

该分析项目的步骤说明

一共分为四个步骤：

1. 数据预处理
2. LDA建模分析-一级特征建模
3. LDA建模分析-二级特征建模
4. 数据分析

## 1、数据预处理

---

数据预处理的步骤说明：

1. 把JSON文件转换为dataframe格式
2. 对content列进行去重处理，确保所有文本内容的唯一性
3. 去掉标点符号和表情包这种会分词无效的内容
4. 对文本进行机械压缩，把一些例如：哈哈哈哈哈，这种变成 哈哈，减少很多意义相同的词
5. 进行分词处理，分词采用jieba分词，用词性来判断是否保留该词，只有符合名词 形容词 动词的词才会被保留下来，并且该词的长度要大于等于，单个词不保留，从而提高分词的准确率和意义

## 2、LDA建模

---

本次LDA建模首先根据清洗好的文本进行分词处理，并获取一级特征的主题，获取一级建模的主题概率和主题类型，并用pyLDAvis.save\_html函数保留下对应的奇葩图

后根据lda.print\_topics模型来获取权重前30的词，把这些词的权重保存下来，后续把权重用平均值的方式来获取一级特征中，不同主题类型，他们对应的一级维度关注度，接着再去获取每个主题类型，他们的占比是多少，用Excel表的方式保存下来

二级建模和一级建模的过程一样，只是二级建模是先根据一级建模对应的主题类型，选定对应的文本内容后，再去进行建模，也就是二次聚类，使得二级维度更加详细

## 3、数据分析

---

数据分析，一共有3个分析文件产生

- 1、APP-各特征维度情感分析.html
- 2、APP-（一级特征名称）.html
- 3、特征数据.xlsx

# 百词斩分析

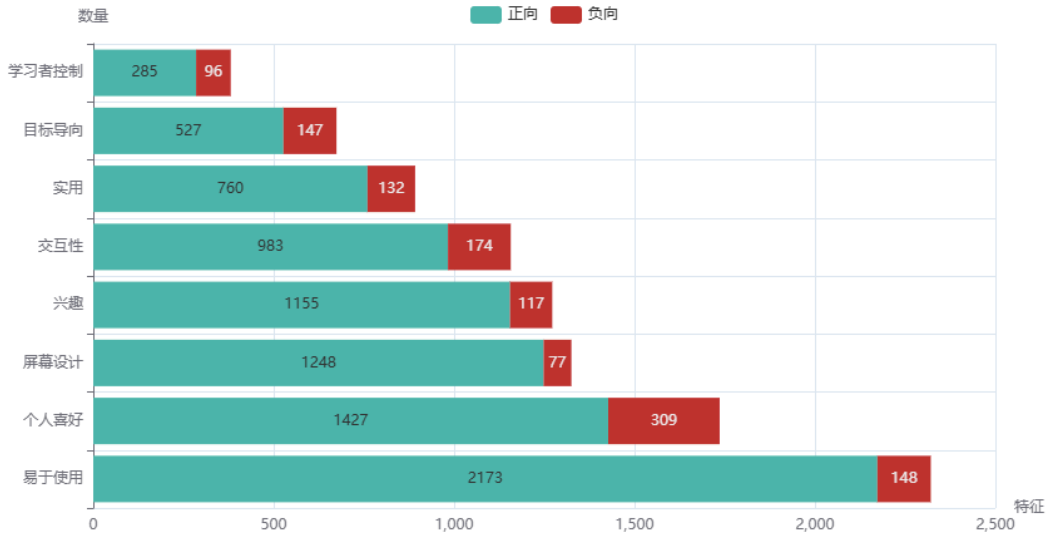
## 一、用户关注度核心结论

### 1. 维度优先级显著差异

- **艺术审美** (67.51%) 的关注度是**技术设计** (32.49%) 的 **2.08倍**，说明用户更关注APP的视觉体验、美学设计及个性化功能。
- 艺术审美下细分维度
  - **美学** (36.96%) > **实用** (32.47%) > **个人喜好** (21.86%) > **学习者控制** (8.71%)
  - **核心需求**：用户期待界面美观与实用功能并重，但对学习自主权（如进度控制）需求较低。
- 技术设计下细分维度
  - **目标导向** (33.08%) > **易于使用** (25.76%) ≈ **导航** (25.26%) > **交互性** (15.90%)
  - **核心需求**：用户希望APP能明确帮助达成学习目标，同时操作简单、路径清晰。

### 2. 高关注维度与情感反馈的关联

APP-各特征维度情感分析



- 正向情感集中领域
  - **易于使用** (正向2173) 和**屏幕设计** (正向1248) 的高正向情感印证了其高关注度，说明用户对基础体验认可。
  - **美学** (图表未直接对应，但关注度第一) 需结合文本分析进一步验证设计满意度。
- 潜在改进点
  - **交互性** (关注度最低，正向983) 和**学习者控制** (正向285) 的正向情感均偏低，需优化交互逻辑与个性化功能。
  - **目标导向** (正向527) 虽关注度高，但情感正向率较低 (527/总讨论量)，可能存在功能与目标匹配不足的问题。

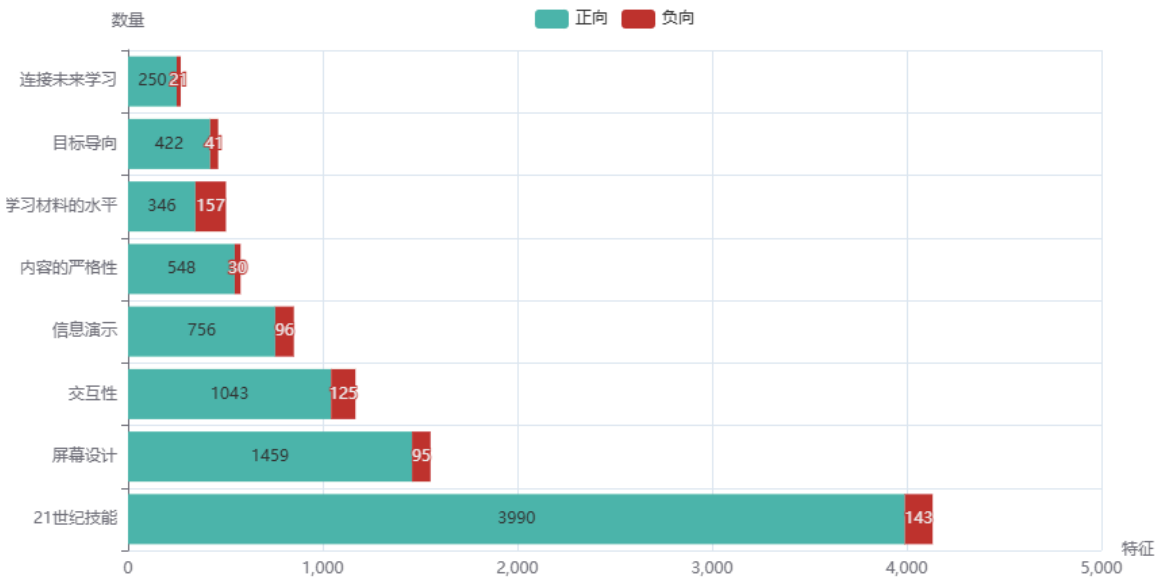
一、用户关注度核心结论

1. 教育导向占据主导地位

- 一级维度对比
  - **面向教育**（57.60%）的关注度显著高于**技术设计**（42.40%），表明用户更重视APP的教育功能与学习价值。
  - 面向教育下细分维度
    - **21世纪技能**（75.35%）占据绝对核心，说明用户期待APP能培养批判性思维、数字素养等现代能力。
    - **学习材料水平**（9.17%）和**内容严格性**（10.54%）关注度较低，可能反映用户对现有内容质量基本满意。
  - 技术设计下细分维度
    - **屏幕设计**（38.49%）和**交互性**（28.93%）是关键，用户对界面操作体验要求较高。
    - **目标导向**（11.47%）关注度最低，暗示学习目标引导功能未被充分感知。

2. 情感反馈的两极分化

APP-各特征维度情感分析



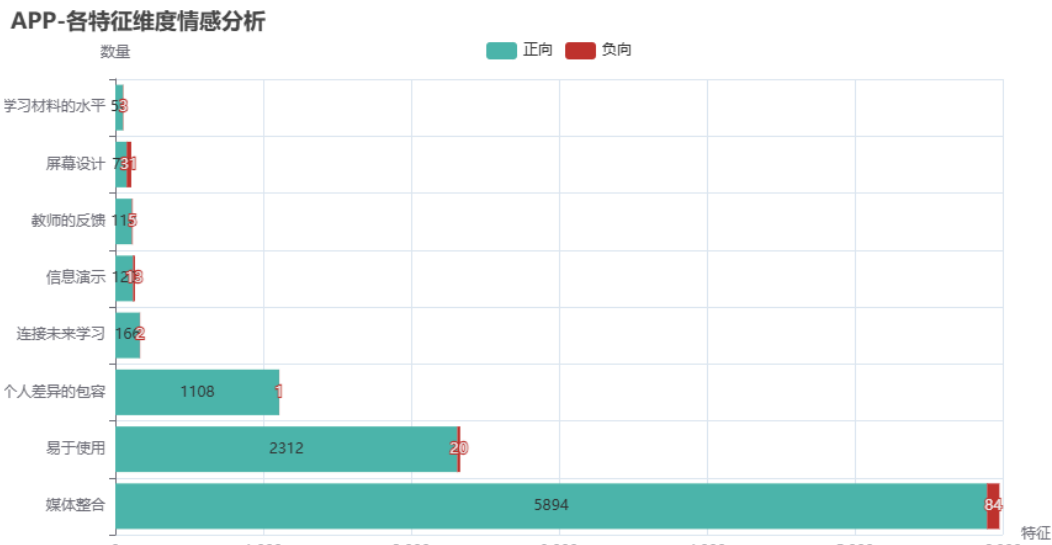
- 正向情感集中领域
  - **易于使用**（正向2173）的正向认可度最高，说明APP的基础操作体验流畅。
  - **屏幕设计**（正向1248）和**个人喜好**（正向1427）表现良好，用户对视觉设计和部分个性化功能满意。
- 负向情感风险点
  - **个人喜好**（负向309）的负向情绪最突出，用户可能对主题定制、学习模式灵活性不满。
  - **交互性**（负向174）和**实用**（负向132）的负向反馈提示功能细节需优化（如手势操作、生词本管理）。

核心结论

1. 技术设计主导用户关注

- 一级维度对比：
  - **技术设计**（85.51%关注度）是用户最核心的关注领域，远超**面向教育**（14.49%）。
  - 技术设计下细分维度
    - **媒体整合**（69.91%）和**易于使用**（27.27%）是绝对重点，用户对功能整合与操作便捷性需求强烈。
    - **屏幕设计**（1.26%）和**信息演示**（1.56%）关注度极低，但情感反馈负面集中（负向31/18），存在显著体验短板。
- 面向教育下细分维度：
  - **个人差异的包容**（76.54%）占据绝对主导，用户高度期待个性化学习支持（如难度分级、学习路径定制）。
  - **连接未来学习**（11.59%）和**教师的反馈**（8.07%）关注度低，但前者情感正向率100%（正向162），后者负向率高达83%（负向5/总6），反映隐性矛盾。

2. 情感反馈的两极分化



- 高认可领域
  - **易于使用**（正向2312，负向20）：操作流畅性广受好评，但少量负向反馈可能涉及特定场景（如离线功能卡顿）。
  - **个人差异的包容**（正向1108，负向1）：个性化功能满意度极高，是核心竞争力。
- 高风险领域
  - **屏幕设计**（负向31，正向7）：界面设计引发大量不满，可能涉及布局混乱或视觉疲劳。
  - **教师的反馈**（负向5，正向1）：用户对人工辅导或智能答疑功能失望，需紧急优化。

、关键结论

1. 艺术审美主导用户体验

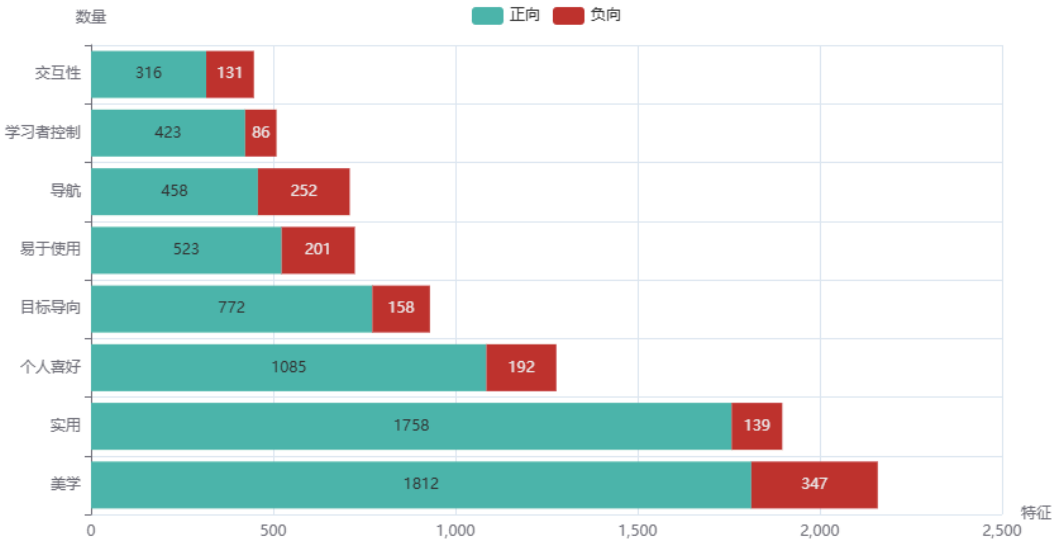
- 一级维度对比
  - **艺术审美**（67.51%）的关注度是**技术设计**（32.49%）的 **2.08倍**，说明用户更关注界面美感与功能实用性。
- 艺术审美下细分维度
  - **美学**（36.96%）和**实用**（32.47%）是核心，用户既要求视觉设计（如配色、图标），又重视功能实用性（如生词本导出）。
  - **个人喜好**（21.86%）的较高占比表明用户期待个性化设置（如主题切换、学习模式定制）。

2. 技术设计需优化交互逻辑

- 技术设计下细分维度
  - **目标导向**（33.08%）关注度最高，但正向情感仅772（远低于美学1812），说明用户对学习目标达成功能（如智能计划）满意度不足。
  - **交互性**（15.90%）关注度最低且负向情感131（技术设计中最高），反映复杂操作（如复习设置路径）引发不满。

3. 情感数据矛盾点

APP-各特征维度情感分析



- 高关注≠高满意
  - **美学**正向1812（最多）但负向347（最高），用户对视觉设计整体认可，但部分用户认为界面过时或过于花哨。
  - **导航**负向情感252（技术设计中最高），提示功能入口隐蔽或跳转逻辑混乱（如“我的课程”层级过深）。