

CCL24-Eval 任务6系统报告：人类思维指导下大小模型协同决策的中文修辞识别与理解方法

王雯¹, 汤思怡¹, 于东^{1,*}, 刘鹏远^{1,2}

1.北京语言大学信息科学学院, 北京, 100083

2.国家语言资源监测与研究平面媒体中心, 北京, 100083

wangwenblcu@gmail.com, tangsiyi0805@gmail.com

yudong@blcu.edu.cn, liupengyuan@blcu.edu.cn

摘要

CCL24-Eval任务6提出了一个多层次、细粒度中小学作文修辞识别与理解任务。针对任务特点, 本文提出了人类思维指导下大小模型协同决策的中文修辞识别与理解方法。该方法根据人类在面对修辞识别和理解任务时的处理思路, 将任务顺序重新定义, 并分别选取大小语言模型, 使每个步骤的实现效果均达到局部最优, 以局部最优达到整体任务的最优效果。结果表明, 本文提出的方法能够有效对修辞进行识别与理解, 在三个赛道上相较于Baseline方法分别提升了13.54、4.03、57.11。

关键词: 大语言模型; 修辞识别; 成分抽取

System Report for CCL24-Eval Task6: A Chinese Rhetoric Recognition and Comprehension Approach for Collaborative Decision-Making in Large and Small Models Guided by Human Thinking

Wen Wang¹, Siyi Tang¹, Dong Yu^{1,*}, Pengyuan Liu^{1,2}

1.Faculty of Computer Science, Beijing Language and Culture University, Beijing, 100083

2.National Language Resources Monitoring and Research Center for Print Media, Beijing, 100083

wangwenblcu@gmail.com, tangsiyi0805@gmail.com

yudong@blcu.edu.cn, liupengyuan@blcu.edu.cn

Abstract

CCL24-Eval Task 6 presents a multi-level, fine-grained elementary and middle school composition rhetoric recognition and comprehension task. Aiming at the task characteristics, this paper proposes a Chinese rhetoric recognition and comprehension method with collaborative decision-making between large and small models under the guidance of human thinking. The method redefines the order of the task according to the human thinking when facing the task of rhetoric recognition and comprehension, and selects the large and small language models respectively, so that the realization effect of each step reaches the local optimization, and the local optimization achieves the optimal effect of the overall task. The results show that the method proposed in this paper can effectively recognize and comprehend rhetoric, and improves 13.54, 4.03, and 57.11 compared with the baseline method on the three tracks, respectively.

Keywords: LLM, Rhetorical Recognition, Component Extraction

*为通讯作者

基金项目: 教育部人文社科规划项目 (23YJAZH184); 北京语言大学梧桐创新平台(中央高校基本科研业务费)(21PT04); 北京语言大学研究生创新基金(中央高校基本科研业务费专项资金)项目成果(24YCX114)

©2024 中国计算语言学大会根据《Creative Commons Attribution 4.0 International License》许可出版

1 引言

修辞识别和理解是一个具有挑战性的任务，其目的是在探究和理解文本的修辞结构和语义信息后识别文本使用的修辞类型，并对其中的修辞成分进行抽取。修辞识别和理解任务主要应用于机器阅读理解、作文自动评分等具体问题中，旨在对语文和作文教学提供帮助。在《义务教育语文课程标准》中，对中小学作文训练提出了明确要求，要求学生能够对修辞手法进行恰当地运用(李娜, 2017)。这一要求的出现，使得修辞自动识别和抽取的意义更加显著。然而，之前的研究工作大多集中在对特定修辞大类的识别(武闾闾 et al., 2023)，粒度较粗且泛用性差。同时，部分将修辞理解任务与成分抽取任务等同的研究工作(赵琳玲 et al., 2021; 郭英豪, 2024)，缺少对修辞全面的定义，无法对作文评价给出全面的指导。

CCL24-Eval任务6提出了一个覆盖细粒度、多层次的中小学作文修辞识别与理解任务。对于层次多且复杂的任务，传统的修辞识别方法并不能很好解决问题。随着ChatGPT等大语言模型的出现，科研人员发现使用海量数据训练出的大语言模型可以有效理解文本的语义，并且对逻辑关系具有良好的分析能力(崔亿萍, 2023)。然而，大语言模型在某些特定领域上存在认知缺陷(Zhang et al., 2023)，知识幻象问题也较为突出(Li et al., 2024)。基于此，针对CCL24-Eval任务6，我们提出了一种人类思维指导下大小模型协同决策的中文修辞识别与理解方法。

2 任务定义

CCL24-Eval任务6旨在对修辞进行多层次、细粒度的识别以及对修辞对象及内容的抽取。针对当前修辞识别泛用性差、粒度较粗、定义不明确等缺点，该任务提出了三个赛道：修辞形式类型识别、修辞内容类型识别、修辞成分抽取。三个赛道分别对比喻、比拟、夸张和排比修辞的共性内容进行统一的识别与抽取。每个赛道的评价标准将不同粒度、不同成分的识别效果进行调和，使结果更贴近中小学生作文修辞识别的真实需求。

然而，人类通过自己的认知能力和思维对修辞进行判断的顺序并非按照赛道1、赛道2、赛道3的顺序依次进行。人类拥有分层级的个人语言能力系统，最基础最核心的是认知力，为观念意识到达了某层级的程度；然后是思维力，对应具体事件策划及分析的程度；再上层是逻辑力，即语言要素组成的逻辑程度(张凯, 2024)。廖巧云的研究为语义修辞的生成(廖巧云, 2018b)和理解机制(廖巧云, 2018a)构建了不同框架，同样表明了修辞需要读者调动相应的语言能力来理解修辞要素之间的关系，进而才能理解语义。因此，人类面对修辞识别等任务时的思维过程也基本符合上述语言能力系统中从低到高的层级顺序，即先对修辞手法进行粗粒度的分类，然后在对修辞成分的抽取和分析的同时，对修辞成分的关系、性质等方面有一个具体的认识。根据人类的思维过程，我们重新定义CCL24-Eval任务6三个赛道中的任务顺序，如图1所示，蓝色代表处理顺序1，绿色代表处理顺序2，黄色代表处理顺序3。



图 1: 重定义后的任务

3 数据扩充

CCL24-Eval任务6发布的数据集⁰中包含375条比喻句、148条比拟句、103条夸张句、107条排比句和125条无修辞句，为提高评测数据集的多样性和规模，提升模型的训练效果和泛化能力，我们利用Li et al. (2022)公开的CLGC中文文采评估语料库¹，并爬取了部分网络数据，采用大语言模型自动标注和人工校对的方法，扩充了基础评测数据集，扩充过程如图2所示。最终形成了比喻、比拟、夸张、排比四种修辞各1000条语料的扩充评测数据集，数据集情况如附录表5所示。

⁰<https://github.com/cubenlp/CERRU>

¹<https://github.com/bicunlp/CLGC/>

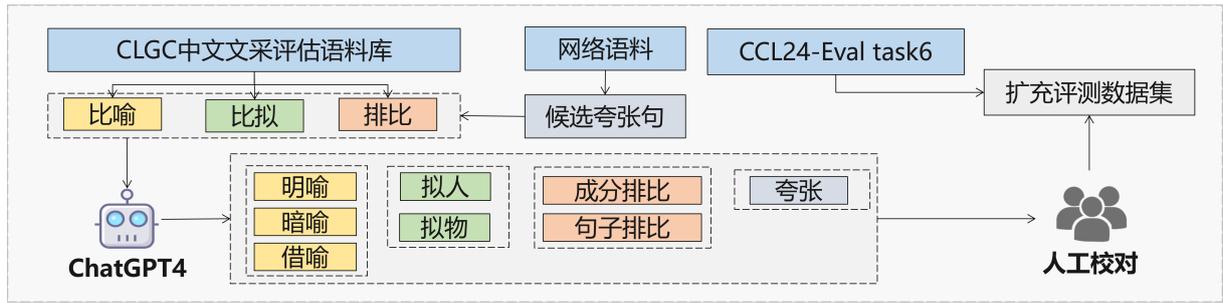


图 2: 评测数据扩充过程

在图2中可以看到，我们从CLGC中文文采评估语料库中抽取比喻、比拟、排比标签的句子；并以“夸张句”为关键词爬取夸张句网站²³⁴⁵，再从小说阅读网⁶爬取1000篇中文小说，抽取文本中带有“夸张”关键词的句子及该句子的前后两句作为候选夸张句。选取ChatGPT4作为自动标注模型，输入抽取的比喻、比拟、夸张、排比句以及候选夸张句进行标注。如果句子具有比喻修辞，则按照成分显隐性细分为明喻、暗喻、借喻；如果句子具有比拟修辞，则按照喻体性质细分为拟人和拟物；如果句子具有排比修辞，则按照排比项成分细分为成分排比和句子排比；如果句子具有夸张修辞，则纳入夸张句语料。ChatGPT4标注完毕后，招募三名语言学硕士研究生进行人工校对，提高评测数据质量。

4 方法

4.1 方法概述

修辞手法的识别和成分抽取任务的多个层次之间有着极高的关联性，即特定的细粒度分类只存在于相应的粗粒度分类下，上一层级的识别准确程度极大程度影响了下一层级的识别与抽取的效果。因此，我们根据该任务层次关联的特点提出了如图3的处理流程，在流程的每个节点上都试图达到局部最优的效果，并以此提高任务整体的性能。此外，任务的处理顺序可能会对效果产生影响，我们调整了任务顺序进行探究，具体情况如A.11所示。

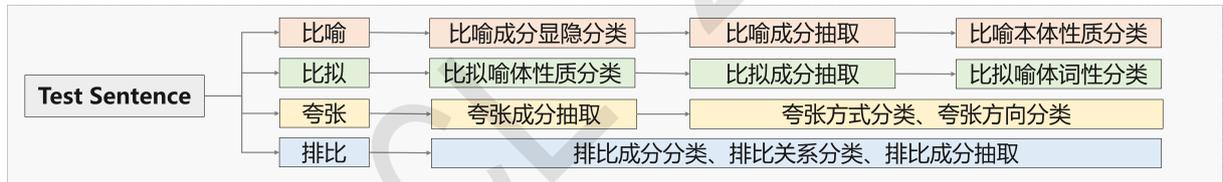


图 3: 方法概况

4.2 修辞的粗粒度识别

我们使用章节3中对CCL24-Eval任务6数据集扩充后带有比喻、比拟、夸张、排比四种修辞粗粒度标签的数据集训练了BERT模型。然而，一个测试句子可能会使用一种或多种修辞手法。因此，我们将BERT输出的概率作为Softmax输入到ChatGPT4中，指导大语言模型对句子中包含的修辞手法作出最终的判断。对于每个测试句子，通过上述BERT-LLMSoftmax方法预测后会被分入到其相应的粗粒度修辞分类中，以待进行后续的识别与抽取任务。BERT-LLMSoftmax方法使用的Prompt模板详见附录A.1。

4.3 比喻的识别与理解

比喻修辞的识别与理解流程如图4所示。首先，利用扩充评测数据集训练BERT模型，如果测试句子在通过LLM Softmax后被标注为比喻句，则再利用扩充评测数据集中具有明喻、

²<https://www.pinyudu.com/jiaju/2341>

³<https://www.baihuawen.cn/yuedu/zhaichao/20125.html>

⁴<https://www.ruiwen.com/word/duanyidiandekuazhangju.html>

⁵<https://wenku.baidu.com/view/264da81551ea551810a6f524ccbff121dd36c53e.html>

⁶<https://www.readnove1.com/>

暗喻、借喻标签的语料训练BERT模型，将测试句子分为明喻、暗喻、借喻三类。接着，构建比喻-形式样本库，样本库中包括明喻、暗喻、借喻三种形式的句子各10条，每条均标注了本体、喻词、喻体等相应修辞成分。采用Few-shot提示策略，设计比喻修辞成分抽取Prompt模板并随机采样，利用ChatGPT4模型，以抽取出比喻的不同修辞成分。抽取明喻类别修辞成分的Prompt模板如附录A.2所示，其他细粒度分类的Prompt模板与之类似。然后，构建比喻-本体分类样本库，其中包含实在物、动作、抽象概念三个类别的样本各10条。采用Few-shot策略，设计Prompt模板并随机采样，利用ChatGPT4对本体内容的分类进行判断。针对比喻句本体内容进行分类的Prompt模板，具体内容如附录A.3所示。

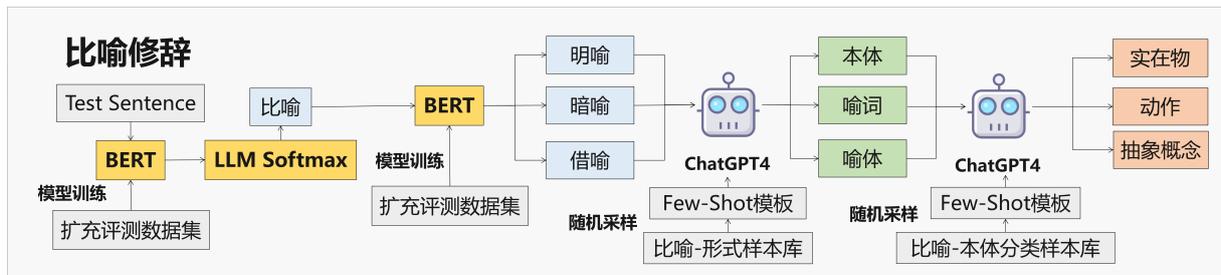


图 4: 比喻修辞处理流程

4.4 比拟的识别与理解

图5展示了比拟修辞的识别与理解流程。首先，利用扩充评测数据集训练BERT模型，如果测试句子在通过LLM Softmax后被标注为比拟句，则再利用扩充评测数据集中具有拟人、拟物标签的语料训练BERT模型，将测试句子分为拟人、拟物两类。接着，构造比拟-成分样本库，其中拟人、拟物两种喻体性质的句子各10条，每个样本均标注了比拟对象和比拟内容。在比拟修辞成分抽取步骤中，采用Few-shot提示策略，设计Prompt模板并随机采样，将该Prompt模板注入ChatGPT4进行成分抽取。附录A.4展示了拟人句子进行成分抽取的Prompt示例，拟物分类的Prompt与之类似。同样，对于比拟内容和其对应的词性分类也构建比拟-喻体词性样本库，包含了标有名词、动词、形容词、副词四种分类的比拟内容示例各10条，使用随机采样的Few-shot策略的Prompt模板，在大语言模型的辅助下将比拟内容分类，对比拟内容进行分类的Prompt模板则如附录A.5所示。

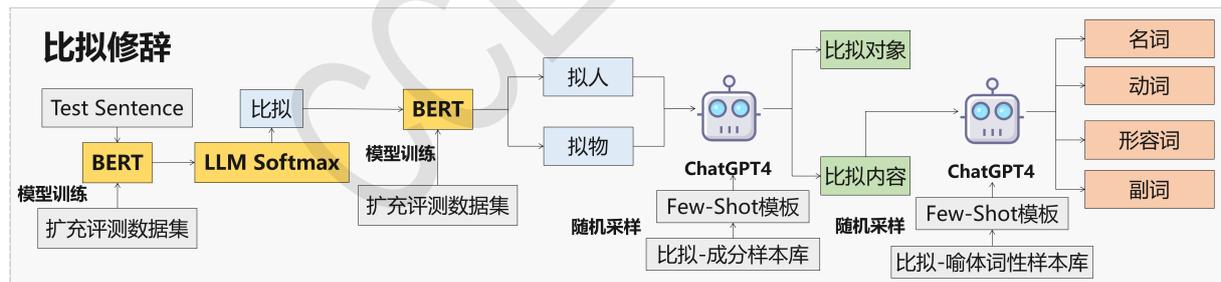


图 5: 比拟修辞处理流程

4.5 夸张的识别与理解

夸张修辞识别的任务特点与其他修辞存在差异，根据人类思维指导，先进行夸张修辞的成分抽取，再根据其成分的性质和形式进行对应的分类，故夸张的识别与理解流程与其他修辞手法有所不同。基于夸张修辞分析的客观情况，夸张的识别与理解任务均使用大语言模型完成。如图6所示，夸张修辞的样本库共有3个，夸张-成分样本库中包含标注了夸张对象和夸张内容的夸张句子样本共10条，夸张-形式样本库中包含直接夸张、间接夸张、融合夸张三种夸张形式的样本各10条，夸张-方向样本库中则包含扩大夸张、缩小夸张、超前夸张三种夸张方向的样本各10条，其中后两个样本库中同时标注了句子中的夸张对象和夸张内容。在夸张成分抽取、夸张形式识别、夸张方向识别三个任务中，分别从其对应的样本库中随机采样出样本，作

为Few-shot策略中的示例样本，设计Prompt模板，使用ChatGPT4完成各个阶段的任务。三个任务所使用的Prompt模板，具体如附录A.6、A.7和A.8所示。

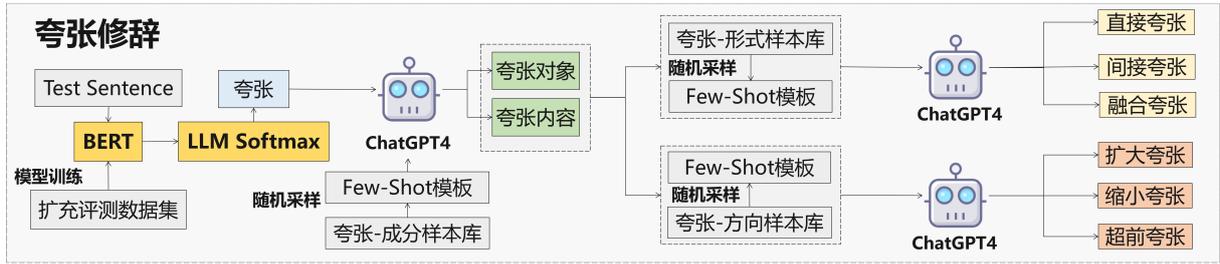


图 6: 夸张修辞处理流程

4.6 排比的识别与理解

排比修辞的识别与理解流程如图7所示。首先，利用扩充评测数据集训练BERT模型，如果测试句子在通过粗粒度识别后被标注为排比句，则可以同时进行排比成分分类、排比关系分类、排比成分抽取。利用标有成分排比、句子排比标签的扩充评测数据集语料训练BERT模型进行排比成分分类。构建排比-成分和排比-关系样本库，排比-成分样本库中包含10条标有排比项的排比句，排比-关系样本库中则包含并列、承接、递进三种排比项关系的句子各10条。采用Few-shot策略，利用ChatGPT4对排比项成分进行抽取和排比句内的逻辑关系判断。具体步骤如图7所示。其中针对排比成分抽取、排比关系判断的Prompt模板可见附录A.9、A.10。

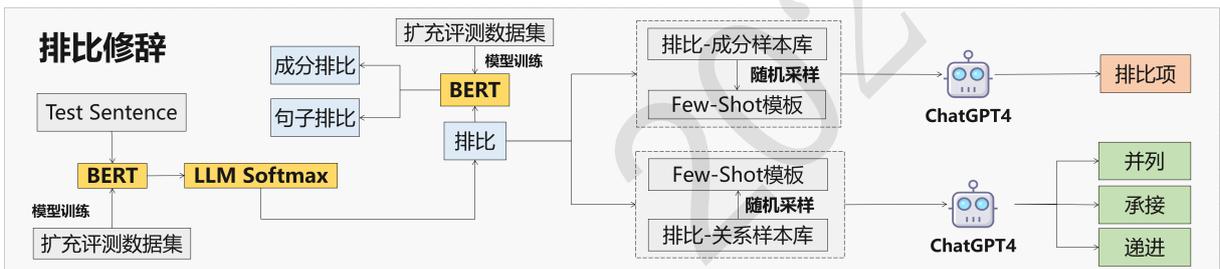


图 7: 排比修辞处理流程

5 实验与结果

5.1 修辞粗粒度识别

在修辞的粗粒度（比喻、比拟、夸张、排比）识别中，对比使用CCL24-Eval任务6训练数据训练BERT模型进行分类、使用大语言模型直接进行分类、使用扩充评测数据集训练BERT模型进行分类三种方法，其中使用扩充评测数据集训练的BERT模型识别效果最好，我们使用该模型输出的概率作为Softmax对ChatGPT4进行指导，以判定句子最终的修辞标签，识别效果较直接分类有显著提高。不同方法的具体识别效果如表1所示。

任务	方法	训练数据	测试数据	P	R	F1
修辞粗粒度识别	BERT	training.json	validation.json	0.37	0.38	0.38
	LLM	-	validation.json	0.50	0.45	0.42
	BERT	扩充评测数据集	validation.json	0.62	0.65	0.63
	BERT.LLMSoftmax	扩充评测数据集	validation.json	0.70	0.73	0.70

表 1: 不同方法在修辞粗粒度识别上的效果

5.2 修辞细粒度识别

每个测试句子经过粗粒度的修辞识别过程后，会被分入到比喻、比拟、夸张、排比四种修辞分类中。在此之后，对比喻、比拟、排比三种修辞类型，进行修辞的细粒度识别，即

对于比喻修辞，按照修辞成分的显隐性分为明喻、暗喻、借喻；对于比拟修辞，按照喻体的性质分为拟人、拟物；对于排比修辞，按照排比项的成分分为成分排比和句子排比。针对每种修辞的细粒度分类识别，本文分别进行了实验，对比了扩充评测数据集训练的BERT模型与ChatGPT4的识别效果。如表2所示，使用BERT的F1值均优于大语言模型。

任务	粗粒度分类	细粒度分类	LLM(P/R/F1)	BERT(P/R/F1)
修辞成分抽取	比喻	明喻/暗喻/借喻	0.31/0.30/0.30	0.74/0.72/0.72
	比拟	拟人/拟物	0.85/0.74/0.78	0.80/ 0.88/0.83
	排比	成分排比/句子排比	0.64/0.56/0.51	0.57/ 0.57/0.57

表 2: 不同方法在修辞细粒度识别上的效果

5.3 修辞成分抽取

依据章节4中介绍的方法，本文使用ChatGPT4进行四种修辞的成分抽取。对于比喻修辞，抽取本体、喻词、喻体；对于比拟修辞，抽取比拟对象、比拟内容；对于夸张修辞，抽取夸张对象、夸张内容；对于排比修辞，抽取排比项。在上述步骤中，我们均对比了Few-shot和思维链(Chain-of-thought, CoT)两种不同策略的抽取效果。具体效果可见表3，使用Few-shot策略的效果均优于CoT策略。

任务	粗粒度分类	连接词	描写对象	描写内容	CoT(F1)	Few-shot(F1)
修辞成分抽取	比喻	喻词	本体	喻体	0.39	0.73
	比拟	-	比拟对象	比拟内容	0.18	0.53
	夸张	-	夸张对象	夸张内容	0.27	0.58
	排比	排比项	-	-	0.16	0.96

表 3: 不同策略在修辞内容抽取上的效果

5.4 修辞成分分类

在抽取出修辞成分之后，进行修辞成分的分类。对于比喻本体，分为实在物、动作、抽象概念；对于比拟成分，分为名词、动词、形容词、副词；对于夸张形式，分为直接夸张、间接夸张、融合夸张；对于夸张方向，分为扩大夸张、缩小夸张、超前夸张；对于排比项之间的关系，分为并列、承接、递进。在修辞成分分类的步骤中，采用Few-shot和CoT策略分别设计相应的Prompt。对比两种策略，可以发现，Few-shot有更高的F1值，具体情况可见表4。

粗粒度分类	成分	成分分类	CoT(P/R/F1)	Few-shot(P/R/F1)
比喻	本体	实在物/动作/抽象概念	0.60/0.58/0.56	0.63/0.67/0.62
比拟	比拟成分	动词/名词/形容词/副词	0.71/0.83/0.69	0.75/0.92/0.78
夸张	形式	直接/间接/融合	0.28/0.29/0.29	0.57/0.54/0.54
夸张	方向	扩大/缩小/超前	0.98/0.75/0.82	0.83/ 0.97/0.89
排比	关系	并列/承接/递进	0.39/0.44/0.25	0.61/0.89/0.66

表 4: 不同方法在修辞成分分类上的效果

6 结论

针对CCL24-Eval任务6中小学作文修辞识别与理解任务，本文提出了人类思维指导下大小模型协同决策的中文修辞识别与理解方法。该方法通过人类在开展修辞识别和理解任务时的思考思路，重新定义了任务顺序。在完成的过程中使用大小模型协同决策的策略，充分利用了大小模型各自的优势，以每个步骤的局部最优来达到整体最优的效果。在本次评测任务中，本文的方法在赛道一、赛道二、赛道三上的分数分别为59.20、60.92、77.96，相比于Baseline结果高出了13.54、4.03、57.11。未来，我们将进一步优化模型的架构，以提升协同效果和模型的综合性能，并将其他类型修辞手法纳入自动识别与理解范围，适应更多样化的应用场景。

参考文献

- Yi Li, Dong Yu, and Pengyuan Liu. 2022. Clgc: A corpus for chinese literary grace evaluation. In *Language Resources and Evaluation Conference*.
- Yangning Li, Shirong Ma, Xiaobin Wang, Shen Huang, Chengyue Jiang, Hai-Tao Zheng, Pengjun Xie, Fei Huang, and Yong Jiang. 2024. Ecomgpt: Instruction-tuning large language models with chain-of-task tasks for e-commerce. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 38, pages 18582–18590.
- Yue Zhang, Yafu Li, Leyang Cui, Deng Cai, Lemao Liu, Tingchen Fu, Xinting Huang, Enbo Zhao, Yu Zhang, Yulong Chen, et al. 2023. Siren’s song in the ai ocean: a survey on hallucination in large language models. *arXiv preprint arXiv:2309.01219*.
- 崔亿萍. 2023. 大模型在自然语言处理的应用和研究. 中国科技期刊数据库工业A, pages 57–61.
- 廖巧云. 2018a. 语义修辞的生成机制研究. 外语教学, 39(03):10–15.
- 廖巧云. 2018b. 语义修辞的识解机制. 现代外语, 41(01):1–11+145.
- 张凯. 2024. 中文修辞能力核心要素计算与智能评价标准构建——基于中小学语文教材修辞语料. 昆明学院学报, 46(02):23–34.
- 李娜. 2017. 小学高年级语文作文教学的评价研究. 小学生(教学实践), (05):1.
- 武闯闯, 宋子尧, 韩旭, 程苗苗, 巩捷甫, 王士进, and 宋巍. 2023. 学生议论文中的比喻论证作用分析. 中文信息学报, 37(10):158–166.
- 赵琳玲, 王素格, 陈鑫, 王典, and 张兆滨. 2021. 基于词性特征的明喻识别及要素抽取方法. 中文信息学报, 35(01):81–87+95.
- 郭英豪. 2024. 面向高考鉴赏类问题的比喻与夸张识别及要素抽取技术研究. Master’s thesis, 山西大学.

A 附录

类别	数量	平均句长	示例
比喻	1000	38.02	他沉默着，像一个木偶似的站立在林青史的面前。
比拟	1000	29.18	这会儿，你看，小草含着泪珠儿，在泣在愁。
夸张	1000	30.22	一个芝麻大的官放个屁，到了你们这儿也是8级地震！
排比	1000	44.50	夏天的夜晚，是那么的宁静，那么的美丽，那么凉爽。
无修辞	1000	27.01	他是一名善良而又喜欢帮助同学的少先队员。

表 5: 扩充评测数据集情况

A.1 BERT LLMSoftmax Prompt模板

##任务描述##

请识别测试句子所具有的修辞，其中可能包含比喻、比拟、夸张、排比和无修辞中的一种或多种修辞。

-比喻：抓住和利用不同事物的相似点，用另一个事物来描绘所要表现的事物。特征为1.一般由三部分组成，即：本体、喻词、喻体。2.本体和喻体必须是性质不同的两类事物。3.在比喻句中，喻体必须出现。4.本体和喻体之间必须有相似点。

-比拟：把物当作人写，或把人当作物写，或把甲物当作乙物来写。特征是1.把物“人化”，或把人“物化”，或“甲物乙物化”；2.存在本体和拟体，其中本体必须出现，而拟体一般不出现。

-夸张：故意言过其实，通过对客观的人、事物做扩大或缩小或超前的描述。特征为：1.故意把客观事物说得“小、少、低、弱、浅”；2.故意把事物说得“大、多、高、深、强”；3.把后出现的说成先出现，把先出现的说成后出现。

-排比：把三个或三个以上结构相同或相似、语气一致、意思密切关联的句子或句子成分排列起来，使内容和语势增强。特征为：1.排比必须三个或三个以上结构相同或相似、意义相关、语气一致的词组或句子排列成串，形成一个整体。2.排比结构、长度等大致相似、没有类似的对称效果，字数不做严格要求。

-无修辞：即不具有任何修辞手法。

##参考分类##

按照BERT模型分类，测试句子具有的修辞标签概率为：{预测1}>{预测2}>{预测3}>{预测4}>{预测5}，该预测概率仅作参考，并非真实确定标签。

##测试句子##

{test sentence}

A.2 比喻修辞抽取明喻形式成分Prompt模板

##任务描述##

测试句子是比喻中的明喻形式，请从中抽取出描写对象（本体）、喻词和描写内容（喻体）。喻词有像、好像、好比、好似、恰似、如、有如、犹如、仿佛、如同。

##参考分类##

【示例1】

Sentence: 其中赛龙舟最有特色，那争先恐后的激烈场面就像屈原强烈的爱国精神一样，几千年来一直激励着我们。

Tenor: 争先恐后的激烈场面

Conjunction: 像

Vehicle: 屈原强烈的爱国精神

【示例2】

Sentence: 回家后我们围绕着她聚成环，但假期结束后，我们又如同蒲公英一般随风而去，飞向远方。

Tenor: 我们

Conjunction: 如同

Vehicle: 蒲公英

【示例3】

Sentence: 一片片枯叶好似仙女下凡，飘然而至。

Tenor: 一片片枯叶

Conjunction: 好似

Vehicle: 仙女

##测试句子##

请按照以上结构，抽取以下测试句子，如果有多个喻词、描写对象（本体）和描写内容（喻体）需要一并写出：

Sentence: {test sentence}

A.3 比喻修辞本体分类Prompt模板

##任务描述##

测试句子是比喻句，本体即被比喻的事物，请判定该句子的本体属于实在物、动作、抽象概念。

-实在物：本体是可见、可触、可想的实在物体；

-动作：本体是某种动作、行为或事件；

-抽象概念：本体是抽象概念，如爱、时间、勇气等。

##参考分类##

【示例1】

Sentence: 他是一个男生，大大的眼睛仿若桃花竞相开放，明媚缠绵。

tenor: 大大的眼睛

Content: 实在物

解释：大大的眼睛是可见的实在物。

【示例2】

Sentence: 人生是一颗多味的“秀逗”糖，会越来越甜！

tenor: 人生

Content: 抽象概念

解释：人生是抽象概念。

【示例3】

Sentence: 雨水流在它脸上流过，像是少女在哭泣。

tenor: 雨水流在它脸上流过

Content: 动作

解释：雨水流在它脸上流过是动作。

【示例4】

Sentence: 我们之间仿佛隔了厚厚的屏障。

tenor: 没有明确指出

Content: 抽象概念

解释：句子的本体没有直接说明，需要进一步理解语义，本体应该是两个人之间的情感或沟通上的隔阂。

##测试句子##

Sentence: {test sentence}

A.4 比拟修辞抽取拟人形式成分Prompt模板

##任务描述##

测试句子是比拟中的拟人形式，请从中抽取出修辞成分，包括描写对象（比拟对象）和描写内容（比拟内容）。

##参考示例##

【示例1】

Sentence: 当暴风雨来临时，任凭风吹雨打，把小草打得东倒西歪，它都没有屈服，依然牢固地扎根在泥土里，顽强地生长着。

Tenor: 小草

Vehicle: 屈服

Tenor: 小草

Vehicle: 顽强地

【示例2】

Sentence: 天上已有七八个星在闪烁, 一切显得那么神秘, 渐渐的, 月亮不再羞答答的了。

Tenor: 月亮

Vehicle: 羞答答的

【示例3】

Sentence: 人们只看见电火的愤怒和山洪的疯狂。

Tenor: 电火

Vehicle: 愤怒

Tenor: 山洪

Vehicle: 疯狂

##测试句子##

请按照以上结构, 抽取以下测试句子, 如果有多个描写对象(比拟对象)和描写内容(比拟内容)需要一并写出:

Sentence: {test sentence}

A.5 比拟修辞比拟成分词性标注Prompt模板

##任务描述##

测试句子是比拟句, 请判断标志比拟成分(vehicle)的词性为名词、动词、形容词和副词中的哪一种。

##参考示例##

【示例1】

Sentence: 我静静地哭着, 寒风无情地吹着你的绒毛, 也让我感到了一丝寒意。

tenor: 寒风

vehicle: 无情地吹着你的绒毛

词性: 副词

【示例2】

Sentence: 你骄傲自满, 尾巴翘上天了。

tenor: 你

vehicle: 尾巴翘上天了

词性: 名词

【示例3】

Sentence: 我们家原来有一盆仙人掌, 长得很丑陋。

tenor: 仙人掌

vehicle: 丑陋

词性: 形容词

【示例4】

Sentence: 她似乎衣袋里全装着天真, 一掏出来就可以用。

tenor: 天真

vehicle: 一掏出来就可以用

词性: 动词

注意: 只需回答比拟成分词性分类, 不需要进行解释。

##测试句子##

Sentence: {test sentence}

Tenor: {tenor}

Vehicle: {vehicle}

A.6 夸张修辞抽取成分Prompt模板

##任务描述##

测试句子是夸张句, 请从中抽取出修辞成分, 包括描写对象(夸张对象)和描写内容(夸张内

容)。

##参考示例##

【示例1】

Sentence: 可怕的雷鸣震裂了天空。

Tenor: 雷鸣

Vehicle: 震裂了天空

【示例2】

Sentence: 常德盛恨不得把一天掰成两天来过。

Tenor: 常德盛

Vehicle: 恨不得把一天掰成两天来过

【示例3】

Sentence: 听到这句话,我仿佛感到血管里的血一下子凝固了。

Tenor: 血管里的血

Vehicle: 一下子凝固了

##测试句子##

请按照以上结构,抽取以下测试句子,如果有多个描写对象(夸张对象)和描写内容(夸张内容)需要一并写出:

Sentence: {test sentence}

A.7 夸张句判断夸张方向Prompt模板

##任务描述##

测试句子是夸张句,请判断句子夸张方向属于扩大夸张、缩小夸张还是超前夸张。

-扩大夸张:故意把事物说得“大、多、高、深、强”,对事物形象、性质、特征、作用、程度等加以扩大;

-缩小夸张:故意把客观事物说得“小、少、低、弱、浅”,对事物形象、性质、特征、作用、程度等加以缩小;

-超前夸张:把后出现的说成先出现,把先出现的说成后出现。

##参考示例##

【示例1】

Sentence: 可怕的雷鸣震裂了天空。

Tenor: 雷鸣

Vehicle: 震裂了天空

Content: 扩大夸张

【示例2】

Sentence: 这个足球场只有巴掌那么大。

Tenor: 这个足球场

Vehicle: 只有巴掌那么大

Content: 缩小夸张

【示例3】

Sentence: 还没回家就闻到香味了。

Tenor: 还没回家

Vehicle: 就闻到香味了

Content: 超前夸张

##测试句子##

Sentence: {test sentence}

Tenor: {tenor}

Vehicle: {vehicle}

A.8 夸张句判断夸张形式Prompt模板

##任务描述##

测试句子是夸张句,请判断句子夸张形式属于直接夸张、间接夸张还是融合夸张。

- 直接夸张: 直接对某物进行夸张
- 间接夸张: 夸大另一样东西来夸大某事物
- 融合夸张: 借助其他修辞进行夸张

##参考示例##

【示例1】

Sentence: 可怕的雷鸣震裂了天空。

Tenor: 雷鸣

Vehicle: 震裂了天空

Form: 直接夸张

解释: 在这句话中, “雷鸣”被直接夸张地描述为能“震裂天空”, 以强调雷声的巨大和恐怖。

【示例2】

Sentence: 常德盛恨不得把一天掰成两天来过。

Tenor: 常德盛

Vehicle: 恨不得把一天掰成两天来过

Form: 间接夸张

解释: 在这句话中, 常德盛因为某种原因(如工作繁忙或时间不够用)而希望把一天掰成两天来过。这种夸张的表达间接地表现出他对时间的极度需求或对事情的重视。

【示例3】

Sentence: 不愿意去上学的孩子像蜗牛爬行一样去学校。

Tenor: 不愿意去上学的孩子

Vehicle: 像蜗牛爬行一样去学校

Form: 融合夸张

解释: 在这句话中, “不愿意去上学的孩子”被比作“蜗牛爬行一样去学校”, 这是将比喻和夸张结合起来使用。通过比喻, 形象地夸张了孩子走得非常慢, 生动地表达了孩子对上学的抗拒和拖延。

注意: 只需回答夸张形式, 不需要进行解释原因。

##测试句子##

Sentence: {test sentence}

Tenor: {tenor}

Vehicle: {vehicle}

A.9 排比修辞抽取排比项成分Prompt模板

##任务描述##

测试句子是排比句, 请抽取该句子的排比项。

-排比项: 多次重复出现的那些共同词、句式等。

##参考示例##

【示例1】

Sentence: 微风吹过, 花瓣随风飘落, 有的落在头上, 有的落在肩上, 有的落在地上。

Conjunction: 有的落在

【示例2】

Sentence: 勿忘昨天的苦难辉煌, 无愧今天的使命担当, 不负明天的伟大梦想。

Conjunction: 天的

【示例3】

Sentence: 劳动, 让我们的双手更灵活, 让我们的身体更健康, 让我们的生活更美好。

Conjunction: 让我们的

##测试句子##

Sentence: {test sentence}

A.10 排比修辞判断排比项关系Prompt模板

##任务描述##

测试句子是排比句, 请判断句子的排比项关系是并列、承接还是递进。

- 并列：排比项顺序改变不影响语义通顺；
- 承接：排比项之间有先后逻辑顺序，如时间、程度、发展状况等，不能改变顺序；
- 递进：各排比项表达的含义情感等层层递进，不能改变顺序。

##参考示例##

【示例1】

Sentence: 微风吹过，花瓣随风飘落，有的落在头上，有的落在肩上，有的落在地上。

Content: 并列

【示例2】

Sentence: 勿忘昨天的苦难辉煌，无愧今天的使命担当，不负明天的伟大梦想。

Content: 承接

【示例3】

Sentence: 劳动，让我们的双手更灵活，让我们的身体更健康，让我们的生活更美好。

Content: 递进

##测试句子##

Sentence: {test sentence}

A.11 任务顺序对于效果的影响

本文提出的人类思维指导下大小模型协同决策的中文修辞识别与理解方法在解决任务过程中，将任务拆解并重新排序。然而，任务顺序的不同也可能对最终的效果产生影响。因此，我们在不同修辞大类下，探究了其他顺序的效果。

对于比喻修辞，本文方法的处理顺序为比喻成分显隐分类、比喻成分抽取、比喻本体性质分类；探究顺序为比喻成分显隐分类、比喻本体性质分类、比喻成分抽取。比喻本体性质分类任务的Prompt依赖于比喻成分抽取任务结果，调换任务顺序后仅对比喻本体分类任务产生影响。对于比拟修辞，本文方法的处理顺序为比拟喻体性质分类、比拟成分抽取、比拟喻体词性分类；探究顺序为比拟喻体性质分类、比拟喻体词性分类、比拟成分抽取。比拟成分词性分类任务的Prompt依赖于比拟成分抽取任务结果，调换任务顺序后仅对比拟成分词性分类任务产生影响。对于夸张修辞，本文方法的处理顺序为在夸张成分抽取后，同时对夸张方式、夸张方向进行分类；探究不依赖夸张成分抽取结果，直接对夸张方式和夸张方向进行分类，调换任务顺序后仅对夸张方式分类和夸张方向分类任务产生影响。对于排比修辞，三个任务之间没有依赖关系，因此不需要调整任务顺序进行探究。修辞的任务顺序对于效果的具体影响如表6所示，可以看出，本文方法中的任务顺序具有更高的P、R、F1值，验证了本文提出方法的有效性。

粗粒度分类	成分	成分分类	探究顺序	本文方法
			P/R/F1	P/R/F1
比喻	本体	实在物/动作/抽象概念	0.27/0.27/0.25	0.63/0.67/0.62
比拟	比拟成分	动词/名词/形容词/副词	0.34/0.50/0.27	0.75/0.92/0.78
夸张	形式	直接/间接/融合	0.30/0.31/0.30	0.57/0.54/0.54
夸张	方向	扩大/缩小/超前	0.65/0.47/0.53	0.83/0.97/0.89

表 6: 调整任务顺序探究结果