

基于聚类算法的地下 钱庄监测分析研究

○ 汇报人：

○ 2024-01-25



| CATALOGUE |

目录

- 引言
- 聚类算法原理及应用
- 地下钱庄交易数据特征分析
- 基于聚类算法的地下钱庄监测模型构建
- 实验结果与分析
- 结论与展望

01

引言

CHAPTER



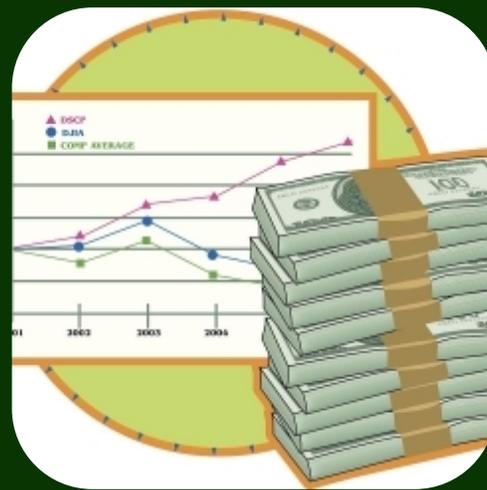
研究背景与意义



地下钱庄活动日益猖獗，严重扰乱金融秩序，危害国家经济安全。



传统监管手段难以应对地下钱庄的隐蔽性和复杂性，需要新的监测分析方法。



基于聚类算法的地下钱庄监测分析可以为监管部门提供有力支持，提高打击地下钱庄的效率和准确性。



国内外研究现状及发展趋势



01

国内研究主要集中在基于规则的方法和基于机器学习的方法，但实际应用中效果有限。

02

国外研究在聚类算法应用于金融领域方面取得了一定进展，但针对地下钱庄的监测分析仍较少。

03

未来发展趋势将更加注重多源数据融合、算法优化和模型可解释性等方面的研究。

研究内容、目的和方法



研究内容

构建基于聚类算法的地下钱庄监测分析模型，包括数据预处理、特征提取、聚类算法选择和模型评估等步骤。



研究目的

提高地下钱庄监测分析的准确性和效率，为监管部门提供有力支持。



研究方法

采用文献综述、实证分析和对比实验等方法，对聚类算法在地下钱庄监测分析中的应用进行深入探讨。

02

聚类算法原理及应用

CHAPTER





聚类算法基本概念

01

聚类

将数据集划分为若干个类或簇的过程，使得同一类内的数据尽可能相似，而不同类间的数据尽可能不同。

02

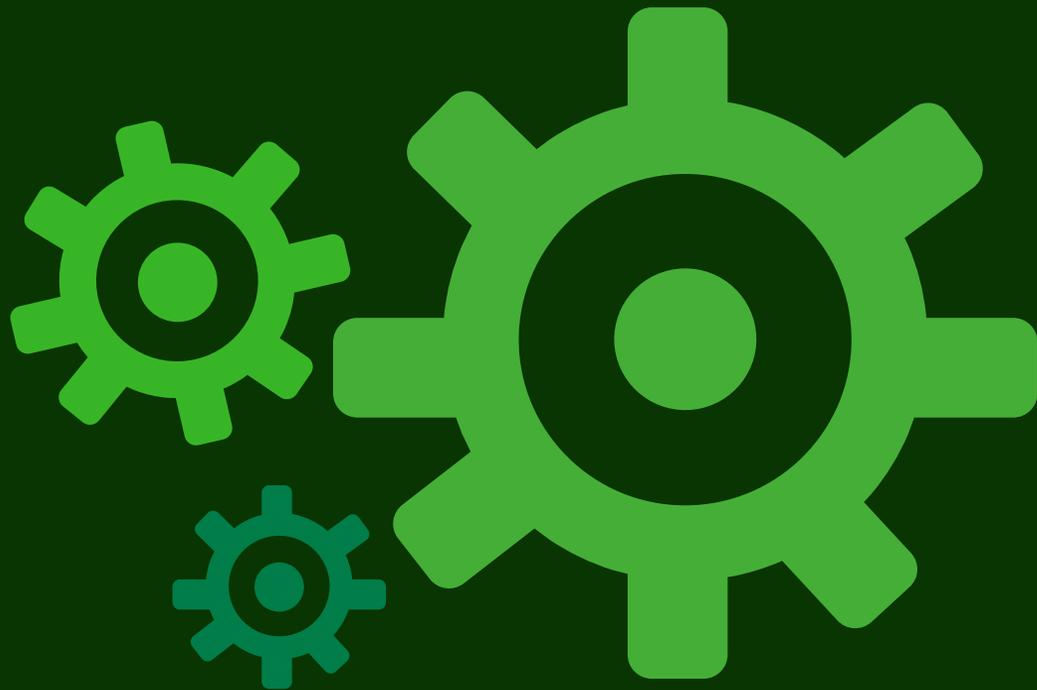
相似度度量

用于衡量数据对象间相似程度的指标，如欧氏距离、余弦相似度等。

03

聚类中心

表示一个类或簇的中心位置的数据对象，通常通过计算类内所有对象的均值或中位数得到。





常见聚类算法原理及特点



K-means聚类

通过迭代将数据对象分配到K个类中，使得每个对象与其所属类的聚类中心之间的相似度最小。特点：简单、快速，但需要指定类别数K且对初始聚类中心敏感。



层次聚类

通过逐层构建嵌套的类簇来实现聚类，可分为凝聚法和分裂法。特点：可发现不同层次的类簇结构，但计算复杂度较高。



DBSCAN聚类

基于密度进行聚类，将密度足够高的区域划分为一类，并可发现任意形状的类簇。特点：对噪声数据鲁棒性强，但需要指定密度阈值和邻域半径。



聚类算法在地下钱庄监测中的适用性

异常检测

地下钱庄交易往往呈现出异常的资金流动模式，通过聚类算法可以识别出这些异常交易行为。

模式识别

聚类算法可用于识别地下钱庄交易中的常见模式或特征，如频繁大额交易、资金快进快出等。

可视化分析

通过将地下钱庄交易数据进行降维处理并应用聚类算法，可以实现交易数据的可视化展示和分析，便于监管部门及时发现可疑交易行为。

03

地下钱庄交易数据特征 分析

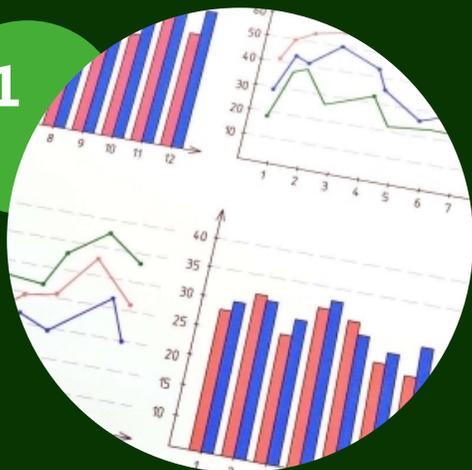
CHAPTER





数据来源及预处理

01



数据来源

从金融机构、监管部门等渠道获取地下钱庄交易数据。

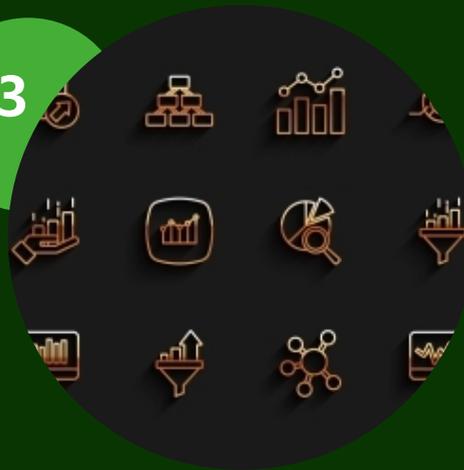
02



数据清洗

去除重复、无效和异常数据，保证数据质量。

03



数据转换

将数据转换为适合聚类算法处理的格式，如数值型、标准化等。



交易数据特征提取

交易金额特征

提取交易金额、交易频率等特征，反映交易规模。

交易时间特征

提取交易时间、交易间隔等特征，反映交易活跃度。



交易网络特征

提取交易对手方、交易关联度等特征，反映交易网络结构。

以上内容仅为本文档的试下载部分，为可阅读页数的一半内容。如要下载或阅读全文，请访问：
<https://d.book118.com/348110141122006100>