数据标注说明

数据集分为训练、验证、测试A与测试B四部分。数据集中的评价对象按照粒度不同划分为两个层次，层次一为粗粒度的评价对象，例如评论文本中涉及的服务、位置等要素；层次二为细粒度的情感对象，例如“服务”属性中的“服务人员态度”、“排队等候时间”等细粒度要素。评价对象的具体划分如下表所示。

The dataset is divided into four parts: training, validation, test A and test B. This dataset builds a two-layer labeling system according to the evaluation granularity: the first layer is the coarse-grained evaluation object, such as “service” and “location”; the second layer is the fine-grained emotion object, such as “waiter’s attitude” and “wait time” in “service” category. The specific description is shown in the following table.

|  |  |
| --- | --- |
| **层次一(The first layer)** | **层次二(The second layer)** |
| 位置(location) | 交通是否便利(traffic convenience) |
| 距离商圈远近(distance from business district) |
| 是否容易寻找(easy to find) |
| 服务(service) | 排队等候时间(wait time) |
| 服务人员态度(waiter’s attitude) |
| 是否容易停车(parking convenience) |
| 点菜/上菜速度(serving speed) |
| 价格(price) | 价格水平(price level) |
| 性价比(cost-effective) |
| 折扣力度(discount) |
| 环境(environment) | 装修情况(decoration) |
| 嘈杂情况(noise) |
| 就餐空间(space) |
| 卫生情况(cleaness) |
| 菜品(dish) | 分量(portion) |
| 口感(taste) |
| 外观(look) |
| 推荐程度(recommendation) |
| 其他(others) | 本次消费感受(overall experience) |
| 再次消费的意愿(willing to consume again) |

每个细粒度要素的情感倾向有四种状态：正向、中性、负向、未提及。使用[1,0,-1,-2]四个值对情感倾向进行描述，情感倾向值及其含义对照表如下所示：

There are four sentimental types for every fine-grained element: Positive, Neutral, Negative and Not mentioned, which are labelled as 1, 0, -1 and-2. The meaning of these four labels are listed below.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 情感倾向值  (Sentimental labels） | 1 | 0 | -1 | -2 |
| 含义  （Meaning） | 正面情感  (Positive) | 中性情感  (Neutral) | 负面情感（Negative） | 情感倾向未提及  （Not mentioned） |

数据标注示例如下：

An example of one labelled review:

|  |
| --- |
| "味道不错的面馆，性价比也相当之高，分量很足~女生吃小份，胃口小的，可能吃不完呢,。环境在面馆来说算是好的，至少看上去堂子很亮，也比较干净，一般苍蝇馆子还是比不上这个卫生状况的。中午饭点的时候，人很多，人行道上也是要坐满的，隔壁的冒菜馆子，据说是一家，有时候也会开放出来坐吃面的人。 |

|  |  |  |
| --- | --- | --- |
| **层次一(The first layer)** | **层次二(The second layer)** | **标注 (Label)** |
| 位置(location) | 交通是否便利(traffic convenience) | -2 |
| 距离商圈远近(distance from business district) | -2 |
| 是否容易寻找(easy to find) | -2 |
| 服务(service) | 排队等候时间(wait time) | -2 |
| 服务人员态度(waiter’s attitude) | -2 |
| 是否容易停车(parking convenience) | -2 |
| 点菜/上菜速度(serving speed) | -2 |
| 价格(price) | 价格水平(price level) | -2 |
| 性价比(cost-effective) | 1 |
| 折扣力度(discount) | -2 |
| 环境(environment) | 装修情况(decoration) | 1 |
| 嘈杂情况(noise) | -2 |
| 就餐空间(space) | -2 |
| 卫生情况(cleaness) | 1 |
| 菜品(dish) | 分量(portion) | 1 |
| 口感(taste) | 1 |
| 外观(look) | -2 |
| 推荐程度(recommendation) | -2 |
| 其他(others) | 本次消费感受(overall experience) | 1 |
| 再次消费的意愿(willing to consume again) | -2 |