENT

This research was supported by the National Social Science Fund of China "Research on the Generation Mechanism and Consensus Construction of Digital and Intelligent Social Prejudice from the Perspective of Memetic Communication" (grant number 24BXW052, principal investigator: Yang Changzheng), and 2024 Qingdao Social Science Planning Research Project "Research on the Group Spread and Control Mechanism of Internet Rumors in Major Emergencies in Qingdao from the Perspective of Multi-Modal Integration" (principal investigator: Yang Changzheng).

We also received funding from Renmin University of China 2022 Central universities to build a world-class university (discipline) and characteristic development guidance special project "User-Centered Internet Public Opinion Governance: Based on the relationship between Internet values and communication behavior" (No. 22RXW186).

PROCESS DATES

March 12, 2025
Received: November 26, 2024, Revision: January 6, 2025, Accepted: February 4, 2025

CORRESPONDING AUTHOR

Correspondence should be addressed to Mengmeng Guo (China, mengmengguomonica@outlook.com)

ACKNOWLEDGMENTS

We would like to thank the anonymous reviewers whose comments and suggestions helped improve this manuscript.

REFERENCES

Alsu Bari, S. N., Deshmukh, S. N., Al-Adhaileh, M. H., Alsaade, F. W., & Aldhyani, T. H. (2021). [Retracted] Development of integrated neural network model for identification of fake reviews in e-commerce using multidomain datasets. *Applied Bionics and Biomechanics, 2021*, 5522574. (Retraction published February 2023, *Applied Bionics and Biomechanics, 2023*(1), 9816851.)DOI: 10.1155/2021/5522574

Alzahrani, M. E., Aldhyani, T. H. H., Alsubari, S. N., Althobaiti, M. M., & Fahad, A. (2022). Developing an intelligent system with deep learning algorithms for sentiment analysis of e-commerce product reviews. *Computational Intelligence and Neuroscience*, *3840071*, 1–10. Advance online publication. DOI: 10.1155/2022/3840071 PMID: 35669644

Basiri, M. E., Nemati, S., Abdar, M., Cambria, E., & Acharya, U. R. (2021). ABCDM: An attention-based bidirectional CNN-RNN deep model for sentiment analysis. *Future Generation Computer Systems*, *115*, 279–294. DOI: 10.1016/j.future.2020.08.005

Cambria, E., Das, D., Bandyopadhyay, S., & Feraco, A. (2017). Affective computing and sentiment analysis. In Cambria, E., Das, D., Bandyopadhyay, S., & Feraco, A. (Eds.), *A practical guide to sentiment analysis. Socio-affective computing* (*Vol. 5*, pp. 1–10). Springer., DOI: 10.1007/978-3-319-55394-8_1

Chang, C.-C., & Lin, C.-J. (2011). LIBSVM: A library for support vector machines. *ACM Transactions on Intelligent Systems and Technology*, *2*(3), 1–27. DOI: 10.1145/1961189.1961199

Devika, M. D., Sunitha, C., & Ganesh, A. (2016). Sentiment analysis: A comparative study on different approaches. *Procedia Computer Science*, *87*, 44–49. DOI: 10.1016/j.procs.2016.05.124

Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2019). BERT: Pre-training of deep bidirectional transformers for language understanding. In *Proceedings of NAACL-HLT*, pp. 4171–4186. Association for Computational Linguistics. https://aclanthology.org/N19-1423.pdf

Dong, Y., Fu, Y., Wang, L., Chen, Y., Dong, Y., & Li, J. (2020). A sentiment analysis method of capsule network based on BiLSTM. *IEEE Access: Practical Innovations, Open Solutions*, *8*, 37014–37020. DOI: 10.1109/ACCESS.2020.2973711

Farimani, S. A., Jahan, M. V., Fard, A. M., & Tabbakh, S. R. K. (2022). Investigating the informativeness of technical indicators and news sentiment in financial market price prediction. *Knowledge-Based Systems*, *247*, 108742. DOI: 10.1016/j.knosys.2022.108742

Grover, V. (2022). Exploiting emojis in sentiment analysis: A survey. *Journal of The Institution of Engineers (India): Series B, 103*(1), 259–272. DOI: 10.1007/s40031-021-00620-7

Guo, Y. (2022). [Retracted] Financial market sentiment prediction technology and application based on deep learning model. *Computational Intelligence and Neuroscience, 2022*(1), 1988396. (Retraction published October 2023. *Computational Intelligence and Neuroscience*, *2023*, 9790721. DOI: 10.1155/2023/9790721

Huang, C., Zhao, J., & Yin, D. (2021). Purchase intent forecasting with convolutional hierarchical transformer networks. In *Proceedings of the 2021 IEEE 37th International Conference on Data Engineering (ICDE)*, pp. 2488–2498. IEEE. DOI: 10.1109/ICDE51399.2021.00281

Huang, J.-C., Shao, P.-Y., & Wu, T.-J. (2018). The study of purchase intention for men's facial care products with K-nearest neighbour. *International Journal of Computer Science and Information Technologies*, *10*(4), 95–104. https://ssrn.com/abstract=3302570

Jain, S., & Roy, P. K. (2024). E-commerce review sentiment score prediction considering misspelled words: A deep learning approach. *Electronic Commerce Research*, *24*(3), 1737–1761. DOI: 10.1007/s10660-022-09582-4

Joshi, R., Gupte, R., & Saravanan, P. (2018). A random forest approach for predicting online buying behavior of Indian customers. *Theoretical Economics Letters*, *8*(3), 448–475. DOI: 10.4236/tel.2018.83032

Kavatagi, S., & Adimule, V. (2021). Bi-GRU model with stacked embedding for sentiment analysis: A case study. In Kumar, R., Wang, Y., Poongodi, T., & Imoize, A. L. (Eds.), *Internet of Things, artificial intelligence and blockchain technology* (pp. 259–275). Springer Cham., DOI: 10.1007/978-3-030-74150-1_12

Li, X., Lei, Y., & Ji, S. (2022). BERT- and BiLSTM-based sentiment analysis of online Chinese buzzwords. *Future Internet*, *14*(11), 332. DOI: 10.3390/fi14110332

Liao, S., Wang, J., Yu, R., Sato, K., & Cheng, Z. (2017). CNN for situations understanding based on sentiment analysis of twitter data. *Procedia Computer Science*, *111*, 376–381. DOI: 10.1016/j.procs.2017.06.037

Liu, Y., Lu, J., Yang, J., & Mao, F. (2020). Sentiment analysis for e-commerce product reviews by deep learning model of Bert-BiGRU-Softmax. *Mathematical Biosciences and Engineering*, *17*(6), 7819–7837. DOI: 10.3934/mbe.2020398 PMID: 33378922

Ma, X., Li, Y., & Asif, M. (2024). E-commerce review sentiment analysis and purchase intention prediction based on deep learning technology. *Journal of Organizational and End User Computing*, *36*(1), 1–29. DOI: 10.4018/JOEUC.335122

Murthy, G. S. N., Allu, S. R., Andhavarapu, B., Bagadi, M. B. M., & Belusonti, M. (2020). Text based sentiment analysis using LSTM. *International Journal of Engineering Research & Technology (Ahmedabad)*, *09*(05). https://www.ijert.org/research/text-based-sentiment-analysis-using-lstm-IJERTV9IS050290.pdf

Nguyen, N. K., Le, A.-C., & Pham, H. T. (2016). Deep bi-directional long short-term memory neural networks for sentiment analysis of social data. In V.-N. Huynh, M. Inuiguchi, Le, B., B. N. Le, & Denoeux, T. (Eds.), *Integrated uncertainty in knowledge modelling and decision making: 5th international symposium, IUKM 2016* (pp. 255–268). Springer International Publishing. DOI: 10.1007/978-3-319-49046-5_22

Pan, G., Li, J., & Li, M. (2024). Multi-channel multi-step spectrum prediction using transformer and stacked Bi-LSTM. arXiv (preprint). arXiv:2405.19138 [eess.SP]. DOI: 10.48550/arXiv.2405.19138

Pondel, M., Wuczyński, M., Gryncewicz, W., Łysik, Ł., Hernes, M., Rot, A., & Kozina, A. (2021). Deep learning for customer churn prediction in e-commerce decision support. *Business Information Systems*, *1*, 3–12. DOI: 10.52825/bis.v1i.42

Pota, M., Ventura, M., Catelli, R., & Esposito, M. (2021). An effective BERT-based pipeline for Twitter sentiment analysis: A case study in Italian. *Sensors (Basel)*, *21*(1), 133. DOI: 10.3390/s21010133 PMID: 33379231

Rehman, A. U., Malik, A. K., Raza, B., & Ali, W. (2019). A hybrid CNN-LSTM model for improving accuracy of movie reviews sentiment analysis. *Multimedia Tools and Applications*, *78*(18), 26597–26613. DOI: 10.1007/s11042-019-07788-7

Sakar, C. O., Polat, S. O., Katircioglu, M., & Kastro, Y. (2019). Real-time prediction of online shoppers' purchasing intention using multilayer perceptron and LSTM recurrent neural networks. *Neural Computing & Applications*, *31*(10), 6893–6908. DOI: 10.1007/s00521-018-3523-0

Saranya, G., Gopinath, N., Geetha, G., Meenakshi, K., & Nithya, M. (2020). Prediction of customer purchase intention using linear support vector machine in digital marketing. *Journal of Physics: Conference Series*, *1712*(1), 012024. DOI: 10.1088/1742-6596/1712/1/012024

Sarker, K. U., Saqib, M., Hasan, R., Mahmood, S., Hussain, S., Abbas, A., & Deraman, A. (2022). A ranking learning model by k-means clustering technique for web scraped movie data. *Computers*, *11*(11), 158. DOI: 10.3390/computers11110158

Shan, Y. (2023). Social network text sentiment analysis method based on CNN-BiGRU in big data environment. *Mobile Information Systems*, *2023*(1), 8920094. DOI: 10.1155/2023/8920094

Sheil, H., Rana, O., & Reilly, R. (2018). Predicting purchasing intent: Automatic feature learning using recurrent neural networks. arXiv (preprint). arXiv:1807.08207 [cs.LG]. DOI: 10.48550/arXiv.1807.08207

Sun, Y. (2022). Design and purchase intention analysis of cultural and creative goods based on deep learning neural networks. *Computational Intelligence and Neuroscience*, *2022*(1), 3234375. DOI: 10.1155/2022/3234375 PMID: 36072728

Wagh, R., & Punde, P. (2018). Survey on sentiment analysis using twitter dataset. In *Proceedings of the 2018 Second International Conference on Electronics, Communication and Aerospace Technology (ICECA)*, pp. 208–211. IEEE. DOI: 10.1109/ICECA.2018.8474783

Wankhade, M., Annavarapu, C. S. R., & Abraham, A. (2024). CBMAFM: CNN-BiLSTM multi-attention fusion mechanism for sentiment classification. *Multimedia Tools and Applications*, *83*(17), 51755–51786. DOI: 10.1007/s11042-023-17437-9

Wu, H., & Li, B. (2022). Customer purchase prediction based on improved gradient boosting decision tree algorithm. In *Proceedings of the 2022 2nd International Conference on Consumer Electronics and Computer Engineering (ICCECE)*, pp. 795–798. IEEE. DOI: 10.1109/ICCECE54139.2022.9712779

Xu, W., Chen, J., Ding, Z., & Wang, J. (2024). Text sentiment analysis and classification based on bidirectional gated recurrent units (GRUs) model. arXiv (preprint). arXiv:2404.17123 [cs.CL]. DOI: 10.48550/arXiv.2404.17123

Yu, J., Liu, X., Gao, Y., Zhang, C., & Zhang, W. (2022). Deep learning for channel tracking in IRS-assisted UAV communication systems. *IEEE Transactions on Wireless Communications*, *21*(9), 7711–7722. DOI: 10.1109/TWC.2022.3160517

Yu, Q., Wang, Z., & Jiang, K. (2021). Research on text classification based on BERT-BiGRU model. *Journal of Physics: Conference Series*, *1746*(1), 012019. DOI: 10.1088/1742-6596/1746/1/012019

Yuan, J., Li, Z., Zou, P., Gao, X., Pan, J., Ji, W., & Wang, X. (2022). Community trend prediction on heterogeneous graph in e-commerce. In *WSDM '22: Proceedings of the Fifteenth ACM International Conference on Web Search and Data Mining*, pp. 1319–1327. Association for Computing Machinery. DOI: 10.1145/3488560.3498522

Zhan, T., Shi, C., Shi, Y., Li, H., & Lin, Y. (2024). Optimization techniques for sentiment analysis based on LLM (GPT-3). arXiv (preprint). arXiv:2405.09770 [cs.CL]. DOI: 10.48550/arXiv.2405.09770

Zhou, S., & Hudin, N. S. (2024). Advancing e-commerce user purchase prediction: Integration of time-series attention with event-based timestamp encoding and graph neural network-enhanced user profiling. *PLoS One*, *19*(4), e0299087. DOI: 10.1371/journal.pone.0299087 PMID: 38635519

Changzheng Yang is a professor at the College of Liberal Arts, Journalism and Communication, at Ocean University of China.