

1. 首先统计5星和1、2星评论的词频，分别取词频最高的那几个形容词

a. 吹风机

i. 5星: love great well perfect super new good like

ii. 1星和2星: disappoint less old wast doesnt didnt fall wouldnt dont

b. 微波炉

i. 5星: great fit well love easi like good perfect nice perfectli

ii. 1星和2星: problem dont not didnt doesnt error

c. 奶嘴

i. 5星: love like great cute good well soothi best perfect favorit nice soft

ii. 1星和2星: didnt doesnt dont disappoint better never

2. 构造正面评价词汇集、负面评价词汇集、否定词汇集

a. positive = ['love

','great','well','perfect','super','good','like','cute','easi','fit','nice','perfectli','soothi','best','favorit ','soft']

b. negtive = ['disappoint ','less','wast ','error','problem','better','never','fall']

c. not_word = ['doesnt','didnt','dont','wouldnt','not']

3. 根据词汇集判断评论的属于正面评论or负面评论

a. 对于一条评论，如果任何一个单词位于负面词汇集中，则认为该评论为负面评论

b. 对于一条评论，如果任何一个单词位于正面词汇集中，且所有单词不位于否定词汇集中，则认为该评论为正面评论

c. 其他情况忽略

4. 初始化 $n=0$ ，对于3中的正面评论和负面评论，若正面评论的真实星级为4和5，则 $n+=1$ ；若负面评论的真实星级为1和2，则 $n+=1$ 。准确率= $n / (\text{正面评论数} + \text{负面评论数})$

实验结果：

- 吹风机准确率=0.7627401837928154
- 微波炉准确率=0.727853152434158
- 奶嘴准确率=0.7901683501683502