

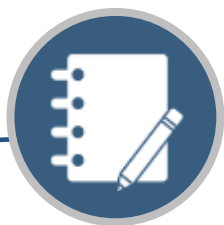
Interpretable Clustering: A Survey

汇报人：韦浩文

日期：2025/1/12



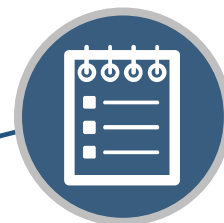
引言



相关工作



方法



文献回顾



总结



第一部分 引言



引言-研究背景



近年来，聚类算法研究重点一直放在提升算法的准确性和效率上，往往忽略了算法的可解释性。

但随着这些方法越来越多的应用于高风险领域，确保聚类算法做出的决策能够被清晰理解和合理解释，已经成为一个基本且迫切的需求。



引言-研究意义



1.对可解释聚类算法的现状进行了全面系统的综述

2.提出了一套对可解释聚类方法分类的标准



第二部分 相关工作



相关工作-可解释聚类的必要性



What is cluster analysis?

聚类分析旨在根据数据的内在特征和模式将数据划分为不同的组以揭示数据点之间有意义的结构和关系



What is interpretability?

“以人类可理解的语言解释或呈现的能力”

“人类能够理解决策原因的程度”

“使机器学习系统的行为和预测对人类可理解”



What is interpretable clustering?

传统聚类的特点：强调结果的准确性和效率，但缺乏透明度和可解释性

可解释聚类的特点：在保持聚类质量的同时，提供了对聚类过程和结果的清晰解释



What is a good interpretable clustering method?

一般来说，好的可解释聚类方法应提供简单、清晰、独特的解释

而要确定和量化可解释聚类方法的好坏，需要考虑所使用的具体的可解释模型



第三部分

方法



方法-可解释聚类方法分类法

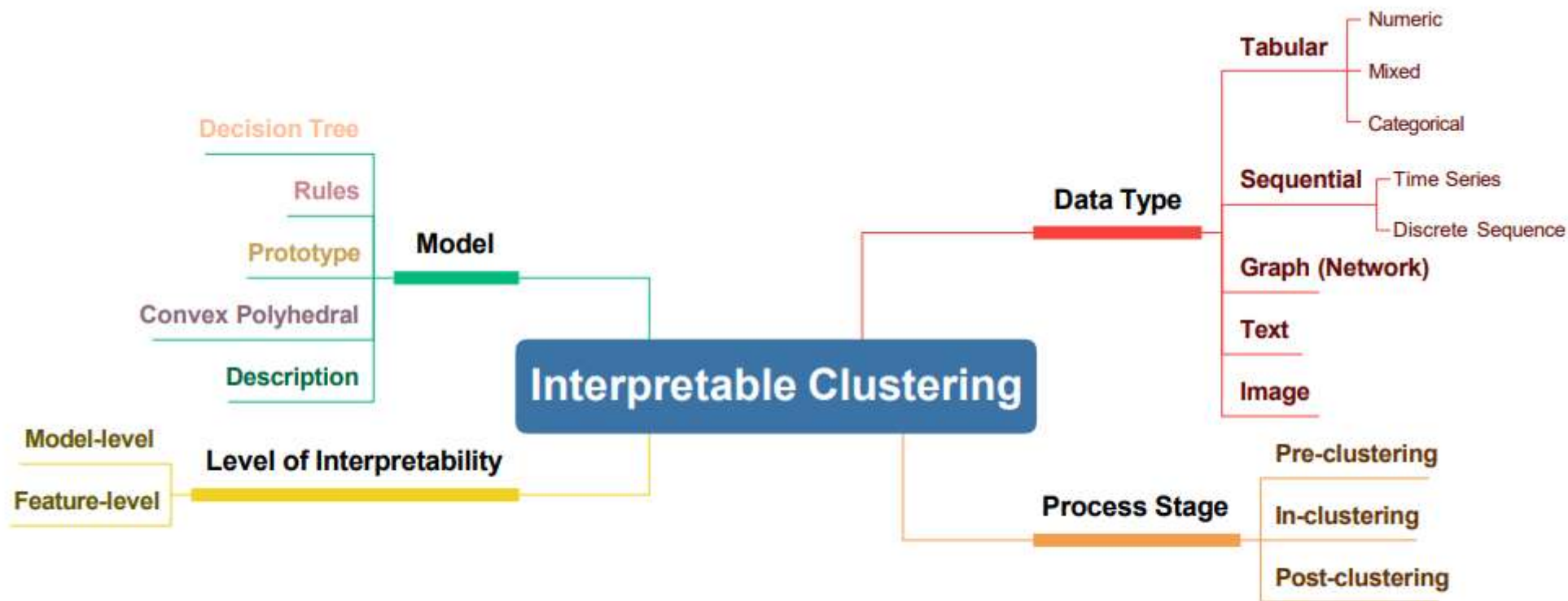


Fig. 2. Interpretable clustering taxonomy categorized by distinct criteria, most existing methods align with a single category per criterion.



方法-可解释模型

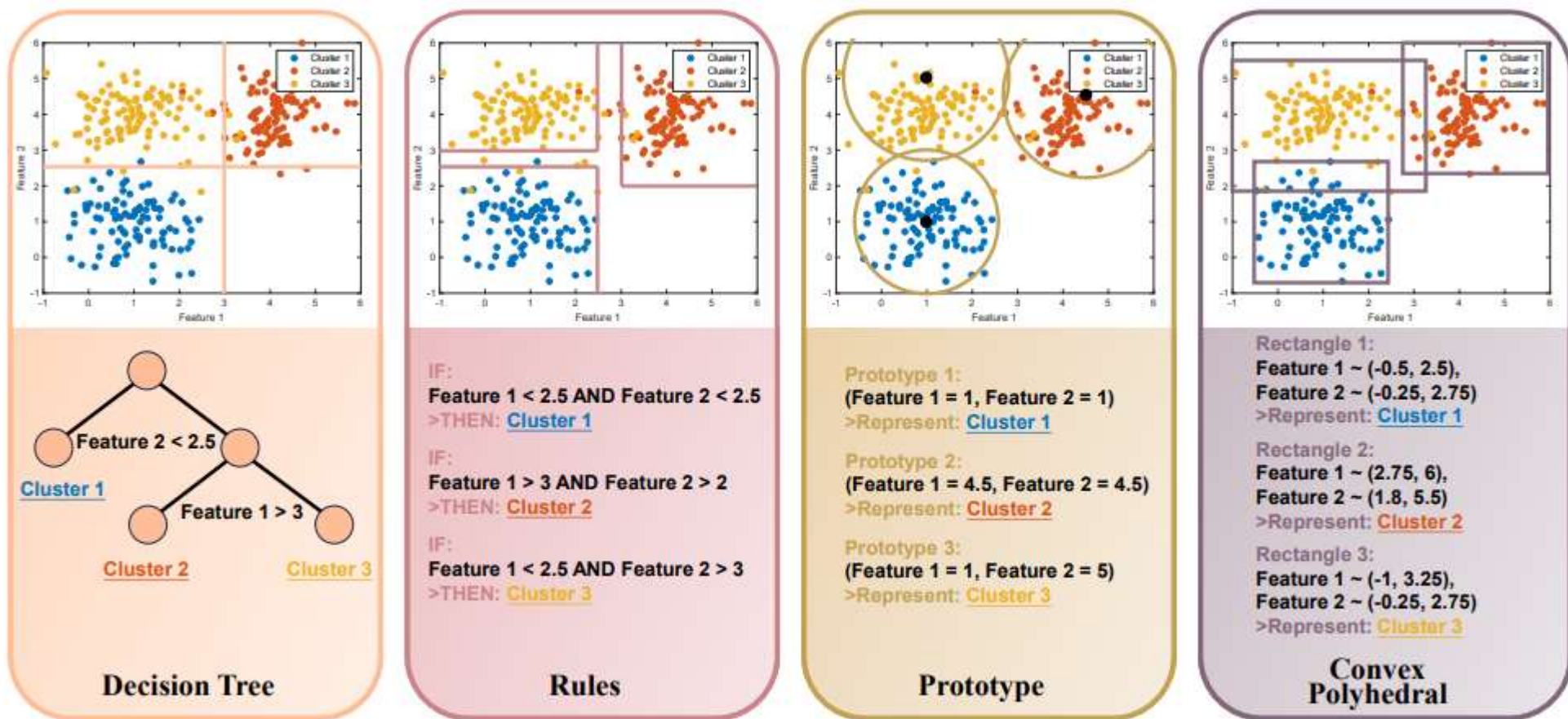


Fig. 1. Illustration of four interpretable clustering models applied to the same two-dimensional dataset with three Gaussian clusters. The upper panels display how each model partitions the feature space, while the bottom panels show the feature values used for interpretability.

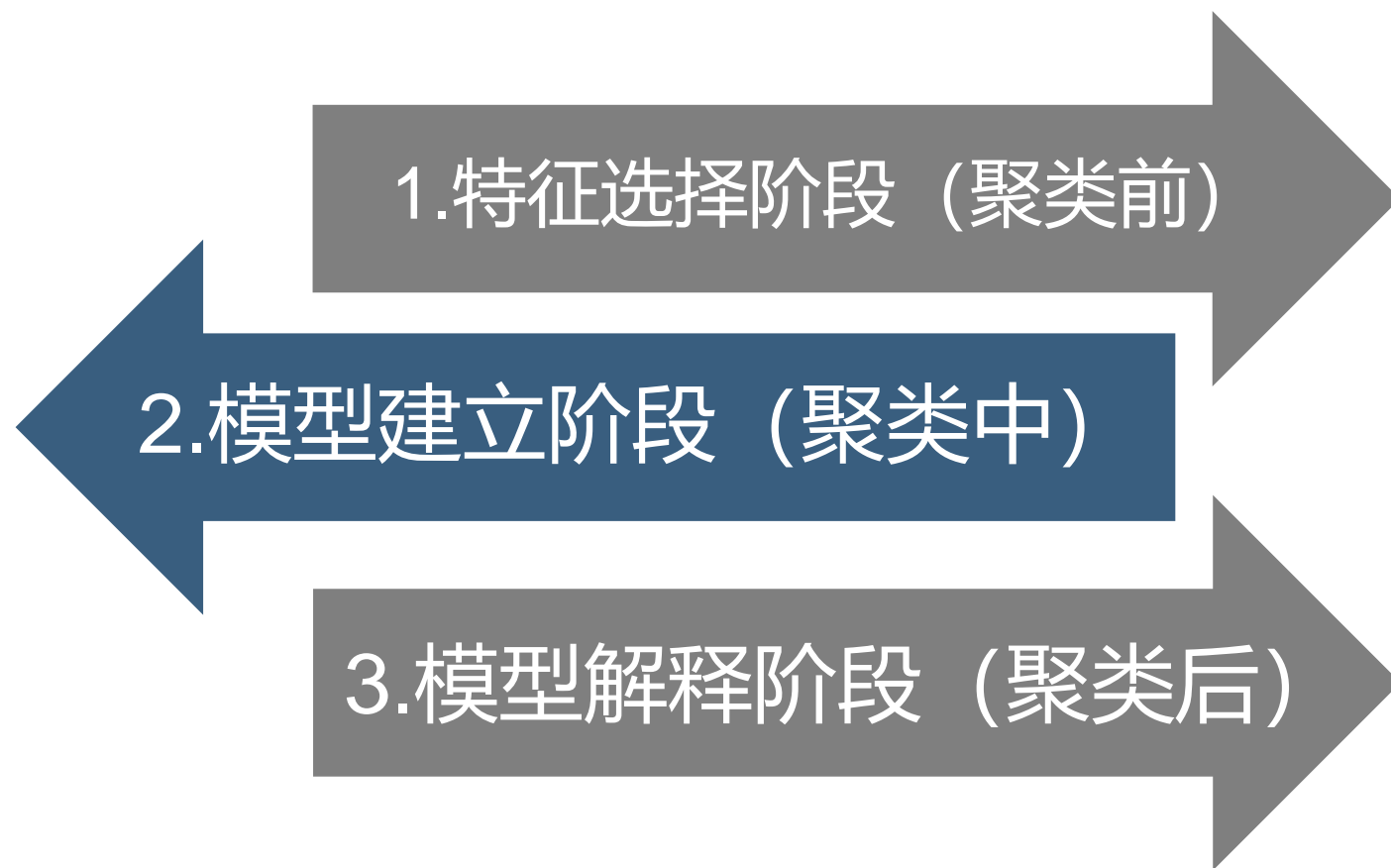


第四部分

文献回顾



文献回顾-总体框架





文献回顾-可解释聚类前方法

1 特征提取

特征提取专注于从原始数据中提取有意义和信息丰富的表示

基于无监督和半监督模式的特征选择(Bonifati et al.,2022)

基于自适应门控机制的特征选择(Salles et al.,2024)

基于格式塔理论的波段选择(Kang et al.,2024)

2 特征选择

特征选择专注于从一组冗余和复杂的特征中选择具有强辨别能力的特征

为每个输入学习特定于样本的稀疏门向量(Svirsky et al.,2024)

结合贝叶斯推理和专家知识识别关键特征(Balabaeva et al.,2020)

使用贪心算法选择最有用的特征(Effenberger et al.,2021)



文献回顾-可解释聚类中方法

1 基于决策树的方法

早期将决策树应用于聚类的尝试

使用合成数据构建监督决策树(Liu et al.,2000)

直接基于原始特征的无监督决策树(Basak et al.,2005)

CUBT(Clustering Using Binary Trees)的引入(Fraiman et al.,2013)

将CUBT扩展到分类数据(Ghattas et al.,2017)

将无监督决策树模型扩展到可解释模糊聚类领域(Jiao et al.,2022)

现代优化技术

混合整数线性优化(MIO)技术(Bertsimas et al.,2021)

树交替优化(TAO)技术(Gabidolla et al.,2022)

单调优化技术(如 BRB算法)(Hwang et al.,2023)



文献回顾-可解释聚类中方法

2

基于规则的方法

好的规则集的两个关键特征

- 1.高频率：表示规则集应尽可能多地覆盖相应聚类内的样本（真阳性）
- 2.高区分能力：表示规则集应尽量减少误覆盖其他聚类样本（假阳性）

用可解释评分量化可解释性(Saisubramanian et al.,2020)

为与聚类相关的每个规则集引入了两个可解释性标准(Carrizosa et al.,2023)

3

其他方法

凸多面体

轴平行边界(Chen et al.,2016)

对角边界(Lawless et al,2022)

原型

最少原型的挑战(Davidson et al.,2024)

新的端到端框架(Pan et al.,2024)



文献回顾-可解释聚类后方法

1 基于决策树的方法

可解释性的代价

解释聚类的成本与最佳聚类成本的比值

近似保证

如果一个算法生成的决策树成本不超过最佳聚类成本的 x 倍，我们就说这个算法有 x 倍的近似保证

研究进展

(Moshkovitz et al., 2020)使用贪婪方法开发决策树实现了对最优 k -medians线性级的近似和最优 k -means平方级的近似

(Laber et al, 2021)通过综合在多个维度上分别构建的决策树来构建最终的决策树实现了更优的近似值

(Makarychev et al., 2021)通过随机化选择分裂特征和值实现更好的近似值

(Esfandiari et al, 2022)通过优化的分裂点选择方法以实现更好的近似值

(Frost et al, 2020)灵活扩展决策树的叶节点数

(Laber et al, 2023)构建深度较小的决策树

(Bandyapadhyay et al, 2023)移除异常点来匹配参考聚类结果



文献回顾-可解释聚类后方法

2

基于规则的方法

(Carrizosa et al.,2023) 通过最大化真阳性样本数，最小化假阳性样本数，同时限制规则长度生成规则

(Weerdt et al.,2014) 通过最佳优先搜索和剪枝构建准确简洁的规则

(Koninck et al.,2017) 基于Weerdt的工作，并通过评估不同的特征集的有效性，生成高质量且可解释的规则

3

其他方法

原型

(Carrizosa et al.,2022)提出了用原型解释聚类的方法，通过解决双目标优化问题来找出原型

凸多面体

(Lawless et al.,2023)使用多面体解释聚类结果

(Chen et al.,2023)使用超立方体覆盖模型解释聚类结果

描述

(Davidson et al.,2018)通过为每个聚类选择一组不重叠的描述来解释聚类结果



第五部分

总结



总结

总结

综述提出了一套对可解释聚类方法分类的标准，并且全面且系统的对各种可解释聚类方法进行了探讨，是首个涵盖聚类分析全生命周期的主题，包括聚类前、聚类中和聚类后阶段。在每个阶段都对相关的可解释聚类方法文献进行了回顾。



未来方向

1. **优化方法**——进一步研究如何将可解释性形式化为优化问题
2. **结构度量的开发**——需要开发更多与可解释性相关的结构度量，以评估新方法的可解释性质量
3. **多阶段方法的结合使用**——结合使用不同聚类阶段的方法
4. **复杂数据的可解释聚类研究**——加强对复杂数据类型（如离散序列、网络图以及多视角和多模态数据）的可解释聚类方法的研究

感 谢 观 看