

# 李佳鑫

i.gaheun46@gmail.com · 📞 (+86)150-8466-4436 · 🌐 gaheunnii.github.io

## 🎓 教育背景

成均馆大学 (QS 123) —— 首尔, 韩国

2021 年 2 月 – 2025 年 2 月

本科 · 学士 数据科学

GPA: 3.63/4.5 ( 91.3/100 )

相关课程: 数据科学与 Python、数据结构与算法、计算机结构与系统、操作系统、机器学习导论、数据挖掘、数据库、计算机网络、软件工程等

## 🛠️ 项目经历

基于主题建模的韩中关系问题分析

2024 年 3 月 – 2024 年 6 月

分析了 2019 年至 2023 年与韩国和中国相关的网络新闻评论, 以揭示两国公民对韩中关系发展的情感态度。关键内容包括:

- 使用网络爬虫技术从 Naver 新闻和新浪微博收集并处理了 28,995 条评论。
- 应用潜在狄利克雷分配 (LDA) 识别突出的讨论主题, 包括政治紧张、经济问题和文化摩擦等领域。
- 韩语评论利用 KNU 韩语情绪词典, 中文评论使用哈工大情感词典, 进行情绪分析以识别积极和消极情绪。
- 使用词云、Matplotlib 等工具将结果可视化, 突出显示常见的讨论主题, 帮助理解公众舆论趋势。

➤ 结果显示, 负面情绪明显上升, 尤其是在经济民族主义、COVID-19 指责和地缘政治冲突等问题上。

使用协同过滤和 PageRank 的啤酒推荐系统

2023 年 9 月 – 2023 年 11 月

开发了基于用户评分的啤酒推荐系统, 以增强个性化推荐。关键组件包括:

- 基于用户的历史评分数据, 利用协同过滤和余弦相似度进行个性化推荐。
- 通过构建评分网络, 应用个性化 PageRank 算法优化啤酒的排名, 提升了推荐结果的相关性。
- 通过 A/B 测试和 t 检验验证了高准确度。
- 通过多元线性回归分析啤酒属性 (如味道、外观、香气), 实现了一个 R 平方值为 0.658 的模型。

➤ 强调通过量身定制的推荐来捕捉用户偏好和提高用户满意度的能力。

消费者情绪分析 - Flipkart 产品评论分析

2024 年 2 月 – 2024 年 4 月

旨在分析 Flipkart 电商平台上的产品评论, 以提取消费者的情绪趋势和关键关注点:

- 自然语言处理 (NLP): 预处理数据, 应用如 TF-IDF 等文本处理技术, 以提取评论中的情感特征。
- 情绪分析: 应用随机森林模型, 识别评论中的情绪, 结合 T-SNE 方法进行可视化分析。
- 数据可视化: 通过 Matplotlib、Seaborn 等工具, 将分析结果以图表形式呈现。

➤ 识别出消费者在 Flipkart 平台上关注的主要问题, 包括产品、质量及价格及等方面。提供了可视化报告, 揭示了不同产品类别的情感分布, 帮助企业优化营销策略并提升用户满意度。

## 🔧 IT 技能

- 技术技能: 熟悉 Python、SQL、C/C++, 熟悉 VS Code、Git、AWS、MySQL、MongoDB 等工具。
- 数据分析与机器学习: 熟练使用 sklearn、numpy、pandas 进行数据分析, 掌握 matplotlib、wordcloud 可视化技术, 擅长 PCA、FA、PIE 等分析方法及 RandomForest 等机器学习算法。

## ♡ 获奖情况

国际学生奖学金 (Top 18%)

2024 秋季, 2023 秋季, 2022 春季

国际学生奖学金 (Top 6%)

2021 秋季

## 📌 其他

- 语言: 英语 (IELTS 6.5), 韩语 (Topik 6)
- 兴趣爱好: 语言学习 (在校间曾运营语言交换社团), 滑冰, 字幕译制 (曾加入字幕组)