

## 基于组合赋权与 KMeans 聚类的商业银行信用评级研究

徐子凯, 俞绍文\*

华东理工大学数学学院 上海

**【摘要】**本文针对我国商业银行信用评级中存在的覆盖率低、模型不透明及小样本建模难等问题, 提出一种融合变异系数法、熵权法与 KMeans 聚类的综合评级模型。首先, 通过变异系数法与熵权法分别对银行指标进行赋权, 并采用成对样本 t 检验验证两者评分结果存在显著差异, 在此基础上引入组合权重优化模型, 融合两种方法优势得出最终评分。随后, 利用一维 KMeans 聚类算法对评分结果进行评级分割, 并通过分析误差函数变化趋势确定最优簇数, 实现灵活有效的评级划分。实证部分以 271 家商业银行 2023 年数据为样本, 选取八项与信用风险密切相关的指标进行建模与评级分析。结果显示, 该方法在小样本下能充分挖掘数据信息, 兼顾模型的客观性与灵活性, 不仅有效区分不同信用风险水平的银行, 也对未来模型的优化方向提出了建议, 具有良好的实际应用与推广价值。

**【关键词】**信用评级; 变异系数法; 熵权法; 组合赋权; KMeans 聚类

**【基金项目】**国家级大学生创新创业训练计划项目 (编号: 202510251082)

**【收稿日期】**2025 年 9 月 19 日 **【出刊日期】**2025 年 10 月 23 日 **【DOI】**10.12208/j.aif.20250032

### A study on credit risk rating methods for commercial banks based on combined weighting and KMeans clustering

Zikai Xu, Shaowen Yu\*

*School of Mathematics, East China University of Science and Technology, Shanghai*

**【Abstract】** This paper addresses key challenges in the credit risk rating of Chinese commercial banks, including low rating coverage, opaque modeling methods, and difficulties arising from small sample sizes. A comprehensive rating model is proposed by integrating multiple objective weighting methods with unsupervised learning techniques. First, the coefficient of variation method and the entropy weight method are used to assign weights to selected indicators, and a paired sample t-test is conducted to confirm the significant difference between the two scoring results. Based on this, a combined weighting scheme is developed to synthesize the strengths of both methods. Then, a one-dimensional KMeans clustering algorithm is applied to categorize the final scores into rating levels, with the optimal number of clusters determined through analysis of the loss function. An empirical study using 2023 data from 271 commercial banks and eight key credit risk indicators demonstrates that the proposed model can effectively extract information from limited data, balancing objectivity and flexibility. The results not only differentiate banks by risk level but also suggest directions for future model enhancements, highlighting the model's practical value and adaptability.

**【Keywords】** Credit rating; Coefficient of variation method; Entropy weight method; Combined weighting; KMeans clustering

#### 1 引言

银行信用风险评级是指运用科学方法对不同银

行的金融实力进行综合评估, 以揭示其信用风险水平的过程。构建科学合理的评级模型, 不仅有助于提

\*通讯作者: 俞绍文

升银行自身的竞争力和信用透明度,还能够增强监管机构的监管效率,促进银行间的合作与信任,同时为公众和投资者提供更具参考价值的银行选择依据。

而我国商业银行的信用评级,主要存在以下两点问题:一是虽有国内外评级机构打分,但是覆盖率太低,超 70%以上银行没有任何评级数据可供查询(来源于国泰安数据库),即使有评级数据,评级机构的数学模型也属于商业机密,无法用于参考,且相互之间无法统一标准;二是小样本问题,国内商业银行数量不到 400 家,且数据完整性差,这使得多数需要大样本的机器学习方法无法运用,增加了解决问题的困难。

在诸多解决方法中,利用层次分析法、熵权法、变异系数法等多方法主客观结合往往受到青睐,李战江和迟国泰(2013)利用变异系数法和最优分割法进行银行信用风险评价<sup>[1]</sup>;李战江(2017)结合了 G1 法(一种主观赋权法)与熵权法,构建了基于最优组合赋权的银行信用风险综合得分的测算模型<sup>[2]</sup>;刘铮铮(2006)通过专家打分运用层次分析法构建信用评级模型<sup>[3]</sup>。赵家敏和黄英婷经过指标优化后进行层次分析法建模<sup>[4]</sup>。这些评级模型不乏科学性,但是存在一定缺陷,比如只要加入了主观打分,就对专家的水平有很高的要求,且某个阶段专家的公式随着时间的发展又可能变动;而类似最优分割法又需要固定分类数目且计算不够灵活。

此外,虽然数据量小,但由于机器学习在拟合上的独特优越性,仍然有研究致力于用机器学习研究评级。戴静霆(2023)利用 Light GBM 模型基于集成学习研究了银行信用评级模型<sup>[5]</sup>;王春峰,万海晖和张维(1999)将神经网络方法运用在信用评级之中<sup>[6]</sup>。蔡艳萍,孙夏(2016)基于 BP 神经网络构建了上市商业银行绩效评价模型<sup>[7]</sup>。郭立仑和周升起(2022)结合 KMV 与随机森林对商业银行信用风险的主要影响因素进行了分析<sup>[8]</sup>。

除了以上两种观点,还有在运筹学角度进行建模分析的模型,也取得了较好的结果,如樊桂岭,笪凤媛(2017)基于对商业银行信用风险外部因素与内部因素的分析,通过构建融资企业和商业银行的博弈模型,发现我国商业银行信用风险的多种产生原因<sup>[9]</sup>。

综合来看,要科学地解决我国商业银行的信用评级问题,势必要合理地选取数据,并有针对性地选

择适合中小样本容量的数学方法,且有选择地使用机器学习方法加以辅助,最大程度地发挥数据的客观性并最大程度保留灵活性。在这种情况下,需要一种“让数据自己说话”的客观性强且灵活的评分模型。

本文采用变异系数法与熵权法分别对各指标进行客观赋权,并通过配对样本 t 检验验证两种方法在评分结果上存在显著差异。在此基础上,进一步对两种权重进行组合赋权,以提升模型稳定性与权重分配的合理性。随后,利用一维 KMeans 聚类方法,在无监督的背景下实现对我国商业银行的信用评级划分。该方法在小样本数据条件下充分挖掘指标信息,兼顾模型的客观性与灵活性,同时实现了对所有数据完备银行的全面评级,具备良好的推广性和与其他评级模型融合潜力。

## 2 评价方法与模型

### 2.1 基于变异系数法与熵权法的组合权重法

#### 2.1.1 样本预处理

选取的样本银行中的指标包含正向与负向评级指标,需要分别进行标准化,以便后续评分。正向评级指标的标准化公式为:

$$q_{ij} = \frac{x_{ij} - \min_{1 \leq j \leq n} x_{ij}}{\max_{1 \leq j \leq n} x_{ij} - \min_{1 \leq j \leq n} x_{ij}} \quad (1)$$

负向评级指标的标准化公式为:

$$q_{ij} = \frac{\max_{1 \leq j \leq n} x_{ij} - x_{ij}}{\max_{1 \leq j \leq n} x_{ij} - \min_{1 \leq j \leq n} x_{ij}} \quad (2)$$

其中,  $q_{ij}$ ——第  $j$  个样本银行的第  $i$  个评级指标的标准化取值;  $x_{ij}$ ——第  $j$  个样本银行的第  $i$  个评级指标的原始数据;  $n$ ——样本银行的个数。

#### 2.1.2 变异系数法

变异系数法是通过指标的均值和方差衡量数据的离散程度并依此赋予权重的方法,具有无量纲、客观等优点。权重计算公式如下:

$$v_i = \frac{s_i/q_i}{\sum_{j=1}^n s_j/q_j} \quad (3)$$

$$s_i = \sqrt{\frac{1}{n-1} \sum_{j=1}^n (q_{ij} - q_i)^2} \quad (4)$$

$$q_i = \frac{1}{n} \sum_{j=1}^n q_{ij} \quad (5)$$

其中,  $v_i$  代表第  $i$  个评级指标的变异系数法权重,

$s_i$  代表第  $i$  个评级指标下数据的标准差,  $q_i$  代表第  $i$  个标准化后的评级指标下数据的平均值。

### 2.1.3 熵权法

熵权法是一种基于信息熵原理的客观赋权方法, 常用于综合评价与多指标决策问题中。其核心思想在于通过各指标在样本中分布的信息差异程度来确定其权重。具体而言, 若某一指标在不同样本中的取值差异越大, 则其所包含的信息量越丰富, 对综合评价的区分作用也就越显著, 因而应赋予更高的权重。反之, 若某一指标的数值变化较小, 信息冗余较多, 其权重应相应较低。

熵权法通过对标准化后的各指标计算其信息熵值, 以反映其内部的不确定性, 然后利用 1 减去熵值得到的“冗余度”来度量该指标对系统的有效贡献, 进而确定各指标的客观权重。这种方法避免了人为因素的干扰, 能充分反映数据本身所蕴含的结构特征, 尤其适用于缺乏专家评分基础或需增强评价客观性的研究场景, 能很好地适配商业银行信用评级的场景。具体公式如下:

$$\omega_i = \frac{1 + \frac{1}{\ln n} \sum_{j=1}^n (q_{ij} / \sum_{j=1}^n q_{ij}) \ln (q_{ij} / \sum_{j=1}^n q_{ij})}{\sum_{i=1}^m (1 + \frac{1}{\ln n} \sum_{j=1}^n (q_{ij} / \sum_{j=1}^n q_{ij}) \ln (q_{ij} / \sum_{j=1}^n q_{ij}))} \quad (6)$$

其中,  $\omega_i$  表示第  $i$  个指标的熵权法权重。

### 2.1.4 成对样本 t 检验

成对样本 t 检验是一种用于比较同一组对象在两种不同处理或两个不同时间点下测量值是否存在显著差异的统计方法。该方法基于两组观测值之间的一一对应关系, 通过对差值的均值进行显著性检验, 以判断两组样本是否来源于同一总体或具有相同的均值。其基本假设为差值的总体均值为零, 检验统计量服从 t 分布, 适用于样本量较小且差值近似正态分布的情况。对变异系数法和熵权法打分后的数据进行成对样本 t 检验, 可以十分科学地判断出两评分的均值是否有显著差异, 也即评分的角度是否有较大的区别, 进一步决策是否要组合权重以达到更好效果。

### 2.1.5 组合权重优化方法

本文利用毛定详(2002)提出的一种最小二乘意义下主客观评价一致的组合评价方法<sup>[10]</sup>, 区别在于本文是将两种客观评价法进行组合, 但仍可借用此思路提升评分的科学性。具体的优化模型如下:

$$\min H(u) = \sum_{i=1}^n \sum_{j=1}^m \{ (\nu_j - u_j) q_{ij} \}^2 + \{ (\omega_j - u_j) q_{ij} \}^2 \quad (7)$$

$$\text{Subject to } \sum_{j=1}^m u_j = 1, \quad u_j \geq 0 \quad (j = 1, \dots, m) \quad (8)$$

利用 Lagrange 乘数法计算出最优解, 可算出各大商业银行信用评分的组合权重, 并进而作出评级分割。最终评分为:

$$x_i = \sum_{j=1}^m u_j p_{ij} \quad (i = 1, \dots, n) \quad (9)$$

其中  $x_i$  代表第  $i$  个银行样本的最终评分。

## 2.2 基于一维 KMeans 法的评级分割

KMeans 法是一种机器学习中的无监督算法, 其优点是不需要标签数据, 较为灵活, 且中小样本量下表现较为健壮。其基本思路是给定 K 值, 再生成 K 个质心, 通过迭代过程不断优化簇的划分。算法首先随机初始化 K 个簇中心点, 然后将每个样本分配到距离其最近的质心所对应的簇中。接着, 根据当前簇中所有样本的均值重新计算质心位置。该过程不断重复, 直到簇分配不再变化或达到预设的迭代次数为止。KMeans 算法追求的是簇内样本之间的相似度最大, 簇间样本之间的差异性最大, 最终实现对数据结构的有效划分。在本文评级模型下 KMeans 的优化函数为:

$$\min J = \sum_{k=1}^K \sum_{x_i \in C_k} \|x_i - \mu_k\|^2 \quad (10)$$

其中, K 是聚类的个数 (即簇的数量),  $C_k$  表示第 k 个聚类的样本集合,  $x_i$  是属于  $C_k$  的样本点 (即属于  $C_k$  的样本评分),  $\mu_k$  是  $C_k$  中所有样本评分的均值,  $\|x_i - \mu_k\|^2$  表示样本点  $x_i$  到其簇质心的欧氏距离平方。

在实际应用中, K 值的选择对聚类效果影响较大, 通常通过轮廓系数、肘部法则等指标来辅助确定合适的 K 值。综合来看, KMeans 法十分适合用于我国商业银行的评级分割, 不仅样本量适宜, 同时能根据不同经济环境选择恰当的 K 值, 兼顾客观性与灵活性。

### 2.3 商业银行评级模型流程

本文首先收集相关数据, 并对样本进行预处理同时标准化, 分别利用变异系数法与熵权法计算权重, 并进一步利用成对样本 t 检验确定两者打分有显著性差异, 在此基础上利用优化方法计算最优组合权重, 并得出各商业银行最终得分, 运用一维 KMeans 进行聚类分析, 得出最优 K 值以及在最优 K 值下的分类, 最终确定评级。具体评级模型流程如图 1 所示。

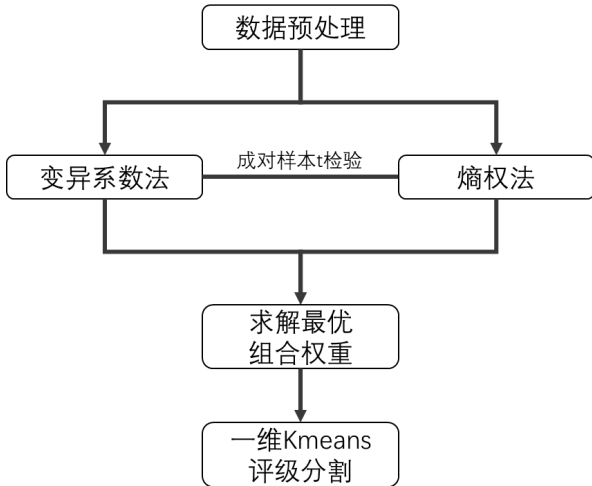


图1 评级模型流程图

### 3 实证研究

#### 3.1 样本选取与数据预处理

本文基于国泰安数据库, 选取我国 271 家商业

银行 2023 年度的相关信用风险指标数据, 并提取出其中 8 个与信用风险强相关的评级指标, 包括资本充足率、核心资本充足率、流动性比例(本外币)、拨备覆盖率、资产利润率、不良贷款率、风险加权资产对总贷款比率、成本收入比例, 其中前五个指标为正向评级指标, 后三个指标为负向评级指标, 空缺数据以均值填充。根据各指标的正负向属性, 分别代入公式(1)与(2)进行标准化处理, 并将结果汇总如表1所示。

#### 3.2 评分数据计算

将表1的数据带入公式(3)(4)(5)(6), 分别得到变异系数法和熵权法的权重, 整理得表2。

将得到的系数与样本数据进行逐样本相乘再累加, 分别得到样本银行在变异系数法和熵权法权重下的得分, 对这两组数据进行成对样本  $t$  检验, 得出结果如表3所示。

表1 样本数据标准化表

银行中文简称	资本充足率	核心资本充足率	流动性比例	拨备覆盖率	资产利润率	不良贷款率	风险加权资产对总贷款比率	成本收入比例
盘谷中国	1	1	1	0.10	0.24	0.96	0.99	0.84
瑞穗中国	0.33	0.36	0.36	1	0.25	1	0.99	0.84
中国农业银行	0.68	0.71	0.70	0.08	0.23	0.97	0.99	0.87
...	...	...	...	...	...	...	...	...
兰州银行	0.047	0.02	0.26	0.05	0.21	0.96	0.99	0.89
四川天府银行	0.04	0.002	0.26	0.05	0.19	0.94	0.99	0.85
营山农商银行	0	0.03	0.27	0.03	0.24	0.90	0.99	0.81

表2 变异系数法和熵权法的权重系数表

银行样本评级指标	变异系数法	熵权法
资本充足率	0.223788	0.240732
核心资本充足率	0.270950	0.387938
流动性比例	0.114329	0.081822
拨备覆盖率	0.252697	0.248855
资产利润率	0.071485	0.024114
不良贷款率	0.019753	0.004689
风险加权资产对总贷款比率	0.018856	0.004470
成本收入比例	0.028141	0.007380

表3  $t$  检验结果表

银行样本评级指标	均值(变异系数法)	均值(熵权法)	均值差	$t$ 值	$p$ 值	显著性
资本充足率	0.2276	0.1769	-0.0507	45.903	0.000*	显著

Note: \* indicates significant difference at  $P < 0.05$ .

结果显示, 两组评分均值分别为 0.2276 与 0.1769, 平均差值为 -0.0507,  $t$  值为 45.903, 对应的  $p$  值小于 0.001。在 95% 的置信水平下, 差异具有统计学显著性, 说明两种方法在加权计算中对样本评分的影响存在显著差异。因此, 组合权重可以综合两种评分法的优势, 将两种有显著差异的方法进行整合。

将表 2 数据带入目标函数 (7), 在约束 (8) 下求解非线性规划。借助计算机, 可求得组合权重法系数表 4。

表 4 组合权重法系数表

银行样本评级指标	组合权重法
资本充足率	0.232222
核心资本充足率	0.329450
流动性比例	0.098104
拨备覆盖率	0.250761
资产利润率	0.047795
不良贷款率	0.012244
风险加权资产对总贷款比率	0.011651
成本收入比例	0.017761

再将表 4 数据带入公式 (9), 基于上述优化权

重的结果, 进一步计算出各商业银行的组合评分, 详见表 5。

表 5 组合权重法评分表

银行中文简称	组合权重法评分
盘谷中国	0.733680885
瑞穗中国	0.532878856
中国农业银行	0.531111226
...	...
兰州银行	0.107050565
四川天府银行	0.094636746
营山农商银行	0.091912071

综上所述, 全部样本银行的信用评分已依据组合权重法计算完成。

### 3.3 一维 KMeans 评级分割

KMeans 算法得样本特征向量可以有多个维度, 但是在此仅取一维, 作为一种求分割的算法使用。在算法中, 需要指定簇的数量  $K$ , 本文设定  $K$  值的上下界, 并结合各  $K$  值下误差函数的变化趋势, 综合判断最优聚类数, 从而达到灵活性的要求, 即未必按照固定的 9 个评级进行分割。此处选取  $K \in [2, 9]$  进行计算, 得到损失函数图图 2:

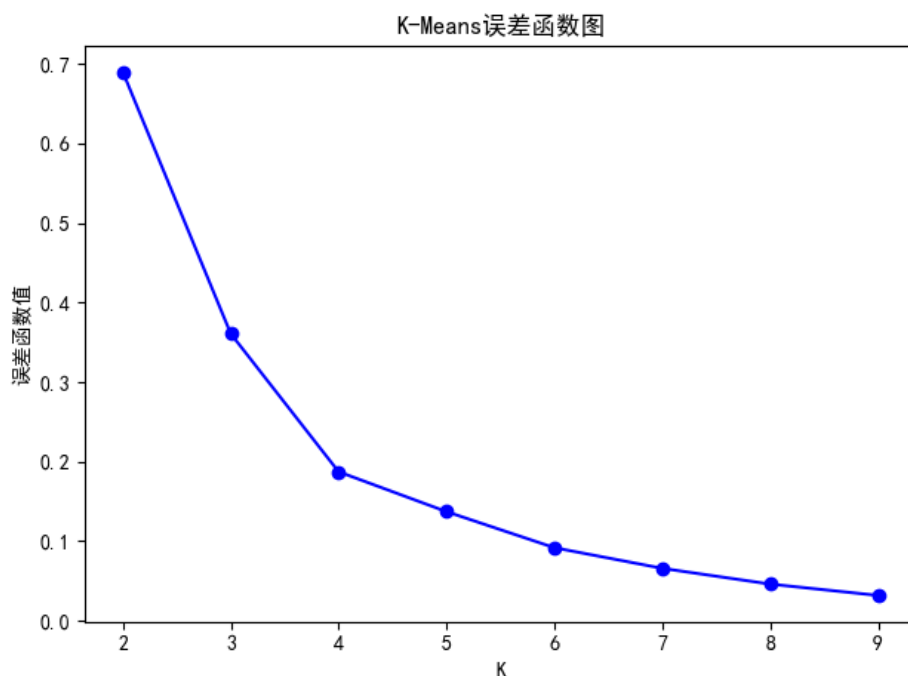


图 2 KMeans 误差函数图

分析图中误差函数趋势, 若按照差分值进行分类, 可以认为有 $K \in [2, 3]$ ,  $K \in [4, 6]$ ,  $K \in [7, 9]$ 三种区间, 在 $K \in [2, 3]$ 中, 误差函数的值还有很大下降空间, 则分类会不够精确, 因此不选; 在 $K \in [7, 9]$ 中, 斜率变化非常缓慢, 表明该区间内聚类数过多, 可能导致分组过细, 降低评级解释性, 故不予采用。故 $K \in [4, 6]$ , 经分析可知, 当  $K=5$  时, 聚类结果在精度与分组合理性之间达到最优平衡。按照 $K = 5$ 进行计算, 得到评级 (4 为最高, 0 为最低, 3、2、1 依次递减), 记录在表 6。

表 6 银行信用评级表

银行中文简称	信用评级
盘谷中国	4
瑞穗中国	4
中国农业银行	4
...	...
兰州银行	0
四川天府银行	0
营山农商银行	0

至此, 本文对商业银行的信用评级已完成。

### 3.4 模型结果分析

#### 3.4.1 模型可靠性评价

对模型可靠性进行评价, 可得如下结论:

一, 将评级结果的关键指标均值绘制为热图, 记为图 3。图 3 清晰地展示了各评级等级在关键财务指标上的均值差异。例如, 高评级银行在“资本充足率”、“核心资本充足率”和“拨备覆盖率”等风险缓冲类指标上表现突出, 而低评级银行在“成本收入比例”这一负面指标上均值偏高, 进一步验证了评级结果的合理性。

二, 为了进一步验证 KMeans 聚类的有效性, 本文绘制了组合评分在不同评级标签下的分布图图 4。从图中可以清晰看到, 评分在各等级间存在明显分界, 组间差异显著, 说明聚类结果较好地实现了同类银行归类。此外, 通过不同评级组的中心评分与银行数量分析图可以看出, 高评级组的中心评分显著高于其他组, 且银行数量集中于中间评级, 进一步印证了本模型在评分与分组合理性之间的匹配性。

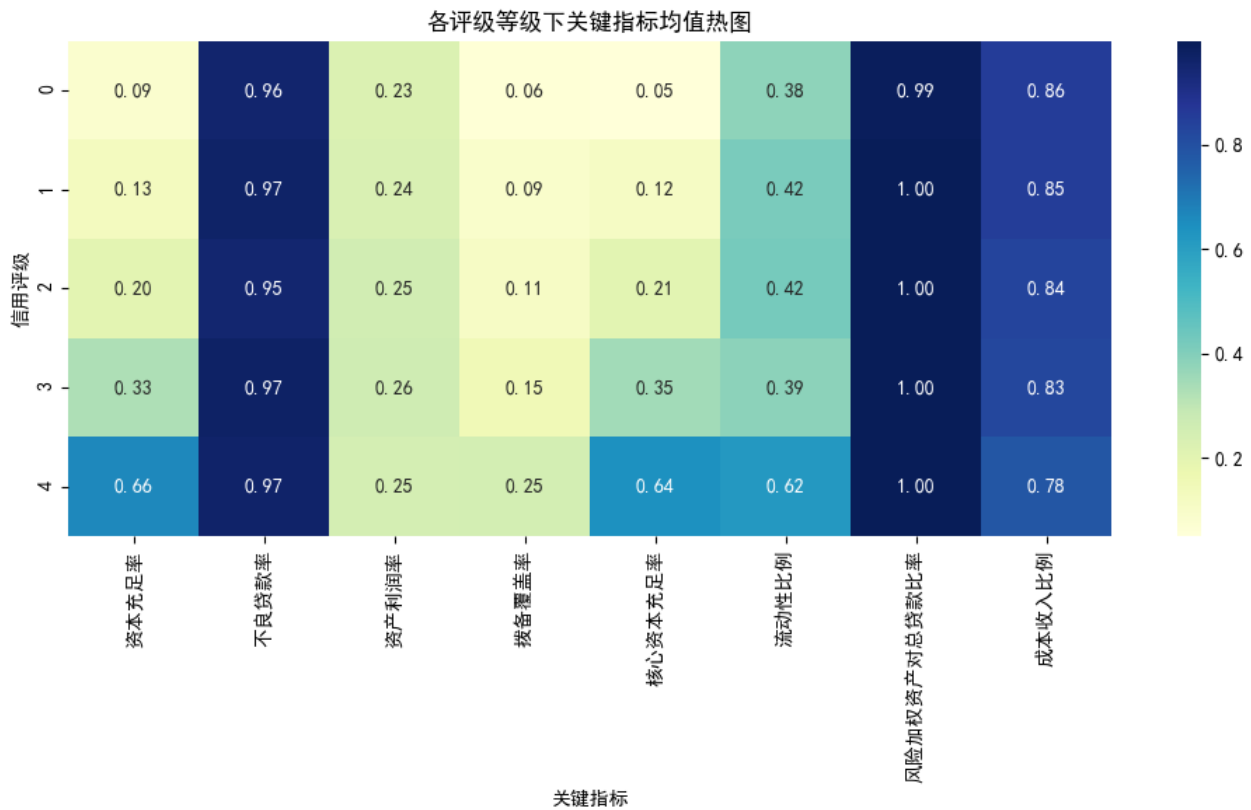


图 3 各评级等级下关键指标均值热图



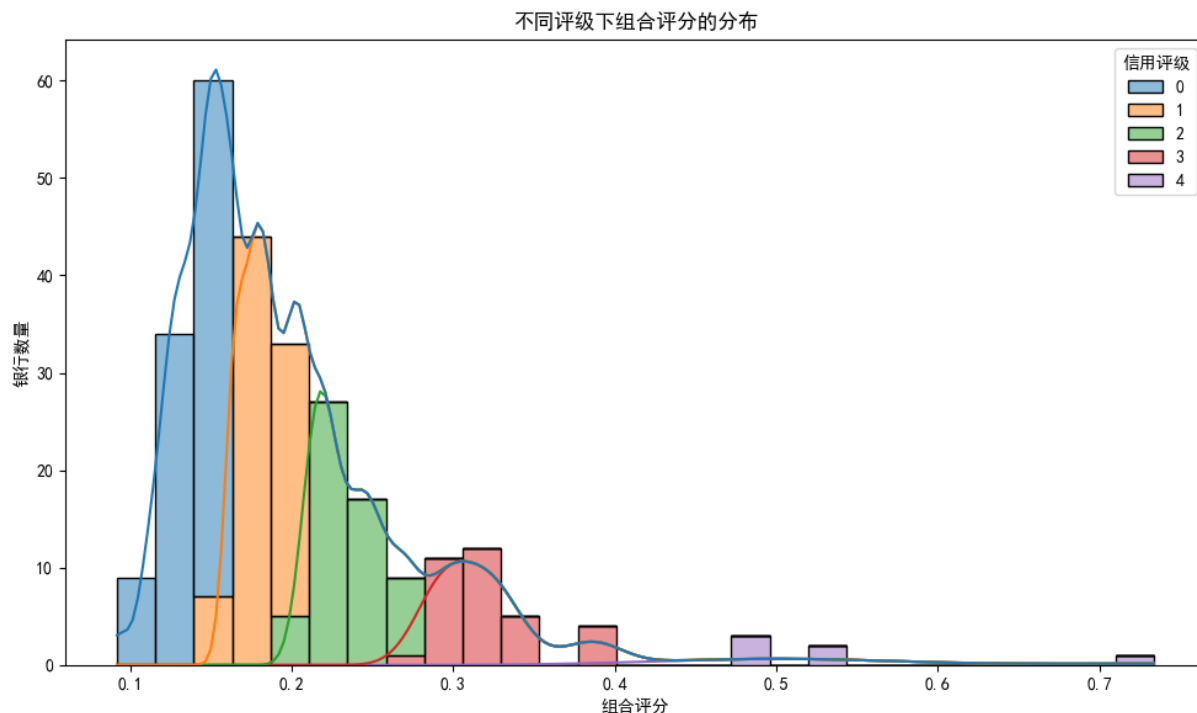


图 4 不同评级下组合评分分布图

### 3.4.2 模型结果综合分析

对模型结果综合分析, 可得如下结论:

一, 在本模型中得到最优评级 (4 级) 的银行, 分别是盘谷中国、瑞穗中国、中国农业银行、屏南联社、澳新 (中国)、重庆黔江银座村镇银行, 由此看出, 外资银行在信用风险控制方面表现良好, 主要得益于其审慎的业务策略、较高的资产质量以及严格的国内外监管要求; 隶属于农村的商业银行信用风险也较低, 这主要得益于其 ‘小额分散’ 的信贷策略, 通过 ‘网格化服务’ 与 ‘整村授信’ 等手段, 有效提升客户覆盖面并分散风险; 中国农业银行作为四大国有独资商业银行之一, 处于服务国家战略的领军地位, 不仅是乡村振兴主力军, 也是实体经济支撑者, 获得高评级理所应当。

二, 在本模型中得到中间等级的银行 (3 和 2 级), 以农村商业银行、农村信用合作联社、以及各大上市城市银行为主, 这类机构中表现较优的已接近高评级标准, 但由于所处区域经济基础相对薄弱、客户结构集中度高, 整体信用风险仍存在一定波动。尤其是部分农村中小银行在风险控制、资本补充及信息系统建设等方面仍存在不足, 影响整体评级表现。相比之下, 前文提到的屏南联社、重庆黔江银座

村镇银行等农村类银行则通过因地制宜的业务模式、精准的客户画像和扎实的基层金融服务, 实现了风险的有效分散, 属于农村银行中的优秀代表。城市商业银行在该等级中占据较大比重, 整体经营稳健, 风控体系较为成熟, 但受区域经济周期波动、信贷集中度较高等因素影响, 部分资产质量出现阶段性下滑, 从而拉低整体评级水平。

三, 在本模型中得到低评级 (1 级和 0 级) 的银行普遍存在资本补充能力较弱、不良贷款占比偏高、业务结构单一等问题, 部分银行对地方经济依赖度高, 抗风险能力较差。此外, 在信息披露规范性、内部控制机制健全性及合规管理水平方面, 部分银行仍需进一步优化。整体来看, 这一评级反映了我国银行业在不同层级、不同类型机构间信用风险管理水平的差异, 也提示相关银行应增强资本充足能力, 优化资产结构, 并全面提升风险管理水平, 以实现稳健经营与风险可控。

此外, 针对与模型本身而言, 由于样本年份单一、指标维度有限, 模型可能在部分样本的风险识别上存在偏差, 导致评级结果偏高或偏低。后续可引入年份趋势变量、银行规模与控股结构等高价值指标, 进一步完善模型结构与评级精度。

#### 4 结论

(1) 利用求最优组合权重的方法结合了变异系数法与熵权法, 实现了数据本身信息的客观量化, 得出有价值的信用评分。

利用机器学习中的 KMeans 算法进行信用评分的评级分割, 观察分析误差函数图像分析选取最优 K 值, 从而得到合理的评级分割, 实现了在小样本条件下的客观、灵活评级, 具备较强的推广性。未来可进一步引入时序数据与更多结构性指标, 以提升模型的时序适应性与风险识别精度。

#### 参考文献

- [1] Li Z J, Chi G T. Rating model of credit risk of commercial bank based on differential weighting and optimum partition[C]. 2013 6th International Conference on Information Management, Innovation Management and Industrial Engineering, 2013: 328-331.
- [2] 李战江. 最优策略下的商业银行信用风险的小样本评级模型[J]. 系统工程, 2017, 35(09): 8-17.
- [3] 刘铮铮. 基于层次分析法的商业银行信用评级模型研究[D]. 西北工业大学, 2006.
- [4] 赵家敏, 黄英婷. 我国商业银行中小企业信用评级模型研究[J]. 金融论坛, 2006, (04): 16-20.
- [5] 戴静霆. 基于集成学习的银行信用评级模型研究[D]. 山西财经大学, 2023.
- [6] 王春峰, 万海晖, 张维. 基于神经网络技术的商业银行信用风险评估[J]. 系统工程理论与实践, 1999, (09): 24-32.
- [7] 蔡艳萍, 孙夏. 基于 BP 神经网络的上市商业银行绩效评价体系[J]. 系统工程, 2016, 34(12): 24-30.
- [8] 郭立仑, 周升起. 商业银行信用风险主要影响因素来自内部还是外部? ——基于 KMV 及随机森林模型的实证研究[J]. 会计与经济研究, 2022, 36(01): 105-124.
- [9] 樊桂岭, 笄凤媛. 商业银行信用风险产生机制与管理研究 ——基于博弈论视角[J]. 技术经济与管理研究, 2017, (10): 72-76.
- [10] 毛定祥. 一种最小二乘意义下主客观评价一致的组合评价方法[J]. 中国管理科学, 2002, (05): 96-98.

**版权声明:** ©2025 作者与开放获取期刊研究中心(OAJRC)所有。本文章按照知识共享署名许可条款发表。

<https://creativecommons.org/licenses/by/4.0/>



**OPEN ACCESS**