

分词和分字的差异

1.分词存在效果差异,可能存在不同字所组合的词意义不同的情况,可能需要对特殊词语进行处理,分字不存在,分字是将字符一个一个分开不需要对特殊词语进行处理,但是可能会存在语义丢失的情况;

2.用分词的方法对应的onehot矩阵维度远远高于分字所对应的onehot矩阵;