



# 关键词提取

## 功能描述

**词语的重要性与它在文件中出现的次数成正比，但同时会随着它在语料库中出现的频率成反比下降。**如果一个词在所有的文档中都多次出现，那么这个词就没有什么价值了。或者说，**如果某个词或者短语在一个文档中出现多次，但是在其他文档中很少出现，就可以认为这个词或短语具有很好的区分性，适合用来对文档进行分类。**

## 业务流程图

