

游客偏好度分析模型

一、模型概述

本模型着重分析游客对 X 市不同类型旅游资源的偏好程度，聚焦于古镇、度假休闲、文博院落、其他历史文化、其他休闲娱乐这 5 类旅游资源，通过科学的数据分析与可视化呈现，以星状图的形式直观展现各类资源在游客心中的受青睐程度，为旅游管理部门、旅游企业等制定针对性的旅游产品开发与营销策略提供有力的数据支撑，从而更好地满足游客需求，提升 X 市旅游产业的吸引力和竞争力。

二、数据来源与整理

游客调查问卷数据：通过在 X 市各旅游景区、交通枢纽、酒店等地，面向游客发放调查问卷收集一手数据。问卷内容涵盖游客对不同类型旅游资源的喜好程度（例如采用李克特量表形式，设置非常喜欢、比较喜欢、一般、不太喜欢、非常不喜欢等选项）、过往游览经历、再次游览意愿等相关问题，这是了解游客主观偏好的直接且重要的数据来源。

在线旅游平台评论与评分数据：各大在线旅游平台（如携程、去哪儿、马蜂窝等）上游客针对 X 市旅游景点、酒店、旅游活动等所发表的评论以及给出的评分信息，其中往往会提及对不同类型旅游资源的感受和评价，通过文本挖掘技术提取相关内容，并结合评分情况，可以从更广泛的游客群体角度，补充完善游客对各类旅游资源的偏好反馈数据。

旅游消费数据：从旅游企业（如景区票务系统、酒店预订系统、旅行社业务记录等）收集的游客消费数据，分析游客在不同类型旅游资源相关项目上的消费频次、消费金额等情况，一定程度上能反映游客对各类资源的偏好倾向，例如在度假休闲类项目上消费较高的游客，大概率对该类资源有较高的偏好度，以此从消费行为层面为分析游客偏好提供客观依据。

数据整理流程：首先，将从上述多渠道收集到的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，比如剔除调查问卷中明显随意作答的数据、纠正在线旅游平台评论中因误解导致的错误评价信息等。然后，针对不同类型的数据进行分类整理，将涉及古镇、度假休闲、文博院落、其他历史文化、其他休闲娱乐这 5 类旅游资源的相关信息提取出来，并按照统一的格式进行编码和量化处理（例如将调查问卷中的选项转化为对应的数值分数，将评论中的积极、消极评价进行分类并赋予相应分值等），为后续的分析计算做好准备。

三、核心算法与分析逻辑

偏好得分计算：

对于调查问卷数据，根据事先设定的赋值规则（如非常喜欢赋值 5 分，比较喜欢赋值 4 分，以此类推），将游客对每类旅游资源的选择转化为具体的分值，然后对所有调查问卷中同一类旅游资源的得分进行求和，并除以有效调查问卷的数量，得到该类旅游资源基于调查问卷的平均得分。

针对在线旅游平台评论与评分数据，运用自然语言处理技术中的情感分析算法，判断评论的情感倾向（积极、消极或中性），并为积极评论赋予正向分值（如 1 分），消极评论赋予负向分值（如 -1 分），中性评论赋予 0 分，再结合景点的评分情况（如满分 5 分制，按照一定权重，如 0.6 的权重与情感分析得分加权求和），统计涉及每类旅游资源相关景点的综合得分，最后求平均得到该类资源基于在线平台数据的得分。

依据旅游消费数据，按照游客在不同类型旅游资源相关项目上的消费金额或消费频次进行排名，根据排名情况赋予相应的分值（例如消费频次排名靠前的赋予较高分值），再通过加权平均等方法（结合消费金额等因素确定权重），计算出每类旅游资源基于消费数据的得分。

最后，将上述三种数据渠道得到的每类旅游资源得分，按照一定权重（可通过专

家评估、历史数据验证等方式确定权重比例，如调查问卷数据权重 0.4，在线平台数据权重 0.3，消费数据权重 0.3）进行加权求和，得出每类旅游资源最终的偏好得分，以此量化体现游客对不同类型旅游资源的偏好程度。

得分标准化（可选，用于统一得分尺度便于可视化展示）：为了使不同类型旅游资源的偏好得分能在同一尺度下更直观地对比，可对计算得出的得分进行标准化处理，例如将各类资源的偏好得分映射到 0 - 100 的区间内（可通过线性变换等方法实现），确保得分在合理且可比的范围内，便于后续星状图的绘制展示。

四、模型输出与可视化

星状图展示：利用专业的可视化工具（如 Python 的 matplotlib 库或其他可视化软件），绘制以古镇、度假休闲、文博院落、其他历史文化、其他休闲娱乐这 5 类旅游资源为顶点的星状图。在星状图中，每个顶点代表一类旅游资源，以顶点到星状图中心的距离来表示该类资源的偏好得分（得分越高，距离中心越远，通过合理的比例尺进行设置），不同类别资源用不同颜色的线条连接顶点与中心，形成类似星星的形状，清晰直观地展示出游客对各类旅游资源的偏好程度对比情况，整个星状图简洁明了，方便快速查看哪类资源更受游客青睐。

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），为星状图添加交互功能。当鼠标指针悬停在某一类别旅游资源对应的图示（顶点到中心的连线或顶点区域）上时，在星状图上方或旁边弹出信息框，清晰展示该种类旅游资源的具体名称以及对应的偏好得分数值，方便用户进一步了解详情，深入洞察游客的偏好倾向。

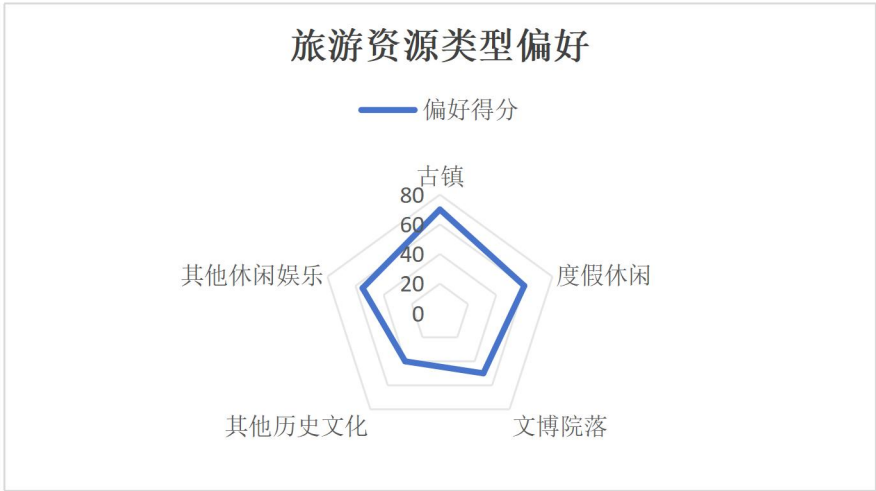
五、模拟数据示例与可视化展示

假设通过上述计算方法，得到以下关于游客对 X 市 5 类旅游资源的偏好得分模拟数据（已标准化到 0 - 100 区间）：

旅游资源类型	偏好得分
--------	------

旅游资源类型	偏好得分
古镇	70
度假休闲	60
文博院落	50
其他历史文化	40
其他休闲娱乐	55

对应的可视化展示如下：



以下是使用 pyecharts 和 selenium 库来生成星状图的图片的示例代码：

一、代码实现

```
python

from pyecharts import options as opts
from pyecharts.charts import Polar
from selenium import webdriver
from selenium.webdriver.chrome.options import Options
import os

# 提取数据
```

```
categories = df['旅游资源类型'].tolist()
values = df['偏好得分'].tolist()

# 将数据转换为适合的格式
data = []
for i, value in enumerate(values):
    data.append([i, value])

# 创建极坐标图表函数
def create_polar_chart():
    polar = (
        Polar()
        .add_schema(
            angleaxis_opts=opts.AngleAxisOpts(
                type_="category",
                boundary_gap=False,
                start_angle=0,
                split_number=len(categories),
                is_clockwise=True,
                data=categories,
            )
        )
        .add(
            series_name="偏好得分",
            data=data,
            type_="effectScatter",
            label_opts=opts.LabelOpts(is_show=False),
        )
    )
```

```
        .set_global_opts(
            title_opts=opts.TitleOpts(title="旅游资源类型偏好得分星状图"),
            toolbox_opts=opts.ToolboxOpts(is_show=True),
        )
    )
    return polar
```

生成图表图片的函数

```
def save_chart_as_image(output_path):
    # 创建极坐标图表
    polar_chart = create_polar_chart()

    # 渲染图表到临时 HTML 文件
    temp_html_path = "temp_polar.html"
    polar_chart.render(temp_html_path)

    # 设置 Chrome 浏览器选项
    chrome_options = Options()
    chrome_options.add_argument('--headless')
    chrome_options.add_argument('--disable-gpu')

    # 启动 Chrome 浏览器驱动
    driver = webdriver.Chrome(options=chrome_options)
    driver.get(f"file:/// {os.path.abspath(temp_html_path)}")

    # 等待页面加载完成
    driver.implicitly_wait(10)
```

```
# 保存截图
driver.save_screenshot(output_path)

# 关闭浏览器和删除临时 HTML 文件
driver.quit()
os.remove(temp_html_path)

# 主程序
if __name__ == "__main__":
    output_path = "/mnt/polar_effect_scatter.png"
    save_chart_as_image(output_path)
```

二、代码解释

创建极坐标图表函数 `create_polar_chart`:

该函数根据数据构建一个极坐标图表对象，设置了角度轴的属性、数据系列以及全局选项，如标题和工具箱。

生成图片函数 `save_chart_as_image`:

首先调用 `create_polar_chart` 函数创建图表对象，然后将图表渲染到一个临时的 HTML 文件中。

使用 `selenium` 的 `webdriver.Chrome` 启动一个无头 Chrome 浏览器，加载临时 HTML 文件，并等待页面加载完成。

调用 `driver.save_screenshot` 方法保存当前页面的截图到指定的输出路径，最后关闭浏览器并删除临时 HTML 文件。

主程序部分:

调用 `save_chart_as_image` 函数，指定输出图片的路径，从而生成并保存星状图的图片。

确保你已经安装了 `pyecharts` 和 `selenium` 库，并且下载了对应的浏览器驱动(如

ChromeDriver)，并根据实际情况配置驱动的路径。这样就可以将生成的星状图保存为图片文件。

（此处应插入一个简单的星状图示例，以古镇、度假休闲、文博院落、其他历史文化、其他休闲娱乐为顶点，各顶点到中心的距离按上述偏好得分情况体现，不同颜色区分各顶点对应的类别，鼠标悬停在顶点或连线上有信息框展示类别名称及偏好得分示例图）

六、应用场景与价值

旅游产品开发与优化：旅游企业（如景区运营方、旅行社、旅游度假区等）可依据模型输出的游客偏好度情况，精准开发和优化旅游产品。例如，若发现游客对古镇类旅游资源偏好度较高，景区可加大对古镇的保护与开发力度，增加特色游览项目、提升古镇的文化体验活动；旅行社则可以设计更多围绕古镇的旅游线路，满足游客需求，提高产品的吸引力和市场竞争力。

旅游营销策略制定：X 市旅游管理部门和旅游企业能够根据游客对不同类型旅游资源的偏好，制定有针对性的营销策略。针对偏好度高的旅游资源，加大宣传推广力度，选择在目标客源市场重点投放广告、举办主题推介活动等；同时，结合游客偏好特点，制定差异化的宣传口号和营销文案，提高营销效果，吸引更多游客来 X 市旅游，提升 X 市旅游品牌的知名度和美誉度。

旅游资源整合与协同发展：了解游客对各类旅游资源的偏好程度，有助于 X 市对现有旅游资源进行整合与协同发展。例如，可以将游客偏好度较高且关联性强的旅游资源进行组合，打造综合性的旅游线路或旅游区域，实现资源共享、优势互补，促进不同类型旅游资源之间的协同合作，提升整个 X 市旅游产业的发展水平，为游客提供更丰富、优质的旅游体验。