Innovation and Application Research of Text Sentiment Analysis Model Based on Attention Mechanism

Ge Youwang

Mongolian National University, Mongolia, 16060;

Abstract: Text sentiment analysis, as an important branch of the field of natural language processing, holds key application value in scenarios such as public opinion monitoring, customer feedback analysis, and intelligent recommendation. Traditional text sentiment analysis models often suffer from insufficient analytical accuracy when processing semantically complex texts with strong contextual dependencies, due to their inability to effectively capture key sentiment information. The attention mechanism, with its ability to selectively focus on important features in text, provides a new approach to solving this problem. This paper focuses on research into text sentiment analysis models based on the attention mechanism. It first reviews the research status and limitations of existing models, then proposes innovative directions regarding model structure and feature fusion methods. Finally, it discusses the application value and future development trends of these models in practical application scenarios, aiming to provide theoretical reference and technical insights for enhancing the accuracy and practicality of text sentiment analysis.

Keywords: Attention Mechanism; Natural Language Processing; Model Innovation

DOI:10.69979/3041-0843.25.02.064

Introduction

With the rapid development of internet technology, massive amounts of text data are generated in scenarios such as social media, e-commerce platforms, and news comment sections. This data contains important information such as users' emotional tendencies and attitudes. Text sentiment analysis aims to automatically identify and extract emotional information from text using computer technology, transforming unstructured text data into structured sentiment judgment results, thereby providing data support for enterprise decision-making, public management, and personal services.

Throughout the development of text sentiment analysis, methods based on rules, traditional machine learning, and deep learning have emerged successively. Rule-based methods rely on manually constructed sentiment lexicons and grammatical rules, making it difficult to adapt to complex and varied text language styles. Traditional machine learning methods (such as Support Vector Machines, Naive Bayes, etc.) require manual extraction of text features, where the quality of feature engineering directly affects model performance, and their ability to capture deep semantic information is limited. Deep learning methods (such as Recurrent Neural Networks RNN, Convolutional Neural Networks CNN) automatically learn text features through neural networks, significantly improving sentiment analysis accuracy. However, RNNs are prone to gradient vanishing problems when processing long texts, and CNNs are insufficient in capturing the temporal context of text. Both fail to dynamically assign weights based on the semantic importance of text, leading to insufficient focus on key emotional information.

The attention mechanism originates from the simulation of human visual attention. Its core idea is to prioritize parts of the information that are more important for the task when processing information. Introducing the attention mechanism into text sentiment analysis models enables the model to automatically identify and focus on words, phrases, or sentences that play a key role in emotional expression during the process of learning text features, thereby enhancing the model's ability to capture deep semantic and emotional information. In recent years, text sentiment analysis models based on the attention mechanism have become a research hotspot. However, existing models still have room for improvement in terms of the rationality of attention distribution, the fusion of multimodal emotional information, and model generalization ability. Therefore, in-depth research on the innovation directions and application scenarios of attention mechanism-based text sentiment analysis models holds significant theoretical and practical value.

1 Research Status of Text Sentiment Analysis Models Based on Attention Mechanism

1.1 Application of Basic Attention Mechanism in Text Sentiment Analysis

Early research introducing the attention mechanism into text sentiment analysis primarily involved adding attention layers to models like RNN and Long Short-Term Memory networks (LSTM) to optimize the feature extraction process. For example, introducing a soft attention mechanism at the encoder output layer of an LSTM model, by calculating the association weight between the output vector at each time step and the sentiment classification target, performs a weighted sum of features from different time steps to obtain a text feature vector that is more representative of sentiment. Compared to traditional LSTM models, such models can effectively highlight key information related to sentiment in the text and have achieved good results in simple text sentiment analysis tasks such as movie reviews and product evaluations.

Furthermore, the proposal of the self-attention mechanism provided a new technical path for text sentiment analysis. The self-attention mechanism does not rely on recurrent structures; instead, it directly captures global semantic relationships in text by calculating the attention weights between each word and all other words within the text. Models based on the self-attention mechanism (such as the encoder part of Transformer) can process text data in parallel, improving model training efficiency, while effectively solving the long-range dependency problem, showing advantages in long-text sentiment analysis tasks.

1.2 Limitations of Existing Models

Although basic attention mechanisms have supported the performance improvement of text sentiment analysis models, existing models still have the following limitations:

First, one-sidedness of attention distribution. Existing models mostly use a single-dimensional approach to attention calculation (e.g., based solely on semantic similarity), ignoring multi-dimensional information such as syntactic structure, part-of-speech features, and sentiment intensity, leading to unreasonable allocation of attention weights. For example, when processing texts containing transitional relationships like "Although this phone looks nice, its battery life is terrible," the model might over-focus on positive words like "nice" and underestimate the emotional weight of the key negative information "battery life is terrible," resulting in sentiment judgment deviation.

Second, insufficient fusion of multimodal emotional information. Emotional expression in practical applications often involves multimodal information such as text, emojis, voice, and images. Existing attention mechanism-based text sentiment analysis models predominantly focus on pure text data and lack effective fusion mechanisms for multimodal emotional information, making it difficult to comprehensively capture users' emotional tendencies.

Third, weak model generalization ability. Existing models are mostly trained and validated on specific domain datasets (e.g., e-commerce product review datasets, movie review datasets). When applied to other domains (e.g., medical consultation reviews, financial news comments), model performance significantly decreases due to differences in text style and emotional expression habits between domains. The root cause of this problem lies in the fact that the attention patterns and feature representations learned by the model overly rely on the domain characteristics of the training data, lacking capture of universal emotional expression patterns.

Fourth, the balance between model complexity and efficiency. To improve the precision of attention allocation, some studies optimize the model by increasing the number of attention heads or constructing multi-layer attention structures. However, this also leads to an expansion of model parameters, increased computational complexity, and raises the difficulty of applying the model in resource-constrained scenarios such as mobile terminals and embedded devices.

2 Innovation Directions for Text Sentiment Analysis Models Based on Attention Mechanism

2.1 Design of Multi-Dimensional Fusion Attention Calculation Mechanism

To address the one-sided attention distribution in existing models, a multi-dimensional fusion attention calculation mechanism can be designed, incorporating text semantics, syntax, and emotional features into the weight calculation. On one hand, introduce syntax-aware attention, combining dependency parsing to extract syntactic structure information, allowing the model to adjust attention weights based on syntax to accurately capture key emotional information in complex sentences like transitions. On the other hand, integrate part-of-speech and sentiment intensity features by constructing a sentiment intensity lexicon and a part-of-speech weight matrix to quantify relevant features, enabling the model to prioritize words with high sentiment intensity and direct expression, reducing judgment deviations caused by single-dimensional calculations, and improving the accuracy of complex text analysis.

2.2 Construction of Attention Fusion Model for Multimodal Emotional Information

To solve the problem of insufficient multimodal information fusion, construct a multimodal attention fusion model. Design feature extraction modules separately for text, emojis, and images: use Transformer encoder for text to capture global semantics; map emojis to vectors combined with sentiment lexicons to assign features; utilize CNN for images to extract visual features and learn emotional patterns. The core involves using a cross-modal attention mechanism to calculate a similarity matrix of modal features, combined with modal importance weights to obtain fusion weights, generating a comprehensive feature vector for sentiment judgment. This fully utilizes the complementarity of multimodality to overcome the limitations of pure text models.

2.3 Attention Transfer Learning Strategy Based on Domain Adaptation

To enhance model generalization ability, adopt a domain adaptation attention transfer learning strategy. First, train a base model on large-scale source domain data to learn general attention patterns. Then, through parameter transfer, migrate parameters from the attention layer and feature extraction layer to the target domain model, retraining only the classifier to reduce the need for annotated data. Finally, introduce a domain adaptation adjustment mechanism, constructing a joint loss function for domain discrepancy and sentiment classification, allowing the model to adapt to the textual characteristics of the target domain while retaining general patterns, thereby reducing dependence on specific domain data.

2.4 Optimal Design of Lightweight Attention Models

To balance performance and efficiency, optimize lightweight models from three aspects: First, prune redundant attention heads and share parameters between layers to reduce the parameter scale. Second, adopt local or sparse attention mechanisms to simplify attention calculation and reduce complexity. Third, use model quantization to convert parameters to low-precision integers, combined with knowledge distillation to allow the lightweight model to learn from the high-precision model's knowledge, ensuring analysis accuracy while reducing resource consumption to meet the needs of resource-constrained scenarios like mobile terminals.

3 Application Scenarios of Text Sentiment Analysis Models Based on Attention Mechanism

3.1 Public Opinion Monitoring and Public Management

In public opinion monitoring, this model can rapidly process massive texts from social media and news comments, identifying the public's emotional tendencies towards events or policies in real-time. Through attention weights, it can locate key emotional information, providing data support for governments to adjust strategies. Simultaneously, the model can analyze the emotional characteristics and propagation paths of online rumors, capturing inflammatory statements within rumors, assisting regulatory authorities in identifying rumor sources and curbing their spread, thereby maintaining cyberspace order and social stability.

3.2 E-commerce Platform Customer Feedback Analysis

In e-commerce scenarios, the model can deeply analyze customer review texts, not only judging the overall sentiment tendency but also locating specific aspects of customer satisfaction or dissatisfaction (e.g., product quality, service, logistics) through attention weights. Based on the analysis results, merchants can optimize products targetedly, strengthen customer service training, or improve logistics efficiency. Meanwhile, the model can construct customer sentiment profiles to assist product recommendation, enhancing customer satisfaction and repurchase rate.

3.3 Intelligent Customer Service and Emotional Interaction

In the field of intelligent customer service, the model can analyze the emotional state of user inquiry texts in real-time, helping customer service adjust interaction strategies. For example, when user dissatisfaction is identified, customer service can prioritize reassurance before solving the problem, reducing complaints. Furthermore, the model can be applied to emotional companion intelligent systems. By analyzing the emotional tendency of dialogue texts and adjusting the tone and content of replies, it achieves humanized interaction, helping users alleviate negative emotions.

3.4 Media Content Emotional Orientation Analysis

The model can analyze the emotional orientation of news, self-media articles, and short video captions, assisting media organizations in optimizing content creation strategies. For news media, they can adjust topic selection based on this, increasing the dissemination of positive emotional content with high attention. For self-media creators, it can help them identify emotional expression patterns associated with high traffic, enhancing clicks, likes, and shares of their works. Simultaneously, the model can monitor whether the emotional orientation of content aligns with social values, avoiding excessively negative or inflammatory content, and guiding a healthy emotional communication atmosphere.

4 Summary and Outlook

This paper conducted research on text sentiment analysis models based on the attention mechanism, reviewed the research status and limitations of existing models, proposed model innovation directions from four aspects: multi-dimensional attention calculation, multimodal information fusion, domain adaptive transfer learning, and lightweight model optimization, and discussed the application value of the model in scenarios such as public opinion monitoring, e-commerce customer feedback analysis, intelligent customer service, and media content analysis. Research shows that the attention mechanism can effectively enhance the ability of text sentiment analysis models to capture key emotional information. Through innovations in model structure and strategies, issues existing in current models regarding the rationality of attention distribution, generalization ability, and efficiency can be further addressed, promoting the application of models in more practical scenarios.

In the future, text sentiment analysis models based on the attention mechanism can be further explored in the following directions:

First, optimize the attention mechanism by integrating cognitive linguistics theories. Starting from the patterns of human cognition and emotional expression, incorporate cognitive features such as emotional metaphors and emotional associations into the attention calculation process, making the model's attention allocation more aligned with human emotional understanding logic.

Second, explore cross-lingual text sentiment analysis. By constructing cross-lingual attention mechanisms, achieve effective transfer and analysis of emotional information in texts of different languages, meeting the needs of sentiment analysis in globalized scenarios.

Third, strengthen research on model interpretability. Although existing attention mechanisms can provide attention weight distributions, the interpretability of the model's sentiment judgment results remains insufficient. In the future, use visualization techniques, causal inference methods, etc., to enhance model interpretability, allowing users to understand the basis for the model's sentiment judgments more clearly.

Fourth, integrate dynamic sentiment analysis. Existing models mostly focus on sentiment analysis of static text. In the future, combine temporal attention mechanisms to analyze the dynamic change process of text sentiment over time (e.g., the change in user sentiment towards an event from concern to worry to satisfaction), providing support for scenarios such as dynamic public opinion monitoring and user sentiment change prediction.

With the continuous development of natural language processing technology, text sentiment analysis models based on the attention mechanism will continue to improve in theoretical research and practical applications, providing strong support for mining the emotional value of text more accurately and efficiently.

References

[1] Yang Jianhao, Zeng Biqing, Deng Huimin, et al. Research on Text Micro-sentiment Analysis Based on Multi-task and Self-attention Mechanism [J]. Computer and Digital Engineering, 2023, 51(12): 2863-2866.
[2] Liu Chao. Research on GRU Short Text Sentiment Analysis Based on Attention Mechanism and SVM [D]. Tianjin University of Finance and Economics, 2019.
[3] Wang Tao. Construction and Application Research of Stock News Sentiment Tendency Analysis Model [D]. Shanghai Jiao Tong University, 2020.
[4] Chen Liyuan, Wu Tao. Short Text Sentiment Analysis Method Fusing Topic Model and Self-attention Mechanism [J]. Foreign Electronic Measurement Technology, 2021, 40(11): 6.
[5] Qian Yanjun. Research on Text Implicit Sentiment Analysis Technology Based on Attention Mechanism [D]. Yunnan University, 2022.