**领域情感词典构建方法步骤**

参考《基于语义规则和词向量的台风灾害网络情感》（地理信息科学 2022）

Step1

分词提取所有文本中的所有词语（共67064个），并与大连理工大学情感词典对比，找出其中重合的词作为各类别的种子词（共4944个）；

对于剩下的62120个词语，选取其中词频大于50的词语作为待选词（共609个）；

Step2

计算待选词与各类种子词之间的余弦相似度，阈值设置为0.7，筛选出符合条件的词作为领域情感词汇的候选词共598个；

Step3

对于筛选出的候选词，先人工筛选出其中有意义的词语共219个，再根据该词汇与不同类别的情感种子词的余弦相似度，取余弦相似度相对较大的情感种子词作为该词汇的所述类别，情感强度与对应的情感种子词相同