

[Code 码农网](#)

- [码农网](#)
- [码农文章](#)
- [码农社区](#)
- [码农教程](#)
- [码农网分类](#)
  - [码农软件](#)
  - [码农书籍](#)
  - [码农日报](#)
  - [码农工具](#)
- [码农网登录](#)
- [码农网导航](#)
  - [关于码农网](#)
  - [联系码农网](#)
  - [码农网导航](#)
- [码农软件](#)
- [码农书籍](#)
- [码农日报](#)
- [码农工具](#)

## 搜索引擎算法之关键词类目预测

栏目: [编程工具](#) · 发布时间: [1年前](#)

来源: [blog.csdn.net](#)

←

Google 提供的广告

发送反馈

为什么显示此广告? ▸

内容简介: 在搜索算法中, 关键词类目是非常重要的一个话题, 是搜索排序中的一个重要模块。搜索排序可以简单的分成几个模块: 文本相关性、质量分、转化率。文本相关性可以由粗到细分解为类目相关性、属性相关性、语义相关性几个部分。本文重点解析一下类目相关性。例如用户搜索“连衣裙”, 最相关的类目是“女装-连衣裙”、“童装-连衣裙”。而其他类目中包含连衣裙就不是那么相关, 例如帽子类目的商品“适合沙滩连衣裙的帽子”。所以我们可以简单的把商品区分为类目相关和类目不太相关的两个部分。而如何区分是否相关就需要用到关键词的类目预测。给个例子

本文转载自: <https://blog.csdn.net/poson/article/details/89672802>, 本站转载出于传递更多信息之目的, 版权归原作者或者来源机构所有。

在搜索算法中, 关键词类目是非常重要的一个话题, 是搜索排序中的一个重要模块。搜索排序可以简单的分成几个模块: 文本相关性、质量分、转化率。文本相关性可以由粗到细分解为类目相关性、属性相关性、语义相关性几个部分。本文重点解析一下类目相关性。

例如用户搜索“连衣裙”，最相关的类目是“女装-连衣裙”、“童装-连衣裙”。而其他类目中包含连衣裙就不是那么相关，例如帽子类目的商品“适合沙滩连衣裙的帽子”。所以我们可以简单的把商品区分为类目相关和类目不太相关的两个部分。而如何区分是否相关就需要用到关键词的类目预测。

给个例子：student bag 的预测结果。

笔袋	0.0214844
女士钱包	0.0390625
其它女包	0.3125
女双肩背包	0.324219
其他双肩包	0.173828
斜跨包	0.0273438
单肩包	0.0449219

关键词类目预测的算法很多，从发展路径上来说分为几种方式：

1. 通过pair <商品标题词,类目>，统计关键词和类目的共现关系
2. 统计用户搜索Query，然后在结果集中点击的商品的集合，统计商品类目的分布。注意这里的Query的term要全部在命中的商品标题中出现，而不是部分出现。
3. 用分类算法的方式，样本是<商品的标题词，类目>，。使用适合多分类的算法分类：最大熵、FastText、TextCNN、BI-LSTM + attention等算法。

在关键词做类目预测之前可以做一个预处理提高准确率。如query归一化、纠错、去除价格区间词、中英文翻译对照等等。

以上就是本文的全部内容，希望对大家的学习有所帮助，也希望大家多多支持 [码农网](http://www.coderc.to)



关注我们，获取更多IT资讯^\_^

## 为你推荐:

- [在搜索引擎广告关键词生成上，算法可以做什么？](#)
- [电商搜索引擎-算法选型](#)
- [海量数据搜索——搜索引擎](#)
- [海量数据搜索---搜索引擎](#)

## 相关软件推荐:

- [Web浏览器引擎 WebKit](#)
- [可嵌入的脚本引擎 Jx9](#)