

酒店实时关注度模型

一、模型概述

本模型聚焦于分析 X 市各酒店的实时关注度情况，旨在通过收集和分析相关数据，以直观可视化的方式呈现各酒店受关注程度，并且支持点击查看具体搜索次数，为酒店经营者了解自身热度、制定营销策略以及游客选择住宿等提供及时且有价值的数据参考，助力优化 X 市酒店市场的资源配置与竞争态势。

二、数据来源与整理

搜索引擎数据：借助主流搜索引擎（如百度、搜狗等）提供的公开数据接口或者相关搜索指数工具，获取包含“X 市酒店名称”等关键词的搜索记录，这些记录能够反映大众在不同时段对各酒店的主动搜索情况，是了解酒店实时关注度的重要基础数据源之一，从中可以提取出各酒店被搜索的频次信息。

在线旅游平台数据：诸如携程、去哪儿、飞猪等在线旅游平台上，用户对 X 市酒店的浏览量、点击量以及预订相关操作等数据，也能从侧面体现酒店受关注的程度。用户频繁浏览和点击某家酒店页面，意味着对其兴趣较高，通过平台的数据接口收集这些信息，与搜索引擎数据相互补充，进一步完善酒店关注度的数据来源，使分析结果更具全面性。

社交媒体平台数据：微博、微信公众号、抖音、小红书等社交媒体平台上，用户发布的与 X 市酒店相关的话题、动态、评论以及点赞、转发等互动信息，包含了大量关于酒店特色、优惠活动、入住体验等方面的内容，通过网络爬虫技术（在遵循平台规则和合法合规的前提下）收集这些文本数据，统计各酒店在社交平台上的曝光度和话题热度，丰富关注度评价的数据维度，更全面地反映酒店在公众视野中的受关注情况。

数据整理流程：首先，将从上述多渠道收集到的海量数据汇总到统一的数据仓库中。接着，运用数据清洗工具和人工核查相结合的方法，去除重复、错误以及逻辑不清晰的数据记录，比如剔除搜索引擎数据中因关键词误匹配导致的错误搜索

记录、纠正在线旅游平台上因系统故障产生的异常浏览数据等。然后，按照酒店名称对数据进行分类整理，针对每个酒店，提取并整合其在搜索频次、平台浏览点击量、社交平台热度等各方面的数据，为后续的关注度综合计算和可视化展示做好准备。

三、核心算法与分析逻辑

关注度综合评估指标构建：为了全面衡量酒店的实时关注度，构建一个综合评估指标。选取多个与关注度相关的数据指标（如上述提到的搜索频次、在线旅游平台浏览量、社交媒体平台热度等），并分别为其赋予合理的权重（权重的确定可通过专家打分法、层次分析法或者基于历史数据进行回归分析等方式来确定，例如搜索频次权重设为 0.4，在线旅游平台浏览量权重设为 0.3，社交媒体平台热度权重设为 0.3）。然后，针对每个酒店，将各指标的实际数值进行标准化处理（例如将搜索频次、浏览量等数值映射到 0 - 1 之间，使不同量级的数据具备可比性），再按照权重进行加权求和，得到每个酒店的实时关注度综合得分，以此来量化酒店的实时关注度情况。

字号大小设定依据：根据计算得出的各酒店实时关注度综合得分，将得分进行归一化处理（例如通过线性变换等方法，把得分映射到一个适合设定字号大小的区间，如 10 - 30 号字体区间），得分越高的酒店，对应的字号越大，以此在可视化展示中通过字号大小直观体现各酒店关注度的高低差异。

四、模型输出与可视化

词云图展示（整体关注度展示）：利用专业的可视化工具（如 Python 的 wordcloud 库或者在线词云生成平台等），根据酒店的实时关注度综合得分以及对应的字号大小设定，生成以 X 市酒店为主题的词云图。在词云图中，将各个酒店名称以不同的字号大小、颜色等形式呈现，酒店的实时关注度综合得分越高，其在词云图中显示的字号越大、颜色越突出（可根据设计自行设定突出显示的方式），整体形成直观且形象的词云效果，让人一眼就能看出哪些酒店是当前实时关注度较高的热点酒店。

交互功能实现：借助前端可视化开发技术（如 JavaScript 结合 Echarts 可视化库等），为词云图添加交互功能。当用户点击某一酒店名称时，在词云图上方或旁边弹出信息框，清晰展示该酒店的具体名称以及对应的具体搜索次数（可同时展示综合得分及各指标原始数据等详细信息），方便用户进一步了解详情，深入洞察各酒店背后所反映的关注度情况。

五、模拟数据示例与可视化展示

假设通过收集和整理数据后，得到以下部分 X 市酒店及对应的实时关注度模拟数据（仅为示例示意，已进行相关计算和处理）：

酒店名称	实时关注度综合得分	搜索次数（次）
X 市国际大酒店	0.8	500
X 市海滨度假酒店	0.7	400
X 市商务中心酒店	0.6	300
X 市温馨民宿	0.5	250
X 市花园酒店	0.4	200



以下是使用 Python 的 wordcloud 和 pandas 库来生成满足要求的词云图的示例

代码：

一、代码实现

python

```
from wordcloud import WordCloud
```

```
import pandas as pd
```

```
import matplotlib.pyplot as plt
```

```
from colormap import rgb2hex
```

```
# 读取 Excel 文件
```

```
df = pd.read_excel('/mnt/39.xlsx')
```

```
# 提取酒店名称和实时关注度综合得分
```

```
hotel_names = df['酒店名称'].tolist()
```

```
scores = df['实时关注度综合得分'].tolist()
```

```
# 创建一个字典，将酒店名称和得分对应起来
```

```
data_dict = dict(zip(hotel_names, scores))
```

```
# 自定义颜色函数，根据得分生成不同颜色
```

```
def color_func(word, font_size, position, orientation,  
random_state=None, **kwargs):
```

```
    max_score = max(data_dict.values())
```

```
    min_score = min(data_dict.values())
```

```
    score = data_dict[word]
```

```
    # 将得分映射到 0 - 1 之间
```

```
    normalized_score = (score - min_score) / (max_score - min_score)
```

```
    # 这里简单地根据得分设置颜色从浅到深（例如蓝色调）
```

```
    r = int(0 + normalized_score * 200)
```

```
    g = int(0 + normalized_score * 200)
```

```
        b = int(255 - normalized_score * 200)
        return rgb2hex(r, g, b)

# 生成词云
wordcloud = WordCloud(background_color='white', width=800,
height=400).generate_from_frequencies(data_dict)

# 重新着色词云
wordcloud.recolor(color_func=color_func)

# 显示词云图
plt.figure(figsize=[10, 6])
plt.imshow(wordcloud, interpolation='bilinear')
plt.axis('off')
plt.show()
```

二、代码解释

数据读取：

使用 `pd.read_excel('/mnt/39.xlsx')` 读取上传的 Excel 文件，假设文件中包含酒店名称和实时关注度综合得分两列。

数据整理：

将酒店名称和对应的得分整理成一个字典 `data_dict`，以便 `wordcloud` 库使用。

颜色函数定义：

`color_func` 函数根据酒店的得分来生成不同的颜色。首先找到得分的最大值和最小值，然后将每个酒店的得分映射到 0 - 1 之间，根据这个归一化的得分来生成从浅到深的颜色（这里是简单的蓝色调变化）。

词云生成：

创建一个 `WordCloud` 对象，并使用 `generate_from_frequencies(data_dict)` 根据酒店名称和得分生成词云。

重新着色词云：

调用 `wordcloud.recolor(color_func=color_func)` 方法，根据定义的颜色函数对词云进行重新着色。

显示词云图：

使用 `matplotlib` 库显示生成的词云图，设置图片大小并关闭坐标轴。

请确保安装了所需的库，可以使用 `pip install wordcloud pandas colormap matplotlib` 命令进行安装。根据实际的数据文件路径和列名，可能需要调整代码中的相关部分。

三、注意事项

此代码中的颜色映射只是一个简单的示例，你可以根据需求自定义更合适的颜色映射方案。

如果需要将生成的词云图保存为图片文件，可以使用 `wordcloud` 库的 `to_file` 方法，例如 `wordcloud.to_file('hotel_wordcloud.png')`。

（此处应插入一个简单的词云图示例，图中“X 市国际大酒店”字号最大、颜色最突出，“X 市海滨度假酒店”次之，以此类推，各酒店名称按照实时关注度综合得分情况在词云图中呈现不同字号大小和颜色，鼠标点击酒店名称有信息框展示酒店名称、搜索次数等详细信息示例图）

六、应用场景与价值

酒店营销策略调整：酒店经营者通过观察模型展示的本酒店及其他竞争对手的实时关注度情况，可及时调整营销策略。如果发现自身酒店关注度较低，可分析原因，比如是否宣传不到位、缺乏特色亮点等，进而加大在搜索引擎、在线旅游平台以及社交媒体上的推广力度，推出有吸引力的优惠活动或特色服务，提升酒店的曝光度和关注度，吸引更多客源。

游客住宿参考：游客在计划前往 X 市出行时，参考该模型呈现的酒店实时关注度信息，能够了解当前哪些酒店比较热门，再结合自身需求（如价格、位置、酒店类型等），更合理地选择心仪的酒店，提高住宿安排的效率和满意度。

旅游市场动态监测：X 市旅游管理部门借助该模型，可以全面掌握全市酒店行业的实时关注热点，了解不同酒店在市场中的受青睐程度变化情况，进而根据市场动态，引导酒店合理竞争，规范市场秩序，促进 X 市酒店业的健康、均衡发展。