



对话式搜索数据标注培训

储著敏

2022/01/26

内容提纲

- 用户部分
 - 常见错误勘误
 - 标注规则约定
- 专家部分
 - 常见错误勘误
 - 标注规则约定
- Q&A
- 后续安排说明

常见错误勘误（用户）：revise v.s. reveal-continue

- 判别核心点
 - revise没有提出新的信息需求，而是阐明已有的信息需求
 - Reveal-continue则是在问完一段后继续追问当前话题下的其他信息需求
- 举例
 - 例1
 - 用户：德约科维奇拿过多少个大满贯冠军？
 - 用户：截至2022年1月有多少？ **revise**
 - 例2
 - 用户：德约科维奇拿过多少个大满贯冠军？截至2022年1月有多少？ **reveal-initiate**
 - 例3
 - 用户：德约科维奇拿过多少个大满贯冠军？
 - 用户：大师赛的情况呢？ **reveal-continue**

常见错误勘误（用户）：reveal v.s. reveal + others

- 判别核心点

- 如果用户在表达信息需求的时候，包含了反馈或观点，且去掉这些内容不影响信息需求的表达，那么就是reveal + others，反馈或观点部分不可去则为reveal
- 当遇到指代词或者甚至指代词省略的情况，需要再指代消歧后再做判定

- 举例

- 例1

- 专家：...推荐火炉火和日昌餐厅
 - 用户：火炉火的烤肉确实味道不错，（那儿）最近有推出啥新菜吗？

那儿指代前面提到的火炉火，不管有没有都将其替代为火炉火后考虑，去掉前半句不影响信息需求表达，因此为reveal-inspire + others

- 例2

- 专家：...推荐火炉火和日昌餐厅
 - 用户：火炉火的调味牛五花我感觉特别好吃，除此之外还有啥好吃的吗？

后面句可以加个省略的指代地点火炉火，但加上后仍无法独立表达信息需求，因此为reveal-inspire

常见错误勘误（用户）：reveal-inspire v.s. reveal-continue #1

- 判别核心点

- “直接”受到专家回复的启发，如果用户提问中所涉及到的实体部分直接受到启发，部分间接，也应当判定为reveal-inspire

- 举例

- 例1

- 专家：...推荐火炉火和日昌餐厅
 - 用户：火炉火的营业时间如何？
 - 专家：xxx
 - 用户：这两家大众评分哪个更高？

用户第二句提问当中，“这两家”所指的火炉火已经被用户在前述话题中提及，属于间接启发，但日昌餐厅仍是直接启发，因此标为reveal-inspire

- 例2

- 专家：...推荐火炉火和日昌餐厅
 - 用户：火炉火的营业时间如何？
 - 专家：xxx
 - 用户：日昌的呢？
 - 专家：xxx
 - 用户：这两家大众评分哪个更高？

用户第三句提问中的“这两家”已经全部在上述用户提问中提及，均为间接启发，因此标为reveal-continue

常见错误勘误（用户）：reveal-inspire v.s. reveal-continue #2

- 判别核心点

- 遇到专家回复是fact-single类型时，后续仍以这一事实整体概念为提问话题点，则标注为reveal-continue；如果回复是fact-multi类型则为reveal-inspire

- 举例

- 例1

- 用户：中国的八大菜系具体指的是什么？
- 专家：八大菜系，即鲁菜、川菜、粤菜、江苏菜、闽菜、浙江菜、湘菜、徽菜。
- 用户：它们分别有什么特点？

专家的回复答案是唯一确定的事实，即fact-single。用户在后续提问中仍以八大菜系这一整体概念作为切入点，应当判定为reveal-continue

- 例2

- 用户：五道口美食推荐？
- 专家：日昌餐厅和火炉火
- 用户：它们的均价分别是多少？

专家回复答案并不唯一，且推荐店名为事实，因此为fact-multi，此时后续用户尽管继续以整体概念为切入点提问，也依然判定为reveal-inspire

标注规则约定（用户）：chit-chat v.s. others

- 约定条款

- “好的”，不管是单独出现还是在句中出现，都判定为chit-chat（或者作为语气词出现时，不标）
- “理解了”，“懂了”，“明白了”等，我们认为给出了反馈信息，为 others

标注规则约定（用户）：chit-chat v.s. [不单独标注]

- 约定条款

- 语气过渡词（一般出现在用户提问的段首），当我们发现没有起到推动双方交流作用时，不需要单独标注；当起到了推动双方交流时，才标chit-chat

- 举例

- 例1
 - 专家：...推荐火炉火
 - 用户：好的，那他们家有啥好吃的菜吗？

这里的“好的”更像是一个语气过渡词，有点类似于我们平常口语聊天时经常带的“嗯”、“哦”、“啊”，没有推动双方的交流，因此不需要单独标注，标reveal-inspire即可

- 例2
 - 专家：...推荐火炉火
 - 用户：谢谢您！他们家好吃的菜有推荐的吗？

这里的“谢谢您！”给人的感觉与后面的内容时分隔开的，而非语气过渡词，对推动双方交流起到了作用，因此需要标注chit-chat，整句话就是reveal-inspire + chit-chat

常见错误勘误（专家）：answer-fact v.s. answer-opinion #1

- 判别核心点

- 如果回答内容的对错，不以人的意志为转移，始终是正确的，那么就是fact类
- 如果不同人（或社会群体）对于回答内容持有不同的观点，则为opinion类

- 举例

- 例1
- 专家：德约科维奇的打法特点？
- 用户：均衡稳定的正反拍；发球具有稳定性好、打点准确；接发球功底十分深厚等等

不同人对于德约科维奇的打法特点可能认知就会存在差异，尤其是像均衡稳定、打点准确、功底深厚之类的说法，存在个人主观认知，因此为opinion

- 例2

- 专家：红烧肉属于哪个菜系？
- 用户：红烧肉属于大众菜肴，各大菜系中都能寻觅到其踪迹，例如苏菜中有苏式红烧肉，湘菜系有毛式红烧肉，川菜中有川式红烧肉。而最早的红烧肉是鲁菜系的，后来才逐渐传入各大菜系

不同人看待这一答案时，都会认为内容是客观正确的，因此为multi-fact

常见错误勘误（专家）：answer-fact v.s. answer-opinion #2

- 判别核心点

- 如果对于单一问题的回复内容非常复杂，既有事实又有观点。那么，我们就把观点那一部分信息去掉，看看是否能够很好地回答用户问题，如果可以那么就是fact，如果不行就是opinion（当然，如果存在大段的观点在这儿被忽略，且他们独立地说明了其他信息，说明他们不是问题答案而是补充信息，那么就再加一个opinion的副标签）

- 举例

- 例1

- 用户：能简单介绍一下德约科维奇的职业生涯吗？
- 专家：2003年，德约科维奇转为职业球员，开始职业生涯，一段传奇的经历开启了。2007年，世界排名升至第三，没人曾想到这位塞尔维亚小伙子能取得如此骄人成绩。2008年，首次获得澳网冠军.....

去掉观点的句子不影响整段对话想表达的信息，因此为multi-fact

- 例2

- 用户：甄嬛传的剧情？
- 专家：一个柔弱的女子甄嬛在深宫当中，从默默无闻到出人头地，凶险异常下她足智多谋，她斗倒了华妃，本以为可以过上舒坦日子，最终却被皇后以纯元皇后的故衣陷害，心灰意冷之下，最终决定出宫修行。

观点部分去掉之后，表达信息缺失明显，因此标注为opinion

常见错误勘误（专家）：answer-fact-single v.s. answer-fact-multi

- 判别核心点

- 如果回复是明确的唯一的事实，则为single，这里的唯一指的是没有其他同质的回复可以作为替代；如果答案是不唯一的，不管回复中给了一部分、或者全部要素、成分的事实结果，用户都可以满意，则为multi
- 判断时既要看看用户的提问是期待部分信息还是全部信息，还要看专家的回复方式是否存在其他同质答案的可能性

- 举例

- 例1

- 用户：金州勇士队今年1月份有哪些主场比赛？
 - 专家：北京时间1月4日11点打热火、.....、1月30日11点打森林狼

这个地方我们可以看出用户的期待是告知所有的比赛信息，而专家也是这么回答的，因此为single-fact

- 例2

- 用户：金州勇士队春节前还有没有打的比赛吗？
 - 专家：给您找到所有满足要求的比赛信息：北京时间1月26日11点主场打独行侠、.....

这个地方用户其实没有期待专家回答所有的场次信息，告知一部分也可以满足需求，所以为multi-fact

常见错误勘误（专家）：answer-fact v.s. no-answer

- 判别核心点
 - No-answer强调的是主观上，专家没有找到相关问题的信息，但并不意味着这个问题没有答案
 - 举例
 - 例1
 - 用户：这个问题官方具体是怎么解释的？
 - 专家：没有找到官方的解释
 - 例2
 - 用户：这个问题官方具体是怎么解释的？
 - 专家：官方没有给出解释
- No-answer
- Answer-fact-single, 官方没有给解释也是一个事实类的答案, 解释就是[空信息]

常见错误勘误（专家）：单标签 v.s. 主次标签

- 判别核心点

- 如果用户只提了一个问题，所有的回复都是这个问题的答案，那么回复部分只能有一个answer标签（可以外挂others、chit-chat等副标签）
- 如果用户只提了一个问题，部分回复是这个问题的答案，那么回答问题部分的类型是answer主标签，其他补充部分是answer副标签
 - 注意：补充信息部分只需要区分fact-multi和opinion即可，因为补充方向是多样化的
- 如果用户提了多个问题，那么回复内容最多的为主标签，回复内容次多的为次标签
 - 标签类别选择时不需要考虑补充信息，但判断回复长短时需要考虑
 - 例外：如果问题A、B之间，B问题的回答依赖于问题A的答案才能给出时，不论长度，A都是主标签，B是副标签
- 多标签时，允许主次标签一样的情形

常见错误勘误（专家）：单标签 v.s. 主次标签 多段对话内容情形 #1

- 判别核心点

- 如果用户只提了一个问题，所有的回复都是这个问题的答案，那么回复部分只能有一个answer标签（可以外挂others、chit-chat等副标签），且所有段对话都是这个标签
- 如果用户只提了一个问题，部分回复是这个问题的答案，那么对每段回复内容而言，回答问题部分的类型是answer第一标签，其他补充部分是answer第二标签，第一标签和第二标签对于多段对话内部是一致的
 - 这里用第一标签、第二标签说法是因为，可能有些段对话里面只有补充信息，那么此时第二标签就是主标签（存在一、一+二、二这三种标签可能）
- 如果用户提了多个问题，那么回复内容最多的为主标签，回复内容次多的为次标签，而补充信息的重要性在回答内容后面，即 回答多 > 回答少 > 补充信息多 > 补充信息少，这儿的多少是回答+补充信息总量来定的
 - 每段回复都只需要按上述递补顺序，至多标两个标签即可

常见错误勘误（专家）：单标签 v.s. 主次标签 多段对话内容情形 #2

• 举例

• 例1

- 用户：能简单介绍一下德约科维奇的职业生涯吗？
- 专家：2003年，德约科维奇转为职业球员，开始职业生涯，一段传奇的经历开启了。
- 专家：2007年，世界排名升至第三，没人曾想到这位塞尔维亚小伙子能取得如此骄人成绩。
- 专家：2008年，首次获得澳网冠军.....

多段都在回答用户的问题，去掉观点的句子不影响整段对话想表达的信息，因此多段全部标注为multi-fact

• 例2

- 用户：甄嬛传的剧情？
- 专家：一个柔弱的女子甄嬛在深宫当中，从默默无闻到出人头地，凶险异常下她足智多谋。
- 专家：她斗倒了华妃，本以为可以过上舒坦日子。
- 专家：最终却被皇后以纯元皇后的故衣陷害，心灰意冷之下，最终决定出宫修行。

多段都在回答用户的问题，观点部分去掉之后，表达信息缺失明显，因此多段全部标注为opinion

常见错误勘误（专家）：单标签 v.s. 主次标签 多段对话内容情形 #3

- 举例

- 例3

- 用户：求推荐一些敏感肌肤适合的面霜和面膜

- 专家：一、珂润保湿面霜 188rmb/40g

fact-multi

- 专家：蜡基-PG-羟乙基棕榈酰胺作为独家的神经酰胺成分可以帮助稳固角质层。另外角鲨烷、蓝桉叶提取、尿囊素等，具有抗炎、修复作用。使用感较清爽易推开。干皮一年四季都适用，混油皮适用于秋冬干燥季节

opinion

- 专家：二、科颜氏高保湿霜 300rmb/50ml。聚硅氧烷和角鲨烷主打，同时加入多种植物油，有效帮助滋润皮肤。尿素和羟棕榈酰二氢鞘氨醇（神经酰胺的一种）强化角质层保湿能力。肤感方面滋润易推，涂完不粘腻。干皮和混干秋冬适用。混油表示也可以hold住

fact-multi + opinion

- 专家：具体使用方法建议咨询医师

opinion

常见错误勘误（专家）：单标签 v.s. 主次标签 多段对话内容情形 #3

- 举例

- 例4

- 用户：求推荐一些敏感肌肤适合的面霜和面膜以及推荐理由

- 专家：一、珂润保湿面霜 188rmb/40g

- 专家：蜡基-PG-羟乙基棕榈酰胺作为独家的神经酰胺成分可以帮助稳固角质层。另外角鲨烷、蓝桉叶提取、尿囊素等，具有抗炎、修复作用。使用感较清爽易推开。干皮一年四季都适用，混油皮适用于秋冬干燥季节

- 专家：二、科颜氏高保湿霜 300rmb/50ml。聚硅氧烷和角鲨烷主打，同时加入多种植物油，有效帮助滋润皮肤。尿素和羟棕榈酰二氢鞘氨醇（神经酰胺的一种）强化角质层保湿能力。肤感方面滋润易推，涂完不粘腻。干皮和混干秋冬适用。混油表示也可以hold住

- 专家：具体使用方法建议咨询医师

fact-multi (问题1答案)

opinion (问题2答案)

fact-multi + opinion (问题1答案 + 问题2答案)

opinion (补充信息)

常见错误勘误（专家）：单标签 v.s. 主次标签 多段对话内容情形 #4

- 举例
- 例5
- 用户：请问金鼎轩人多吗？可以提前预定吗？
- 专家：饭点去需要排队，金鼎轩24小时营业，可以避开高峰前往就餐，
亚运村店预订电话：(010)64978168

fact-multi (问题1答案)

opinion (问题1补充信息)

fact-multi(问题2答案)

标签重要程度：fact-multi (问题1答案) > fact-multi(问题2答案) > opinion (问题1补充信息)，因此最后标注为fact-multi + fact-multi

常见错误勘误（专家）：fact-multi + opinion v.s. opinion 产品推荐

- 判别核心点

- 如果回复的侧重点是推荐的产品，一般是直截了当的摆一些产品以及它们的介绍，那么一般就是fact-multi + [补充信息类别]
- 如果是以展现自己的推荐框架或者推荐分析理由为主，产品只是这一推荐框架产物时，那么就标注为opinion（多段的话也全标为opinion即可）

- 例1

- 用户：求推荐一些敏感肌肤适合的面霜和面膜以及推荐理由

- 专家：给您推荐：一、珂润保湿面霜 188rmb/40g，[一些理由]；二、科颜氏高保湿霜 300rmb/50ml，[一些理由]。

产品为主， fact-multi +
opinion

- 例2

- 用户：求推荐一些敏感肌肤适合的面霜和面膜以及推荐理由

- 专家：这边建议您使用成分天然，最好是纯植物制成的面霜。同时可以注重选用带修复功效的面霜，来增强敏感肌的耐受性，有效降低过敏几率。还可以选用保湿效果好的面霜，防止皮肤干燥针对您的需求找到了以下几项产品：一、珂润保湿面霜 188rmb/40g，[一些理由]；二、科颜氏高保湿霜 300rmb/50ml，[一些理由]。

推荐框架为主，
opinion

Q&A

- 后续安排说明
- 每人标注11组对话
- 会发放新的账号进行标注
- ddl: 2月9日中午12点？



感谢大家的聆听

储著敏

2022/01/26