
开源社区运营活动评估

目录

- 开源社区运营活动评估..... 1
 - 1 项目信息..... 1
 - 1.1 项目名称..... 1
 - 1.2 方案描述..... 1
 - 1.2.1 数据获取..... 2
 - 1.2.2 特征计算..... 2
 - 1.2.3 因果推断..... 3
 - 1.2.4 评价算法..... 4
 - 1.3 时间规划..... 4
 - 2 运营评估体系的构建..... 5
 - 2.1 运营活动效果评估..... 5
 - 2.1.1 五步走策略： 5
 - 2.1.2 当前商业分析团队评估方法： 5
 - 2.2 指标体系..... 6
 - 2.1.2 评价指标构建原则： 6
 - 2.1.3 指标构建方法： 6
 - 2.1.4 指标表..... 6
 - 3 评价算法及结果..... 9
 - 3.1 评价算法..... 9
 - 3.1.1 层次分析法： 9
 - 3.1.2 模糊评价法： 11
 - 3.1.3 秩和比法： 11
 - 3.1.4 综合指数法： 11
 - 3.1.5 逼近理想解排序法： 12
 - 3.2 评价结果..... 12
 - 3.3 指标优化结果..... 15
 - 4 项目总结..... 17
 - 4.1 项目产出..... 17
 - 4.2 方案进度..... 17
 - 4.3 遇到的问题及解决方案..... 17
 - 4.4 项目完成质量..... 17
 - 4.5 与导师沟通及反馈情况..... 17
 - 参考文献..... 18

1 项目信息

1.1 项目名称

名称：开源社区运营活动评估

项目描述：活动是推广的一种手段，活动的目的就是让更多的人了解我们的产品，使用我们的产品，为我们社区做出贡献，成为我们社区的一员。什么样的活动才能吸引更多的人才？举办此次活动的到底有多大的影响力？我们就需要评价这次活动的价值。活动效果的评估，不仅仅是为了体现此次活动的价值，也要能够指导我们进行下一次活动的方案优化。什么样的活动更吸引人，什么类型的人来参加更合适，在什么地方举办活动更好。活动包括：B 站、抖音、线下 meetup、各种会议、比赛等等 本任务会提供部分活动的样本数据。

1.2 方案描述

针对每场运营活动，活动都需要有进行有针对性的分析，来给之后的工作做指导。对本项目的主要产出要求是：

- 输出活动评价的方案设计文档，详细的代码实现。
- 输出此次活动的优势、缺点分析。
- 对下一次活动的活动效果预测准确率达到 50%以上。

运营活动在本项目中主要分为线上活动与线下活动两种，两类项目有不同的指标类型。本项目建立了运营地图，如图 1-1 所示，作为项目的大纲及指导；同时，运营地图也有单独的文件“活动运营地图”，可做进一步学习使用。



图 1-1，活动运营地图

以线上活动，2020 年高校开发者大赛为例，项目分析方案主要分为四部分：

1.2.1 数据获取

数据获取的来源可分为三部分，分别是：

- 参赛者：包括活动报名名单、参赛人员信息。
- 组织者：包括活动举办方历史信息、历史活动活跃人数信息。
- 活动类型：包括活动举办的时间、领域、范围等信息。

对于本项目中所需要的数据，赛题主题、参赛人数等可以通过赛方网站得到，社区代码相关的数据，包括社区历史活跃数可以由爬虫提取，其他一些无法得到的信息，包括学校信息等，可以有进行噪声模拟。

1.2.2 特征计算

获取数据后，再利用特征工程的方法，进行数据的计算。特征工程的框架图如图 1-2 所示：

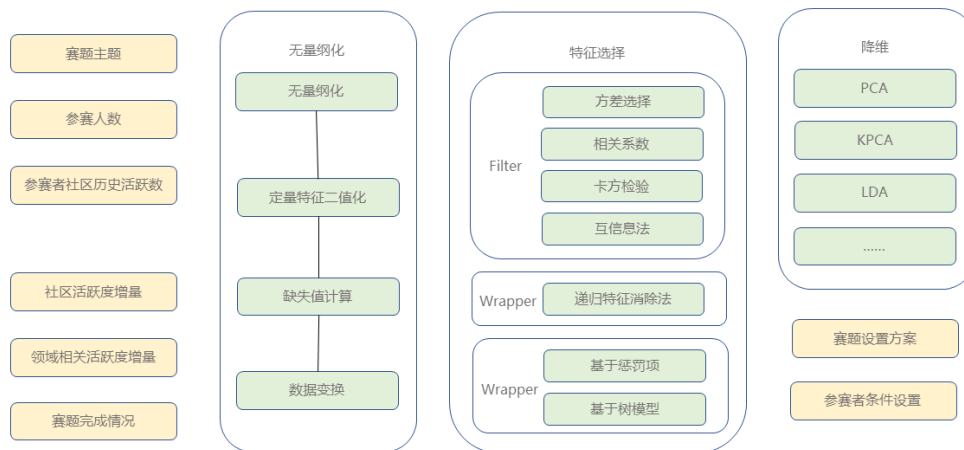


图 1-2，特征工程框架

特征处理是特征工程的核心部分，`sklearn` 提供了较为完整的特征处理方法，包括数据预处理，特征选择，降维等。因此本项目主要使用 `sklearn` 进行特征处理。框架图图 1 中，绿色部分是算法部分，可以直接复用。黄色的指标可以依靠数据量来改变。目前已经得到的数据包括黄色指标部分，其他指标之后可以进行进一步挖掘或者模拟得到。

首先进行无量纲化，无量纲化使不同规格的数据转换到同一规格。常见的无量纲化方法有标准化和区间缩放法。标准化的前提是特征值服从正态分布，标准化后，其转换成标准正态分布。区间缩放法利用了边界值信息，将特征的取值区间缩放到某个特点的范围，例如`[0, 1]`等。同时优于部分数据信息不能直接使用，还需要将定性数据转化为定量数据，进行定量特征二值化处理。最后，再对缺失的数据进行筛选。

当数据预处理完成后，我们需要选择有意义的特征输入机器学习的算法和模型进行训练。通常来说，从两个方面考虑来选择特征：

- 特征是否发散：如果一个特征不发散，例如方差接近于 0，也就是说样本在这个特征上基本上没有差异，这个特征对于样本的区分并没有什么用。
- 特征与目标的相关性：这点比较显见，与目标相关性高的特征，应当优选选择。除方差法外，本文介绍的其他方法均从相关性考虑。

根据特征选择的形式又可以将特征选择方法分为 3 种：

- **Filter:** 过滤法，按照发散性或者相关性对各个特征进行评分，设定阈值或者待选择阈值的个数，选择特征。
- **Wrapper:** 包装法，根据目标函数（通常是预测效果评分），每次选择若干特征，或者排除若干特征。
- **Embedded:** 集成法，先使用某些机器学习的算法和模型进行训练，得到各个特征的权值系数，根据系数从大到小选择特征。类似于 **Filter** 方法，但是是通过训练来确定特征的优劣。

当特征选择完成后，可以直接训练模型了，但是可能由于特征矩阵过大，导致计算量大，训练时间长的问题，因此降低特征矩阵维度也是必不可少的。常见的降维方法除了以上提到的基于 L1 惩罚项的模型以外，另外还有主成分分析法（PCA）和线性判别分析（LDA），线性判别分析本身也是一个分类模型。PCA 和 LDA 有很多的相似点，其本质是要将原始的样本映射到维度更低的样本空间中，但是 PCA 和 LDA 的映射目标不一样：PCA 是为了让映射后的样本具有最大的发散性；而 LDA 是为了让映射后的样本有最好的分类性能。所以说 PCA 是一种无监督的降维方法，而 LDA 是一种有监督的降维方法。

通过特征工程计算后，已经得到某些指标的贡献以及余结果的相关关系。例如参赛赛题的领域与社区活跃度的相关关系等。但是要想得到进一步的关系，还需要进行因果推断计算。

1.2.3 因果推断

因果推断的框架如图 1-3 所示：

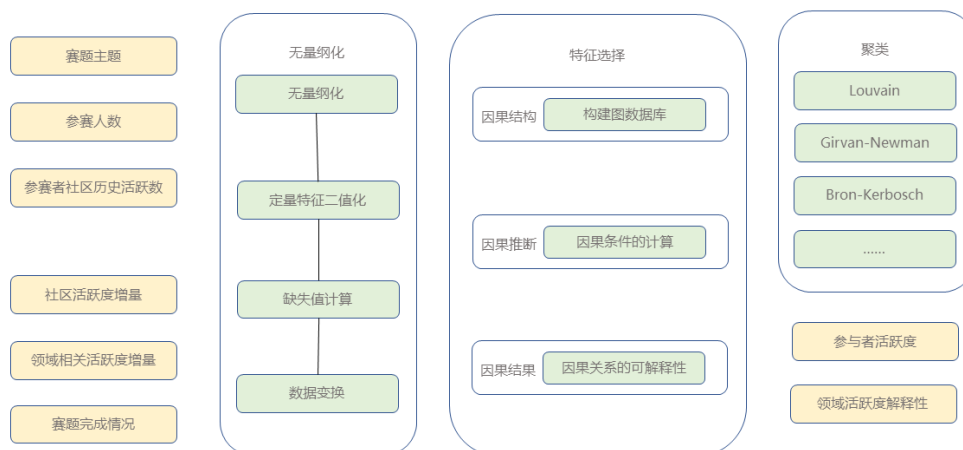


图 1-3，因果推断框架

- **构建知识图谱**

基于数据类型及其相互之间的关系，利用 neo4j 图数据库，构建对应的知识图谱。从原始的数据到形成知识图谱，需要经历知识抽取、知识融合（实体对齐）、数据模型构建、质量评估等步骤。

- **图聚类**

通过社区发现算法，挖掘出具有相似的用户群体。基于群体，可以进行数据补全与预测等操作。图聚类的主要算法包括：

- ✓ **Louvain 算法**
- ✓ **Girvan-Newman 算法**

- ✓ 层次聚类
- ✓ 谱聚类
- ✓ Bron-Kerbosch 算法

● 因果推断及验证

基于前门准则与后门准则等推断原理，基于已有数据，推断出用户行为、信息等导致结果的混杂因素、直接影响因素等。因果推断的两个主要的优势结果在于：

- ✓ 可解释性：可使运营者对指标之间的关系一目了然。
- ✓ 知识推断：基于已有数据，可对图谱进行某些指标的预测与补全。

1.2.4 评价算法

从特征工程得到指标数据后，可再根据数据类型、适应场景，进行活动阶段评价。活动主要分为五个阶段，包括知晓、了解、获取、参与、成长。每个阶段对应的指标不同，得到的特征数据不同，所适应的评价算法也不同。适合的活动运营评价算法主要包括有层次分析法、模糊评价法、秩和比法、综合指数法和逼近理想解排序法。本文第三节将介绍具体的算法内容与案例结果。最后通过可视化方法，展示各阶段评价结果。

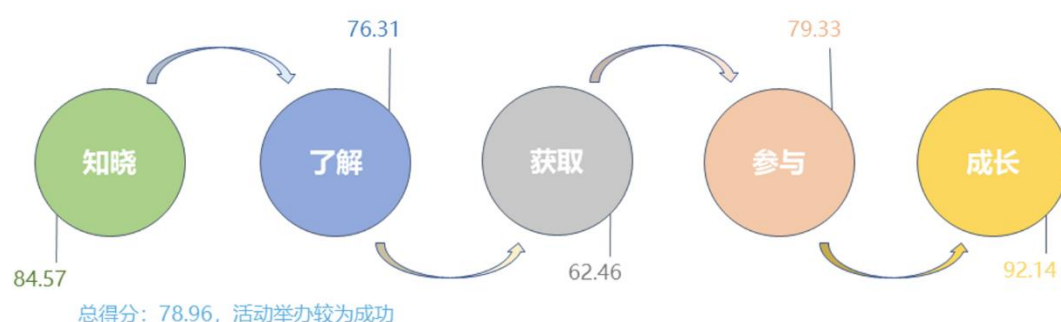


图 1-4，各阶段评价分数

1.3 时间规划

2021.6.30-2021.8.15:

完成指标体系、评价方案的构建。

2021.8.16-2021.8.30:

完善结果展示代码。

2021.8.31-2021.9.30:

项目测试与改进。

2 运营评估体系的构建

2.1 运营活动效果评估

2.1.1 五步走策略：

在当前的活动运营中，主要遵循的是五步走策略：

- 第一步，明确活动的预算和目标。
- 第二步，明确活动涉及的部门和活动的执行流程。

在这两个步骤中，需要和发起活动的部门做深入沟通，对活动的预算、活动的目标、活动的流程和参与的部门做到详细的了解。

- 第三步，设置准确科学的结果指标和过程指标。

根据业务的目标和流程去进行指标的合理设置，这既包括结果指标的设置，也包括过程指标的设置；运营活动的目标不同，设置的指标也是不同的。

- 第四步，选择合理的评估方法，利用步骤三的指标数值或变化对效果进行评估。

选择合理的评估方法对步骤三的指标进行监控和评估，通过指标的数值或变化情况判断活动效果的好坏。

- 第五步，通过评估和分析发现活动的问题点和启发点，提出合理的建议，为下次活动或其他活动提供策略和建议。

2.1.2 当前商业分析团队评估方法：

对于活动指标变化的评价，当前商业团队主要采用一些简单的前后对比，采用定性的方法居多，定量手段较为缺乏。主要的一些评估方法有：

- 纵向对比法

通过对比活动前和活动后结果指标的变化，判断运营活动是有效果。纵向对比一定要保证活动前后除运营活动外，其他的影响因素都是稳定或相同的；另外，还要注意前后的对比，在周期上一定要跨满一个周期。

- 横向对比法

通过实验组和对照组的设置和对比，判断运营活动是否有效果。只要样本选取的正确，横向对比的效果要好于纵向对比。

- 综合对比法

同时结合横向对比和纵向对比，判断运营活动是否有效果。

- 目标对比法

和活动刚开始设定的目标对比，判断运营活动的效果是否达到了预期。

- 活动与活动对比法

利用相似活动的横向对比，判断运营活动的效果，其实也是一种横向对比的方法。难点在于：很难定位相似的活动，有的活动其目标就是不一样的；有的即使目标一样，可能活动针对的用户群体、活动的时间点、活动的流程、活动的地点不一样，这导致很难 apple to apple 的对比，会产生较大的误差。

- 投入产出比法

利用投入的预算和产出计算 ROI，判断投入产出比是否合理^[1]。

2.2 指标体系

2.1.2 评价指标构建原则：

- 科学性和客观性原则

只有保障了评价指标的科学性和客观性，评价指标才能真正具有实际价值。所以在构建评价指标体系的过程中，时刻以科学的理论为指导，主要表现在使用科学理论的基础上选取评价指标，确定科学的评价方法。

- 系统性原则

活动的成果，包括像影响力的形成等是一个系统的过程，在这个过程中的各环节之间有着紧密的联系。因此在构建评价模型时需要坚持系统论的观点，将运营活动结果作为一个系统进行考察，而不是只单纯地以结果性指标来评价，是由促成结果的原因性指标和结果性指标共同构成评价模型。

- 简明性与可行性原则

运营活动的评价模型中的指标还应该避免重复、冗繁，并且综合考虑指标如何去量化、资料获取是否存在局限等因素。保证整体活动的评估易于操作^[2]。

2.1.3 指标构建方法：

- 文献调查法

本研究搜集网上已有文献资料，对已有的活动运营学术研究、研究成果进行搜集，梳理和分析现有的相关评价指标体系，从而为本研究构建出运营活动评价指标体系提供来源依据。

- 德尔菲法

本研究综合多位经验丰富的运营行业从业人员及专家，综合考虑他们的意见，减少个人主观因素的影响，从而筛选出初始指标^[2]。

2.1.4 指标表

依托上文方法、搜寻得到的文献，并根据实际活动抽象出模型，按照指标来源、活动对应阶段，可以得到得到如表 1 所示的指标表。

表 1. 按阶段构建的活动指标

	一级指标	二级指标	三级指标
	知晓	媒体接触度	媒体级别
			活动发文数量
			发文时间
			关注数量
		媒体认知度	内容针对性
			最高阅读量
			平均阅读量
			最高点赞数
			平均点赞数

社区运营活动指标		说服度	文章收藏数
			文章留言数
			留言点赞数
		二次传播度	被转发数
			被转载数
	了解	渠道	官网
			公众号
			头条
			B 站
			搜索引擎
		学习	文档访问量
			慕课观看量
		互动	新闻
			博客
		接触度、认知度、说服度、二次传播度	知晓相关指标
	获取	下载	不同版本软件包下载量
			不同版本软件包讨论量
			镜像仓库访问量
		互动	直播参与人数
			直播回顾人数
			直播讨论数
			沙龙参与人数
			会议参与人数
	参与	开源社区项目状态	Pull request 数 ^[3]
			Commit 数
			Issue 提交数与评论数
			Star 数
			Fork 数
		活动参与情况	参与人数
			参与范围
			活动主题
		开源社区贡献者趋势	贡献者活跃时间
			活动时间内贡献者增量
			活动贡献者停止贡献量
	成长	技术交流与输出情况	国际论文
			讲座
		企业/组织分布与劳动力投入情况	投入成本
		贡献情况	仓库总量
			仓库新增量
			CLA 签署量 ^[4]
			CLA 签署增量

			开源社区贡献者趋势
			开源社区状态

按照运营活动指标的来源阶段，可以将指标分为五类。其中活动前期的知晓、了解阶段你，主要涉及到活动的推广，与媒体推广的很多指标息息相关，包括有媒体的接触度^[5]，接触度主要涉及媒体级别、活动发文数量、发文时间、关注数量等，会使用户与活动信息更有机会直接接触。此外认知度（内容针对性、最高阅读量、平均阅读量）、说服力（最高点赞数、平均点赞数、文章收藏数、文章留言数、留言点赞数）、二次传播度（被转发数、被转载数），则是媒体的传播属性。了解阶段还涉及流量通道的问题，即用户是通过哪个渠道来参与活动的，还有学习、互动平台的指标划分。

获取、参与和成长阶段，则主要与社区自身的建设有关。获取阶段用户在下载（不同版本软件包下载量、不同版本软件包讨论量、镜像仓库访问量）、互动（直播参与人数、直播回顾人数、直播讨论数、沙龙参与人数、会议参与人数）方面的指标^[6]，体现了社区在资源获取方面的建设情况。参与阶段，主要从开源社区项目状态（Pull request 数、Commit 数、Issue 提交数与评论数、Star 与 Fork 数）、活动参与情况（参与人数、参与范围、活动主题）、开源社区贡献者趋势（贡献者活跃时间、活动时间内贡献者增量、活动时间贡献者停止贡献量）三方面进行指标评价。最后在成长阶段，用户已经参与到社区活动中，并具有一定经验，则需从技术交流与输出情况（国际论文、讲座）、企业/组织分布与劳动力投入情况、贡献情况（仓库总量、仓库新增量、CLA 签署量、CLA 签署增量、开源社区贡献者趋势、开源社区状态）等社区成果展现方面进行评价。

3 评价算法及结果

3.1 评价算法

3.1.1 层次分析法:

层次分析法以系统分层分析为手段，对评价对象总目标进行连续性分解，通过两两比较确定各层子目标权重，并以最下层目标组合权重定权，加权求出综合指数，再以综合指数大小评定目标实现情况。将待评价对象构造为一个多层次分析评价模型，要求专家根据经验对同层指标分 1~9 个重要程度进行比较打分，构造判断矩阵确定各指标间相对重要度，由此计算各指标权重并进行一致性检验，充分利用了专家经验。

算法步骤:

(1) 因素层次分析结构模型建立

因素层次分析结构模型的建立是在活动指标提取的基础上，把提取出的指标分成若干个组，然后根据模型层次结构模型的建立规则，形成不同层次。在层次结构模型中，一般地，最底层为因素层，包含所有要进行排序的活动指标；中间层为准则层，对下层元素起支配作用，同时又受到上层元素的支配，来确定元素排序的准则；最高层只有一个元素，称为目标层。

针对活动指标的特点，可以建立如下的层次分析结构模型：

- 1) 目标层：表示活动指标要达到的目标。
- 2) 准则层：一般从指标发生概率和指标的结果影响程度作为准则。
- 3) 因素层：与活动阶段相关的因素。

(2) 构造互补判断矩阵

本文因素间的两两对比，采用的是一个因素和另一个因素相比的重要程度的定量比较，从而可以得到判断矩阵 A：

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix} \quad (3-1)$$

对矩阵各行求和：

$$r_i = \sum_{k=1}^n a_{ik} \quad (3-2)$$

做数学变换：

$$r_{ij} = \frac{r_i - r_j}{2(n-1)} + 0.5 \quad (3-3)$$

从而得到模糊一致性矩阵 R，其中元素为 r_{ij} 。对 R 进行行和归一化处理，得到因素排序向量 W：

$$W_i = \frac{\sum_{j=1}^n a_{ij} - 1 + n/2}{n(n-1)} \quad (3-4)$$

计算出特征矩阵：

$$W_{ij} = \frac{w_i}{w_i + w_j} \quad (3-5)$$

(3) 模糊互补判断矩阵一致性检验

判断得到的因素的重要性权重值是否合理，还应该对模糊互补判断矩阵进行一致性检验。本文选用相容性指标 I 进行检验。其中 I 的表达式为：

$$I(A, B) = \frac{1}{n^2} \sum_{i=1}^n \sum_{j=1}^n |a_{ij} + b_{ij} - 1| \quad (3-5)$$

当 $I \leq \alpha$ (α 为决策者态度) 时，则认为判断矩阵满足一致性。

(4) 排序

1) 层次单排序

根据排序向量 W 得到某层上的元素相对上一层某相关元素的相对权重。

2) 层次总排序

风险因素层次总排序指计算所有元素相对于目标层的相对重要性的排序权重，这一过程从最高层开始逐层向最低层进行。上一层为 A ，下一层次为 B ， A 层中风险因素在本层所占的权重为 a_{ij} ， B 层中相对于风险因素层次单排序重要性权重为 b_{ij} ，则 B 层总权重可以按照下面式子来计算：

$$b_j = \sum_{k=1}^n a_k b_{kj} \quad (3-6)$$

最后可以依据权重，进行评价模型总分数的计算。

算法优劣势及适用场景分析：

算法优势：

- 分析方法更为系统。
- 方法简洁且更为客观。
- 所需定量数据信息较少。

算法劣势：

- 不能为决策提供新方案。
- 定量数据较少，定性成分多，不易令人信服。
- 指标过多时数据统计量大，且权重难以确定。

适用场景：

适用于定量数据较难确定的**获取、成长**环节。

3.1.2 模糊评价法:

模糊评价法主要是将模糊因素集对应的模糊权向量集, 依靠单因素评判矩阵采取合适的合成算子, 进行模糊变换, 得到一个模糊综合评价结果, 并对结果进行比较分析, 来评价事物的优劣。一般模糊评价法会和层次分析法、综合指数法等方法联合使用。

算法优劣势及适用场景分析:

算法优势:

- 通过精确数字手段处理模糊评价对象, 量化评价更为合理。
- 评价结果是矢量, 信息更为丰富合理。

算法劣势:

- 计算较为复杂, 对指标权重矢量的确定主观性较强。
- 指标过多时数据统计量大, 且权重难以确定。

适用场景:

适用于权重能定量, 但是不能精确的**知晓、了解**环节。

3.1.3 秩和比法:

秩和比法以秩和法(Rank-sum ratio, RSR)为基础, 取各指标数与个体数秩和的平均值, 得到具有 0~1 连续变量特征的非参数统计量, 即秩和比, 再据此进行排序。

秩和比指的是表中行(或列)秩次合计的平均值或加权平均值, 是一个非参数计量的综合指数, 具有 0~1 区间连续变量的特征, 其所有比较组秩和比之和为 $(n+1)/2$; 如果编秩不按照经典的秩变换方法, 各组秩和比的合计可能不为 $(n+1)/2$ 。其基本思想是在一个 n 行 (n 评价对象) m 列 (m 个评价指标或等级) 矩阵中, 通过秩转换, 获得无量纲的统计量 RSR, 以 RSR 值对评价对象的优劣进行排序, 进而根据比较组数的多少, 进行分档处理 (比较组数较多) 或进行 RSR 平方根反正弦变换值可信区间处理 (比较组数较少)。

算法优劣势及适用场景分析:

算法优势:

- 不引入主观变量, 克服了主观定权的问题。
- 综合能力强, 可作为专门的综合指标进行统计分析。
- 可进行分档排序, 消除异常值干扰, 显示数据间的微小差异。

算法劣势:

- 指标值进行秩代换会损失信息, 导致对信息利用不完全。
- 对离群值不敏感。

适用场景:

适用于只需要利用次序信息的**了解**环节。

3.1.4 综合指数法:

综合指数法将不同性质、不同单位的各种实测指标值通过指数变换, 加权得出综合指数, 对综合指数进行比较与评价。

算法优劣势及适用场景分析:

算法优势:

- 评价过程系统、全面、简便

-
- 数据利用充分，通过对综合指数和个体指数分析，找出薄弱环节，为改进提供依据。

算法劣势：

- 对比较标准依赖太强，同时标准的确定较为困难
- 指标值无上下限，若存在极大值会影响评价结果的准

适用场景：

适用于能简单综合指标的**获取、成长**环节。

3.1.5 逼近理想解排序法：

逼近理想解排序法（TOPSIS 法）是基于归一化后的原始数据矩阵，找出最优与最劣方案，通过计算评价对象与两者距离，得到与两者的接近程度，从而评价各对象优劣。TOPSIS 法其中“理想解”和“负理想解”是 TOPSIS 法的两个基本概念。所谓理想解是一设想的最优的解（方案），它的各个属性值都达到各备选方案中的最好的值；而负理想解是一设想的最劣的解（方案），它的各个属性值都达到各备选方案中的最坏的值。方案排序的规则是把各备选方案与理想解和负理想解做比较，若其中有一个方案最接近理想解，而同时又远离负理想解，则该方案是备选方案中最好的方案。

算法优劣势及适用场景分析：

算法优势：

- 对样本资料无特殊要求
- 比较充分地利用了原有的数据信息，与实际情况较为吻合
- 可对每个评价对象的优劣进行排序

算法劣势：

- 当两个评价对象指标关于最优与最劣方案的连线对称时，无法得出准确的结果
- 只能对每个评价对象的优劣进行排序，不能分档管理，灵敏度不高

适用场景：

适用于能确定最优解的**了解、参与**环节。

3.2 评价结果

在知晓阶段，运用模糊评价法，得到若干个指标的权重，如图 3-1 所示：

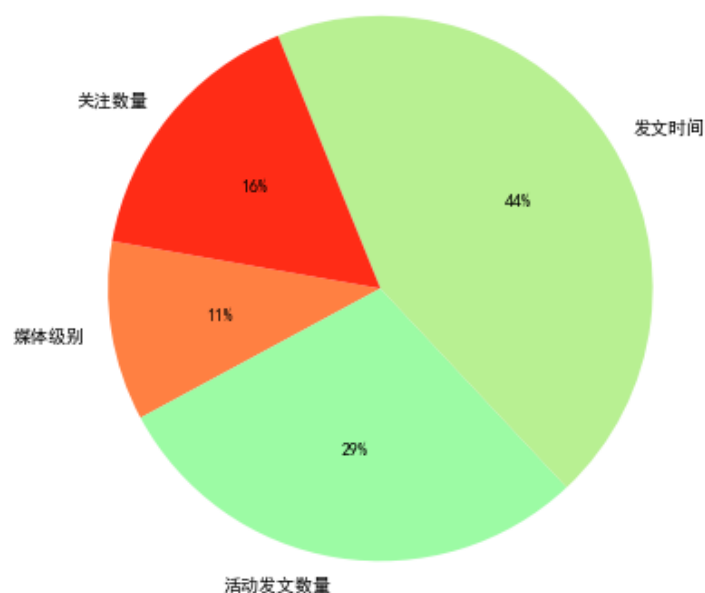


图 3-1.模糊评价法评价指标

在了解阶段，采用秩和比法进行评价。这里模拟了 8 组活动，其中第 1 组为需要计算的活动，可以得到结果：

OLS Regression Results						
Dep. Variable:	y	R-squared:	0.935			
Model:	OLS	Adj. R-squared:	0.913			
Method:	Least Squares	F-statistic:	43.23			
Date:	Thu, 30 Sep 2021	Prob (F-statistic):	0.00716			
Time:	15:08:21	Log-Likelihood:	10.810			
No. Observations:	5	AIC:	-17.62			
Df Residuals:	3	BIC:	-18.40			
Df Model:	1					
Covariance Type:	nonrobust					
	coef	std err	t	P> t	[0.025	0.975]
const	-0.0529	0.093	-0.569	0.609	-0.349	0.243
Probit	0.1098	0.017	6.575	0.007	0.057	0.163
Omnibus:	nan	Durbin-Watson:	2.998			
Prob(Omnibus):	nan	Jarque-Bera (JB):	0.138			
Skew:	0.124	Prob(JB):	0.933			
Kurtosis:	2.226	Cond. No.	33.3			

Notes:

[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.

回归直线方程为: $y = 0.10979548043547699 \text{ Probit} - 0.052939694021058484$

图 3-2.秩和比法评价结果 a

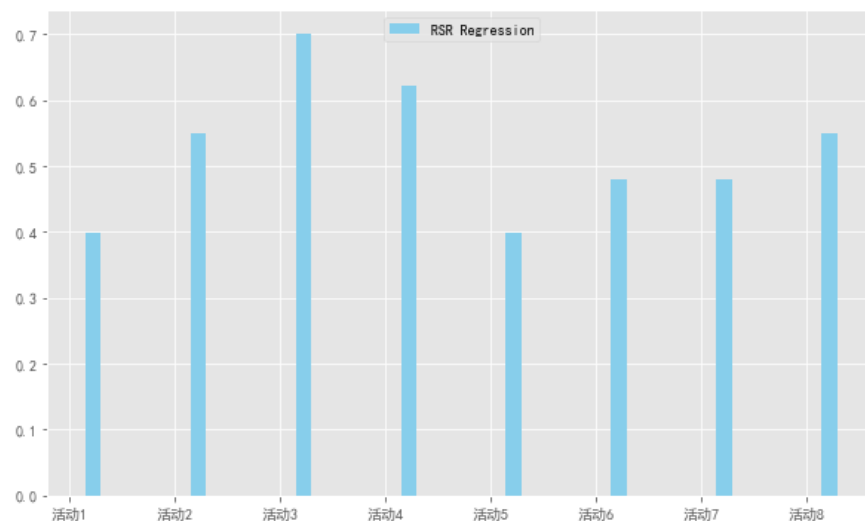


图 3-3.秩和比法评价结果 b

在获取阶段，采用层次分析法进行评价，对该阶段八个指标进行分析，得到指标权重如图 3-4 所示：

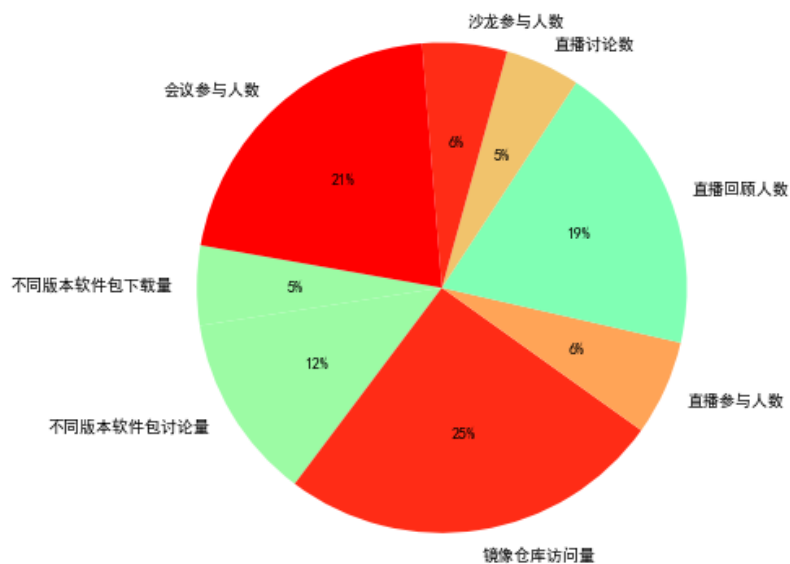


图 3-4.层析分析法指标权重

在参与阶段，采用 TOPSIS 进行评价。这里输入四个活动的数据，其中活动一为需要评价的数据，得到四个活动排序结果为：



图 3-5.TOPSIS 评价结果

成长阶段，则采用综合指数法进行评价。

最后对五个阶段进行综合，得到活动总分。并对各阶段评分进行可视化，如图 3-6 所示：

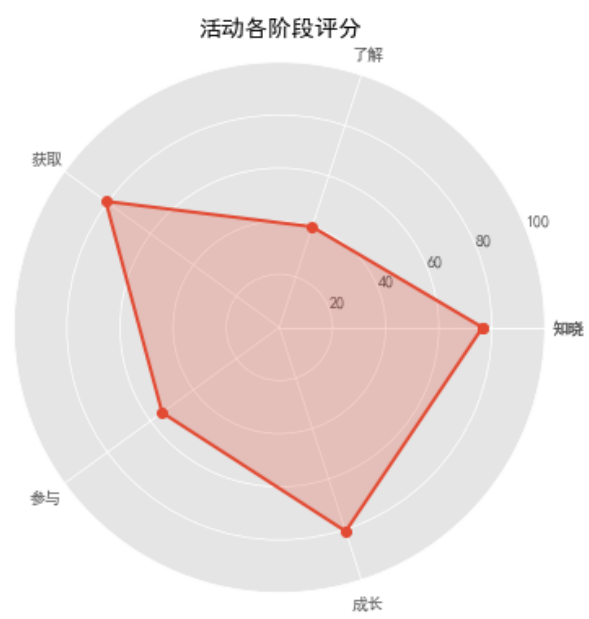


图 3-6.活动阶段评价

3.3 指标优化结果

利用高校开发者大赛的数据，可进行真实案例评价。提取活动期间社区评论数据，如图 3-7 所示：

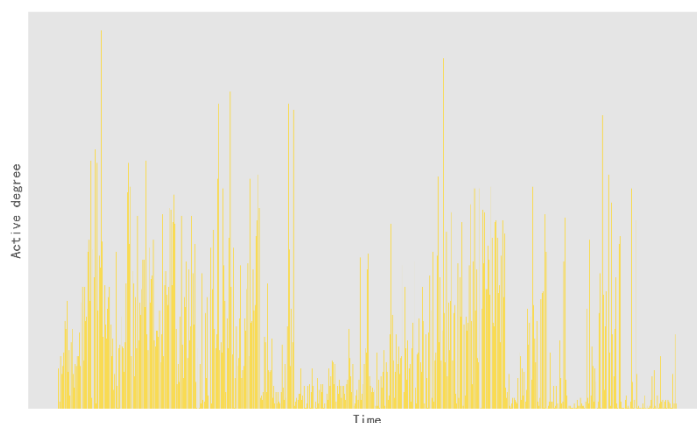


图 3-7.社区评论数据

此外还包括 issue 数据：

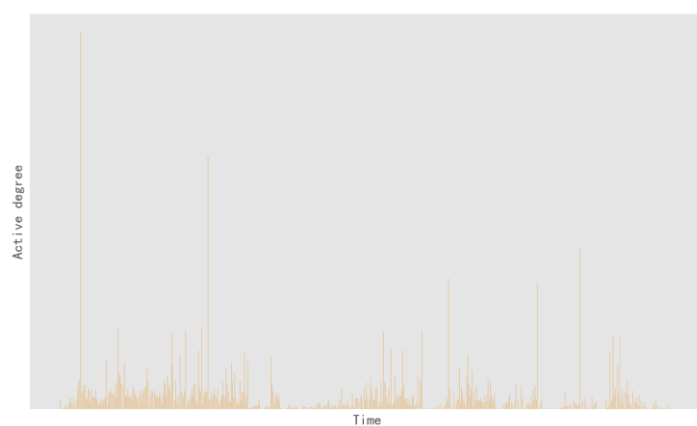


图 3-8.社区 issue 数据

可计算社区活跃度指标的增量：

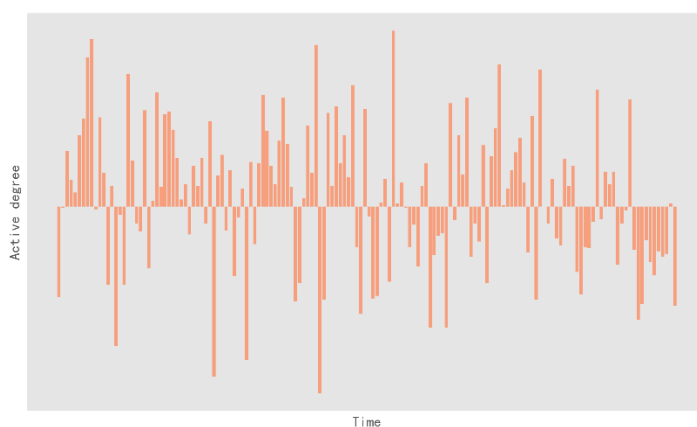


图 3-9.社区活跃度指标增量

最后结合真实数据，采用层析分析法进行分析，可得到活动评价总分。再采用 `causalml` 包进行因果推断计算，可以对指标进行平均处理效应估计。

4 项目总结

4.1 项目产出

已完成项目方案的设计与实现；
已完成项目的数据收集与处理；
已完成特征工程、因果推断、评价方案的框架计算代码；
已完成特征工程、因果推断、评价方案的案例计算及结果展示；

4.2 方案进度

进度为 100%。

4.3 遇到的问题及解决方案

遇到的主要问题为数据收集、数据质控较难实现。因此本项目以模拟数据为主，实现评价方案分析。

4.4 项目完成质量

项目从活动场景入手，分五阶段进行指标构建、方案评价，并能就指标进行阶段化分析，具备较好的活动运营参考价值。

4.5 与导师沟通及反馈情况

导师指导很多，帮助很大，给予了活动运营地图的建设方向，也提供了很关键的思路，十分感谢！

参考文献

- [1] 《如何对运营活动效果进行全面有效地评估？》，<https://zhuanlan.zhihu.com/p/350396288>
- [2] 张琪，出版社微信公众号影响力评价模型构建及其应用研究[D]，武汉大学，2017.
- [3] 王晨，面向大型开源社区的缺陷数据分析与研究[D]，武汉大学，2018.
- [4] Identifying indicators for monitoring and evaluation. 2017.
- [5] 尹秀波，读者视角下高校图书馆阅读推广活动评价指标体系构建研究[J]，大学图书情报学刊，2016，Vol.34 No.1。
- [6] 盛捷，许增雯，李芮，浅谈构建银行类互联网线上营销活动运营效果后评价模型[J]，金融天地，：1671—6728(2021) 21—0078—03。