A review of Chinese sentiment analysis: subjects, methods, and trends

Zhaoxia Wang[1] · Donghao Huang[1,2] · Jingfeng Cui[3] · Xinyue Zhang[1,4] · Seng-Beng Ho[5] · Erik Cambria[6]

Abstract

Sentiment analysis has emerged as a prominent research domain within the realm of natural language processing, garnering increasing attention and a growing body of literature. While numerous literature reviews have examined sentiment analysis techniques, methods, topics and applications, there remains a gap in the literature concerning thematic trends and research methodologies in sentiment analysis, particularly in the context of Chinese text. This study addresses this gap by presenting a comprehensive survey dedicated to the progression of research subjects, methods and trends in sentiment analysis of Chinese text. Employing a framework that combines keyword co-occurrence analysis with a sophisticated community detection algorithm, this survey offers a novel perspective on the landscape of Chinese sentiment analysis research. By tracing the interplay between research methodologies and emerging topics over the past two decades, our study not only facilitates a comparative analysis of their correlations but also illuminates evolving patterns, identifying significant hotspots and trends over time for Chinese language text analysis. This invaluable insight provides a roadmap for researchers seeking to navigate the intricate terrain of sentiment analysis within the context of Chinese language. Moreover, this paper extends beyond the academic realm, offering practical insights into sentiment analysis methodologies and themes while pinpointing avenues for future exploration, technical limitations, and directions for sentiment analysis of Chinese text.

1 Introduction

In today's digitally interconnected world, understanding public sentiment and emotions expressed in textual data has become increasingly important (Abdullah and Rusli 2021). Within the realm of natural language processing (NLP), sentiment analysis stands as a crucial discipline, serving as a linchpin in decoding these nuanced emotions and opinions (Cambria et al. 2024; Wang et al. 2023a). Its applications extend across diverse fields includ-

---

Extended author information available on the last page of the article

ing marketing strategies, brand management, social media monitoring, video analysis, and more, offering invaluable insights (Cui et al. 2023; Cambria et al. 2013; Wang et al. 2023b; Tan et al. 2023; Stappen et al. 2021). The ability to analyze sentiments across languages is especially crucial as the global landscape becomes more diverse and interconnected (Mihalcea et al. 2012; Araújo et al. 2020; Cambria et al. 2023).

Language diversity is an inherent characteristic of global communication, necessitating sentiment analysis techniques that can transcend linguistic barriers (Abdullah and Rusli 2021; He et al. 2022). The intricacies of language nuances and cultural variations add layers of complexity to sentiment interpretation (Wang et al. 2014a). A sentiment-bearing phrase that conveys positivity in one language might carry a different connotation in another (Xu et al. 2022). Moreover, cultural norms influence how emotions are expressed, demanding adaptable models that can decipher sentiments within different cultural contexts (Wang et al. 2020a; Ortony and Turner 1990). Multilingual sentiment analysis (Abdullah and Rusli 2021; Cui et al. 2023; Mihalcea et al. 2012; Araújo et al. 2020; Lo et al. 2017), the process of assessing emotions and opinions in different languages, addresses this need by enabling a comprehensive understanding of global sentiment trends. With multilingual analysis, we can unveil not only linguistic differences but also cultural nuances and contextual variations in sentiment expressions.

In the realm of multilingual sentiment analysis, the Chinese language holds a significant position due to its extensive usage and the unique challenges it poses (Peng et al. 2017; Yan et al. 2014; Zhang and Zhou 2023; Li et al. 2022). Chinese, with its ideographic writing system, tonal complexities, and rich idiomatic expressions, demands specialized approaches for sentiment analysis (Zhang and Zhou 2023; Li et al. 2022; Xu et al. 2019a). Moreover, given China's influence in global markets and digital spaces, comprehending sentiment within Chinese text is essential for businesses and researchers alike. While existing literature reviews have diligently surveyed these facets, there remains an uncharted territory awaiting exploration: the thematic trends, evolution of research methodologies and topics within the realm of Chinese sentiment analysis.

This study aims to bridge the existing gap by elucidating the distinct evolution of sentiment analysis research methodologies and themes within the realm of Chinese sentiment expression. It addresses this intriguing gap by presenting a comprehensive survey that delves into the dynamic progression of sentiment analysis research methods and thematic foci. While prior literature reviews have expounded upon sentiment analysis techniques, this study takes a pioneering step by specifically examining the interplay between evolving research trends, methodologies and emergent themes, with a particular emphasis on the analysis of sentiment within Chinese text.

This paper is structured to provide a comprehensive review of Chinese sentiment analysis. Section 1 highlights the growing significance of sentiment analysis in understanding public sentiment expressed in textual data, with specific attention to the sentiment analysis within the Chinese language. Section 2 reviews the existing work and underscores the extensive exploration of sentiment analysis, with a focus on the gap in understanding sentiment within the unique linguistic and cultural framework of Chinese text, pointing out there is a need for a pioneering comprehensive survey that examines the evolution of sentiment analysis research methods and themes specifically within the realm of Chinese text. Section 3 provides the detailed survey and analysis on the trends, methods, and topics in sentiment analysis of Chinese text, outlining the evolution of methodologies in multilingual sentiment

analysis, especially with a focus on the Chinese language. Section 4 explores the practical applications, limitations, and future prospects of Chinese sentiment analysis. Lastly, Sect. 5 and Sect. 6 conclude the paper and discuss potential future directions.

2 Background and importance of this survey

2.1 Background

Sentiment analysis, a cornerstone of natural language processing, has undergone extensive exploration across a multitude of languages, cultures, and contexts (Abdullah and Rusli 2021; Cui et al. 2023; Mihalcea et al. 2012; Araújo et al. 2020). A growing body of research has shed light on sentiment analysis techniques, methods, and applications, leading to a deeper understanding of human emotions and opinions as conveyed through textual content (Cambria et al. 2013; Wang et al. 2023b; Tan et al. 2023; Stappen et al. 2021). However, a critical gap in the existing literature lies in the investigation of sentiment analysis within the distinct linguistic and cultural framework of Chinese text.

Research in sentiment analysis has predominantly focused on the English language, resulting in a comprehensive understanding of sentiment expressions in this English linguistic domain (Cui et al. 2023; Wang and Lin 2020; Wang et al. 2020b). Numerous studies have explored lexicon-based approaches, machine learning techniques, and deep learning models for sentiment classification and polarity detection in English text (Cui et al. 2023). These efforts have culminated in the development of robust tools and methodologies for deciphering sentiment cues and gauging emotional tone within English textual content (Raghuvanshi and Patil 2016; Nair et al. 2019; Obiedat et al. 2021; Medhat et al. 2014; Prabha and Srikanth 2019; Ligthart et al. 2021).

In contrast, sentiment analysis within the context of Chinese text introduces a host of unique challenges (Peng et al. 2017; Zhou et al. 2022; Yang et al. 2022c; Sun et al. 2016). The rich morphological structure of Chinese characters, combined with the intricate interplay between ideographic meanings, necessitates tailored approaches to sentiment classification (Sun et al. 2016). Additionally, the tonal nature of Chinese writing and the reliance on context for accurate sentiment interpretation further complicate the sentiment analysis process (Peng et al. 2021). The authors presents a pioneering approach that leverages the modern Chinese pronunciation system, known as pinyin, to enhance Chinese sentiment analysis. By exploring phonetic information extracted from pinyin, including audio clips and pinyin token corpus, the authors developed a novel method to disambiguate intonations and integrate this phonetic information with textual and visual features. Such an approach generated innovative multimodal representations for Chinese words, significantly contributing to the improvement of sentiment analysis in Chinese across various datasets (Peng et al. 2021).

Chinese sentiment expression often relies on implicit cultural references, idiomatic phrases, and subtle context cues (Pan et al. 2019). These intricacies demand a deep understanding of Chinese culture, history, and societal norms to accurately decipher the emotional undercurrents of text. Consequently, sentiment analysis models tailored for Chinese must not only account for linguistic peculiarities but also navigate the intricate web of cultural subtext (Tan and Zhang 2008).

A wealth of research in this domain has been dedicated to exploring sentiment analysis specifically in English text, resulting in a plethora of survey papers that comprehensively summarize and analyze the existing methodologies, datasets, and challenges. However, a noticeable discrepancy arises when considering sentiment analysis in the context of Chinese text. Despite the substantial importance of Chinese as one of the world's most widely spoken languages, the literature landscape reveals a distinct lack of comprehensive survey papers addressing sentiment analysis in this linguistic context. While numerous studies have investigated sentiment analysis techniques for Chinese text (Xu et al. 2022; Huai and Voorde 2022; Huang et al. 2022), the scarcity of surveys hampers the consolidation of knowledge and the identification of trends and gaps in this area. This glaring disparity underscores the need for further scholarly endeavors to bridge this gap, providing insights that can guide future research and advancements in Chinese text sentiment analysis. By addressing this imbalance, researchers can contribute to a more inclusive understanding of sentiment analysis across diverse languages and cultures. However, a comprehensive investigation into the evolution of research methodologies and thematic trends within this realm is conspicuously absent.

This paper fills this void by presenting a comprehensive survey that specifically probes the evolution of sentiment analysis research trends, methods, and thematic foci within the context of Chinese text.

This research not only builds upon the foundation of the previous survey paper (Cui et al. 2023) but also expands its horizons to unravel the intricate tapestry of sentiment analysis within the realm of Chinese text. By employing innovative techniques such as keyword co-occurrence analysis and community detection algorithms, our study not only pioneers an exploration of sentiment analysis within the realm of Chinese but also provides a roadmap for understanding the nuanced interplay between linguistic, cultural, and technological factors in sentiment expression.

2.2 Importance of Chinese sentiment analysis

Chinese sentiment analysis remains crucial despite the availability of advanced language models like ChatGPT and other large language models (LLMs) for several reasons:

- Language Nuances: Chinese language and sentiments entail unique cultural and linguistic nuances that may not be fully captured by generalized models (Du et al. 2022). Cultural context, idiomatic expressions, and subtle linguistic nuances specific to Chinese language require specialized models for accurate sentiment analysis (Liao et al. 2022; Bhaskaran et al. 2022; Sun et al. 2022). For instance, idiomatic expressions such as "坐井观天" (literally "sitting in the well and looking at the sky," which actually means "having a limited view of the world" and carries a negative connotation) and "画蛇添足" (literally "drawing legs on a snake," which means "to overdo something unnecessary" and also carries a negative connotation) are deeply rooted in Chinese culture and require specialized models for accurate sentiment analysis. If we translate such Chinese idioms directly into English without considering their true meanings, the models may miss these subtleties, leading to incorrect sentiment interpretations.
- Domain Specificity: Specialized models for Chinese sentiment analysis can cater to industry-specific jargon, domains, and nuances, providing more accurate analysis for

various sectors like finance, healthcare, or technology (Du et al. 2022; Guo et al. 2022). For example, in the finance sector, terms like "牛市" (meaning "bull market") directly refer to a financial market characterized by rising prices and positive investor sentiment, while "熊市" (meaning "bear market") denotes a market trend where prices are declining and investor sentiment is generally negative. Such domain-specific terms are particular to Chinese financial discourse and require precise understanding for sentiment analysis. Similarly, medical terms used in healthcare, such as "+" ("正号", which means "positive" in general but in the healthcare domain it would actually mean negative), illustrate how different technical domain jargons in various sectors demand domain-specific models for accurate Chinese sentiment analysis.

- Data Diversity: Chinese sentiment analysis demands diverse and extensive datasets reflecting various dialects, regions, and cultural sentiments within the Chinese language (Peng et al. 2021; Guo et al. 2022; Álvarez-Carmona et al. 2022; Yang et al. 2020). Specialized models focus on accumulating and understanding this diverse dataset for accurate analysis. For example, a comprehensive Chinese restaurant review dataset was designed for Aspect Category Sentiment Analysis (ACSA) and rating prediction (RP) (Li et al. 2023a). It contains 46,730 genuine reviews, meticulously annotated with sentiment polarities across 18 fine-grained aspect categories and 5-star ratings, making it one of the largest and most detailed Chinese sentiment datasets. While Ciron (Xiang et al. 2020) stands as the first Chinese benchmark for irony detection, featuring 8.7 K microblog posts annotated for irony by native speakers, facilitating the exploration of irony detection in Chinese using machine learning models. Additionally, Chinese EmoBank (Lee et al. 2022) offers resources for dimensional sentiment analysis, with valence-arousal ratings for 5512 words, 2998 multi-word phrases, 2582 single sentences, and 2969 multi-sentence texts, supporting the development of fine-grained sentiment models for the Chinese language.
- Local Applications: Local applications and platforms in China often require sentiment analysis tailored to the Chinese language (Yang et al. 2020). For instance, the phrase "贼好" (literally "thief good") is a very positive comment often used in Northern China, while "侬好" (meaning "Hello") is a greeting in the Shanghai dialect. These regional variations necessitate specialized models that can accurately interpret sentiments across different Chinese dialects and regions. Such applications might demand models that are specifically trained and optimized for local user interactions and content. For example, social media platforms like Weibo and WeChat have unique user interactions and slangs that are specific to the Chinese context. Models trained specifically on content from these platforms can better understand and analyze sentiments expressed by local users, enhancing the relevance and accuracy of the analysis.
- Performance Enhancement: While ChatGPT and similar LLMs are powerful, dedicated sentiment analysis models for Chinese can be fine-tuned and optimized specifically for sentiment-related tasks, potentially providing better performance and accuracy for sentiment analysis tasks in the Chinese language context. For example, a model trained exclusively on Chinese customer reviews from e-commerce platforms like Taobao and JD.com can provide better performance in identifying customer satisfaction and feedback trends than a generalized model. Therefore, fine-tuning and optimization of models for different platforms and application areas is particularly important.

While generalized language models like ChatGPT offer significant capabilities, dedicated Chinese sentiment analysis models are essential to address language nuances, cultural specificity, domain requirements, data diversity, local applications, and performance enhancement within the Chinese language sentiment analysis domain.

In addition, China's rise as a global economic leader and its role in initiatives like the Belt and Road highlight its ascent (Tombari 2021; Zou et al. 2022). Positioned to become the world's largest economy, China's strategic initiatives, especially the Belt and Road, showcase its influence, shaping global trade, connectivity, and geopolitics significantly.

3 Detailed survey on Chinese sentiment analysis

This section offers an in-depth survey on Chinese sentiment analysis including Data Acquisition of Scientific Publications, Subjects, Methods, Trends, and Hot Topics within the realm of sentiment analysis of Chinese text. It also highlights the progression of methodologies in multilingual sentiment analysis, with a specific emphasis on the Chinese language.

3.1 Data acquisition of scientific publications

In alignment with previous research methodologies (Cui et al. 2023), this study's data collection involved sourcing research publications from the Web of Science across four primary databases: Conference Proceedings Citation Index—Social Sciences & Humanities (CPCI-SSH), Conference Proceedings Citation Index—Science (CPCI-S), Science Citation Index Expanded (SCI-Expanded), and Social Sciences Citation Index (SSCI). To establish a comprehensive dataset, specific keywords including "sentiment analysis," "sentiment mining," and "sentiment classification" were employed. This targeted keyword-based approach facilitated the identification and extraction of pertinent papers that constituted our data pool. After screening, a total of 711 literature pieces relevant to "Chinese" or the "Chinese language" were identified.

The search spanned from January 2002 to December 2022, encompassing publication types such as "article" and "conference paper". Following the screening process, a total of 711 papers were identified. During this time period, the earliest related research appeared in 2005. The distribution of the publications from 2005 to 2022 is shown in Fig. 1.

Figure 1 illustrates the number of publications in the realm of Chinese sentiment analysis spanning the years 2005 to 2022. Evidently, the publication count follows an exponential growth trajectory over time. The data demonstrates a notable peak in 2016, followed by a minor decline and a subsequent peak in 2018. There is a slight decrease in 2019, after which the count experiences another exponential surge, culminating in the highest point by 2022.

Keywords were extracted from all the collected publications using the same method as previous research (Cui et al. 2023), in which KeyBERT was used for keyword extraction. Extracted keywords were compared and merged with the original keywords in the publications. Subsequently, merging and screening were conducted, excluding terms like "sentiment analysis," "sentiment classification," "Chinese language," and "Chinese sentiment analysis", which are used as keywords for publication collection. A total of 4062 keywords were selected with a word frequency of 6172. Upon observation, keywords with a frequency below 3 were mostly devoid of substantive meaning. Therefore, 258 keywords
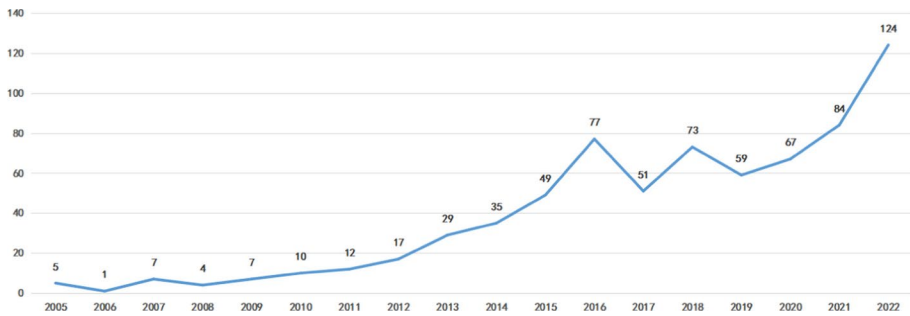
Fig. 1 Number of publications related to sentiment analysis in Chinese from 2005 to 2022. This data collection in this research involved sourcing research publications from the Web of Science (see details in Sect. 3.1) and the search range was set to include publications from a 20-year period between 2002 and 2022. However, since the earliest relevant literature within this range appears in 2005, the results shown in the figures all start from 2005 rather than 2002

with a frequency greater than 2 were chosen, totaling 2025 in frequency, accounting for roughly one-third of the total frequency count. Keywords with a frequency of 10 or more are listed in the Table 1.

High-frequency keywords often serve as indicators of research focal points. For example, "microblog" is a high-frequency keyword, and it indicates that "microblog" analysis is one important focus for Chinese sentiment analysis. "Deep learning" is also a high-frequency keyword, and it indicates that the "deep learning" method is one important focus in Chinese sentiment analysis. Such discoveries are reasonable, and consistent with previous research (Jia 2022; Wu et al. 2022). We have extracted these prominent keywords to form the foundation for our subsequent analysis.

3.2 Subject analysis

Table 1 showcases the high-frequency keywords which essentially encapsulate the primary research focus within the realm of Chinese sentiment analysis. "Microblog" and "deep learning" secures the top positions, trailed by "social medium," and "text mining". These high-frequency keywords encompass a wide range, including the study's subject, content, and employed techniques and methods. With these keywords as a foundation, we established a keyword co-occurrence network, effectively visualizing the interplay between research methodologies and subjects.

Using the same drawing approach as described in Cui et al. (2023), we utilized Pajek software to create a visualization of the keyword co-occurrence network. Pajek's Louvain community detection algorithm aids in partitioning the keyword co-occurrence network into sub-communities, each representing different sub-fields within keyword analysis (Cui et al. 2023; Blondel et al. 2008; Leydesdorff et al. 2014; Rotta and Noack 2011). The characteristics of this keyword co-occurrence network are elaborated upon in Fig. 2. The keyword co-occurrence network characteristics within the six sub-communities are described in Table 2. The count of nodes within each community reflects the quantity of keywords represented, while the connections, depicted as links, denote the correlations established among these keywords.

Table 1 Keywords with a frequency of 10 or more

Keywords	Frequency	Keywords	Frequency
Microblog	75	Sentiment dictionary	19
Deep learning	52	Neural network	19
Social medium	50	Short text	17
Text mining	46	Word2vec	17
Natural language processing	44	BERT	15
Opinion mining	44	Naive Bayes	13
Word embedding	43	Text classification	13
Chinese text	40	Product review	13
Sentiment lexicon	40	Dimensional sentiment	13
Topic model	37	Movie review	12
Chinese review	35	Feature selection	12
Convolutional neural network	35	Latent Dirichlet allocation	12
SVM	33	Feature extraction	11
user review	32	Domain sentiment	11
Machine learning	31	Word vector	11
Weibo	31	Semantic orientation	10
LSTM	26	Hotel review	10
COVID-19	24	Semi-supervised learning	10
Online review	23	Affective computing	10
Social network	23	Text analysis	10
Attention mechanism	22	Review sentiment	10
Aspect-based	21	Semantic	10
Text sentiment	20	Transfer learning	10

Figure 2 shows the keyword community network obtained using all the keywords related to sentiment analysis in Chinese text from 2005 to 2022 (Due to the limited size of the image, not all keywords can be fully displayed here). C1 community centers on keywords related to Chinese sentiment in the context of "microblog" analysis, like "sentiment lexicon", "sentiment dictionary", "polarity classification" and so on. C2 and C3 communities delve into techniques for sentiment analysis, with C2 community primarily exploring the application of deep learning in NLP, and C3 community focusing predominantly on machine learning approaches and traditional sentiment analysis methods, such as Naive Bayes method, rule-based methods, etc. C4 and C5 communities revolve around specific themes within sentiment analysis. C4 community encompasses subjects linked to "weibo," "COVID-19," and "social medium," while C5 community encompasses topics concerning "user reviews" and "opinions." C6 community encompasses keywords pertinent to various intricacies within the sentiment analysis process, encompassing terms like "sentiment score," "dimensional sentiment," "Word2vec," and "word segmentation." Among them, Dimensional sentiment analysis is an important research topic, with existing studies primarily focusing on various analytical methods. The methods used include position-based character-enhanced word embedding and FastText as word vectors (Du and Zhang 2016), the Community-Based Weighted Graph Model (Wang et al. 2016a), a semi-supervised approach for DSA based on the variational autoencoder model (Wu et al. 2019), a tree-structured regional CNN-LSTM
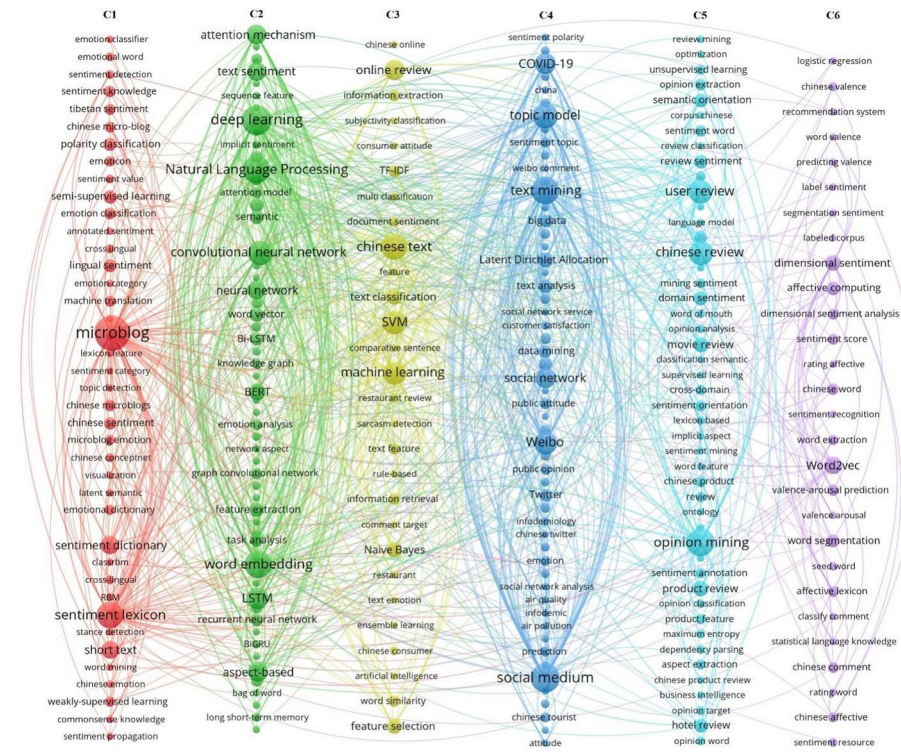
Fig. 2 Keyword community network obtained using all the keywords related to sentiment analysis in Chinese text from 2005 to 2022

Table 2 Network-wide attributes of sub-communities on a global scale

Community	Connections amongst or within the communities						Nodes
	C1	C2	C3	C4	C5	C6	
C1	116	115	45	99	58	27	41
C2	115	326	102	111	127	55	59
C3	45	102	95	61	66	28	28
C4	99	111	61	262	75	20	55
C5	58	127	66	75	187	38	47
C6	27	55	28	20	38	101	28
Global network							258

model (Wang et al. 2020c), the BiLSTM model (Cheng et al. 2021), and the BERT model (Deng et al. 2023), etc.

For evaluating the effectiveness of these methods, F1-score (Li et al. 2023a; Xiang et al. 2020; Hercigt et al. 2016; Bu et al. 2021; Zhao et al. 2023), Accuracy (Xiang et al. 2020; Hercigt et al. 2016; Bu et al. 2021), Mean Absolute Error (MAE), and Pearson Correlation Coefficient (r) (Lee et al. 2022; Xie et al. 2021) are commonly used.

An intriguing discovery emerges from the comparison with the original keyword "Twitter" in the survey for sentiment analysis of English text (Cui et al. 2023; Wang and Parth

2016; Chin and Wang 2016): "Twitter" is replaced by "microblog". This substitution appears both logical and congruent with prior research, given that "microblog" predominantly pertains to the Chinese context, whereas "Twitter" is primarily associated with English discourse. This adjustment maintains consistency with previous scholarly undertakings.

Table 2 presents the quantity of nodes denoting keywords within individual communities, along with the corresponding number of interconnecting links. It is observed that there are 326 connections within C2 (C2–C2 connections) and 102 connections between C2 and C3 (C2–C3 connections). Thus, the number of connections within C2 exceeds the number of connections between C2 and C3. The inter-community linkages highlight robust correlations existing among them. Notably, sub-communities C1, C3, C4, C5, and C6 exhibit significant associations with the C2 community, underscoring a collective emphasis on deep learning methodologies within this research domain. The C2 community has a strong correlation with the C5 community, reflecting the wide application of deep learning methods in the fields of user review and opinion mining. This is consistent with our findings in Cui et al. (2023), indicating that researchers are paying more attention to sentiment analysis technology, especially in deep learning methods. The article will analyze the application of sentiment analysis methods in Chinese texts in the next section.

3.3 Method analysis

Table 3 shows the four main methods of Chinese multilingual sentiment analysis. Categorizing the methods in this way helps provide a high-level understanding of the types of approaches available for sentiment analysis and how they relate to each other. It is worth noting that the boundaries between these categories can sometimes be blurred, as certain methods might incorporate aspects from more than one category.

3.3.1 Lexicon-based approaches

In the context of Chinese sentiment analysis, lexicon-based approaches face unique challenges due to the language's complex structure. Unlike alphabetic languages, Chinese characters often carry multiple meanings depending on the context, and the language lacks spaces between words, making tokenization and segmentation crucial (Xu et al. 2010; Gu et al. 2022). As shown in Table 3, lexicon-based approaches are rooted in the analysis of individual words and their associated sentiment scores (Liu et al. 2009; Xu et al. 2010; Qiang et al. 2005; Du and Tan 2010; Zhang et al. 2008a; Lu et al. 2010; Gravano 2010). The lexicon-based approaches include using lexicons and rules for sentiment analysis. These methods rely heavily on Chinese sentiment lexicons specifically tailored to capture the nuances of Chinese expressions and leverage sentiment lexicons, which are lists of words with predefined sentiment values, to assign sentiment scores to text (Wan 2011; Wang et al. 2016b). The lexicons often require extensive customization and expansion to handle the diverse and context-dependent meanings of Chinese words, especially in Chinese microblogs where slang and abbreviations are prevalent.

With the help of lexicons, researchers can analyze the part of speech and emotional tendency of words in the sentences. As can be seen from Fig. 2, keywords such as "sentiment lexicon," "sentiment dictionary," and "emotional dictionary" in the C1 community have a strong connection with "microblog". Because of the irregular expressions on the microblog,

Table 3 The main categories of the methods, and their characteristics, merits and disadvantages

Methods	Characteristics	Merits	Disadvantages
Lexicon-based approaches including rule-based methods (Du et al. 2022; Liu et al. 2009, 2012, 2014; Xu et al. 2010; Gu et al. 2022; Qiang et al. 2005; Du and Tan 2010; Zhang et al. 2008a, 2018, 2012; Lu et al. 2010; Gravano 2010; Wan 2011; Wang et al. 2016b, 2016c; Wu et al. 2016; Jiang et al. 2013; Xianghua et al. 2013; Yang et al. 2013; Hu and Li 2011; Lipenkova 2015; Shi et al. 2013; Wen and Li 2014; Yan et al. 2018)	Utilize lexicons (lists of words with associated sentiment scores). Assign sentiment scores to words and calculate an overall sentiment score based on the words present in the text e.g., Lexicon-Based Approaches (e.g., using sentiment lexicons like HowNet, National Taiwan University Sentiment Dictionary (NTUSD), and WordNet) and Rule-based approaches (e.g., defining rules based on linguistic patterns)	Simple to implement, interpretable, can handle domain-specific sentiment, and work well for short texts	• Limited Capability for Complex Text Handling: Lexicon-based methods often struggle with understanding context. Texts with ambiguous or multi-meaning words can confuse the lexicon-based approach • Challenges arise when handling extensive amounts of text data, particularly with sentences and paragraphs. The method might falter in comprehending complex semantic contexts and implicit sentiment expressions within lengthy texts • Difficulty in Capturing Diverse Expressions: The method's coverage of sentiment expressions may be restricted, especially in discerning nuances like sarcasm and contextual variations
Traditional machine learning algorithms (Yan et al. 2014; Tan and Zhang 2008; Wang and Parth 2016; Qiang et al. 2005; Wan 2011; Li and Sun 2007; Wang et al. 2007, 2014b; Zhang et al. 2008b, 2011, 2015; Zhai et al. 2009; Zheng and Ye 2009; He 2011; Yao et al. 2011; Yang and Yu 2013; Su et al. 2014; Bai and Yu 2016; Zhang and Zheng 2016; Zheng et al. 2018)	Train supervised traditional machine learning models, e.g., Naive Bayes Classifiers, Support Vector Machines (SVM), Decision Trees, Random Forest, etc., using labeled multilingual sentiment datasets	Can capture complex relationships in data, adaptable to various domains, and suitable for medium-sized datasets	• Reliance on Annotated Data: The method necessitates labeled training data, which could pose a challenge when acquiring or preparing datasets with accurately annotated sentiments for training purposes • Need for Appropriate Feature Selection: The process requires careful consideration in choosing relevant and effective training features. Incorrect or inadequate feature selection might impact the model's performance • Difficulty in Interpreting Sarcasm: The method may struggle to accurately interpret or detect sarcasm, a nuanced form of expression, thereby affecting the precision of sentiment analysis outcomes

Table 3 (continued)

Methods	Characteristics	Merits	Disadvantages
Deep learning models including transfer learning methods (Zhou et al. 2022; Sun et al. 2016; Sun et al. 2016; Li et al. 2014, 2016; Tai et al. 2015; Yanmei and Yuda 2015; Peng et al. 2018; Xu et al. 2019b; Wang et al. 2020d; Liu and Shen 2020; Yang et al. 2021; Feng and Cheng 2021; Zhao et al. 2022; Liu et al. 2022; Hu et al. 2023; Teo et al. 2023; Poria et al. 2018)	Utilize deep learning architectures such as Recurrent Neural Networks (RNNs) based methods, and Transformer based models (like BERT) to capture complex patterns in sentiment expression in the text Fine-tune pre-trained sentiment analysis models on general text in Chinese and other languages. This approach can leverage knowledge from a source language to improve performance in the target language E.g., Deep Learning Models (e.g., LSTM, CNN) Transfer Learning Methods (e.g., using pre-trained models like BERT, GPT)	Capture sequential patterns in text Capture contextual information and achieve state-of-the-art performance. Such as Transformer based methods (e.g., BERT, GPT)	• Necessity for Abundant Annotated Data: Deep learning methods often require substantial volumes of data for effective training, which can be challenging and resource-intensive to obtain or create • High Computational Demands: Deep learning models are known for their computational hunger, often demanding significant processing power and resources, potentially limiting their practicality in resource-constrained environments • Black-box Nature and Lack of Transparency in Decision-Making: These methods often function as black-box models, lacking transparency in how they arrive at conclusions or make predictions, which can hinder interpretability and understanding of their inner workings • Difficulty in Handling Out-of-Domain Information: Deep learning models may struggle when confronted with data outside their trained domain, leading to potential inaccuracies or misinterpretations, especially when applied to unfamiliar or diverse datasets
Hybrid approaches (Xu et al. 2019a; Yang et al. 2020; Jia 2022a; Wu et al. 2022, 2019b; Jiang et al. 2013; Xianghua et al. 2013; Yan et al. 2018; Wang et al. 2007; Yang and Yu 2013; Pan et al. 2011; Day and Lin 2017; Fu et al. 2017, 2018; Xiong et al. 2018; Fang et al. 2018; Zhou et al. 2019; Long et al. 2019; Ahmed et al. 2020; Gan et al. 2021; Zhang et al. 2022)	Combine multiple methods, such as using a lexicon-based approach to initialize sentiment scores and then fine-tuning using machine learning or deep learning techniques e.g., Hybrid Approaches (combining the strengths of multiple methods, like lexicon-based and machine learning or deep learning)	Combine strengths of different methods, potentially improving accuracy and robustness	• Augmented Complexity in Implementation: Hybrid methods often introduce increased complexity due to the integration of multiple techniques or approaches. This augmented complexity might pose challenges in implementation and understanding • Need for Integration Optimization: Integrating distinct methods within a hybrid framework demands meticulous optimization to ensure seamless cooperation and effectiveness among the combined components. This optimization process can be intricate and time-consuming • Persistence of Individual Method Limitations: Despite their integration, hybrid methods may still retain or inherit limitations from the individual techniques or approaches they amalgamate. These limitations could impact the overall performance or efficiency of the hybrid system

the basic lexicons cannot contain all the meaning or the part of speech of words (Wu et al. 2016; Jiang et al. 2013). Many Out-Of-Vocabulary (OOV) words are commonly used in Chinese microblogs, so constructing a new lexicon can help identify phrases (n-grams) more accurately and reduce semantic ambiguity. This application is critical in handling the rich emotional expressions found in Chinese social media. Some researchers will expand

based on dictionaries built by professional organizations (Xu et al. 2010; Zhang et al. 2018, 2012; Xianghua et al. 2013), while others will build a new dictionary based on the language characteristics of user comments or opinions in specific scenarios (Wu et al. 2016; Yang et al. 2013).

Rule-based approaches, on the other hand, rely on predefined linguistic patterns to determine sentiment in Chinese. These approaches often involve sentence dependency analysis, emotional word recognition, and other feature extraction techniques, which are based on predefined sentiment values or rule templates (Liu et al. 2009, 2012, 2014; Wang et al. 2016c). In Chinese sentiment analysis, the effectiveness of rule-based methods can be particularly strong when the rules account for the language's unique characteristics, such as its rich use of idioms, homophones, and the absence of word boundaries (Liao et al. 2022; Bhaskaran et al. 2022; Sun et al. 2022). If this method can exhaust the rule template as much as possible, it can achieve good sentiment analysis results. However, the limitations of rule-based methods become apparent when dealing with irregular or large-scale corpora in Chinese, especially given the diversity and complexity of Chinese expressions. The challenge is exacerbated by the frequent use of colloquialisms and new terms on Chinese social media platforms like Chinese Weibo, which are not always captured by traditional rule templates. As shown in Fig. 2, the "rule-based" keyword appears in the C3 community. This reflects the close relationship between rule-based methods and machine learning approaches in Chinese sentiment analysis. To overcome the limitations inherent in rule-based methods, these techniques are often combined with machine learning or deep learning methods to enhance the accuracy and robustness of Chinese sentiment analysis (Wen and Li 2014; Yan et al. 2018).

3.3.2 Traditional machine learning algorithms

As more and more users or consumers express their opinions and experiences about products or services online, a large number of comments or opinion texts are in urgent need of sentiment analysis and mining (Zhang et al. 2011; Wang et al. 2014b). While lexicon-based methods have traditionally been used, they struggle to handle the sheer volume and diversity of online content, especially in the context of big data. Similarly, rule-based methods often fall short in accuracy when applied to large-scale Chinese text corpora, which are characterized by linguistic complexity and the frequent emergence of new terms (Liu et al. 2009; Xu et al. 2010; Qiang et al. 2005).

To address these challenges, traditional machine learning algorithms have been increasingly employed in Chinese sentiment analysis. These algorithms, such as Support Vector Machines, Naive Bayes, Decision Trees, and Random Forest, excel at capturing patterns in data without requiring extensive computational resources (Wang et al. 2014b). By using features derived from Chinese text data, these models can effectively distinguish between different sentiment classes, making them well-suited for processing the vast amounts of textual data generated in the Chinese language (Li and Sun 2007; Zhang et al. 2008b; Zheng and Ye 2009).

Researchers have explored various text features to enhance the accuracy of sentiment analysis. Commonly used features include TF-IDF (Zhang and Zheng 2016; Zheng et al. 2018; Li and Sun 2007), continuous bag-of-words and continuous skip-gram models (Bai and Yu 2016), n-grams (Zhang et al. 2008b; Zheng et al. 2018), Word2Vec (Su et al. 2014;

Zhang et al. 2015), DF value (which indicates the frequency of features across reviews) (Zheng et al. 2018), as well as syntactic and semantic features (Yang and Yu 2013). Additionally, features that reflect the unique characteristics of Chinese text, such as those with category-distinguishing ability and sentiment orientation, have been identified for improving model performance (Wang et al. 2007).

In Fig. 2, keywords related to text features (e.g., "feature selection," "TF-IDF," "information extraction") and those related to rules (e.g., "rule-based") are both prominent in the C3 community. This grouping highlights the integration of algorithmic rules with statistical analysis to enhance Chinese sentiment prediction. However, the focus that remains predominantly on traditional machine learning methods have proven particularly effective for Chinese sentiment analysis (Sun et al. 2022; Yang and Yu 2013; Su et al. 2014; Zhang et al. 2015).

3.3.3 Deep learning models

With the escalating prominence of Web 2.0, microblogging has become integral in our daily lives. Platforms like Chinese Weibo facilitate users to freely share reviews and opinions on diverse subjects, including products and events. Leveraging this public opinion data offers valuable insights into societal perspectives. This resource is particularly beneficial for businesses seeking to comprehend user sentiments toward their products, thereby aiding in the development of enhanced offerings (Li et al. 2016).

While traditional machine learning models have showcased proficiency in Chinese sentiment analysis, their effectiveness dwindles when processing vast volumes of data. This decline in performance often stems from their heavy reliance on feature extraction methods, which struggle with the brevity of Chinese microblog texts and the continuous emergence of new words in the Chinese language (Xu et al. 2019b). The brevity of Chinese microblog texts restricts feature extraction possibilities. Furthermore, the continuous emergence of new words and extensive network data complicate the extraction and training of text features for traditional machine learning approaches (Sun et al. 2016).

Deep learning methods present a compelling solution by bypassing manual feature extraction. These models utilize neural networks to autonomously discern intricate patterns and relationships within text data (Zhao et al. 2023). Their capacity to automatically learn complex patterns alleviates the burden of manual annotation, thereby reducing overall costs associated with sentiment analysis tasks.

As shown in C2 community of Fig. 2, evident is the prevalence of deep learning methods surpassing traditional machine learning approaches in sentiment analysis. This C2 community stands out with its extensive network of nodes and connections, indicating the significance of deep learning in the sentiment analysis domain. Notably, keywords from other sub-communities exhibit strong ties to those within the prominent C2 community. Recent advancements have seen a surge in the adoption of deep learning models for sentiment analysis. Researchers have employed diverse models, including but not limited to:

- Long short-term memory (LSTM) (Day and Lin 2017),
- Tree-LSTM (Tai et al. 2015),
- Bidirectional long short-term memory (Bi-LSTM) (Wang et al. 2016d),
- Recursive neural deep model (RNDM) (Li et al. 2014),

- Convolutional neural networks (CNN) (Yanmei and Yuda 2015; Li et al. 2016; Xu et al. 2019b),
- Deep neural network (DNN) stacked with Restricted Boltzmann machine (RBM) layers (Sun et al. 2016),
- Bidirectional encoder representations from transformers (BERT) (Wang et al. 2020d),
- Gated alternate neural network (GANN) (Liu and Shen 2020),
- Innovative models derived from popular architectures (Peng et al. 2018; Yang et al. 2021; Feng and Cheng 2021).

Notably, the graph convolutional network (GCN) has gained prominence in recent years due to its ability to encode both graph structures and node features (Bruna et al. 2014). Its widespread adoption spans various domains within natural language processing, including sentiment analysis (Zhao et al. 2022; Liu et al. 2022). Categorizing these models underlines the significance of neural networks in capturing contextual information crucial for understanding subtle sentiment nuances.

3.3.4 Hybrid approaches

Hybrid methodologies in Chinese sentiment analysis integrate multiple techniques to enhance accuracy and robustness, addressing the specific linguistic and cultural complexities of the Chinese language. Two primary hybrid approaches have emerged in current research.

One approach combines lexicon-based methods with traditional machine learning or deep learning models, particularly to tackle the challenges posed by Chinese text, such as character ambiguity and the lack of word boundaries. For instance:

- Dictionary-based hybrid methods: Researchers like Xu et al. extended Chinese sentiment lexicons, integrating them with Naive Bayes (NB) (Wang et al. 2016b), while Wang and Jiang selected Chinese informative words combined with Support Vector Machine (SVM) models to improve sentiment classification accuracy (Jiang et al. 2013; Wang et al. 2007).
- Leveraging LDA and HowNet: Fu et al. used Latent Dirichlet Allocation (LDA) for topic identification, employing HowNet for sentiment polarity classification (Xianghua et al. 2013). Similarly, Day et al. extracted multi-feature words from Chinese sentiment resources, like NTUSD, HowNet, and iSGoPaSD, combining them with Bi-LSTM models to enhance sentiment prediction (Day and Lin 2017).
- Emotional Dictionary with Neural Networks: Yang et al. merged Chinese emotional dictionaries with CNN and attention-based Bidirectional Gated Recurrent Unit (BiGRU) models to develop the SLCABG model, which effectively captures the intricacies of Chinese emotional expressions (Yang et al. 2020). Ahmed et al. crafted Chinese domain-specific sentiment dictionaries, integrating them with LSTM models for aspect-level sentiment analysis (Ahmed et al. 2020).
- Emotion Identification Based on Rules: Yan et al. identified emotional subjects using rule-based syntactic dependencies, integrating this with SVM models for automatic emotion analysis in Chinese social media (Yan et al. 2018).
- The Ortony-Clore-Collins (OCC) Model Combining Rules and Convolutional Neural

Network (CNN): Wu et al. proposed an OCC model and a CNN-based opinion summarization method. They tested their hybrid method using real-world Chinese microblog data and compared the accuracy of manual sentiment annotation with the accuracy achieved using the OCC-based sentiment classification rule library. Their study highlights the potential of the proposed hybrid method for sentiment analysis of Chinese social media data (Wu et al. 2020).

These examples demonstrate that lexicon-based approaches remain vital in Chinese sentiment analysis, especially when combined with machine learning techniques to form hybrid approaches to address the specific challenges of the Chinese language.

The other approach involves merging diverse machine learning models, leading to innovative methods for Chinese sentiment analysis. For instance:

- BT-CNN-ATT Model: Jia developed the BT-CNN-ATT model, amalgamating BERT, CNN, and attention mechanisms. This model not only extracts global features from Chinese Weibo contexts but also captures local features, such as words, to effectively mine emotional information (Jia 2022).
- Bi-LSTM with Multi-head Attention (MHAT): Long et al. explored sentiment analysis of Chinese social media texts by integrating Bi-LSTM networks with the Multi-head Attention (MHAT) mechanism. This fusion aims to address the limitations of traditional machine learning approaches (Long et al. 2019).
- Combining CNN with LSTM or Bi-LSTM: Several studies have merged CNN with LSTM (Zhang et al. 2022) or Bi-LSTM (Gan et al. 2021) models, aiming to enhance the effectiveness of Chinese sentiment analysis.
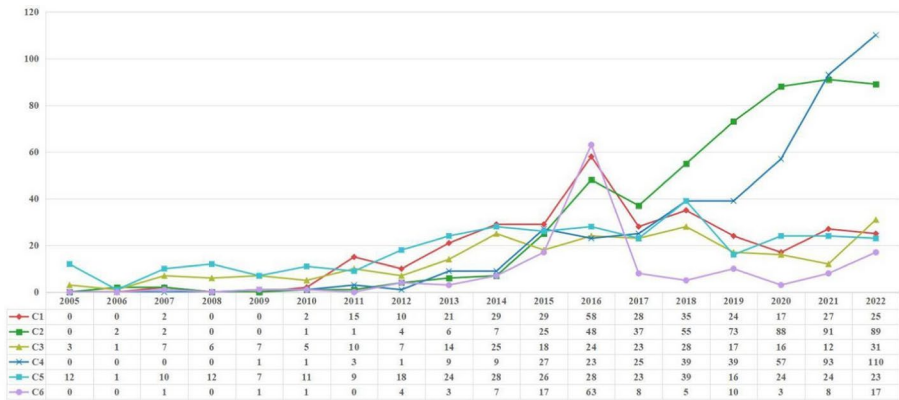
This categorization framework provides a structured understanding of the various approaches used in sentiment analysis. It emphasizes the distinctive traits of each category, elucidating the differences in their principles and methodologies. The above method classification scheme provides a structured framework for understanding the various methods used in sentiment analysis, highlights the unique characteristics of each category, and helps to clarify how these methods differ in their underlying principles and methodologies.

### 3.4 Trend analysis based on the number of keywords

Sections 3.2 and 3.3 delve into the distinct facets of Chinese sentiment analysis research: Subjects and methods. The analysis involves tracking annual shifts in keyword frequency, serving as a reflection of the evolutionary trajectory of research methods and topics within this field (Cui et al. 2023). This section leverages the keyword community network (referenced in Fig. 2) together with the analysis achievement from Sects. 3.2 and 3.3 to quantify the annual word frequency across sub-communities. By discerning fluctuations in the number of keywords over time, these sections aim to elucidate the evolving research trends within the realm of Chinese sentiment analysis. Furthermore, the visual representation of the keyword community's evolution is presented in Fig. 3, serving as a graphical aid to complement and illustrate the changes in research trends over the years.

To elucidate the evolutionary patterns of keywords within each community over the years, we meticulously curated high-frequency keywords representative of each category.

	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022
C1	0	0	2	0	0	2	15	10	21	29	29	58	28	35	24	17	27	25
C2	0	2	2	0	0	1	1	4	6	7	25	48	37	55	73	88	91	89
C3	3	1	7	6	7	5	10	7	14	25	18	24	23	28	17	16	12	31
C4	0	0	0	0	1	1	3	1	9	9	27	23	25	39	39	57	93	110
C5	12	1	10	12	7	11	9	18	24	28	26	28	23	39	16	24	24	23
C6	0	0	1	0	1	1	0	4	3	7	17	63	8	5	10	3	8	17

Fig. 3 Evolution diagram of keywords obtained using all the keywords related to sentiment analysis in Chinese text from 2005 to 2022. Communities C1–C6 in figure correspond to different community categories within the keyword co-occurrence network. Figure illustrates the changes in the number of keywords within each community over the years. By examining figure, one can discern whether the research themes represented by each community are gaining attention or gradually declining in popularity, thus revealing trends in research topics and subjects for Chinese sentiment analysis
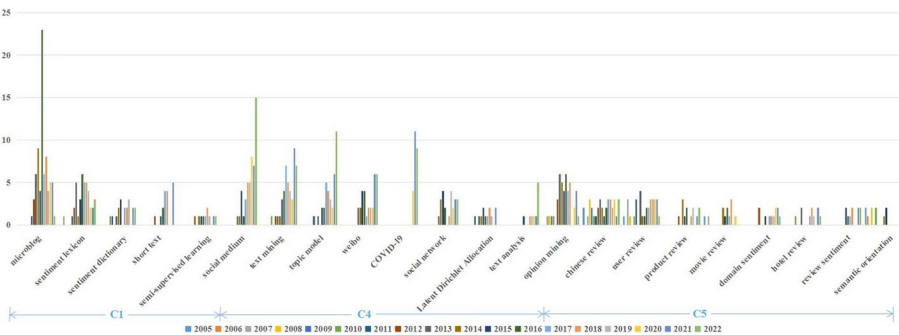
Fig. 4 Keyword evolution diagram of C1, C4, and C5 communities. Figure shows the high-frequency keywords for communities C1, C4, and C5, displaying the frequency counts of these keywords over the years. This visual representation allows for a clear observation of the evolving trends of high-frequency keywords within each community. Figure also reflects shifts in research focus and methodologies, providing deep insights into the dynamic changes within the field by highlighting the rise of emerging trends and the decline of older topics. For instance, it can be observed that sudden events, such as "COVID-19," triggered a temporary surge in attention to related keywords starting from 2019 in the context of Chinese sentiment analysis

The fluctuations in word frequency for these selected keywords across different years were carefully plotted and are vividly presented in Figs. 4 and 5. These figures illustrate the changes in the number of high-frequency keywords within each category, reflecting how certain keywords, such as "deep learning" and "attention mechanism," have gradually gained attention from researchers over time. Conversely, interest in some keywords, such as "SVM" and "hotel review," has decreased. Additionally, some keywords experienced a temporary surge in attention due to emerging events, such as "COVID-19." The variations
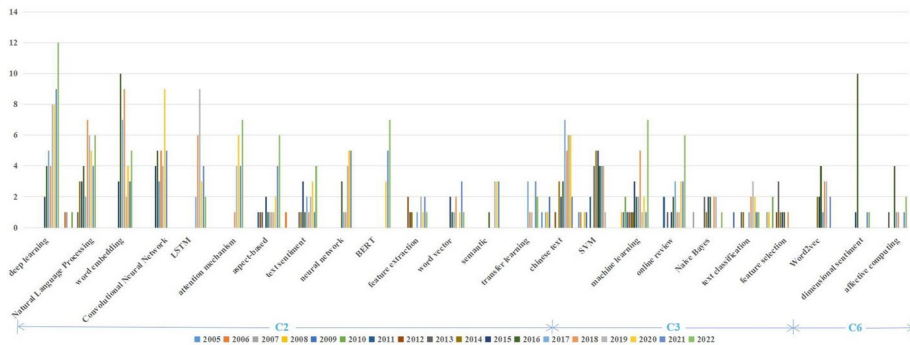
Fig. 5 Keyword evolution diagram of C2, C3, and C6 communities. Similarly, figure shows the high-frequency keywords for communities C2, C3, and C6, along with their frequency counts over the years. This visual representation allows for a clear observation of the evolving trends of high-frequency keywords within each community. For instance, the keyword "deep learning" has gradually gained more attention since 2015, whereas it was not a prominent topic before that year

in keyword counts shown in the figures, combined with textual analysis, highlight the trends in evolving topics and methods in the field of Chinese sentiment analysis.

Combining the data shown in Figs. 3, 4 and 5, we can see that the earliest research mainly started with keywords in the C5 community. Keywords in the C5 community mainly involve "opinion mining," "user review," "Chinese review," etc. The number of keywords has continued to grow over the years, reflecting researchers' focus on mining user opinions and comments. The keywords of the C1 community predominantly revolve around sentiments and emotions related to "microblog," "sentiment lexicon," and "sentiment dictionary". The frequency of keywords has been growing rapidly before 2016, and reached a peak in 2016. It was mainly affected by the "microblog" keyword, and also showed the importance of sentiment-related dictionaries in sentiment analysis. After 2016, the number of keywords exhibited a declining trend, indicating that scholars' research inclination has shifted towards exploring technology (C2) and specific trending topics (C4). Keywords in the C2 and C3 communities are mainly related to sentiment analysis technology (see Fig. 5). The keywords of the C2 community mainly involve words related to deep learning methods such as "Deep learning", "Natural language Processing", "word embedding", "convolutional neural network", "LSTM" and so on. Since 2014, the frequency of keywords has increased significantly, highlighting researchers' high attention on deep learning methods in the field of sentiment analysis.

In addition, with the advancement of technology, the granularity of sentiment analysis has gradually changed from document-level and sentence-level to more fine-grained aspect-level (Peng et al. 2018; Fu et al. 2017; Wu et al. 2012; Zhang et al. 2014; Zhang and Lan 2015; Wang et al. 2017; Chen et al. 2022). Document-level sentiment analysis typically provides a general sentiment score for an entire document, while sentence-level analysis focuses on using individual sentences to determine sentiment. Wu et al. (2012) address the challenges of analyzing sentiment in comparative sentences within Chinese documents, proposing a method that enhances document-level sentiment analysis. Zhang et al. (2014) question the reliability of star ratings in capturing review sentiment and present a framework that improves phrase-level sentiment labeling through unsupervised review-level classification. Zhang and Lan (2015) present an innovative approach to enhancing sentiment analysis

at both the word and sentence levels by incorporating sentiment information directly into word embeddings. Fu et al. (2017) enhance sentence-level sentiment analysis by integrating the HowNet lexicon with a Phrase Recursive Autoencoder.

However, as the demand for more precise insights has grown, aspect-level sentiment analysis has emerged as a powerful tool. This approach involves dissecting text into various aspects or features, allowing for a nuanced understanding of sentiment toward specific elements (Peng et al. 2018; Wang et al. 2017; Chen et al. 2022). For instance, in the context of product reviews, aspect-level sentiment analysis can identify and analyze sentiments related to different facets such as the product's design, functionality, durability, and customer support. Peng et al. (2018) introduce the Aspect Target Sequence Model for Chinese aspect-based sentiment analysis, capturing multi-grained text representations and improving sequential information analysis. Chen et al. (2022) develop a language-agnostic, discrete latent opinion tree model for aspect-based sentiment analysis, offering improved interpretability and robustness in low-resource scenarios. Collectively, these studies underscore the ongoing advancement of sentiment analysis in NLP, particularly within the Chinese language.

Moreover, the advancement of aspect-level sentiment analysis is supported by sophisticated techniques such as dependency parsing, attention mechanisms in neural networks, and transfer learning. These innovations enable systems to better capture contextual relationships and subtle variations in sentiment. As a result, the analysis becomes more accurate and actionable, offering a detailed understanding of customer opinions and enhancing decision-making processes across various domains. Researchers are actively tackling challenges such as complex sentence structures, multi-word aspect targets, and domain-specific nuances. The shift towards finer-grained, aspect-level analysis highlights the growing need for nuanced sentiment understanding in various applications, driving the evolution of sentiment analysis research and enhancing its practical impact.

From Fig. 5, we can see that the word frequency of "aspect-based" keywords reached the highest in 2022. The C3 community mainly includes keywords related to machine learning and text features such as "machine learning," "SVM," "NB," and "feature selection". The frequency of keywords in the C3 community had been increasing before 2018. However, after 2018, the number of keywords showed a downward trend, reflecting the fact that those researchers paid more attention to the application of deep learning methods in sentiment analysis in the C2 community. The keywords of the C4 community mainly involve social media topics related to "Weibo," "social medium," "topic model," and "COVID-19". The number of keywords experienced two significant increases between 2015 and 2020, indicating that researchers are increasingly interested in social media research, especially during COVID-19, with special attention paid to users' emotional expressions on social media during the epidemic (Gu et al. 2022; Wang et al. 2020d; Yang et al. 2022a; Gao et al. 2021). The C6 community contains keywords related to various complexities in the sentiment analysis process. C6 reached its peak in 2016 in terms of the number of keywords. This is attributed to topics such as "multi-dimensional sentiment analysis," "word segmentation," "affective lexicon," "Word2vec," and "affective computing."

From the perspective of trend analysis, the sentiment analysis of user opinion mining and comments has always been the focus of research. However, in recent years, research topics have gradually tended to explore hot topics on social media platforms, especially Weibo, China's largest social media platform, where the most discussed topic is COVID-19. This

is similar to the results obtained from a review of sentiment analysis in English literature, indicating that the world has paid attention to sentiment analysis research on social media platforms in recent years, especially on the topic of COVID-19 (Cui et al. 2023). In terms of sentiment analysis technology, the focus of research has gradually shifted from machine learning to deep learning, and hybrid methods that can mutually compensate for the short-comings of the model have also been receiving attention. The conclusion is also similar to our research on English sentiment analysis. However, we found that in Chinese sentiment analysis research, sentiment lexicon and sentiment dictionary have received greater atten-tion than English sentiment analysis research. In addition, the graph convolutional network is gradually applied to the field of sentiment analysis (Zhao et al. 2022; Liu et al. 2022; Wang et al. 2022; Yang et al. 2022b). It reflects the complexity of Chinese text language sentence patterns, word meanings, etc. Integrating lexicons and knowledge graphs into sen-timent analysis research can help improve the accuracy of sentiment analysis.

3.5 Chinese-language datasets and evaluations

Chinese sentiment analysis has made notable strides, fueled by the development of high-quality datasets, various dimensional sentiment analysis methods, and performance eval-uations. This subsection provides a concise overview of key Chinese-language datasets, different dimensional sentiment analysis methods supported by these corpora, and their per-formance evaluations.

3.5.1 Chinese sentiment datasets

As shown in Fig. 2, among the high-frequency keywords, several of them are related to cor-pora, such as "corpus" in community C1, "corpus Chinese" in community C5, and "labeled corpus" in community C6. This reflects the fact that Chinese sentiment corpora is a key focus for Chinese sentiment analysis. Numerous Chinese sentiment corpora have been established, including the NLPCC 2023 Task 4 dataset (DiaASQ dataset, which includes both Chinese and English languages) (Li et al. 2023a), the SemEval-2016 Task 5 dataset (Hercigt et al. 2016), a large-scale Chinese restaurant review dataset (Bu et al. 2021), the Chinese EmoBank (Lee et al. 2022), and a Chinese three-dimensional corpus with valence-arousal-irony (VAI) ratings (Xie et al. 2021). Different research scenarios, such as sarcasm detection—a challenging task involving non-literal expressions—also require specialized corpora such as the Chinese Irony Corpus for support (Xiang et al. 2020; Tang and Chen 2014). Other corpora, such as HowNet (Xianghua et al. 2013; Fu et al. 2017) and NTUSD (Day and Lin 2017) are Chinese sentiment lexicons used to support lexicon-based or hybrid methods.

Chinese sentiment corpora are foundational to the study of sentiment analysis in Chinese, providing essential datasets for training and evaluating different methods. These corpora encompass various aspects of sentiment analysis, including Aspect-Based Sentiment Analy-sis (ABSA) (Li et al. 2023a), irony detection (Xiang et al. 2020; Tang and Chen 2014), and fine-grained sentiment analysis (Zhou et al. 2022; Xu et al. 2010; Gu et al. 2022), among oth-ers, each addressing different facets of how sentiment is expressed in the Chinese language (Zhao et al. 2023; Xu et al. 2016). For instance, a comprehensive Chinese restaurant review dataset was designed for Aspect Category Sentiment Analysis (ACSA) and rating prediction

(RP) (Li et al. 2023a). It contains 46,730 genuine reviews, meticulously annotated with sentiment polarities across 18 fine-grained aspect categories and 5-star ratings, making it one of the largest and most detailed Chinese sentiment datasets. Ciron (Xiang et al. 2020) stands as the first Chinese benchmark for irony detection, featuring 8.7 K microblog posts annotated for irony by native speakers, facilitating the exploration of irony detection in Chinese using machine learning models. Additionally, Chinese EmoBank (Lee et al. 2022) provides resources for dimensional sentiment analysis, with valence-arousal ratings for 5512 words, 2998 multi-word phrases, 2582 single sentences, and 2969 multi-sentence texts, supporting the development of fine-grained sentiment models for the Chinese language. C-STANCE (Zhao et al. 2023) focuses on zero-shot stance detection (ZSSD), presenting a dataset of 48,126 annotated text-target pairs that span various domains and introduce complex target-based and domain-based ZSSD tasks. These Chinese sentiment corpora support different dimensional sentiment analysis in Chinese.

3.5.2 Dimensional sentiment analysis

Chinese sentiment datasets (also known as Chinese sentiment corpora) support different levels of dimensional sentiment analysis in Chinese, which is an advanced approach extending beyond traditional binary or categorical sentiment classification (such as positive, negative, or neutral) (Lee et al. 2022; Xie et al. 2021). Instead of merely classifying text as positive or negative, dimensional sentiment analysis assesses various dimensions, such as emotions (joy, sadness, anger), intensity, and specific aspects of sentiment related to different topics or entities within the text (Lee et al. 2022; Zhao et al. 2023). It also involves fine-grained sentiment analysis, which evaluates not only the presence of positive or negative sentiment but also the degree of sentiment (e.g., strongly positive, slightly positive, strongly negative, and slightly negative) (Zhou et al. 2022; Xu et al. 2010; Gu et al. 2022).

Dimensional sentiment analysis in Chinese has gained increasing attention as researchers aim to capture the nuanced emotions expressed in the Chinese language, which often carries cultural and linguistic subtleties that differ from other languages (Du et al. 2022; Liao et al. 2022; Sun et al. 2022; Guo et al. 2022; Yang et al. 2020; Xiang et al. 2020; Lee et al. 2022; Bu et al. 2021; Zhao et al. 2023). Chinese dimensional sentiment analysis is an important research topic, with existing studies primarily focusing on various analytical methods. These methods include position-based character-enhanced word embedding and FastText as word vectors (Du and Zhang 2016), the Community-Based Weighted Graph Model (Wang et al. 2016a), a semi-supervised approach for DSA based on the variational autoencoder model (Wu et al. 2019), a tree-structured regional CNN-LSTM model (Wang et al. 2020c), and deep learning-based methods (Cheng et al. 2021; Deng et al. 2023).

3.5.3 Performance evaluations

For evaluating the effectiveness of these Chinese sentiment corpora and methods, different evaluation metrics such as F1-score (Li et al. 2023a; Xiang et al. 2020; Hercigt et al. 2016; Bu et al. 2021; Zhao et al. 2023), accuracy (Xiang et al. 2020; Hercigt et al. 2016; Bu et al. 2021), mean absolute error (MAE), and Pearson correlation coefficient (r) (Lee et al. 2022; Xie et al. 2021) are commonly used. A review might be positive overall but have different sentiments toward various aspects like service quality, product features, or pricing (Lee et

al. 2022; Xie et al. 2021). For examining different Chinese dimensional sentiment analysis, a sufficient size of Chinese corpus (e.g., Chinese sentiment dataset) is necessary. For instance, a Chinese EmoBank (Lee et al. 2022), and a Chinese three-dimensional corpus with valence-arousal-irony (VAI) ratings (Xie et al. 2021) provide the Chinese copus support for such dimensional sentiment analysis.

Evaluations on these datasets highlight the effectiveness of various models, including different dimensional sentiment analysis methods. In Aspect Category Sentiment Analysis (ACSA) tasks, BERT-based models have consistently outperformed traditional methods on datasets like ASAP and the SemEval-2016 RESTAURANT dataset in terms of accuracy (Hercigt et al. 2016), with further enhancements achieved through joint learning models that combine ACSA and rating prediction (RP) in terms of F1-score (Bu et al. 2021). For zero-shot stance detection (ZSSD) on C-STANCE, transformer-based models such as RoBERTa and XLNet have set new benchmarks, surpassing RNN-based models and underscoring the complexity of claim target prediction in terms of F1-score as well (Zhao et al. 2023). BERT has also demonstrated superior performance in irony detection tasks on the Ciron dataset (Xiang et al. 2020), as well as in predicting valence and arousal ratings across different text granularities in Chinese EmoBank in terms of Pearson correlation coefficient (Lee et al. 2022).

These evaluations underscore the significant advancements in Chinese sentiment analysis, driven by the availability of rich Chinese sentiment analysis datasets, different dimensional sentiment analysis methods, and the applications of such methods including powerful deep learning models.

4 Practical applications and future prospects of Chinese sentiment analysis

In Sect. 3, we analyze the themes, methods and trends of existing Chinese sentiment analysis. The practical applications, limitations, and future prospects of Chinese sentiment analysis hold significant implications for industries, communication, and technology.

Despite extensive exploration within existing studies, no publications have explored the use of ChatGPT for sentiment analysis when we conduct this survey research. While there might be such studies available, our search across prominent databases, including the Conference Proceedings Citation Index—Social Sciences & Humanities (CPCI-SSH), Conference Proceedings Citation Index—Science (CPCI-S), Science Citation Index Expanded (SCI-Expanded), and Social Sciences Citation Index (SSCI), yielded no publications related to ChatGPT in this context.

4.1 ChatGPT for Chinese sentiment analysis

Language models like ChatGPT can be categorized under the broader category of Deep Learning Models. These models, including transformer-based architectures like ChatGPT, belong to the deep learning category because they leverage neural network architectures with multiple layers to understand and generate human-like text. Deep learning models are known for their ability to capture complex patterns and relationships in data, making them highly suitable for tasks like sentiment analysis.

Among advanced language models, ChatGPT, powered by transformer architectures, has emerged as a versatile tool for sentiment analysis. In a recent breakthrough, GPT-4 showcased remarkable "theory of mind" abilities by effectively tackling 95% of false-belief tasks (Brunet-Gouet et al. 2023), This achievement hints at the exciting possibility of GPT-like models evolving to acquire cognitive empathy, potentially revolutionizing their role in understanding human emotions and intentions. ChatGPT is pre-trained on a diverse range of text sources, enabling it to capture linguistic variations, idiomatic expressions, and cultural nuances. Its capacity for transfer learning empowers it to adapt to different languages, including Chinese. ChatGPT's capabilities in multilingual sentiment analysis include:

- Contextual "Understanding": ChatGPT's ability to "understand" context aids sentiment analysis by capturing the meaning behind sentiment-bearing phrases. In a recent study encompassing affective computing tasks, ChatGPT excelled in sentiment analysis, underscoring its prowess in grasping emotional nuances in text (Amin et al. 2023). Polgan et al. underscored the pivotal role of ChatGPT in decoding sentiment and emotions, especially in the context of making informed decisions within the ever-evolving specific landscape (Polgan et al. 2023). With its proficient text analysis capabilities, ChatGPT serves as a valuable tool, equipping users with actionable insights to navigate and gain a deeper understanding of sentiments. Notably, ChatGPT's ability to capture linguistic nuances and adapt to the Chinese sentiment analysis domain positions it as a promising asset for future applications in this field.
- Cross-Lingual Adaptation: The transfer learning approach of ChatGPT enables it to generalize sentiment understanding from one language to another. Previously, Zhang et al. introduced an innovative approach known as hybrid-tuning, aimed at addressing the challenge of catastrophic forgetting (Zhang et al. 2023a). This method seamlessly merges both general and domain-specific knowledge, while also harmonizing the stages of pre-training and fine-tuning. The result is a system that excels in delivering precise and contextually fitting responses within the specific Chinese domain. It is worth noting that such a method holds significant promise for extending its applicability to emerging emotional tasks.
- Handling Idiomatic Expressions: ChatGPT's exposure to a wide array of linguistic contexts equips it with the ability to handle idiomatic expressions and nuanced sentiments. For idiom fill-in-the-blank and idiom understanding, Li et al. observed that ChatGPT which provides a straightforward chain-of-thought prompt can enhance performance in intricate reasoning (Li et al. 2023b).

Amin et al. evaluated the ability of ChatGPT on English affective computing problems, and the results showed that ChatGPT is a good generalist model that can achieve good results on a variety of problems without any professional training (Amin et al. 2023). A few studies have also confirmed the feasibility of ChatGPT in the field of Chinese question answering and event extraction (Shao et al. 2023; Zhu et al. 2023). The utilization of LLMs, such as ChatGPT, for sentiment analysis in a Chinese-centric context will be one of the future research directions and applications (Fu et al. 2024).

4.2 Global business insights

For multinational businesses, understanding sentiment in the Chinese language is invaluable. As the world's most widely spoken language with over 1.4 billion native speakers, Chinese market advantage enables multinational businesses to effectively target Chinese consumers and scale their offerings to a massive population. China's economic influence calls for precise sentiment analysis in a Chinese-centric context, aiding market strategies and decision-making. In a thorough evaluation using a dataset of 7165 financial questions, Ren et al. found that ChatGPT demonstrated higher levels of professionalism and accuracy compared to human services in Chinese financial conundrums, resulting in increased efficiency, cost savings, and improved customer satisfaction, ultimately enhancing the competitiveness and profitability of financial institutions (Ren et al. 2023). Cross-cultural diplomacy is enhanced through accurate sentiment analysis. Chinese-centric sentiment understanding fosters better cross-border communication, essential for diplomatic relations and cultural exchanges. The Chinese market's magnitude makes sentiment analysis pivotal for e-commerce. Chinese-centric analysis unveils consumer preferences, leading to tailored marketing strategies that resonate with Chinese consumers.

4.3 Enhanced Chinese language models

Advancements in Chinese-specific pre-trained models will bolster sentiment analysis accuracy, addressing linguistic intricacies and cultural subtleties. Models like ChatGPT can further adapt to Chinese-specific linguistic nuances, catering to the dynamic expressions of sentiment across regions. Wang et al. exhibited that GPT-4's proficiency is on par with Chinese participants who passed a specific exam, demonstrating its potential for discharge summarization, group learning, and strong verbal fluency in human–computer interactions (Wang et al. 2023c). ChatGPT significantly enhances various capabilities, particularly in Chinese-specific task comprehension, while also addressing issues like hallucinations, legal risks, and ethical concerns. This suggests that ChatGPT has the potential to serve as an essential component of the Chinese-centric sentiment analysis. Besides, creating more labeled data in Chinese is vital for robust sentiment analysis.

Hassani et al. explored the potential impact of ChatGPT on data science, presenting opportunities and challenges, that emphasized ChatGPT's role in data augmentation, citing an example where ChatGPT was employed to generate synthetic radiology reports, enhancing the training data for a machine learning model in radiology report classification at the University of California, San Diego (Hassani and Silva 2023). This potentiality has spurred the application of ChatGPT as a means of data augmentation in various interdisciplinary fields. Van Nooten et al. utilized ChatGPT 3.5 to create authentic anti-vaccination tweets in Dutch, aiming to balance a skewed vaccine hesitancy classification dataset (Van Nooten and Daelemans 2023). By augmenting the gold standard data with these generated examples, the study demonstrated notable enhancements for underrepresented classes, overall recall improvement, and a slight decline in precision for more prevalent classes, while also assessing the synthetic data's generalization to human-generated data in the classification task. Collaborative efforts in data augmentation can drive more accurate Chinese-centric models. Future research should prioritize ethical multilingual sentiment analysis, accounting for cultural sensitivity and fairness in sentiment interpretation.

Buscemi and Proverbio evaluate the performance of ChatGPT 3.5, ChatGPT 4, Gemini Pro, and LLaMA2 7b in multilingual sentiment analysis, including Mandarin Chinese (Buscemi and Proverbio 2024). Their findings indicate that while ChatGPT 4 and Gemini Pro generally exhibit a positive bias towards Mandarin, showing improvements over earlier versions, they still face notable inconsistencies and biases, particularly in handling nuanced texts such as irony and sarcasm. LLaMA2, on the other hand, shows minimal variation across languages but tends to be overly optimistic. Comparisons with human responses reveal that native Mandarin speakers provided more positive feedback, somewhat aligning with the models' outputs. However, significant gaps remain, underscoring the need for further refinement to ensure cultural sensitivity and reliability in sentiment analysis across different languages.

Dong et al. conducted a detailed comparative analysis of various pre-trained language models (PLMs) for Chinese sentiment analysis in human–computer negotiation dialogues (Dong et al. 2024). They examined the advanced models such as GPT-3.5, BERT, Claude, ELECTRA, NEZHA, ERNIE 3.0, BART, and XLNet. The study assesses their effectiveness in sentiment detection using a large, diverse Chinese dataset with sentiment annotations. The findings highlight significant performance differences among the models, offering valuable insights for enhancing automated negotiation systems and improving sentiment analysis in negotiation scenarios.

Zhang et al. explored the effectiveness of three large language models (LLMs)—a baseline GPT-3.5, a Chinese language-specific Erlangshen-RoBERTa-110M-Sentiment, and a financial domain-specific Chinese FinBERT—in extracting sentiment from Chinese financial news to predict stock price movements (Zhang et al. 2023b). Using a standardized procedure, the study analyzes 394,426 news summaries, developing quantitative trading strategies based on the sentiment factors extracted by each model. Results show that the Erlangshen model, despite its smaller size, outperforms the others in annual return, risk-adjusted return, and excess return, highlighting the importance of language-specific pre-training for enhanced performance in sentiment analysis and quantitative trading in the Chinese market.

Ihnaini et al. investigated the application of advanced large language models (LLMs) for sentiment analysis of classical Chinese literature, focusing on Song Dynasty poetry (Song Ci) (Ihnaini et al. 2024). The study acknowledges the complexities of the linguistic structures and unique emotional expressions inherent in Song Ci, which pose significant challenges to traditional sentiment analysis methods. Utilizing fine-tuning techniques, the research navigates and interprets the nuanced language and emotional content of Song Ci with improved accuracy. The evaluation of the fine-tuned Chinese LLaMA 2 and Qwen models highlights their performance in detecting subtle emotional shifts within these poems. Specifically, the ChatGLM-7B (8-bit) model stands out, achieving an F1 Score of 0.840, showcasing its ability to merge ancient literary analysis with modern computational technology.

The fusion of Chinese-centric sentiment analysis and advanced language models has the potential to revolutionize cross-cultural communication and understanding. As we conclude this exploration, we reflect on the trajectory of sentiment analysis in a globalized world and envision a future where Chinese sentiment analysis paves the way for richer, more accurate cross-lingual emotional assessment.

5 Future directions

Based on the results of exploring the prospects of Chinese sentiment analysis discussed in Sects. 3–4, this section outlines several future directions that warrant further exploration.

5.1 Ethical multilingual sentiment analysis and understanding

Research in multilingual contexts introduces specific ethical challenges related to language dynamics, data handling, and presentation (Abdullah and Rusli 2021; Schembri and Jahić Jašić 2022). It is crucial to address these ethical considerations systematically. In the study by Janusch et al., an examination of interviews with Chinese teachers conducted initially in English and later in Chinese highlighted the impact of language dominance on interpretation. The researchers observed that the perspective brought by researchers from another cultural orientation and a position of power can influence interpretation (Janusch 2011). Switching to Chinese empowered participants, fostering more accurate and expressive communication, resulting in richer, more valid data and robust research outcomes. Holmes et al. recognized the widespread ethical concerns in multilingual research, they advocated for proactive ethical practices, emphasizing the importance of researcher reflexivity (Hassani and Silva 2023). Researchers are encouraged to challenge monolingual state and institutional practices. Therefore, efforts to minimize biases and promote cultural sensitivity in multilingual sentiment analysis will lead to more reliable and universally applicable models. Ethical considerations in model training and deployment are essential for fostering responsible AI development in the future.

5.2 Multilingual emotion recognition

Emotion recognition holds a pivotal role in enhancing human–computer interaction. While numerous studies have historically focused on speech emotion recognition utilizing various classifiers and feature extraction methods, the majority have primarily tackled this challenge within the confines of a single language. Numerous studies take a significant leap by expanding monolingual speech emotion recognition to encompass emotions expressed in multiple languages simultaneously, thereby constituting a comprehensive system (Lee 2019; Heracleous and Yoneyama 2019; Ma et al. 2024). This departure from the norm represents an exploration into uncharted territory, as the domain of multilingual emotion recognition extends beyond the well-established realm of sentiment analysis. Traditionally, emotions such as joy, sadness, anger, and fear have been explored in the context of single-language studies. However, the current research broadens the scope to comprehend and recognize these emotions across diverse languages (Zhang et al. 2024). This pioneering approach has the potential to revolutionize human–computer interactions by fostering a more inclusive and culturally sensitive understanding of emotional expressions. By acknowledging and responding to emotions expressed in various languages, this study strives to enhance the effectiveness and adaptability of computer systems, ultimately contributing to a more nuanced and responsive interaction between humans and machines.

5.3 Interpretable sentiment understanding

Chinese-centric sentiment analysis involves the identification, examination, quantification, and retrieval of implicit emotions and subject-related information. Its impact spans a wide range of domains, including assessing the mental health of individuals and detecting fraud in the financial sector (Jawale and Sawarkar 2020). As the volume of social media data continues to surge, there is an increasing demand for automated sentiment analysis. Deep learning, although offering high accuracy, often operates with an opaque decision-making strategy. To bolster decision-making integrity, trust, belief, fairness, reliability, and impartiality become paramount. It is essential to move beyond mere accuracy and address the interpretability of the models. Developing models that not only deliver accurate results but also offer clear explanations for their decisions is crucial for enhancing transparency and accountability in sentiment analysis. This approach ensures that the decision-making process is not only accurate but also comprehensible, fostering trust and confidence in the results generated by sentiment analysis models.

5.4 Fine-tuning and domain adaptation

Analyzing sentiment in Chinese-centric content is essential for extracting user sentiments related to various events or topics, be it in tweets or on Weibo (Zhao et al. 2015; Wan 2008). This enables a more nuanced understanding of user sentiments, providing a complementary perspective to sentiment analysis in other languages. While there has been considerable advancement in sentiment analysis technology, there is a noticeable gap in research focusing on Chinese-centric sentiments. To address this gap, future work could involve pretraining models on large-scale datasets specifically for object recognition, laying the groundwork for effective transfer learning. Customizing models for distinct domains and languages through fine-tuning and domain adaptation emerges as a crucial strategy to enhance performance (Zhai et al. 2024; Lossio-Ventura et al. 2024). This approach ensures that sentiment analysis models are attuned to the intricacies of the Chinese language and its unique cultural contexts, ultimately leading to more accurate and contextually relevant sentiment predictions. The exploration of such tailored approaches will contribute significantly to advancing the field of sentiment analysis in the Chinese language.

5.5 Multimodal understanding

In addition to Chinese-centric textual content, other modalities, such as images and videos, represent straightforward media through which individuals express emotions on social networking sites (Yang et al. 2024; Miah et al. 2024). Social media users are progressively turning to images and videos to articulate opinions and share experiences. Conducting sentiment analysis on this extensive visual content can significantly improve the extraction of user sentiments related to events or topics, creating a complementary aspect to textual sentiment analysis. While substantial advancements have been achieved in this technology, there remains a dearth of research focusing on multi-modal Chinese-centric sentiments. To address this gap, future research could delve into incorporating visual and contextual cues into sentiment analysis models. Recognizing the prevalence of visual content in expressing emotions, integrating these cues can enhance the models' overall understanding of senti-

ment, leading to more nuanced and accurate predictions. Exploring the multi-modal aspects of sentiment analysis in the Chinese context is imperative to fully grasp the diverse ways in which users convey emotions across different mediums, ultimately advancing the capabilities of sentiment analysis technology in the realm of visual content.

6 Conclusion

In conclusion, this comprehensive survey on the progression of sentiment analysis in Chinese text unveils essential insights into the thematic trends, methodologies, and emerging patterns within this domain. The utilization of a unique framework combining keyword co-occurrence analysis and sophisticated community detection algorithms has illuminated the landscape of Chinese sentiment analysis research.

Throughout the past two decades, this study has traced the dynamic interplay between research methodologies and evolving topics, revealing correlations and shedding light on significant hotspots and trends in Chinese language text analysis. The comparative analysis presented here not only highlights correlations but also signifies evolving patterns, providing invaluable insights into the intricate terrain of sentiment analysis within the Chinese language context.

Our investigation of multilingual sentiment analysis highlights the amplification of challenges in sentiment interpretation due to language diversity. The Chinese language, with its intricate characters and cultural nuances, requires specialized techniques for accurate sentiment assessment. ChatGPT has emerged as a transformative technology, showcasing its adaptability to languages, including Chinese, and its ability to conduct sentiment analysis in context.

Moreover, beyond its academic contribution, this paper serves as a practical guide, offering insights into sentiment analysis methodologies and thematic trends. It lays the groundwork for future explorations, delineates technical limitations, and outlines promising directions for the advancement of sentiment analysis in Chinese text.

The roadmap provided by this study is poised to aid researchers, practitioners, and stakeholders navigating the complexities of sentiment analysis in the Chinese language. The identified trends and methodologies serve as a cornerstone for future investigations and advancements in this burgeoning field, fostering continued progress and innovation in sentiment analysis within the context of Chinese text.

Author contribution Prof. Zhaoxia Wang, Dr. Seng-Beng Ho, and Prof. Erik Cambria prepared the main manuscript text. Visiting PhD student Xinyue Zhang prepared Table 3, while Visiting PhD student Jingfeng Cui prepared Figs. 1–5. Both visiting PhD students also helped enhance the explanations of the tables and figures. Jingfeng Cui also handled the data collection and visualization. Current doctoral student Donghao Huang contributed to the revision of the manuscript, validated all analyses and visualizations and added new sections to the revised manuscript. All co-authors contributed to the manuscript.

Data availability The data that support the findings of this study are available from the corresponding author upon reasonable request.

Declarations

Conflict of interest The authors declare that they have no conflict of interest or competing interest in this article.

Ethical approval This article does not contain any studies with human participants or animals performed by any of the authors.

References

Abdullah NAS, Rusli NIA (2021) Multilingual sentiment analysis: a systematic literature review. Pertanika J Sci Technol 29(1):445–470. https://doi.org/10.47836/pjst.29.1.25

Ahmed M, Chen Q, Li Z (2020) Constructing domain-dependent sentiment dictionary for sentiment analysis. Neural Comput Appl 32(18):14719–14732. https://doi.org/10.1007/s00521-020-04824-8

Álvarez-Carmona M et al (2022) Natural language processing applied to tourism research: a systematic review and future research directions. J King Saud Univ Comput Inf Sci 34(10):10125–10144. https://doi.org/10.1016/j.jksuci.2022.10.010

Amin MM, Cambria E, Schuller BW (2023) Will affective computing emerge from foundation models and general artificial intelligence? A first evaluation of ChatGPT. IEEE Intell Syst 38(2):15–23. https://doi.org/10.1109/MIS.2023.3254179

Araújo M, Pereira A, Benevenuto F (2020) A comparative study of machine translation for multilingual sentence-level sentiment analysis. Inf Sci (NY) 512:1078–1102. https://doi.org/10.1016/j.ins.2019.10.031

Bai H, Yu G (2016) A Weibo-based approach to disaster informatics: incidents monitor in post-disaster situation via Weibo text negative sentiment analysis. Nat Hazards 83(2):1177–1196. https://doi.org/10.1007/s11069-016-2370-5

Bhaskaran R et al (2022) Intelligent machine learning with metaheuristics based sentiment analysis and classification. Comput Syst Sci Eng 44(1):235–247. https://doi.org/10.32604/csse.2023.024399

Blondel VD, Guillaume JL, Lambiotte R, Lefebvre E (2008) Fast unfolding of communities in large networks. J Stat Mech Theory Exp 2008(10):P10008. https://doi.org/10.1088/1742-5468/2008/10/P10008

Bruna J, Zaremba W, Szlam A, LeCun Y (2014) Spectral networks and deep locally connected networks on graphs. In: 2nd Int. Conf. Learn. Represent. ICLR 2014—Conf. Track Proc., pp 1–14

Brunet-Gouet E, Vidal N, Roux P (2023) Can a conversational agent pass theory-of-mind tasks? A case study of ChatGPT with the hinting, false beliefs, and strange stories paradigms. In: International conference on human and artificial rationalities. Springer, Cham

Bu J et al (2021) ASAP: A chinese review dataset towards aspect category sentiment analysis and rating prediction. In: NAACL-HLT 2021—2021 Conf. North Am. Chapter Assoc. Comput. Linguist. Hum. Lang. Technol. Proc. Conf., pp 2069–2079. https://doi.org/10.18653/v1/2021.naacl-main.167

Buscemi A, Proverbio D (2024) ChatGPT vs Gemini vs LLaMA on multilingual sentiment analysis, pp 1–11. http://arxiv.org/abs/2402.01715

Cambria E, Schuller B, Liu B, Wang H, Havasi C (2013) Statistical approaches to concept-level sentiment analysis. IEEE Intell Syst 28(3):6–9. https://doi.org/10.1109/MIS.2013.68

Cambria E, Mao R, Chen M, Wang Z, Ho S-B (2023) Seven pillars for the future of artificial intelligence. IEEE Intell Syst 38(6):62–69

Cambria E, Zhang X, Mao R, Chen M, Kwok K (2024) SenticNet 8: fusing emotion AI and commonsense AI for interpretable, trustworthy, and explainable affective computing. In: International conference on human–computer interaction (HCII), Washington DC, USA

Chen C, Teng Z, Wang Z, Zhang Y (2022) Discrete opinion tree induction for aspect-based sentiment analysis. Proc Annu Meet Assoc Comput Linguist 1:2051–2064. https://doi.org/10.18653/v1/2022.acl-long.145

Cheng YY, Chen YM, Yeh WC, Chang YC (2021) Valence and arousal-infused bi-directional lstm for sentiment analysis of government social media management. Appl Sci 11(2):1–14. https://doi.org/10.3390/app11020880

Chin HC, Wang Z (2016) Understanding commuter sentiments from Tweets. In: Proceedings of international conference on urban planning, transport and construction engineering, pp 7–13

Cui J, Wang Z, Ho S-B, Cambria E (2023) Survey on sentiment analysis: evolution of research methods and topics. Artif Intell Rev. https://doi.org/10.1007/s10462-022-10386-z

Day MY, Lin YD (2017) Deep learning for sentiment analysis on google play consumer review. In: Proc.—2017 IEEE Int. Conf. Inf. Reuse Integr. IRI 2017, vol. 2017-Janua, pp 382–388. https://doi.org/10.1109/IRI.2017.79

Deng YC, Wang YR, Chen SH, Lee LH (2023) Toward transformer fusions for chinese sentiment intensity prediction in valence-arousal dimensions. IEEE Access 11:109974–109982. https://doi.org/10.1109/ACCESS.2023.3322436

Dong J, Luo X, Zhu J (2024) A comparative study of different pre-trained language models for sentiment analysis of human–computer negotiation dialogue. In: Knowledge science, engineering and management, pp 301–317

Du W, Tan S (2010) Optimizing modularity to identify semantic orientation of Chinese words. Expert Syst Appl 37(7):5094–5100. https://doi.org/10.1016/j.eswa.2009.12.088

Du Z, Huang AG, Wermers R, Wu W (2022) Language and domain specificity: A Chinese financial sentiment dictionary. Rev Financ 26(3):673–719. https://doi.org/10.1093/rof/rfab036

Du S, Zhang X (2017) Aicyber's system for IALP 2016 shared task: character-enhanced word vectors and boosted neural networks. In: Proc. 2016 Int. Conf. Asian Lang. Process. IALP 2016, vol 39, pp 161–163. https://doi.org/10.1109/IALP.2016.7875958

Fang Y, Tan H, Zhang J (2018) Multi-strategy sentiment analysis of consumer reviews based on semantic fuzziness. IEEE Access 6:20625–20631. https://doi.org/10.1109/ACCESS.2018.2820025

Feng Y, Cheng Y (2021) Short text sentiment analysis based on multi-channel CNN with multi-head attention mechanism. IEEE Access 9:19854–19863. https://doi.org/10.1109/ACCESS.2021.3054521

Fu X, Liu W, Xu Y, Cui L (2017) Combine HowNet lexicon to train phrase recursive autoencoder for sentence-level sentiment analysis. Neurocomputing 241:18–27. https://doi.org/10.1016/j.neucom.2017.01.079

Fu X, Yang J, Li J, Fang M, Wang H (2018) Lexicon-enhanced LSTM with attention for general sentiment analysis. IEEE Access 6:71884–71891. https://doi.org/10.1109/ACCESS.2018.2878425

Fu Z, Hsu YC, Chan CS, Lau CM, Liu J, Yip PSF (2024) Efficacy of ChatGPT in cantonese sentiment analysis: comparative study. J Med Internet Res 26(1):1–14. https://doi.org/10.2196/51069

Gan C, Feng Q, Zhang Z (2021) Scalable multi-channel dilated CNN–BiLSTM model with attention mechanism for Chinese textual sentiment analysis. Futur Gener Comput Syst 118:297–309. https://doi.org/10.1016/j.future.2021.01.024

Gao H, Guo D, Wu J, Zhao Q, Li L (2021) Changes of the public attitudes of china to domestic COVID-19 Vaccination after the vaccines were approved: a semantic network and sentiment analysis based on sina weibo texts. Front Public Heal 9:1–11. https://doi.org/10.3389/fpubh.2021.723015

Gravano A (2010) Turn-taking and affirmative cue words in task-oriented dialogue. Diss Abstr Int B Sci Eng 70(8):4943. https://doi.org/10.1162/COLI

Gu M, Guo H, Zhuang J, Du Y, Qian L (2022) social media user behavior and emotions during crisis events. Int J Environ Res Public Health. https://doi.org/10.3390/ijerph19095197

Guo Y et al (2022) Acceptability of traditional chinese medicine in Chinese people based on 10-year's real world study with mutiple big data mining. Front Public Health. https://doi.org/10.3389/fpubh.2021.811730

Hassani H, Silva ES (2023) The role of ChatGPT in data science: how AI-assisted conversational interfaces are revolutionizing the field. Big Data Cogn Comput. https://doi.org/10.3390/bdcc7020062

He Y (2011) Latent sentiment model for weakly-supervised cross-lingual sentiment classification. In: Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics). LNCS, vol 6611, pp 214–225. https://doi.org/10.1007/978-3-642-20161-5_22

He J, Wumaier A, Kadeer Z, Sun W, Xin X, Zheng L (2022) A local and global context focus multilingual learning model for aspect-based sentiment analysis. IEEE Access 10:84135–84146. https://doi.org/10.1109/ACCESS.2022.3197218

Heracleous P, Yoneyama A (2019) A comprehensive study on bilingual and multilingual speech emotion recognition using a two-pass classification scheme. PLoS ONE 14(8):1–20. https://doi.org/10.1371/journal.pone.0220386

Hercigt T, Brychcín T, Svobodat L, Konkolt M (2016) UWB at SemEval-2016 task 5: aspect based sentiment analysis. In: SemEval 2016—10th Int. Work. Semant. Eval. Proc., pp 342–349. https://doi.org/10.186 53/v1/s16-1055

Hu Y, Li W (2011) Document sentiment classification by exploring description model of topical terms. Comput Speech Lang 25(2):386–403. https://doi.org/10.1016/j.csl.2010.07.004

Hu Z, Wang Z, Wang Y, Tan AH (2023) MSRL-Net: a multi-level semantic relation-enhanced learning network for aspect-based sentiment analysis. Expert Syst Appl 217:119492. https://doi.org/10.1016/j.esw a.2022.119492

Huai S, Van de Voorde T (2022) Which environmental features contribute to positive and negative perceptions of urban parks? A cross-cultural comparison using online reviews and natural language processing methods. Landsc Urban Plan 218:104307. https://doi.org/10.1016/j.landurbplan.2021.104307

Huang W, Lin M, Wang Y (2022) Sentiment Analysis of Chinese E-commerce product reviews using ERNIE word embedding and attention mechanism. Appl Sci. https://doi.org/10.3390/app12147182

Ihnaini B, Sun W, Cai Y, Xu Z, Sangi R (2024) Sentiment analysis of song dynasty classical poetry using fine-tuned large language models: a study with LLMs. In: 2024 7th International conference on artificial intelligence and big data, ICAIBD 2024, pp 590–597. https://doi.org/10.1109/ICAIBD62003.2024.10 604440

Janusch S (2011) Reality, dysconsciousness, and transformations: personal reflections on the ethics of cross-cultural research. TESL Canada J 28:80. https://doi.org/10.18806/tesl.v28i0.1083

Jawale S, Sawarkar SD (2020), Interpretable sentiment analysis based on deep learning: an overview. In: 2020 IEEE pune section international conference (PuneCon), pp 65–70. https://doi.org/10.1109/Pune Con50868.2020.9362361

Jia K (2022) Sentiment classification of microblog: a framework based on BERT and CNN with attention mechanism. Comput Electr Eng 101:108032. https://doi.org/10.1016/j.compeleceng.2022.108032

Jiang F, Cui A, Liu Y, Zhang M, Ma S (2013) Every term has sentiment: learning from emoticon evidences for Chinese microblog sentiment analysis. Commun Comput Inf Sci 400(61073071):224–235. https://d oi.org/10.1007/978-3-642-41644-6_21

Lee S (2019) The generalization effect for multilingual speech emotion recognition across heterogeneous languages. In: 2019 IEEE International conference on acoustics, speech and signal processing (ICASSP), pp 5881–5885. https://doi.org/10.1109/ICASSP.2019.8683046

Lee LH, Li JH, Yu LC (2022) Chinese EmoBank: building valence-arousal resources for dimensional sentiment analysis. ACM Trans Asian Low-Resource Lang Inf Process 21(4):1–18. https://doi.org/10.1145 /3489141

Leydesdorff L, Park HW, Wagner C (2014) International co-authorship relations in the social science citation index: Is internationalization leading the network? J Assoc Inf Sci Technol 65(10):2111–2126. https://d oi.org/10.48550/arXiv.1305.4242

Li J, Sun M (2007) Experimental study on sentiment classification of Chinese review using machine learning techniques. In: IEEE NLP-KE 2007—Proc. Int. Conf. Nat. Lang. Process. Knowl. Eng., pp 393–400. https://doi.org/10.1109/NLPKE.2007.4368061

Li C, Xu B, Wu G, He S, Tian G, Hao H (2014) Recursive deep learning for sentiment analysis over social data. In: Proc.—2014 IEEE/WIC/ACM Int. Jt. Conf. Web Intell. Intell. Agent Technol.—Work. WI-IAT 2014, vol 2, pp 180–185. https://doi.org/10.1109/WI-IAT.2014.96

Li Q, Jin Z, Wang C, Zeng DD (2016) Mining opinion summarizations using convolutional neural networks in Chinese microblogging systems. Knowl Based Syst 107:289–300. https://doi.org/10.1016/j.knosys .2016.06.017

Li X, Lei Y, Ji S (2022) BERT- and BiLSTM-based sentiment analysis of online Chinese buzzwords. Futur Internet. https://doi.org/10.3390/fi14110332

Li B et al (2023a) DiaASQ: a benchmark of conversational aspect-based sentiment quadruple analysis. In: Proc. Annu. Meet. Assoc. Comput. Linguist., pp 13449–13467. https://doi.org/10.18653/v1/2023.find ings-acl.849

Li L, Zhang H, Li C, You H, Cui W (2023b) Evaluation on ChatGPT for Chinese language understanding. Data Intell 5(4):1–19. https://doi.org/10.1162/dint_a_00232

Liao J, Wang M, Chen X, Wang S, Zhang K (2022) Dynamic commonsense knowledge fused method for Chinese implicit sentiment analysis. Inf Process Manag 59(3):102934. https://doi.org/10.1016/j.ipm.2 022.102934

Ligthart A, Catal C, Tekinerdogan B (2021) Systematic reviews in sentiment analysis: a tertiary study. Artif Intell Rev 54(7):4997–5053. https://doi.org/10.1007/s10462-021-09973-3

Lipenkova J (2015) A system for fine-grained aspect-based sentiment analysis of Chinese. In: ACL-IJCNLP 2015—53rd Annu. Meet. Assoc. Comput. Linguist. 7th Int. Jt. Conf. Nat. Lang. Process. Proc. Syst. Demonstr., pp 55–60. https://doi.org/10.3115/v1/p15-4010

Liu N, Shen B (2020) Aspect-based sentiment analysis with gated alternate neural network. Knowl Based Syst 188:105010. https://doi.org/10.1016/j.knosys.2019.105010

Liu X, Zeng D, Li J, Wang F-Y, Zuo W (2009) Sentiment analysis of Chinese documents: from sentence to document level. J Am Soc Inf Sci Technol 60(12):2474–2487. https://doi.org/10.1002/asi.21206

Liu L, Nie X, Wang H (2012) Toward a fuzzy domain sentiment ontology tree for sentiment analysis. In: Proceedings of the 5th international congress on image and signal processing, pp 1620–1624. https://doi.org/10.1109/CISP.2012.6469930

Liu L, Song W, Wang H, Li C, Lu J (2014) A novel feature-based method for sentiment analysis of Chinese product reviews. China Commun 11(3):154–164. https://doi.org/10.1109/CC.2014.6825268

Liu X, Tang T, Ding N (2022) Social network sentiment classification method combined Chinese text syntax with graph convolutional neural network. Egypt Inform J 23(1):1–12. https://doi.org/10.1016/j.eij.2021.04.003

Lo SL, Cambria E, Chiong R, Cornforth D (2017) Multilingual sentiment analysis: from formal to informal and scarce resource languages. Artif Intell Rev 48(4):499–527. https://doi.org/10.1007/s10462-016-9508-4

Long F, Zhou K, Ou W (2019) Sentiment analysis of text based on bidirectional LSTM with multi-head attention. IEEE Access 7:141960–141969. https://doi.org/10.1109/ACCESS.2019.2942614

Lossio-Ventura JA et al (2024) A comparison of ChatGPT and fine-tuned open pre-trained transformers (OPT) against widely used sentiment analysis tools: sentiment analysis of COVID-19 survey data. JMIR Ment Health 11(1):1–17. https://doi.org/10.2196/50150

Lu Y, Kong X, Quan X, Liu W, Xu Y (2010) Exploring the sentiment strength of user reviews. In: Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics). LNCS, vol 6184, pp 471–482. https://doi.org/10.1007/978-3-642-14246-8_46

Ma Z et al (2024) EmoBox: multilingual multi-corpus speech emotion recognition toolkit and benchmark. In: Procceedings of INTERSPEECH, pp 2–6. http://arxiv.org/abs/2406.07162

Medhat W, Hassan A, Korashy H (2014) Sentiment analysis algorithms and applications: a survey. Ain Shams Eng J 5(4):1093–1113. https://doi.org/10.1016/j.asej.2014.04.011

Miah MSU, Kabir MM, Bin Sarwar T, Safran M, Alfarhood S, Mridha MF (2024) A multimodal approach to cross-lingual sentiment analysis with ensemble of transformer and LLM. Sci Rep 14(1):1–18. https://doi.org/10.1038/s41598-024-60210-7

Mihalcea R, Banea C, Wiebe J (2012) Multilingual subjectivity and sentiment analysis. In: Proceedings of the 50th annual meeting of the association for computational linguistics: tutorial abstracts, p 4

Nair RR, Mathew J, Muraleedharan V, Deepa Kanmani S (2019) Study of machine learning techniques for sentiment analysis. In: 2019 3rd International conference on computing methodologies and communication, pp 978–984. https://doi.org/10.1109/ICCMC.2019.8819763

Obiedat R, Al-Darras D, Alzaghoul E, Harfoushi O (2021) Arabic aspect-based sentiment analysis: A systematic literature review. IEEE Access 9:152628–152645. https://doi.org/10.1109/ACCESS.2021.3127140

Ortony A, Turner TJ (1990) What's basic about basic emotions? Psychol Rev 97(3):315–331. https://doi.org/10.1037/0033-295X.97.3.315

Pan J, Xue GR, Yu Y, Wang Y (2011) Cross-lingual sentiment classification via bi-view non-negative matrix tri-factorization. In: Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics), vol 6634 LNAI, no. PART 1, pp 289–300. https://doi.org/10.1007/978-3-642-20841-6_24

Pan D, Yuan J, Li L, Sheng D (2018) Deep neural network-based classification model for sentiment analysis. In: BESC 2019—6th Int. Conf. Behav. Econ. Socio-Cultural Comput. Proc. https://doi.org/10.1109/BESC48373.2019.8963171

Peng H, Cambria E, Hussain A (2017) A review of sentiment analysis research in chinese language. Cognit Comput 9(4):423–435. https://doi.org/10.1007/s12559-017-9470-8

Peng H, Ma Y, Li Y, Cambria E (2018) Learning multi-grained aspect target sequence for Chinese sentiment analysis. Knowl Based Syst 148:167–176. https://doi.org/10.1016/j.knosys.2018.02.034

Peng H, Ma Y, Poria S, Li Y, Cambria E (2021) Phonetic-enriched text representation for Chinese sentiment analysis with reinforcement learning. Inf Fusion 70:88–99. https://doi.org/10.1016/j.inffus.2021.01.005

Polgan JM et al (2023) Understanding sentiment and emotion through ChatGPT to support emotion-based management decision making. J Minfo Polgan 12:1778–1788. https://doi.org/10.33395/jmp.v12i2.13000

Poria S, Majumder N, Hazarika D, Cambria E, Gelbukh A, Hussain A (2018) Multimodal sentiment analysis: addressing key issues and setting up the baselines. IEEE Intell Syst 33(6):17–25. https://doi.org/10.1109/MIS.2018.2882362

Prabha MI, Srikanth GU (2019) Survey of sentiment analysis using deep learning techniques. In: 2019 1st International conference on innovations in information and communication technology, pp 1–9

Qiang YE, Lin B, Yi-Jun LI (2005) Sentiment classification for chinese reviews: A comparison between SVM and semantic approaches. In: 2005 Int. Conf. Mach. Learn. Cybern. ICMLC 2005, pp 2341–2346. https://doi.org/10.1109/icmlc.2005.1527335

Raghuvanshi N, Patil JM (2016) A brief review on sentiment analysis. In: 2016 International conference on electrical, electronics, and optimization techniques (ICEEOT), pp 2827–2831. https://doi.org/10.1109/ICEEOT.2016.7755213

Ren C, Lee S-J, Hu C (2023) Assessing the efficacy of ChatGPT in addressing Chinese financial conundrums: an in-depth comparative analysis of human and AI-generated responses. Comput Hum Behav Artif Humans 1(2):100007. https://doi.org/10.1016/j.chbah.2023.100007

Rotta R, Noack A (2011) Multilevel local search algorithms for modularity clustering. ACM J Exp Algorithmics 16(2):1–27. https://doi.org/10.1145/1963190.1970376

Schembri N, Jahić Jašić A (2022) Ethical issues in multilingual research situations: a focus on interview-based research. Res Ethics 18(3):210–225. https://doi.org/10.1177/17470161221085857

Shao C, Li H, Liu X, Li C, Yang L, Zhang Y, Luo J (2023) Appropriateness and comprehensiveness of using ChatGPT for perioperative patient education in thoracic surgery in different language contexts: survey study, p e46900. https://doi.org/10.2196/46900

Shi W, Wang H, He S (2013) Sentiment analysis of Chinese microblogging based on sentiment ontology: a case study of '7.23 Wenzhou Train Collision.' Conn Sci 25(4):161–178. https://doi.org/10.1080/09540091.2013.851172

Stappen L, Baird A, Cambria E, Schuller BW (2021) Sentiment analysis and topic recognition in video transcriptions. IEEE Intell Syst 36(2):88–95

Su Z, Xu H, Zhang D, Xu Y (2014) Chinese sentiment classification using a neural network tool—Word2vec. In: Proc. 2014 Int. Conf. Multisens. Fusion Inf. Integr. Intell. Syst. MFI 2014, pp 1–6. https://doi.org/10.1109/MFI.2014.6997687

Sun X, Li C, Ren F (2016) Sentiment analysis for Chinese microblog based on deep neural networks with convolutional extension features. Neurocomputing 210:227–236. https://doi.org/10.1016/j.neucom.2016.02.077

Sun A, Wei F, Wang G, Li Y (2022) Chinese sentiment analysis using regularized extreme learning machine and stochastic optimization. In: Proceedings of the 4th international conference on natural language processing (ICNLP), pp 525–529. https://doi.org/10.1109/ICNLP55136.2022.00096

Tai KS, Socher R, Manning CD (2015) Improved semantic representations from tree-structured long short-term memory networks. In: Proceedings of the 53rd annual meeting of the association for computational linguistics and the 7th international joint conference on natural language processing, pp 1556–1566. https://doi.org/10.3115/v1/p15-1150

Tan S, Zhang J (2008) An empirical study of sentiment analysis for chinese documents. Expert Syst Appl 34(4):2622–2629. https://doi.org/10.1016/j.eswa.2007.05.028

Tan YS, Teo N, Ghe E, Fong J, Wang Z (2023) Video sentiment analysis for child safety. In: 2023 IEEE 16th international conference on data mining workshops (ICDMW), pp 783–790. https://doi.org/10.1109/ICDMW60847.2023.00106

Tang YJ, Chen HH (2014) Chinese irony corpus construction and ironic structure analysis. In: COLING 2014—25th international conference on computational linguistics, proceedings of COLING 2014: technical papers, pp 1269–1278

Teo A, Wang Z, Pen H, Subagdja B, Ho S-B, Quek BK (2023) Knowledge graph enhanced aspect-based sentiment analysis incorporating external knowledge. In: 2023 IEEE 16th international conference on data mining workshops (ICDMW), pp 791–798. https://doi.org/10.1109/ICDMW60847.2023.00107

Tombari F (2021) The belt and road initiative. In: China and South Asia, pp 141–156. https://doi.org/10.4324/9780367855413-11

Van Nooten J, Daelemans W (2023) Improving Dutch vaccine hesitancy monitoring via multi-label data augmentation with GPT-3.5. In: Proceedings of the 13th workshop on computational approaches to subjectivity, sentiment, & social media analysis, pp 251–270. https://doi.org/10.18653/v1/2023.wassa-1.23

Wan X (2008) Using bilingual knowledge and ensemble techniques for unsupervised Chinese sentiment analysis. In: Proceedings of the 2008 conference on empirical methods in natural language processing, pp 553–561. https://doi.org/10.3115/1613715.1613783

Wan X (2011) Bilingual co-training for sentiment classification of Chinese product reviews. Comput Linguist 37(3):587–616. https://doi.org/10.1162/COLI_a_00061

Wang Z, Lin Z (2020) Optimal feature selection for learning-based algorithms for sentiment classification. Cognit Comput 12(1):238–248. https://doi.org/10.1007/s12559-019-09669-5

Wang Z, Parth Y (2016) Extreme learning machine for multi-class sentiment classification of tweets. In: Proc. ELM-2015, vol 1, pp 1–11. Springer, Cham. https://doi.org/10.1007/978-3-319-28397-5_1

Wang S, Wei Y, Zhang W, Li D, Li W (2007) A hybrid method of feature selection for chinese text sentiment classification. In: Proc.—Fourth Int. Conf. Fuzzy Syst. Knowl. Discov. FSKD 2007, vol 3, no. Fskd, pp 435–439. https://doi.org/10.1109/FSKD.2007.49

Wang Z, Tong JC, Chan D (2014a) Issues of social data analytics with a new method for sentiment analysis of social media data. In: 2014 IEEE 6th International conference on cloud computing technology and science, pp 899–904. https://doi.org/10.1109/CloudCom.2014.40

Wang Z, Tong JC, Chin HC (2014b) Enhancing machine-learning methods for sentiment classification of web data. In: Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics), vol 8870, pp 394–405. https://doi.org/10.1007/978-3-319-12844-3_34

Wang J, Yu LC, Lai KR, Zhang X (2016a) Community-based weighted graph model for valence-arousal prediction of affective words. IEEE/ACM Trans Audio Speech Lang Process 24(11):1957–1968. https://doi.org/10.1109/TASLP.2016.2594287

Wang Z, Tong JC, Ruan P, Li F (2016b) Lexicon knowledge extraction with sentiment polarity computation. In: IEEE international conference on data mining workshops, ICDMW, pp 978–983. https://doi.org/10.1109/ICDMW.2016.0142

Wang B, Huang Y, Yuan Z, Li X (2016c) A multi-granularity fuzzy computing model for sentiment classification of Chinese reviews. J Intell Fuzzy Syst 30(3):1445–1460. https://doi.org/10.3233/IFS-151853

Wang Y, Feng S, Wang D, Zhang Y, Yu G (2016d) Context-aware chinese microblog sentiment classification with bidirectional LSTM. In: Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics). LNCS, vol 9931, pp 594–606. https://doi.org/10.1007/978-3-319-45814-4_48

Wang W, Tan G, Wang H (2017) Cross-domain comparison of algorithm performance in extracting aspect-based opinions from Chinese online reviews. Int J Mach Learn Cybern 8(3):1053–1070. https://doi.org/10.1007/s13042-016-0596-x

Wang Z, Ho S-B, Cambria E (2020a) A review of emotion sensing: categorization models and algorithms. Multimed Tools Appl 79(47–48):35553–35582. https://doi.org/10.1007/s11042-019-08328-z

Wang Z, Ho S-B, Cambria E (2020b) Multi-level fine-scaled sentiment sensing with ambivalence handling. Int J Uncertain Fuzziness Knowl-Based Syst 28(4):683–697. https://doi.org/10.1142/S0218488520500294

Wang J, Yu LC, Lai KR, Zhang X (2020c) Tree-structured regional CNN-LSTM model for dimensional sentiment analysis. IEEE/ACM Trans Audio Speech Lang Process 28:581–591. https://doi.org/10.1109/TASLP.2019.2959251

Wang T, Lu K, Chow KP, Zhu Q (2020d) COVID-19 sensing: negative sentiment analysis on social media in china via BERT model. IEEE Access 8:138162–138169. https://doi.org/10.1109/ACCESS.2020.3012595

Wang Y, Liu C, Xie J, Yang S, Jia Y, Zan H (2022) Aspect-based sentiment analysis with dependency relation graph convolutional network. In: 2022 Int. Conf. Asian Lang. Process. IALP 2022, pp 63–68. https://doi.org/10.1109/IALP57159.2022.9961321

Wang Z, Hu Z, Ho S-B, Cambria E, Tan AH (2023a) MiMuSA—mimicking human language understanding for fine-grained multi-class sentiment analysis. Neural Comput Appl 35(21):15907–15921. https://doi.org/10.1007/s00521-023-08576-z

Wang Z, Hu Z, Li F, Ho S-B, Cambria E (2023b) Learning-based stock trending prediction by incorporating technical indicators and social media sentiment. Cognit Comput 15(3):1092–1102. https://doi.org/10.1007/s12559-023-10125-8

Wang X et al (2023c) ChatGPT performs on the chinese national medical licensing examination. J Med Syst 47(1):1–9. https://doi.org/10.1007/s10916-023-01961-0

Wen Z, Li T (2014) Emotional element extraction based on CRFs. Adv Intell Syst Comput. https://doi.org/10.1007/978-3-642-54927-4

Wu G, Wu X, Wei J (2012) Sentiment analysis of comparative sentences for Chinese document. Appl Mech Mater 157–158:1079–1082. https://doi.org/10.4028/www.scientific.net/AMM.157-158.1079

Wu F, Huang Y, Song Y, Liu S (2016) Towards building a high-quality microblog-specific Chinese sentiment lexicon. Decis Support Syst 87:39–49. https://doi.org/10.1016/j.dss.2016.04.007

Wu C, Wu F, Wu S, Yuan Z, Liu J, Huang Y (2019) Semi-supervised dimensional sentiment analysis with variational autoencoder. Knowledge Based Syst 165:30–39. https://doi.org/10.1016/j.knosys.2018.11.018

Wu P, Li X, Shen S, He D (2020) Social media opinion summarization using emotion cognition and convolutional neural networks. Int J Inf Manag 51:101978. https://doi.org/10.1016/j.ijinfomgt.2019.07.004

Wu O, Yang T, Li M, Li M (2022) Two-level LSTM for sentiment analysis with Lexicon embedding and polar flipping. IEEE Trans Cybern 52(5):3867–3879. https://doi.org/10.1109/TCYB.2020.3017378

Xiang R et al (2020) Ciron: a new benchmark dataset for chinese irony detection. In: *Lr. 2020*—12th Int. Conf. Lang. Resour. Eval. Conf. Proc., pp 5714–5720

Xianghua F, Guo L, Yanyan G, Zhiqiang W (2013) Multi-aspect sentiment analysis for Chinese online social reviews based on topic modeling and HowNet lexicon. Knowl-Based Syst 37:186–195. https://doi.org/10.1016/j.knosys.2012.08.003

Xie H, Lin W, Lin S, Wang J, Yu LC (2021) A multi-dimensional relation model for dimensional sentiment analysis. Inf Sci (Ny) 579:832–844. https://doi.org/10.1016/j.ins.2021.08.052

Xiong S, Wang K, Ji D, Wang B (2018) A short text sentiment-topic model for product reviews. Neurocomputing 297:94–102. https://doi.org/10.1016/j.neucom.2018.02.034

Xu H, Zhao K, Qiu L, Hu C (2010) Expanding Chinese sentiment dictionaries from large scale unlabeled corpus. In: PACLIC 24—Proc. 24th Pacific Asia Conf. Lang. Inf. Comput., pp 301–310

Xu R, Zhou Y, Wu D, Gui L, Du J, Xue Y (2016) Overview of NLPCC shared task 4: Stance detection in Chinese microblogs. In: *Natural language understanding and intelligent applications: 5th CCF conference on natural language processing and chinese computing, NLPCC 2016, and 24th international conference on computer processing of oriental languages, ICCPOL 2016*, pp 907–916. https://doi.org/10.1007/978-3-319-50496-4_85

Xu G, Yu Z, Yao H, Li F, Meng Y, Wu X (2019a) Chinese text sentiment analysis based on extended sentiment dictionary. IEEE Access 7(1):43749–43762. https://doi.org/10.1109/ACCESS.2019.2907772

Xu F, Zhang X, Xin Z, Yang A (2019b) Investigation on the Chinese text sentiment analysis based on convolutional neural networks in deep learning. Comput Mater Contin 58(3):697–709. https://doi.org/10.32604/cmc.2019.05375

Xu Y, Cao H, Du W, Wang W (2022) A survey of cross-lingual sentiment analysis: methodologies, models and evaluations. Data Sci Eng 7(3):279–299. https://doi.org/10.1007/s41019-022-00187-3

Yan G, He W, Shen J, Tang C (2014) A bilingual approach for conducting Chinese and English social media sentiment analysis. Comput Netw 75:491–503. https://doi.org/10.1016/j.comnet.2014.08.021

Yan D, Hu B, Qin J (2018) Sentiment analysis for microblog related to finance based on rules and classification. In: Proc.—2018 IEEE Int. Conf. Big Data Smart Comput. BigComp 2018, pp 119–126. https://doi.org/10.1109/BigComp.2018.00026

Yang DH, Yu G (2013) A method of feature selection and sentiment similarity for Chinese micro-blogs. J Inf Sci 39(4):429–441. https://doi.org/10.1177/0165551513480308

Yang A, Lin J, Zhou Y, Chen J (2013) Research on building a Chinese sentiment lexicon based on SO-PMI. Appl Mech Mater 263–266(PART 1):1688–1693. https://doi.org/10.4028/www.scientific.net/AMM.263-266.1688

Yang L, Li Y, Wang J, Sherratt RS (2020) Sentiment analysis for e-commerce product reviews in Chinese based on sentiment lexicon and deep learning. IEEE Access 8:23522–23530. https://doi.org/10.1109/ACCESS.2020.2969854

Yang H, Zeng B, Yang J, Song Y, Xu R (2021b) A multi-task learning model for Chinese-oriented aspect polarity classification and aspect term extraction. Neurocomputing 419:344–356. https://doi.org/10.1016/j.neucom.2020.08.001

Yang Q, Luo Z, Li M, Liu J (2022a) Understanding the landscape and propagation of COVID-19 misinformation and its correction on Sina Weibo. Glob Health Promot 29(1):44–52. https://doi.org/10.1177/17579759211035053

Yang Q, Kadeer Z, Gu W, Sun W, Wumaier A (2022b) Affective knowledge augmented interactive graph convolutional network for Chinese-oriented aspect-based sentiment analysis. IEEE Access 10:130686–130698. https://doi.org/10.1109/ACCESS.2022.3228299

Yang Y, Zhang Y, Zhang X, Cao Y, Zhang J (2022c) Spatial evolution patterns of public panic on Chinese social networks amidst the COVID-19 pandemic. Int J Disaster Risk Reduct 70:102762. https://doi.org/10.1016/j.ijdrr.2021.102762

Yang H, Si Z, Zhao Y, Liu J, Wu Y, Qin B (2024) MACSA: A multimodal aspect-category sentiment analysis dataset with multimodal fine-grained aligned annotations, vol 1, no. 1. Association for Computing Machinery. https://doi.org/10.1007/s11042-024-18796-7

Yanmei L, Yuda C (2015) Research on Chinese micro-blog sentiment analysis based on deep learning. In: Proc.—2015 8th Int. Symp. Comput. Intell. Des. Isc. 2015, vol 1, pp 358–361. https://doi.org/10.1109/ISCID.2015.217

Yao J, Wang H, Yin P (2011) Sentiment feature identification from Chinese online reviews. In: Commun. Comput. Inf. Sci., CCIS, vol 201, no. PART 1, pp 315–322. https://doi.org/10.1007/978-3-642-22418-8_44

Zhai Z, Xu H, Li J, Jia P (2009) Sentiment classification for Chinese reviews based on key substring features. In: 2009 Int. Conf. Nat. Lang. Process. Knowl. Eng. NLP-KE 2009, pp 1–8. https://doi.org/10.1109/NLPKE.2009.5313782

Zhai W et al (2024) Chinese MentalBERT: domain-adaptive pre-training on social media for Chinese mental health text analysis. arXiv Prepr. https://app.dimensions.ai/details/publication/pub.1168870420

Zhang Z, Lan M (2016) Learning sentiment-inherent word embedding for word-level and sentence-level sentiment analysis. In: Proc. 2015 Int. Conf. Asian Lang. Process. IALP 2015, no. 1, pp 94–97. https://doi.org/10.1109/IALP.2015.7451540

Zhang X, Zheng X (2017) Comparison of text sentiment analysis based on machine learning. In: Proc.—15th Int. Symp. Parallel Distrib. Comput. ISPDC 2016, pp 230–233. https://doi.org/10.1109/ISPDC.2016.39

Zhang B, Zhou W (2023) Transformer-encoder-GRU (T-E-GRU) for Chinese sentiment analysis on chinese comment text. Neural Process Lett 55(2):1847–1867. https://doi.org/10.1007/s11063-022-10966-8

Zhang ZQ, Li YJ, Ye Q, Law R (2008a) Sentiment classification for chinese product reviews using an unsupervised internet-based method. In: 2008 Int. Conf. Manag. Sci. Eng. 15th Annu. Conf. Proceedings, ICMSE, vol 70771032, pp 3–9. https://doi.org/10.1109/ICMSE.2008.4668885

Zhang C, Zuo W, Peng T, He F (2008b) Sentiment classification for Chinese reviews using machine learning methods based on string kernel. In: Proc.—3rd Int. Conf. Converg. Hybrid Inf. Technol. ICCIT 2008, vol 2, pp 909–914. https://doi.org/10.1109/ICCIT.2008.51

Zhang Z, Ye Q, Zhang Z, Li Y (2011) Sentiment classification of Internet restaurant reviews written in Cantonese. Expert Syst Appl 38(6):7674–7682. https://doi.org/10.1016/j.eswa.2010.12.147

Zhang W, Xu H, Wan W (2012) Weakness finder: Find product weakness from Chinese reviews by using aspects based sentiment analysis. Expert Syst Appl 39(11):10283–10291. https://doi.org/10.1016/j.eswa.2012.02.166

Zhang Y, Zhang H, Zhang M, Liu Y, Ma S (2014) Do users rate or review?: boost phrase-level sentiment labeling with review-level sentiment classification. In: Proceedings of the 37th international ACM SIGIR conference on research & development in information retrieval, pp 1027–1030. https://doi.org/10.1145/2600428.2609501

Zhang D, Xu H, Su Z, Xu Y (2015) Chinese comments sentiment classification based on word2vec and SVM-perf. Expert Syst Appl 42(4):1857–1863. https://doi.org/10.1016/j.eswa.2014.09.011

Zhang S, Wei Z, Wang Y, Liao T (2018) Sentiment analysis of Chinese micro-blog text based on extended sentiment dictionary. Futur Gener Comput Syst 81:395–403. https://doi.org/10.1016/j.future.2017.09.048

Zhang W, Li L, Zhu Y, Yu P, Wen J (2022) CNN-LSTM neural network model for fine-grained negative emotion computing in emergencies. Alex Eng J 61(9):6755–6767. https://doi.org/10.1016/j.aej.2021.12.022

Zhang X, Yang Q, Xu D (2023a) XuanYuan 2.0: a large Chinese financial chat model with hundreds of billions parameters. In: Proceedings of the 32nd ACM international conference on information and knowledge management, pp 4435–4439. https://doi.org/10.1145/3583780.3615285

Zhang H, Hua F, Xu C, Guo J, Kong H, Zuo R (2023b) Unveiling the potential of sentiment: can large language models predict Chinese stock price movements? http://arxiv.org/abs/2306.14222

Zhang X, Mao R, Cambria E (2024) Multilingual emotion recognition: discovering the variations of lexical semantics between languages. In: 2024 Int. Jt. Conf. Neural Networks (IJCNN)

Zhao Y, Qin B, Liu T (2015) Creating a fine-grained corpus for Chinese sentiment analysis. IEEE Intell Syst 30(1):36–43. https://doi.org/10.1109/MIS.2014.33

Zhao M, Yang J, Zhang J, Wang S (2022) Aggregated graph convolutional networks for aspect-based sentiment classification. Inf Sci (NY) 600:73–93. https://doi.org/10.1016/j.ins.2022.03.082

Zhao C, Li Y, Caragea C (2023) C-STANCE: A large dataset for chinese zero-shot stance detection. Proc Annu Meet Assoc Comput Linguist 1:13369–13385. https://doi.org/10.18653/v1/2023.acl-long.747

Zheng L, Wang H, Gao S (2018) Sentimental feature selection for sentiment analysis of chinese online reviews. Int J Mach Learn Cybern 9(1):75–84. https://doi.org/10.1007/s13042-015-0347-4

Zheng W, Ye Q (2009) Sentiment classification of Chinese traveler reviews by support vector machine algorithm. In; 3rd Int. Symp. Intell. Inf. Technol. Appl. IITA 2009, vol 3, pp 335–338. https://doi.org/10.1109/IITA.2009.457

Zhou J, Lu Y, Dai HN, Wang H, Xiao H (2019) Sentiment analysis of Chinese microblog based on stacked bidirectional LSTM. IEEE Access 7:38856–38866. https://doi.org/10.1109/ACCESS.2019.2905048

Zhou F, Zhang J, Song Y (2022) Chinese Fine-grained sentiment classification based on pre-trained language model and attention mechanism. In: LNCS, vol 13202. Springer, Cham. https://doi.org/10.1007/978-3-030-97774-0_4

Zhu Z, Ying Y, Zhu J, Wu H (2023) ChatGPT's potential role in non-English-speaking outpatient clinic settings. Digit Heal 9:1–3. https://doi.org/10.1177/20552076231184091

Zou L, Shen JH, Zhang J, Lee CC (2022) What is the rationale behind China's infrastructure investment under the Belt and Road Initiative. J Econ Surv 36(3):605–633. https://doi.org/10.1111/joes.12427

Authors and Affiliations

Zhaoxia Wang[1] · **Donghao Huang**[1,2] · **Jingfeng Cui**[3] · **Xinyue Zhang**[1,4] · **Seng-Beng Ho**[5] · **Erik Cambria**[6]

✉ Zhaoxia Wang
zxwang@smu.edu.sg

Donghao Huang
dh.huang.2023@engd.smu.edu.sg

Jingfeng Cui
cuijf2019@163.com

Xinyue Zhang
xyzhang@stu.ecnu.edu.cn

Seng-Beng Ho
sengbeng.ho@aiiglobal.ai

Erik Cambria
cambria@ntu.edu.sg

[1]   School of Computing and Information Systems, Singapore Management University, 80 Stamford Rd, Singapore 178902, Singapore

[2]   Research and Development, Mastercard, 3 Fraser Street, Singapore 189352, Singapore

[3]   Fujian Normal University, 1 Keji Rd, Fuzhou 350117, China

[4]   Shanghai Institute of AI for Education, East China Normal University, 3663 Zhongshan Rd, Shanghai 200062, China

[5]   AI Institute Global, 79 Ayer Rajah Crescent, Singapore 139955, Singapore

[6]   College of Computing and Data Science, Nanyang Technological University, 50 Nanyang Avenue, Singapore 639798, Singapore