

Article

Multi-Chain of Thought Prompt Learning for Aspect-Based Sentiment Analysis

Yating He ¹, Zhenzhen He ^{1,*}, Tiquan Gu ², Bowen Gu ², Yaling Wan ¹ and Min Li ³¹ School of Computer Science and Technology, Xinjiang Normal University, Urumqi 830054, China² School of Computer Science and Technology, Xinjiang University, Urumqi 830046, China³ School of Information Management, Xinjiang University of Finance & Economics Copyright, Urumqi 830012, China

* Correspondence: hezhzhenzhen@xjnu.edu.cn

Abstract

Due to their extensive common-sense knowledge and linguistic understanding, large language models (LLMs) have demonstrated remarkable capabilities in text comprehension and logical reasoning for natural language processing tasks. Traditional prompt-based learning methods, which rely on contextual pattern matching, have proven to be effective in extracting knowledge from LLMs. However, these approaches are constrained by training data pattern matching, overlook reasoning processes, and consequently suffer from suboptimal prompt performance and limited interpretability. Moreover, considering that the intermediate steps generated by single-chain reasoning may not effectively assist LLMs in identifying the sentiment polarity of aspect terms, and that multiple reasoning paths often exist for complex reasoning tasks to reach correct conclusions, this paper proposes a Multi-Chain Thought Prompt Learning framework (MT-CPL). Starting from fundamental concepts, this method simulates human multi-path reasoning patterns to progressively construct comprehensive thought processes and deeply explore sentiment cues. Based on syntactic structures and the semantic logic of text, the framework incorporates four distinct perspectives of text comprehension: hierarchical reading, experiential reading, keyword-based reading, and analogical reading. It establishes a multi-chain prompt template and employs voting mechanisms to select correct reasoning path outcomes. The MT-CPL approach aims to guide LLMs in mining multi-dimensional textual information from different perspectives, gradually uncovering hidden contextual sentiment clues, while mitigating issues caused by irrelevant sentiment cues in intermediate reasoning steps. By decomposing main tasks incrementally, the method achieves progressive reasoning, effectively reduces the difficulty of direct analysis, and further enhances model interpretability through the integration of inherent common-sense knowledge.

Keywords: large language model; multi-chain of thought prompting; sentiment analysis; commonsense information



Academic Editors: Jose María Alvarez Rodríguez and Douglas O'Shaughnessy

Received: 15 September 2025

Revised: 28 October 2025

Accepted: 4 November 2025

Published: 18 November 2025

Citation: He, Y.; He, Z.; Gu, T.; Gu, B.; Wan, Y.; Li, M. Multi-Chain of Thought Prompt Learning for Aspect-Based Sentiment Analysis. *Appl. Sci.* **2025**, *15*, 12225. <https://doi.org/10.3390/app152212225>

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

With the rapid development of social media, e-commerce platforms, short videos, and other Internet application platforms, more and more users share their experiences, views, and emotions on the Internet. However, due to the complexity of comment structures and the diversity of forms, sentiment analysis (SA) has emerged. Aspect-based sentiment analysis (ABSA) is a fine-grained sentiment analysis [1–3] that focuses on analyzing the sentiment polarity of different aspects in a comment sentence. For example, as shown

in Figure 1, given comment sentence: “The food in this restaurant is delicious, but the environment could be improved.” The goal of ABSA is to determine the emotional polarity of the aspect words “food” and “environment”, respectively, as positive and negative.

The **food** in this restaurant is **delicious**, but the **environment** **could be improved**.

Aspect: food	Opinion: delicious	Polarity: positive
Aspect: environment	Opinion: could be improved	Polarity: negative

Figure 1. Example of aspect based sentiment analysis task, where red represents aspect words, blue represents sentiment words, and purple represents sentiment polarity.

In recent years, large language models (LLMs) have continuously undergone unsupervised multi-task training on massive text data, accumulating knowledge and gaining rich common-sense and language knowledge and demonstrating excellent text comprehension and logical reasoning abilities [4,5]. Due to its excellent text representation and reasoning abilities, some experts have applied it in the field of sentiment analysis [6,7]. In addition, prompt engineering has been proven to be an effective method for extracting knowledge from LLMs due to its feature of learning alongside context [8,9]. Simple prompt learning relies solely on pattern matching training data without considering the inference process and only generates the final answer. Chain prompt learning can start from basic concepts, gradually construct a complete thinking process to deeply explore emotional clues, and ultimately solve problems. This enables LLMs to simulate human thinking and reasoning step by step when solving problems rather than simply retrieving and mapping answers to similar problems in the training set [10–12]. Among them, Madaan et al. [11] further explored the potential mechanisms behind the effectiveness of Chain of Thought (CoT) on LLMs. Research has found that using Greek letters instead of numbers in examples does not affect the performance of the model, but simply modifying word order or changing grammar will have a huge impact on the performance of the model. They speculate that the effectiveness of CoT may stem from its ability to convey task-specific content to LLMs. Meanwhile, they also found that the intermediate steps in CoT may not necessarily contribute to learning how to solve tasks.

Many scholars have explored how to find the best solution to a problem. Stanovich and West [12] pointed out that, in complex reasoning tasks, there are usually multiple reasoning paths that can lead to the correct answer. Evans [13] found that the more thoughtful thinking and analysis a problem requires, the more likely there are to be diverse paths of reasoning for answers. Based on the above findings, this article will explore how multi-thought chain prompt learning can prompt and think about problem-solving methods from multiple thinking perspectives, alleviating the phenomenon of intermediate steps being unrelated to the final problem solving.

LLMs have extensive knowledge and information but lack the ability to deeply understand and apply this knowledge. This article focuses on aspect-level sentiment analysis tasks, starting from the perspective of reading comprehension, using various reading comprehension techniques (layered reading, experiential reading, keyword reading, and analogical reading) to prompt LLMs to infer emotional clues through multiple thinking modes.

Firstly, considering the multi-level text representation and the grammatical structure and semantic logic characteristics of comment sentences themselves, a hierarchical reading approach was designed to gradually analyze the emotions of aspect words by mining their specific attributes. With step-by-step parsing of the text content, LLMs can gradually establish an overall grasp of the text content, thereby more accurately determining the emotional colors contained therein.

Secondly, in comment sentences with multiple aspects of words and emotions, the emotional tone of the comment sentence itself highlights the commentator's preferences and focus on the aspect, providing important emotional background knowledge for judging the sentiment of aspect words. Based on this, this article designs experiential reading, aiming to simulate the human commenting experience from the perspective of the experiencer and place LLMs in the context described in the text in order to better understand the emotions expressed in the text and more accurately identify the emotional colors contained therein.

Again, mining and modeling text information of different granularities (sentence, phrase, and keyword information) can learn short distance collocations between words, thereby assisting in generating semantically rich contextual vector representations. Meanwhile, the perceptual stage of cognitive theory also suggests that attention selectively points to important key phrases in the context [14–16]. Keyword clues can guide learners to make positive predictions about the content of the text based on their existing knowledge background. Based on this, this paper designs a keyword reading method that guides LLMs to quickly locate emotional information in the text from the perspective of mining keywords or phrases, thereby grasping the main ideas and emotional tendencies of the text more efficiently.

Finally, consider that multi-task learning can utilize shared representation learning to learn semantic features and emotional clues from auxiliary tasks and transfer them to the main task to improve its generalization ability. Based on this, this article designs an analogy reading method to simulate auxiliary tasks by constructing new sentences with similar grammatical structures and emotional elements in order to capture grammatical structures and emotional features and then infer the emotions of the original comment sentence.

Based on LLM dimensional chain prompts, this paper proposes a multi-thought chain reasoning framework (MC-TPL). This article regards the LLM as a child with a wide range of knowledge but lacking professional practical experience. Starting from four different perspectives of text understanding, layered reading, experiential reading, keyword reading, and analogical reading, a multi-thinking layered chain prompt template is established. A voting method is adopted to select the appropriate thinking and reasoning paths. The model utilizes a multi-thinking hierarchical chain reasoning architecture to gradually reveal the emotional clues hidden in the text from multiple thinking perspectives. This method significantly reduces the complexity of directly conducting the overall analysis by decomposing the main task into a series of smaller steps. At the same time, combining the built-in common-sense information not only improves the accuracy of inference but also enhances the interpretability and comprehensibility of the model output results.

Our work contributions can be summarized in the following three aspects:

- (1) This article draws on human thinking patterns to gradually build a complete thinking path and deeply explore implicit emotional clues in the text.
- (2) Given the intermediate steps generated by single-thought chain prompt learning, there may be limitations in terms of word sentiment polarity in assisting LLM recognition. Based on the grammatical structure and semantic logic of the text, this study constructs a multi-thought chain prompt template from four different dimensions of text understanding and conducts multi-dimensional emotional semantic analysis research.
- (3) The MC-TPL model was tested on a benchmark dataset, and the experimental results showed the effectiveness of multi-thought chain prompt learning and external knowledge assistance.

Next, in the Section 2, we will first introduce the relevant work of predecessors; then, in the Section 3, we will provide a detailed description of our proposed MC-TPL model; in the Section 4, we conducted comparative experiments and ablation experiments on public datasets; in addition, in the Section 5, we discussed the shortcomings of the model and future solutions; and in the Section 6, we summarized the characteristics of the model and future work plans.

2. Related Work

In this section, this article first reviews previous aspect-based sentiment analysis tasks and pre-trained model-based methods and then introduces related research on large language models and prompt learning.

2.1. Neural Network Methods in Sentiment Analysis

Aspect-based sentiment analysis tasks are different from traditional sentence-level and document-level tasks, as they are entity-level and fine-grained sentiment analysis tasks. The sentiment analysis method based on machine learning [17,18] has better generalization ability and adaptability compared to the method of manually constructing dictionaries and rules [19]. It can automatically learn text patterns from large-scale corpora, thereby achieving text sentiment analysis. But its generalization ability in the face of complex problems is limited, and it cannot fully learn the contextual information of the text. With the widespread application and development of deep learning in the field of natural language processing (NLP), deep learning methods have gradually become the mainstream of research due to their powerful feature learning and semantic information representation capabilities [20–22]. Considering that different contextual factors have varying impacts on the emotional judgment of aspect words, Tang et al. [23] used two Long Short Term Memory (LSTM) networks to model the left and right contexts of aspect words, respectively, in order to obtain the left and right context vectors. Then, the left and right context vectors are combined to determine the emotional polarity of aspect words. In addition, considering that comment sentences may have long contexts or contain multiple words, the previous use of coarse-grained attention mechanisms may lead to insufficient information capture and the loss of key information. To address this issue, Fan et al. [24] proposed an Aspect-Level Sentiment Analysis Model (MGAN) based on a multi-granularity attention mechanism. They designed a fine-grained attention mechanism that can capture word-level interaction information aimed at evaluating the emotional impact of each contextual word on each aspect word in a sentence.

2.2. Pre-Trained Model Methods

Previous attention-based deep learning algorithms were limited by the distance between aspect words and emotion words. When there are multiple emotional words in a comment sentence, aspect words may not consider the overall meaning and may be influenced by closely related emotional words, leading to misjudgment. Therefore, some scholars adopt the syntactic distance of graph neural networks to solve the dependency problem of remote words. Thomas et al. [25] were the first to propose graph convolutional neural network models, which have been widely applied in multiple NLP tasks such as sentiment analysis, relationship extraction, and stance detection. Zhang et al. [26] constructed a sentence as a dependency tree using grammar information and word dependency relationships, learning the relationships between words based on whether they are connected by edges. Considering the influence of attention between words, Velickovic et al. [27] used the attention mechanism to convolve neighboring node features instead of simple average convolution. Wang et al. [28] considered the influence of different types of edges between word nodes and proposed an aspect specific relation graph attention network. Finally, considering the issues of inaccurate grammar parsing and the incomplete grammar structure of comment sentences, Li et al. [29] proposed a dual-graph convolutional neural network method that considers both grammar structure and semantic association and adopts regularization to overcome the above shortcomings. In addition, Ma et al. [30] used Abstract Meaning Representation (AMR) instead of dependency trees to extract the semantic features of comment sentences and achieved very good results.

2.3. Large Language Model Methods

The big language model has been widely proven to exhibit extraordinary abilities in text comprehension and logical reasoning. Transformer is an attention-based structural framework created by the Google team, which demonstrated powerful language understanding capabilities when first applied in the field of natural language processing (NLP) [31,32]. This model is divided into an encoder and a decoder. The encoder maps the input sequence to a continuous sequence representation, and the decoder generates an output sequence based on the encoder's sequence representation. The T5 (text-to-text) model is a universal text generation framework proposed by the Google team based on the Transformer framework. This model adopts a text-to-text framework, aiming to incorporate multiple tasks into the model to generate target text. This model also adopts the Encoder–Decoder structure, but T5 uses the offset between queries and values in the self-attention mechanism to generate relative position embeddings, which are more sensitive to position [33,34]. In addition, T5 considers multiple pre-training methods to improve the performance of generative models and outperforms Bert and GPT in question answering and generative extraction tasks [35,36]. Based on the T5 model, Deng et al. [37] proposed a bidirectional generation framework based on the T5 model to solve the problem of cross-domain differences caused by labels only existing in the original domain in cross-domain sentiment analysis tasks. This framework trains models in both text to label and label to text directions. The text to label direction converts various tasks into a unified format and can generate noise predictions for unlabeled target data, while the label to text direction utilizes noise predictions to generate natural language containing given labels in order to enhance high-quality training data and enrich model knowledge in the target domain. In addition, Zhang et al. [38] designed a unified generative framework GAS-T5 for Attribute Central Sentiment Extraction (ACSTE), which is based on the T5 model and uses both annotation and extraction paradigms to transform the original task into a text generation task. Excellent experiments have demonstrated the generality and effectiveness of the model.

ChatGPT is an intelligent chat machine created by Google using artificial intelligence and cognitive computing methods. It is built on the foundation of the Transformer and Generative Pre-trained Transformer (GPT) [39] and has significant text generation capabilities that can be used in multiple natural language processing fields [40]. Generative pre-trained models have gradually enhanced their natural language processing and text generation capabilities through multiple versions (GPT-1, GPT-2, GPT-3, InstructGPT, and ChatGPT) by increasing the number of parameters and improving self-supervised methods [41–43]. Wang et al. [44] conducted a detailed study on whether ChatGPT is an excellent sentiment analyzer by combining prompt learning methods in the fields of sentiment cause analysis, sentence-level sentiment analysis, and aspect-level sentiment analysis. Detailed research has shown that ChatGPT exhibits excellent unsupervised sentiment analysis capabilities and performs better in open domains compared to fine-tuned BERT. However, its capabilities in fields such as medicine and social media need to be strengthened.

2.4. Prompt Learning Methods

In addition, prompt learning has been proven to effectively guide large models in logical reasoning and knowledge acquisition. Ping et al. [45] manually designed a prompt template [CLS]polarity[SEP]sentence + template[SEP] in the aspect category sentiment analysis task and combined it with the attention mechanism to guide BERT to filter out important feature information (phrases and words). Fei et al. [6] first applied the combination of CoT and LLMs to implicit sentiment analysis to address the issue of unclear emotional clues. They use a three-hop chain reasoning framework to gradually infer fine-grained aspects, potential opinion words, and, ultimately, emotions. This model has demonstrated excellent performance in both supervised and unsupervised training, providing new ideas

for the research and development of large-scale models in the field of sentiment analysis. In addition, Inaba et al. [46] further explored how to use CoT to call LLMs using external tools. On the chemical dataset, they created the MultiTool-CoT method to call external tools such as calculators, chemical reaction predictors, and molar mass tables, which demonstrated excellent performance in chemical numerical reasoning tasks.

3. Methodology

The aspect-level sentiment analysis task is defined as follows: given a sentence X containing the target term A , the model determines the sentiment polarity Y of aspect term A , which is positive, neutral, or negative. For standard prompt-based methods, we can construct the following prompt template as input for the LLM:

Given the sentence X , what is the sentiment polarity towards t ? (1)

The LLM should return the answer via

$$\hat{y} = \operatorname{argmax}(y|X, t).$$

where $\hat{y}$ is the sentiment polarity predicted by the LLM.

3.1. Multi-Thinking Chain Prompt Template

Given that traditional prompt methods often exhibit deficiencies such as logical leaps, error propagation, and limited interpretability in complex tasks due to the lack of structured guidance, this study introduces a three-step Chain of Thought (CoT) prompting approach [10,47]. By designing explicit task decomposition and step-by-step reasoning chains, this method breaks down multi-step problems into verifiable subtasks, thereby reducing reasoning biases and enhancing result reliability. Furthermore, considering that multiple valid reasoning paths may lead to correct answers in complex tasks, this paper details the construction of four multi-perspective CoT prompting templates for comment sentence S , including progressive reading, experiential reading, keyword-driven reading, and analogical reading, to improve model adaptability and robustness across diverse reasoning scenarios. The model structure diagram is shown in Figure 2.

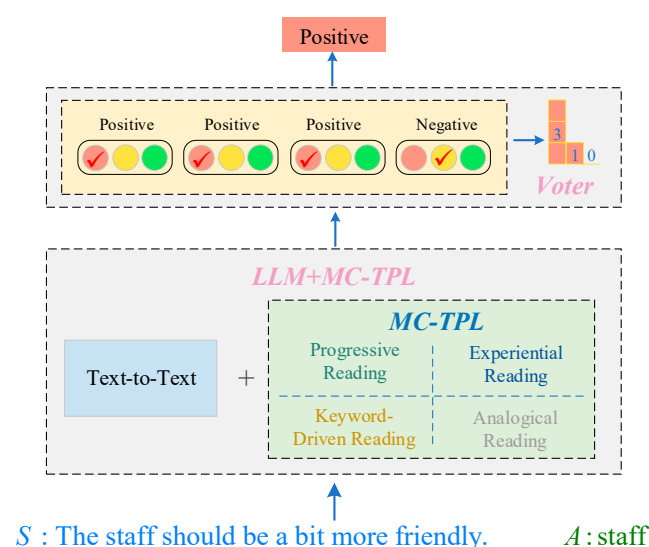


Figure 2. The proposed functional framework.

3.1.1. Progressive Reading

Hierarchical reading aims to identify the emotions of aspect words by mining their specific attributes. With step-by-step parsing of the text content, the LLM can gradually establish an overall grasp of the text content, thereby more accurately determining the emotional colors contained therein. The specific process is shown in Figure 3. This section constructs a chain prompt template in three steps:

Step 1: Inquire with LLM about the specific attribute of aspect word A mentioned in sentence S:

$$P_1[\text{Given sentence } S], \text{ which specific attribute of the aspect } A \text{ is possibly mentioned?} \quad (2)$$

where P_1 is the context of the first step prompt in progressive reading, which can be formulated as follows: $T_p = \text{argmax}_p(t_p|S, A)$, where T_p is the specific attribute that aspect word A may mention.

Step 2: Under the premise that sentence S, aspect word A, and its attribute T_p are known, guide LLM to combine external common-sense knowledge to search for the specific opinion word of attribute T_p , which may be

$$P_2[P_1, T_p]. \text{ Based on the common sense, what is the implicit opinion towards the mentioned aspect of } T, \text{ and why?} \quad (3)$$

where P_2 is the context of the first step prompt in progressive reading, which can be formulated as follows: $O_p = \text{argmax}_p(o_p|S, A, T)$, where O_p is the opinion word of attribute T_p of aspect word A in the comment sentence.

Step 3: Using the entire layer-wise progressive information (S, A, T_p , O_p) as context, inquire about the specific sentiment polarity of the aspect word A from the LLM:

$$P_3[P_2, O_p], \text{ Bases on the opinion, what is the sentiment polarity towards } A? \quad (4)$$

where P_3 is the context of the entire hierarchical chain prompting method, which can be formulated as follows: $\hat{y}_p = \text{argmax}_p(y_p|S, A, T_p, O_p)$.

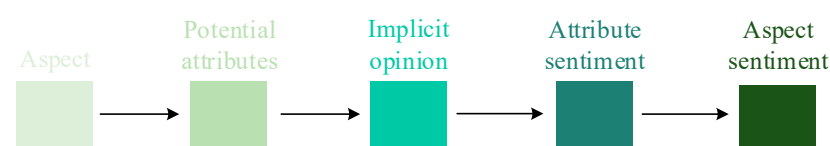


Figure 3. Flowchart of progressive reading.

3.1.2. Experiential Reading

Experiential reading aims to simulate the human commenting experience by assigning LLM role identities in order to better understand the emotions expressed in the text and more accurately identify the emotional colors contained therein. In aspect-level sentiment analysis tasks, commentators often express their opinion through statements that are affirmative, negative, and devoid of emotion. The LLM starts from an experiential perspective, infers the purpose of the commentator's opinion expression, and then infers specific emotional tendencies. The specific process of experiential reading is shown in Figure 4. This section constructs a chain prompt template in three steps:

Step 1: By setting the identity of a reviewer in a specific field for LLM, inquire whether LLM is writing a comment sentence to express affirmation, dissatisfaction, expectation for improvement, and a statement with a ruthless tone regarding aspect A:

E_1 [Given sentence S], This is a review you wrote in the dataset field. For aspect A , what was your purpose when you wrote this review? Such as, to express satisfaction, dissatisfaction, expected improvements, and unemotion statement. (5)

where E_1 is the context of the first step prompt in experiential reading, which can be formulated as follows: $T_e = \text{argmaxp}(t_e|S, A)$, where T_e is the purpose description of aspect word A in the comment sentence.

Step 2: Under the premise that sentence S , aspect word A , and opinion purpose W are known, guide LLM to combine external common-sense knowledge to search for possible emotional tendencies involved in purpose T_e :

$E_2[E_1, T_e]$. Based on the common sense, what do you think is the emotional inclination of T_e , and why? (6)

where E_2 is the context of the second step prompt in experiential reading, which can be formulated as follows: $O_e = \text{argmaxp}(o_e|S, A, T_e)$, where O_e is the emotional tendency described about the purpose of aspect word A .

Step 3: Using the entire experiential information (S, A, T_e, O_e) as the prompt context, ask LLM about the specific emotional polarity of word A :

$E_3[E_2, O_e]$, Based on the opinion, what is the sentiment polarity towards A ? (7)

where E_3 is the context of the entire experiential chain prompting method, which can be formulated as follows: $\hat{y}_e = \text{argmaxp}(y_e|S, A, T_e, O_e)$.

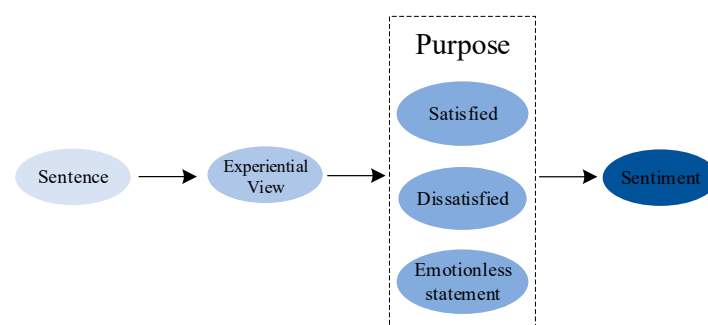


Figure 4. Flowchart of experiential reading.

3.1.3. Keyword-Driven Reading

Keyword-Driven Reading learns short distance collocations between words by grasping the keywords and phrases in comment sentences, allowing the LLM to quickly locate emotional information in the text and more quickly grasp the main idea and emotional tendency of the text. The specific process of keyword reading is shown in Figure 5. This section constructs a chain prompt template in three steps:

Step 1: Inquire about the keywords and phrases related to the emotional expression of aspect word A in the LLM comment sentence:

K_1 [Given sentence S], what are the keywords or phrases relation to emotional expression in the sentence for aspect A ? (8)

where K_1 is the context of the first step prompt in Keyword-Driven Reading, which can be formulated as follows: $T_k = \text{argmaxp}(t_k|S, A)$, where T_k is the purpose description of aspect word A in the comment sentence.

Step 2: Under the premise that sentence S , aspect word A , and key word or phrase T_k are known, guide LLM to combine external common-sense knowledge to identify the possible sentiment polarity associated with the key word or phrase T_k :

$$K_2[K_1, T_k]. \text{ Based on the common sense, what do you think are the emotional inclination of keywords or phrases } T_k, \text{ and why?} \quad (9)$$

where K_2 is the context of the second step prompt in experiential reading, which can be formulated as follows: $O_k = \text{argmaxp}(o_k|S, A, T_k)$, where O_k represents the emotional inclination towards the keyword or phrase T_k .

Step 3: Using the entire experiential information (S, A, T_k, O_k) as the prompt context, ask LLM about the specific emotional polarity of word A :

$$K_3[K_2, O_k], \text{ Bases on the opinion, what is the sentiment polarity towaeds } A? \quad (10)$$

where K_3 is the context of the entire Keyword-Driven chain prompting method, which can be formulated as follows: $\hat{y}_k = \text{argmaxp}(y_k|S, A, T_k, O_k)$.

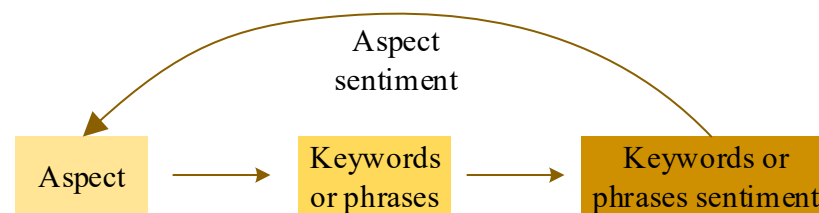


Figure 5. Flowchart of Keyword-Driven Reading.

3.1.4. Analogical Reading

Analogous reading simulates auxiliary tasks by constructing new sentences with similar grammatical structures and emotional elements in order to capture grammatical structural information and emotional features, inferring the emotions of the original comment sentence. The specific process of analogical reading is shown in Figure 6. This section constructs a chain prompt template in three steps:

Step 1: Guide LLM to construct a new sentence with grammatical structure and emotional elements similar to the original comment sentence:

$$C_1[\text{Given sentence } S], \text{ use your imagination to create a new sentence containing aspect word } A, \text{ which needs to be similar in grammar and emotional inclination to the comment sentence } S. \quad (11)$$

where C_1 is the context of the first step prompt in analogical reading, which can be formulated as follows: $T_c = \text{argmaxp}(t_c|S, A)$, where T_c is a sentence with a grammatical structure and emotional elements similar to the comment sentence S .

Step 2: Under the premise that sentence A , aspect word A , and new comment sentence T_c are known, guide LLM to analyze the emotional tendency of aspect word A in the new comment sentence T_c by combining external common-sense knowledge:

$$C_2[C_1, T_c]. \text{ Based on the common sense, what do you think are the emotional inclination of aspect } A \text{ in comment sentence } T_c, \text{ and why?} \quad (12)$$

where C_2 is the context of the second step prompt in experiential reading, which can be formulated as follows: $O_c = \text{argmaxp}(o_c|S, A, T_c)$, where O_c is the emotional tendency of the aspect word A in the new comment sentence T_c .

Step 3: Using the entire analogical information (S, A, T_c, O_c) as the prompt context, ask LLM about the specific emotional polarity of word A:

$$C_3[C_2, O_e], \text{Bases on the opinion, what is the sentiment polarity towards A?} \quad (13)$$

where C_3 is the context of the entire analogical chain prompting method, which can be formulated as follows: $\hat{y}_c = \operatorname{argmaxp}(y_c | S, A, T_c, O_c)$.

Finally, after obtaining four emotions through four types of thought chain prompts, the emotion category with the highest number of votes was counted and determined as the final emotional polarity of the aspect word.

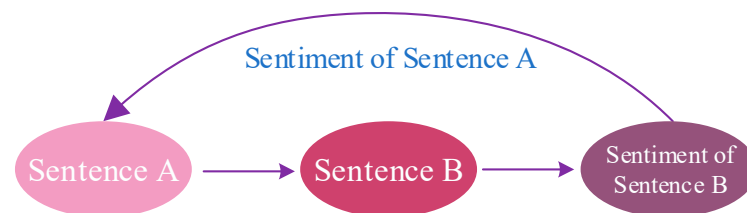


Figure 6. Flowchart of analogical reading.

3.2. Emotional Reasoning Enhancement Strategies

We enhance reasoning by leveraging self-consistency mechanisms [48,49] to ensure reasoning accuracy. For each of the three reasoning steps, the LLM decoder generates multiple answers, which may vary in predictions regarding aspect A, opinion o, and polarity Y. At each step, answers with high-voting consistency for a, o, or y are retained, and the most confident one is used as context for the next step. Additionally, when on-demand training sets are available, we can fine tune THOR using supervised fine tuning. We design a reasoning revising method: at each step, a prompt is constructed by concatenating the initial context, the reasoning answer text, and the final question, which is then fed into the LLM to predict the sentiment label instead of proceeding to the next reasoning step. Under the supervision of gold labels, the LLM learns to generate more accurate intermediate reasoning, which is beneficial for the final prediction.

4. Experiment

4.1. Relevant Parameters

The experimental validation of this article was conducted on the benchmark datasets [50] Restaurants14 and Laptop14. Due to the fact that the BERT model is composed of Transformer encoders and cannot generate text based on the multi-round prompts of CoT, this paper adopts the T5 model with the Encoder–Decoder style as the backbone network of the MC-TPL framework. In addition, this article also uses ChatGPT (GPT-3.5-based) to verify the effectiveness of the MC-TPL framework. Two versions of the T5 model, T5 base (250 M) and T5 large (780 M), were used in the experiment, while the ChatGPT model was accessed through an API. The experiment was conducted on two NVIDIA A40 GPUs, and the results were evaluated using the accuracy and F1 score as metrics.

We conducted experiments on three public sentiment analysis datasets: the review data from the SemEval 2014 task containing reviews of restaurants and laptops [50]. The dataset is divided into a training set and a test set, and a comment sentence will have one or more aspect words, each corresponding to one of the positive, negative, and neutral sentiments. The detailed statistics are shown in Table 1.

Table 1. Statistics of the three datasets for the ABSA task.

Dataset	Positive		Negative		Neutral	
	Train	Test	Train	Test	Train	Test
Restaurant	2164	728	807	196	637	196
Laptop	994	341	870	128	464	169

4.2. Baseline Model

In order to evaluate the effectiveness of the model, this article compares three types of baseline models with models based on the MC-TPL framework.

(1) Pre-trained model methods:

BERT-SPC [51]: Uses the [CLS] + Sentence + [SEP] + Aspect + [SEP] sequence input into BERT to jointly learn contextual information and target representation.

RGAT-BERT [28]: Focusing on the complexity of language and the presence of multiple aspect words in a comment sentence, the original dependency tree is reconstructed and pruned to establish a dependency tree with aspect words as the root node.

HGCN [52]: In order to jointly learn structural range and predict sentiment polarity, a hybrid graph convolutional network (HGCN) is proposed to synthesize information from selection trees and dependency trees, exploring the potential of connecting the two syntactic parsing methods to enrich representations.

KDGN [53]: Considering the influence of entity-related domain knowledge in comment texts, a knowledge-aware dependency graph network model is proposed to capture domain knowledge related to texts.

BERTAsp + CEPT [54]: Considering the problem of insufficient recognition of implicit emotional expressions, a model based on supervised contrastive pre-training (SCAPT) was designed to effectively capture implicit and explicit emotional tendencies in text by learning emotional knowledge using a large-scale emotion annotation corpus.

CABiLSTM-BERT [55]: Considering the problem of implicit feature information loss, the BERT pre-trained model is frozen to extract word vectors and combined with the BiLSTM and multi-head self-attention mechanism, effectively extracting and preserving the implicit feature information of each layer

MambaForGCN [56]: In response to the difficulties in capturing long-distance dependencies and the high complexity of attention mechanisms, the MambaFormer module based on the state space model (SSM) and the Kolmogorov–Arnold Network (KAN) gating fusion mechanism were combined to effectively enhance the model’s ability to capture long-distance dependencies.

HPEP-GCN [57]: Considering the insufficient capture of feature correlation and the limited utilization of external knowledge, the accuracy and robustness of sentiment analysis have been effectively improved by integrating external knowledge, introducing hierarchical and location-aware graph convolutional networks, and maximizing mutual information feature interaction mechanisms.

(2) Prompt learning methods:

Fei et al. [6] proposed BERT + Prompt, T5 base + Prompt, and T5 large + Prompt by combining the LLM with cue learning and conducted comparative experiments and zero-shot learning experiments on benchmark datasets, respectively.

ChatGPT + Prompt [44]: This method combines ChatGPT with cue learning and conducts zero-shot learning experiments on a benchmark dataset using “Sentence:{sentence}what is the sentence polarity of the aspect {aspect}in the sentence?” as the cue.

Fei et al. [6] proposed T5 base + IoT and T5 large + CoT by combining the LLM and CoT. We conducted comparative experiments and zero-shot learning experiments on the benchmark dataset.

(3) The proposed method:

This article proposes T5 base + MC-TPL, T5 large + MC-TPL, and ChatGPT + MC-TPL, respectively, using T5 base, T5 large, and ChatGPT as model skeletons and a multi-thought chain prompt learning framework. On the benchmark dataset, comparative experiments and zero-shot learning experiments were conducted on the T5 base + MC-TPL and T5 large + MC-TPL models, respectively. ChatGPT + MC-TPL conducted zero-shot learning experiments.

4.3. Comparative Experiments and Analysis

Tables 2 and 3, respectively, show the results of comparative experiments and zero-shot learning experiments of the multi-thought chain prompt learning framework on the dataset. The model uses ACC and F1 as evaluation indicators, and the results show the effectiveness of MC-TPL. Through comparative experiments, the following can be further analyzed:

Results on Supervised Fine Tuning

From Table 2, it can be seen that some methods based on prompt learning and CoT (BERT + Prompt, T5 base + Prompt, and T5 base + CoT) perform worse than KDGN. This is mainly because KDGN has demonstrated excellent sentiment analysis capabilities through pre-training in large-scale sentiment analysis corpora. In addition, this section also found that the effect of T5 base + CoT[#] did not perform well, which may be due to the fact that the intermediate steps of single-thinking CoT may not necessarily help learn how to solve aspect sentiment analysis tasks. Under the guidance of the multi-thought chain prompt framework proposed in this article, the LLM has shown remarkable results. This article argues that multi-thought chain prompt learning can prompt and think from multiple perspectives, providing diverse methods for problem solving and effectively guiding the LLM to conduct emotional cue reasoning, demonstrating its unique advantages.

Results on Zero-shot Reasoning

From the zero sample comparison data in Table 3, it can be seen that the method based on prompt learning exhibits significant advantages compared to the method based on BERT, which fully demonstrates the superiority of prompt learning in the field of zero-shot learning. Prompt learning can guide the LLM to effectively extract relevant emotional knowledge, thereby improving the accuracy of aspect-level sentiment analysis. In addition, this article also found that the method based on CoT is better than the method based on prompt learning. This indicates that CoT technology can start from basic concepts, gradually construct a complete thinking process to deeply explore emotional clues, and ultimately solve problems. This enables the LLM to simulate human thinking and perform step-by-step reasoning when conducting sentiment analysis rather than relying solely on simple retrieval and the mapping of similar problems in the training set. Finally, the MC-TPL-based model showed better performance than the CoT-based model. This fully demonstrates the excellent performance of MC-TPL, which can provide prompts for LLMs from multiple perspectives, guiding them to think and solve problems from multiple angles and thinking. MC-TPL not only makes the LLM more accurate in sentiment inference but also demonstrates its enormous potential in improving sentiment analysis performance.

Table 2. Comparative experiment of MC-TPL model.

	Models	Restaurants14		Laptop14	
		Acc	F1	Acc	F1
pre-train	CABiLSTM-BERT	83.75	75.87	77.91	73.04
	SPC-BERT	84.46	76.98	78.99	75.03
	RGAT-BERT	86.33	81.04	78.21	74.07
	HGCN	86.45	80.60	79.59	76.24
	KDGN	87.01	81.94	81.32	77.59
	MambaForGCN	86.68	80.86	81.80	78.56
	HPEP-GCN	86.86	80.65	81.96	79.10
prompt	BERT + Prompt	-	81.34	-	78.58
	T5-base + Prompt	-	81.50	-	79.02
	T5-base + CoT [#]	86.16	79.70	81.03	77.25
	T5-large + CoT [#]	88.04	82.17	83.31	79.55
our	T5-base + MC-TPL	87.86	81.99	83.45	79.62
	T5-large + MC-TPL	88.75	83.73	84.21	80.33

[#] A represents the result of reproducing the code in this article.

Table 3. Experimental results of zero-shot learning for MC-TPL model.

	Models	Restaurants14		Laptop14	
		Acc	F1	Acc	F1
pre-train	SPC-BERT	-	21.76	-	25.34
	RGAT-BERT	-	27.48	-	25.68
	SCAPT-BERT	-	30.02	-	25.77
prompt	BERT + Prompt	-	33.62	-	35.17
	T5-base + Prompt [#]	76.87	53.32	68.49	51.06
	T5-large + Prompt [#]	78.39	55.03	69.28	52.08
	T5-base + CoT [#]	77.14	53.58	68.34	51.34
	T5-large + CoT [#]	79.02	55.66	70.85	53.54
our	T5-base + MC-TPL	77.67	54.31	69.43	51.82
	T5-large + MC-TPL	79.82	56.36	71.32	54.06

[#] A represents the result of reproducing the code in this article.

4.4. Comparative Experiments and Analysis of Implicit Sentiment Analysis

In order to better validate the importance of MC-TPL in implicit sentiment analysis, this section conducted comparative experiments on the benchmark dataset, and the results are shown in Figure 7. The results indicate that the method based on prompt learning achieved better results compared to the method based on pre-training, and implicit emotional reasoning may require more common-sense knowledge to assist reasoning. Compared to simple prompt learning, chain prompt learning can gradually decompose complex reasoning tasks, improve the utilization of implicit knowledge, enhance the interpretability of the model, and achieve better results. Our proposed multi-thought chain prompting method can infer knowledge chains in parallel, more comprehensively capture the complex semantics, contextual dependencies, and subjectivity of implicit emotions, and achieve the best results.

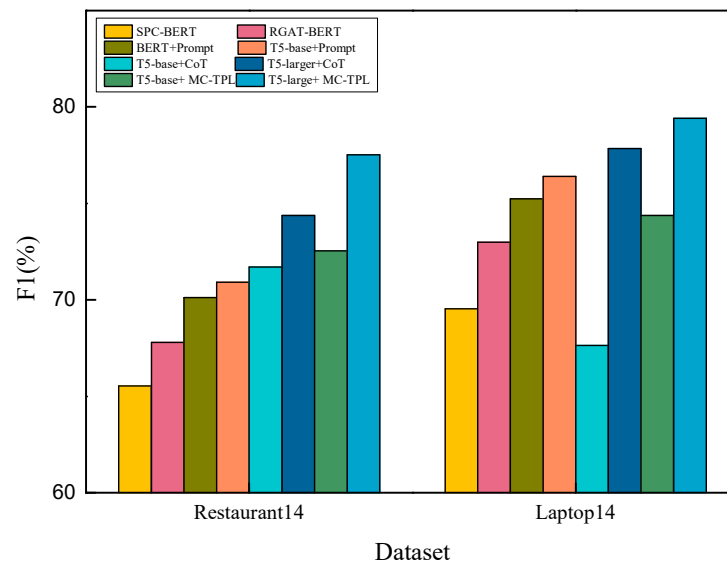


Figure 7. Comparative experiment of MC-TPL model in implicit sentiment analysis.

4.5. Ablation Experiment

In order to better verify the impact of each prompt method in MC-TPL on the model, this section conducted experiments comparing the model based on the T5 base, and the results are shown in Table 4. Among them, progressive reading, experiential reading, keyword reading, and analogical reading are, respectively, represented as progressive, experiential, keyword, and analogical. The results indicate that using the T5 base as the benchmark framework and employing these four prompt learning methods achieved better results compared to most BERT-based methods. This proves that the thinking chain constructed based on sentence structure characteristics and theoretical support in previous articles can achieve good prompting effects. However, compared to the MC-TPL method proposed in this article, the performance of models using a single-thought suggestion method has significantly decreased. A single-thought chain may have intermediate reasoning steps unrelated to emotions, but a multi-thought chain approach can effectively reduce this impact. In addition, complex reasoning tasks often have multiple reasoning paths that can lead to the correct answer. Therefore, combining the results of multiple thinking perspectives can provide a more comprehensive inference of specific aspects of emotional polarity.

Table 4. Impact of single chain hint method on the model.

Models	Restaurant14		Laptop14	
	Acc	F1	Acc	F1
Progressive	86.25	79.79	82.28	78.67
Experiential	87.69	80.75	81.03	76.21
Keyword	86.79	81.67	81.96	78.25
Analogical	86.60	80.31	81.19	76.86
MC-TPL	87.86	81.99	83.45	79.62

4.6. Error Analysis

To further validate the effectiveness of multi-thought chain prompt learning, this section also studied the joint effect of MC-TPL and ChatGPT in zero-shot learning scenarios, as shown in Figure 8. From the table, it can be seen that, although ChatGPT + MC-TPL has an improved performance compared to ChatGPT with only a traditional prompt, the difference is not significant.

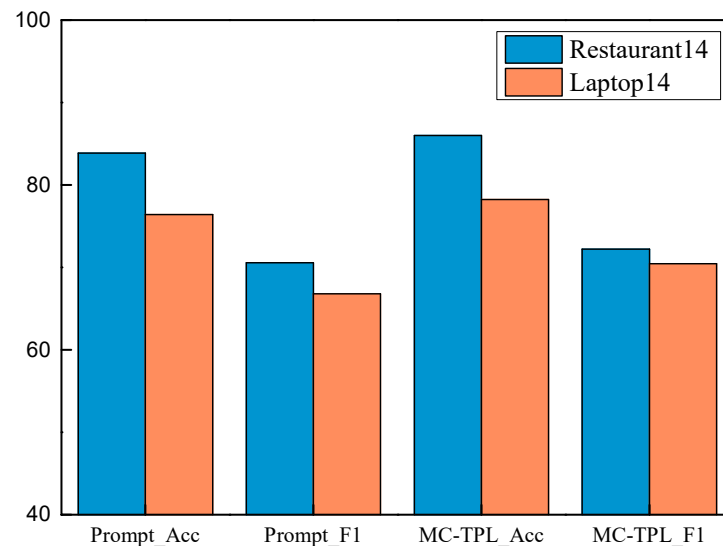


Figure 8. Comparison results of different prompt learning methods.

Therefore, this section analyzed examples of ChatGPT + MC-TPL sentiment analysis errors. The analysis results show that most of the errors stem from MC-TPL guiding the LLM, which leads to the model making excessive emotional inferences based on existing information. For example, in the comment sentence: “Entrees includes classics like lasagna, fettuccine alfredo, and chicken parmigiana.”, the aspect word “Entrees” corresponds to the emotion “Neutral.” The two different reasoning methods are shown in Figure 9. The comment sentence only briefly introduces the classic dishes included in “Entrees” and does not have a clear emotional tendency. The ChatGPT + Prompt method can determine the emotion of “Neutral” based on the aspect words. However, under the guidance of multi-thought chain prompt learning, the LLM may engage in excessive inference and speculation regarding the content of comment sentences, resulting in the allocation of “positive” or “negative” emotions to aspect words that were originally “neutral”.



Figure 9. The process of processing example sentences for ChatGPT + MC-TPL and ChatGPT + Prompt models.

5. Limitations

In aspect-level sentiment analysis, although parallel inference chains (multi-thought chain prompt learning) can improve complex sentiment analysis ability through multi-dimensional perspectives, their shortcomings still need to be addressed. The core contradiction lies in the output conflict between chains: different inference chains (such as the semantic chain and pragmatic chain) may produce contradictory judgments on the same aspect of words and lack interaction and coordination mechanisms due to the independent modeling of each chain. In addition, the problem of computational redundancy increases linearly as the number of inference chains increases, which can lead to memory overflow and inference delay in long-text scenarios. Therefore, future improvements can focus on two aspects: firstly, building interactive inference links and achieving dynamic interaction through inter-chain attention mechanisms or graph fusion modules. Secondly, developing a dynamic chain selection mechanism that adaptively activates relevant inference links based on input complexity and combines gating mechanisms or knowledge distillation techniques to balance efficiency and performance.

6. Conclusions

Based on the characteristics of sentence structure and the effectiveness of common-sense knowledge, this paper innovatively proposes the Multi-Chain of Thought Prompt Learning (MC-TPL) to address the issue that the intermediate steps generated by large language models combined with single-thought chain prompts may deviate from the core of sentiment analysis. This framework starts from four unique dimensions of text understanding, carefully constructs a multi-thinking layer chain prompt template, and guides the big language model to gradually explore and extract hidden contextual emotional clues from different thinking perspectives, achieving incremental reasoning.

With the help of this multi-thinking hierarchical chain reasoning framework, machines can infer contextual emotional clues hidden in the text from multiple thinking perspectives, layer by layer. In the future, we will consider working from the following aspects: (1) We will explore dynamic interactive multi-thought chain prompting methods to enhance the deep mining of implicit information in text and thus construct a model that combines complex semantic reasoning and emotion retrieval functions and has interpretability; (2) we will consider starting with voting methods or using multiple thought chains to gradually adjust feedback and carefully analyze implicit emotional cues; and (3) we plan to explore in depth the specific application methods and interpretability of external common-sense information and multi-thought chains and expand the relevant results to other natural language processing task areas.

Author Contributions: Conceptualization, T.G. and Y.H.; methodology, T.G. and B.G.; software, Y.W. and M.L.; validation, T.G., Z.H., Y.H., and Y.W.; formal analysis, T.G.; investigation, Z.H. and B.G.; resources, Z.H.; data curation, T.G.; writing—original draft preparation, T.G.; writing—review and editing, T.G., Y.H., and Z.H.; visualization, T.G.; supervision, Z.H. and Y.W.; funding acquisition, Z.H., T.G., and B.G. All authors have read and agreed to the published version of the manuscript.

Funding: This study was funded by the Young Scientists Fund of the Natural Science Foundation of Xinjiang Province under Grant No. 2025D01C299, 2025D01B80, 2025D01C294; 2025 Tianchi Talents Young Doctors Fund under Grant No. 601002000103, 51052501823; Doctoral (Postdoctoral) Research Initiation Fund of Xinjiang Normal University (Project No. XJNUZBS2404); Basic Research Funds for Universities in the Autonomous Region (Project No. XJEDU2025P068).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are openly available in <https://github.com/scofield7419/THOR-ISA> (accessed on 28 August 2023).

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Nazir, A.; Rao, Y.; Wu, L.; Sun, L. Issues and challenges of aspect-based sentiment analysis: A comprehensive survey. *IEEE Trans. Affect. Comput.* **2020**, *13*, 845–863. [\[CrossRef\]](#)
2. Poria, S.; Cambria, E.; Gelbukh, A. Aspect extraction for opinion mining with a deep convolutional neural network. *Knowl. Based Syst.* **2016**, *108*, 42–49. [\[CrossRef\]](#)
3. Ren, F.; Feng, L.; Xiao, D.; Cai, M.; Cheng, S. DNet: A lightweight and efficient model for aspect based sentiment analysis. *Expert Syst. Appl.* **2020**, *151*, 113393. [\[CrossRef\]](#)
4. Paranjape, B.; Michael, J.; Ghazvininejad, M.; Zettlemoyer, L.; Hajishirzi, H. Prompting Contrastive Explanations for Commonsense Reasoning Tasks. Findings of the Association for Computational Linguistics: ACL-IJCNLP 2021. *arXiv* **2021**, arXiv:2106.06823.
5. Liu, J.; Liu, A.; Lu, X.; Welleck, S.; West, P.; Bras, R.L.; Hajishirzi, H. Generated knowledge prompting for commonsense reasoning. In Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics, Dublin, Ireland, 22–27 May 2022; pp. 3154–3169.
6. Fei, H.; Li, B.; Liu, Q.; Bing, L.; Li, F.; Chua, T.S. Reasoning implicit sentiment with chain-of-thought prompting. In Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics, Toronto, ON, Canada, 9–14 July 2023; pp. 1171–1182.
7. Přibáň, P.; Šmíd, J.; Steinberger, J.; Mištera, A. A comparative study of cross-lingual sentiment analysis. *Expert Syst. Appl.* **2024**, *247*, 123247. [\[CrossRef\]](#)
8. Reynolds, L.; McDonell, K. Prompt programming for large language models: Beyond the few-shot paradigm. In Proceedings of the CHI Conference on Human Factors in Computing Systems, Online, 8–13 May 2021; pp. 1–7.
9. Achiam, J.; Adler, S.; Agarwal, S.; Ahmad, L.; Akkaya, I.; Aleman, F.L.; McGrew, B. GPT-4 technical report. *arXiv* **2023**, arXiv:2303.08774. [\[CrossRef\]](#)
10. Wei, J.; Wang, X.; Schuurmans, D.; Bosma, M.; Xia, F.; Chi, E.; Zhou, D. Chain-of-thought prompting elicits reasoning in large language models. *Adv. Neural Inf. Process. Syst.* **2022**, *35*, 24824–24837.
11. Madaan, A.; Hermann, K.; Yazdanbakhsh, A. What makes chain-of-thought prompting effective? A counterfactual study. In Proceedings of the Empirical Methods in Natural Language Processing, Singapore, 6–10 December 2023; pp. 1448–1535.
12. Chen, Z.; Zhou, Q.; Shen, Y.; Hong, Y.; Sun, Z.; Gutfreund, D.; Gan, C. Visual chain-of-thought prompting for knowledge-based visual reasoning. In Proceedings of the AAAI Conference on Artificial Intelligence, Vancouver, Canada, 26–27 February 2024; Volume 38, pp. 1254–1262.
13. Stanovich, K.E.; West, R.F.; Alder, J. Individual differences in reasoning: Implications for the rationality debate? Open peer commentary-three fallacies. *Behav. Brain Sci.* **2000**, *23*, 665. [\[CrossRef\]](#)
14. Evans, J.S.B. Intuition and reasoning: A dual-process perspective. *Psychol. Inq.* **2010**, *21*, 313–326. [\[CrossRef\]](#)
15. Sloman, S.A. The empirical case for two systems of reasoning. *Psychol. Bull.* **1996**, *119*, 3. [\[CrossRef\]](#)
16. Scherer, K.R.; Schorr, A.; Johnstone, T. *Appraisal Processes in Emotion: Theory, Methods, Research*; Oxford University Press: Oxford, UK, 2001.
17. Baccianella, S.; Esuli, A.; Sebastiani, F. Sentiwordnet 3.0: An enhanced lexical resource for sentiment analysis and opinion mining. In Proceedings of the International Conference on Language Resources and Evaluation, Valletta, Malta, 19–21 May 2010; Volume 10, pp. 2200–2204.
18. Ding, X.; Liu, B.; Yu, P.S. A holistic lexicon-based approach to opinion mining. In Proceedings of the International Conference on Web Search and Web Data Mining, Palo Alto, CA, USA, 11–12 February 2008; pp. 231–240.
19. Wang, H.; Ester, M. A sentiment-aligned topic model for product aspect rating prediction. In Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing, Doha, Qatar, 25–29 October 2014; pp. 1192–1202.
20. Li, W.; Shao, W.; Ji, S.; Cambria, E. BiERU: Bidirectional emotional recurrent unit for conversational sentiment analysis. *Neurocomputing* **2022**, *467*, 73–82. [\[CrossRef\]](#)
21. Li, X.; Bing, L.; Lam, W.; Shi, B. Transformation networks for target-oriented sentiment classification. In Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics ACL, Melbourne, Australia, 15–20 July 2018; pp. 946–956.
22. Li, W.; Zhu, L.; Shi, Y.; Guo, K.; Cambria, E. User reviews: Sentiment analysis using lexicon integrated two-channel CNN-LSTM family models. *Appl. Soft Comput.* **2020**, *94*, 106435. [\[CrossRef\]](#)
23. Tang, D.; Qin, B.; Feng, X.; Liu, T. Effective LSTMs for target-dependent sentiment classification. In Proceedings of the 26th International Conference on Computational Linguistics: Technical Papers, Osaka, Japan, 11–16 December 2016; pp. 3298–3307.

24. Fan, F.; Feng, Y.; Zhao, D. Multi-grained attention network for aspect-level sentiment classification. In Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing Brussels, Belgium, 31 October–4 November 2018; pp. 3433–3442.
25. Kipf, T.N.; Welling, M. Semi-supervised classification with graph convolutional networks. In Proceedings of the 5th International Conference on Learning Representations, Toulon, France, 24–26 April 2017.
26. Zhang, C.; Li, Q.; Song, D. Aspect-based Sentiment Classification with Aspect specific Graph Convolutional Networks. In Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing, Hong Kong, 3–7 November 2019; pp. 4568–4578.
27. Veličković, P.; Cucurull, G.; Casanova, A.; Romero, A.; Lio, P.; Bengio, Y. Graph attention networks. *arXiv* **2017**, arXiv:1710.10903.
28. Wang, K.; Shen, W.; Yang, Y.; Quan, X.; Wang, R. Relational Graph Attention Network for Aspect-based Sentiment Analysis. In Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics, Online, 5–10 July 2020; pp. 3229–3238.
29. Li, R.; Chen, H.; Feng, F.; Ma, Z.; Wang, X.; Hovy, E. Dual graph convolutional networks for aspect-based sentiment analysis. In Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing, Bangkok, Thailand, 1–6 August 2021; Volume 1: Long Papers, pp. 6319–6329.
30. Bai, X.; Chen, Y.; Zhang, Y. Graph Pre-training for AMR Parsing and Generation. In Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics, Dublin, Ireland, 22–27 May 2022; Volume 1: Long Papers, pp. 6001–6015.
31. Vaswani, A.; Shazeer, N.; Parmar, N.; Uszkoreit, J.; Jones, L.; Gomez, A.N.; Polosukhin, I. Attention is all you need. *Adv. Neural Inf. Process. Syst.* **2017**, *30*, 1–11.
32. Schick, T.; Dwivedi-Yu, J.; Dessì, R.; Lomeli, M.; Hambro, E.; Scialom, T. Toolformer: Language models can teach themselves to use tools. *Adv. Neural Inf. Process. Syst.* **2023**, *36*, 68539–68551.
33. Raffel, C.; Shazeer, N.; Roberts, A.; Lee, K.; Narang, S.; Matena, M.; Liu, P.J. Exploring the limits of transfer learning with a unified text-to-text transformer. *J. Mach. Learn. Res.* **2020**, *21*, 1–67.
34. Yang, Z.; Einolghozati, A.; Inan, H.; Diedrick, K.; Fan, A.; Donmez, P.; Gupta, S. Improving text-to-text pre-trained models for the graph-to-text task. In Proceedings of the 3rd International Workshop on Natural Language Generation from the Semantic Web (WebNLG+), Online, 15–18 December 2020; pp. 107–116.
35. Habu, R.; Ratnaparkhi, R.; Askhedkar, A.; Kulkarni, S. A hybrid extractive-abstractive framework with pre & post-processing techniques to enhance text summarization. In Proceedings of the 13th International Conference on Advanced Computer Information Technologies (ACIT), Wrocław, Poland, 21–23 September 2023; pp. 529–533.
36. Yang, W.; Gu, T.; Sui, R. A faster method for generating Chinese text summaries-combining extractive summarization and abstractive summarization. In Proceedings of the 5th International Conference on Machine Learning and Natural Language Processing, Online, 23–25 December 2022; pp. 54–58.
37. Deng, Y.; Zhang, W.; Pan, S.J.; Bing, L. Bidirectional generative framework for cross-domain aspect-based sentiment analysis. In Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics, Toronto, ON, Canada, 9–14 July 2023; pp. 12272–12285.
38. Zhang, W.; Li, X.; Deng, Y.; Bing, L.; Lam, W. Towards generative aspect-based sentiment analysis. In Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing, Bangkok, Thailand, 1–6 August 2021; pp. 504–510.
39. Radford, A.; Narasimhan, K.; Salimans, T.; Sutskever, I. Improving language understanding by generative pre-training. *Preprint* **2018**.
40. Wang, M.; Wang, M.; Xu, X.; Yang, L.; Cai, D.; Yin, M. Unleashing ChatGPT’s Power: A case study on optimizing information retrieval in flipped classrooms via prompt engineering. *IEEE Trans. Learn. Technol.* **2023**, *17*, 629–641. [[CrossRef](#)]
41. Ouyang, L.; Wu, J.; Jiang, X.; Almeida, D.; Wainwright, C.; Mishkin, P.; Lowe, R. Training language models to follow instructions with human feedback. *Adv. Neural Inf. Process. Syst.* **2022**, *35*, 27730–27744.
42. Brown, T.; Mann, B.; Ryder, N.; Subbiah, M.; Kaplan, J.D.; Dhariwal, P.; Amodei, D. Language models are few-shot learners. *Adv. Neural Inf. Process. Syst.* **2020**, *33*, 1877–1901.
43. Radford, A.; Wu, J.; Child, R.; Luan, D.; Amodei, D.; Sutskever, I. Language models are unsupervised multitask learners. *OpenAI Blog* **2019**, *1*, 9.
44. Wang, Z.; Xie, Q.; Feng, Y.; Ding, Z.; Yang, Z.; Xia, R. Is ChatGPT a good sentiment analyzer? A preliminary study. *arXiv* **2023**, arXiv:2304.04339.
45. Ping, Z.; Sang, G.; Liu, Z.; Zhang, Y. Aspect category sentiment analysis based on prompt-based learning with attention mechanism. *Neurocomputing* **2024**, *565*, 126994. [[CrossRef](#)]
46. Inaba, T.; Kiyomaru, H.; Cheng, F.; Kurohashi, S. MultiTool-CoT: GPT-3 can use multiple external tools with chain of thought prompting. In Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics, Toronto, ON, Canada, 9–14 July 2023; pp. 1522–1532.

47. Fu, Y.; Peng, H.; Sabharwal, A.; Clark, P.; Khot, T. Complexity-based prompting for multi-step reasoning. *arXiv* **2022**, arXiv:2210.00720.
48. Wang, X.; Wei, J.; Schuurmans, D.; Le, Q.V.; Chi, E.H.; Zhou, D. Self-consistency improves chain of thought reasoning in language models. *arXiv* **2022**, arXiv:2203.11171.
49. Li, Y.; Lin, Z.; Zhang, S.; Fu, Q.; Chen, B.; Lou, J.-G.; Chen, W. On the advance of making language models better reasoners. *arXiv* **2022**, arXiv:2206.02336.
50. Pontiki, M.; Galanis, D.; Pavlopoulos, J.; Papageorgiou, H.; Androutsopoulos, I.; Manandhar, S. Semeval-2014 task 4: Aspect-based sentiment analysis. In Proceedings of the 8th International Workshop on Semantic Evaluation (SemEval 2014), Dublin, Ireland, 23–24 August 2014; pp. 27–35.
51. Song, Y.; Wang, J.; Jiang, T.; Liu, Z.; Rao, Y. Attentional encoder network for targeted sentiment classification. *arXiv* **2019**, arXiv:1902.09314. [[CrossRef](#)]
52. Xu, L.; Pang, X.; Wu, J.; Cai, M.; Peng, J. Learn from structural scope: Improving aspect-level sentiment analysis with hybrid graph convolutional networks. *Neurocomputing* **2023**, *518*, 373–383. [[CrossRef](#)]
53. Wu, H.; Huang, C.; Deng, S. Improving aspect-based sentiment analysis with knowledge-aware dependency graph network. *Inf. Fusion* **2023**, *92*, 289–299. [[CrossRef](#)]
54. Li, Z.; Zou, Y.; Zhang, C.; Zhang, Q.; Wei, Z. Learning Implicit Sentiment in Aspect-based Sentiment Analysis with Supervised Contrastive Pre-Training. In Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing, Punta Cana, Dominican Republic, 7–11 November 2021; pp. 246–256.
55. He, B.; Zhao, R.; Tang, D. CABiLSTM-BERT: Aspect-based sentiment analysis model based on deep implicit feature extraction. *Knowl.-Based Syst.* **2025**, *309*, 112782. [[CrossRef](#)]
56. Lawan, A.; Pu, J.; Yunusa, H.; Umar, A.; Lawan, M. Enhancing Long-range Dependency with State Space Model and Kolmogorov-Arnold Networks for Aspect-based Sentiment Analysis. In Proceedings of the 31st International Conference on Computational Linguistics, Abu Dhabi, United Arab Emirates, 19–24 January 2025; pp. 2176–2186.
57. Huang, C.; Li, X.; Zheng, X.; Du, Y.; Chen, X.; Huang, D.; Fan, Y. Hierarchical and position-aware graph convolutional network with external knowledge and prompt learning for aspect-based sentiment analysis. *Expert Syst. Appl.* **2025**, *278*, 127290. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.