**同行评审意见属性级情感标注规范**

**1 标注方式**

首先将评审意见文本进行分句处理，并将其以句为单位导入Excel中以便于后续标注。利用评审属性词与情感词之间的词间关系（参考DP算法原理）等规则抽取出包含有对属性词进行评价的评审句，将（情感词，词间关系，评审属性词，评审句）按行保存于excel文件中，然后结合评审句语义，人工对每句评审句中针对属性的情感倾向进行判断，标注情感极性。

标签分为四类，分别为正面情感、负面情感、中性情感、模糊情感。表1展示了具体的标注方式。

表1 同行评审属性标注方式

|  |  |  |
| --- | --- | --- |
| **属性情感分类** | **定义** | **标注方式** |
| 同行评审负面情感属性 | 针对该属性包含的情感为贬义 | -1 |
| 同行评审中性情感属性 | 针对该属性包含的情感为中性 | 0 |
| 同行评审正面情感属性 | 针对该属性包含的情感为褒义 | 1 |
| 同行评审模糊情感属性 | 无法确定该属性词的情感或确定该属性词在该评审句中不是属性 | 2 |

**2 标注流程**

在标注过程中，由三位标注人员对评审意见文本中的属性情感进行标注，标注人员均同行评审过程。在标注框架中，每位标注人员须针对评审句中存在的属性来标注相应的情感，为每个属性添加其相对应的情感极性标签，标签分为以下四类——‘1’（正面的）、‘-1’（负面的）、‘0’（中立的）和‘2’（模糊的）。首先，我们使用stanfordnlp工具对评审句进行句法分析，利用情感词及评审属性词间的依存关系，抽取出评审句中所包含的属性词，句法依存关系包括subj，conj等。然后，将标注语料三等分，每一份标注语料由两名标注人员分别独立完成评审句中属性的相应情感标注，对两人的标注结果进行标注一致性检验；对标注不一致的数据，由第三人对两人的标注数据进行最终标注。最后，对标注好的数据进行统计。

**3 判断标准**

**3.1 属性的中性情感判断标准**

1、对论文工作的概括描述，其中修饰属性的词为中性

Eg: This manuscript describes a study on the development of a conditional Por-KO, liver humanized chimeric mouse **model** and its **utility** for the assessment of drug metabolism.

2、对属性的修饰词词性为中性，如other、many、some、proposed等

Eg: The authors mention that mutations in pik3ca and tp53 are mutually exclusive in breast cancer; but did not indicate if the analysis showed significance in other/all cancer **types**.

3、句中仅提及但未作评价（即仅作为句子组成部分出现）的属性词情感为中性

Eg: The paper presents a novel method for answering “How many …?” questions in the VQA **datasets**.

**3.2 属性的负面情感判断标准**

1、修饰属性的情感词为贬义词

Eg: The **re-referencing** of previous figures out of order is confusing.

2、属性词本身带有负面情感

Eg: There is a **typo** in expression 2 in fig s1.

3、包含评审人对论文某方面存在的疑问

Eg: Do you mean the **average** of all hindcast available at that year?

4、评审人对某属性表示自己的心情或情绪时用到贬义词

Eg: I'm not sure what these discrepancy values **mean**.

5、包含评审人对论文某方面的期望，认为怎么改进能更好；或评审人认为作者

未解决其所提出的哪些问题，如果解决哪些问题将会更好，则该句中涉及的属性为负面情感

Eg: The inset **plot** would benefit greatly from more explanation.

Eg: These are interesting areas for future research, so i hope you can address my concerns above so that **discussion** on these important issues can progress.

6、对文章的某方面提出建议或要求，比较常见的如句中包含need, must等词

Eg: There are several **aspects** of the current manuscript that were distracting and

need to be addressed.

7、属性所处的子句前有转折词，且前句为正面情感

Eg: The **model** is well described but the authors should pay more effort to walk

through the reader of the dynamic development how the **structure** evolves towards the presented **outcome**.

**3.3 属性的正面情感判断标准**

1、修饰属性的情感词是褒义词

Eg：I agree with the authors that this is a very important new set of vapour **observations** that allows new **insights** into key isotopic processes.

2、属性词本身带有褒义情感

Eg：There are several **findings** and **contributions** reported in this paper

3、评审人对某属性表示自己的心情或情绪时用到褒义词

Eg：I very much like the **data**, the **observations** and i am sure that the authors are right that microbial processes result in the effects they observe.

4、句中对论文某方面的描述包含褒义情感，如性能水平提升等

Eg：They achieved interesting **performance** levels (projected retention time, 1 y; programming voltage range, within +/- 10 v; max strain, 2.8%; etc).

5、句中包含不同的情感信息，整体是正面情感，即属性所处的子句前有转折词，且前句为负面情感

Eg：My personal sense is that this is still mostly a strawman argument, but i appreciate the **effort** that the authors have done to address it.

6、有些情感词虽然未与属性词共同出现，但是其也对某个属性进行了情感表达（句法关系xcomp），比较常见的有

clearly：‘clarity’ +1；

well described 描述的好：‘description’ +1

well written 写得好 writing+1

easy read ,easy follow, easy understand容易读：‘readability’+1；

**3.4 情态动词判断标准**

① Can / could: 表达完成某件事的能力，正面情感

② Can / could +比较级：负面情感

③ Could / would +正面情感词：负面情感

④ Could / would +比较级：负面情感

⑤ Should / shall +比较级：负面情感

⑥ Should / shall +否定词：负面情感

⑦ Should + do / suggest / please / encourage / recommend: 负面情感

⑧ Need+否定词：正面情感；Need+比较级：负面情感；Need：负面情感

⑨ Must+否定词：正面情感；Must+比较级：负面情感

⑩ Have to / had better / better +否定：正面情感

**4 注意事项**

① 首先对句子中的情感进行判断，然后对属性进行情感极性的标记，在不能对该属性做出情感极性判断的情况下，则可以参照整个句子的情感倾向；

② 情感极性的判断只针对评审句表中抽取出来的属性词；

③在不使用转折词的情况下，句子中前后所包含的属性情感极性是一致的；

④ 实体和其属性具有相同的情感极性，如the label of Figure，其中label ，Figure的情感相同；

⑤ 对于一个句子中的某个属性被重复抽出的情况，如果该属性在句中出现多次，且没有明显的情感转折词，则极性相同；如果仅出现一次，则第二次标2

⑥ 对于句中已经提到优点或缺点，该句属性的情感极性与之对应