

基于因子模型的上市公司财务预警——以房地产公司为例

刘佳鑫 张松艳

浙江科技大学，浙江杭州，310000；

摘要：本研究选取 2023 年 A 股上市房地产公司，利用因子分析和机器学习技术（包括逻辑回归、BP 神经网络、XGBoost）建立财务风险预警模型。通过分析 9 项财务指标，提取 4 个主要公因子（资产负债率、流动比率、利息保障倍数、存货周转率），构建了预警模型。XGBoost 模型的预测准确率和 F1-score 分别为 94.44% 和 0.95，均高于逻辑回归和 BP 神经网络。研究发现高风险企业常有短期偿债困难、存货积压和现金流问题。基于此，研究从企业、投资者和监管三个角度提出管理建议，旨在为房地产行业风险防控提供理论和实践指导。

关键词：因子分析；风险预警；房地产公司；逻辑回归模型；BP 神经网络模型；XGBoost 模型

DOI：10.69979/3029-2700.25.09.086

引言

房地产是国民经济的重要部分，2022 年其增加值占 GDP 的 6.1%。然而，由于高杠杆、长周期和政策敏感性，该行业面临频繁的财务风险。尽管 Altman 的 Z-score 模型在制造业和零售业广泛使用，但其在房地产行业的适用性有限。房地产企业资产结构复杂，存货中土地和在建项目占比较大，变现周期长，流动性指标难以准确反映偿债能力。现有研究多集中于单一财务指标，缺乏对多维风险因子的系统整合，预警结果解释力不足。因此，构建一个符合房地产行业特征的财务风险预警模型，既是理论研究的迫切需求，也是实践管理的重要挑战。张荣艳和廖萌（2012）指出，Z-Score 模型在房地产上市公司中的适用性存在一定局限性，因子模型在不同行业的财务风险预警中均表现出良好的适用性。潘丹（2023）在新能源汽车企业中，通过因子分析法提取关键财务指标，并结合功效系数法构建预警模型，能够有效评估企业的财务风险水平。贾玉婷（2016）在煤炭上市公司中，因子分析法同样被用于提取主因子，并结合逻辑回归模型进行财务风险预警。在房地产行业，因子模型同样发挥了重要作用。欧国良等（2018）通过因子分析法提取了房地产企业的关键财务指标，并构建了基于 L

ogistic 模型的财务风险预警模型，验证了因子分析在提高预警模型准确性方面的有效性。综上所述，本研究创新性融合因子分析法与机器学习技术，构建适配房地产行业的动态预警模型，旨在提升风险识别精度与时效性。

1 研究方法与数据

1.1 样本与指标

选取 2023 年 A 股 92 家非 ST 房地产上市公司，数据源自东方财富 Choice 金融终端。在盈利能力、偿债能力、营运能力、成长能力以及和现金流量五大指标中，基于行业特性筛选 9 项指标，主要包括资产负债率（X1）、流动比率（X2）、利息保障倍数（X3）、存货周转率（X4）、应收账款周转率（X5）、毛利率（X6）、净资产收益率 ROE（X7）、净利润现金含量（X8）以及现金净流量与负债总额的比率（X9）。

1.2 主成分分析

1.2.1 主成分分析的适用性检验

由表 1 可知，其中的 KMO 值达 0.592，并且 Bartlett 检验的显著性水平为 0.000，该数值大幅低于 0.01 这一临界值，通过了相关性检验，该指标体系具备开展因子分析的适宜性。

表 1 KMO 和 Bartlett 的检验

取样足够度的 Kaiser-Meyer-Olkin 度量。		.592
Bartlett 的球形度检验	近似卡方	124.568
	df	36

表 1 KMO 和 Bartlett 的检验		
取样足够度的 Kaiser-Meyer-Olkin 度量。		.592
Bartlett 的球形度检验	近似卡方	124.568
	df	36
	Sig.	.000

1.2.2 主成分的提取

公因子数量由特征值和方差贡献率共同决定, 如表 3 的解释总方差表所示, 该表揭示了四个显著的特征值,

前四个因子的方差累计贡献率达到 69.622%, 涵盖了财务预警指标的绝大部分信息, 从而得到公共因子 F1、F2、F3 和 F4。

表 2 解释的总方差

因子编号	特征根			旋转前方差解释率			旋转后方差解释率		
	特征根	方差解释率%	累积%	特征根	方差解释率%	累积%	特征根	方差解释率%	累积%
1	2.438	27.084	27.084	2.438	27.084	27.084	1.757	19.525	19.525
2	1.593	17.697	44.781	1.593	17.697	44.781	1.660	18.448	37.973
3	1.221	13.569	58.349	1.221	13.569	58.349	1.637	18.187	56.160
4	1.015	11.273	69.622	1.015	11.273	69.622	1.212	13.462	69.622
5	0.866	9.625	79.247	-	-	-	-	-	-
6	0.659	7.319	86.566	-	-	-	-	-	-
7	0.535	5.944	92.510	-	-	-	-	-	-
8	0.394	4.374	96.884	-	-	-	-	-	-
9	0.280	3.116	100.000	-	-	-	-	-	-

1.2.3 主成分的解释

由表 3 可得, 使用因子分析法进行权重计算, 基于标准化后, 建立因子和研究项之间的关系等式, 如下:

$$F1 = -0.369X1 + 0.541X2 + 0.273X3 + 0.234X4 - 0.077X5 - 0.103X6 - 0.096X7 - 0.229X8 - 0.181X9$$

$$F2 = -0.125X1 - 0.137X2 - 0.061X3 - 0.243X4 + 0.013X5 + 0.561X6 + 0.515X7 + 0.193X8 + 0.055X9$$

$$F3 = -0.005X1 - 0.104X2 + 0.264X3 + 0.454X4 - 0.047X5$$

$$-0.181X6 + 0.113X7 + 0.329X8 + 0.424X9$$

$$F4 = 0.113X1 - 0.001X2 + 0.383X3 - 0.079X4 + 0.731X5 - 0.107X6 - 0.120X7 - 0.293X8 + 0.149X9$$

$$\text{具体计算公式为: } F = (19.525F1 + 18.448F2 + 18.187F3 + 13.462F4) / 69.622$$

$$\text{最终为: } F = 0.280F1 + 0.265F2 + 0.261F3 + 0.193F4$$

表 3 成分得分系数矩阵

名称	成分			
	成分 1	成分 2	成分 3	成分 4
X1	-0.369	-0.125	-0.005	0.113
X2	0.541	-0.137	-0.104	-0.001
X3	0.273	-0.061	0.264	0.383
X4	0.234	-0.243	0.454	-0.079
X5	-0.077	0.013	-0.047	0.731

名称	成分			
	成分 1	成分 2	成分 3	成分 4
X6	-0.103	0.561	-0.181	0.107
X7	-0.096	0.515	0.113	-0.120
X8	-0.229	0.193	0.329	-0.293
X9	-0.181	0.055	0.424	0.149

本研究分析房地产上市公司财务数据,采用因子分析预测 F 值,并设定风险评估阈值为-0.3。F 值大于-0.3 表明公司财务稳健,风险等级 1; 小于-0.3 则可能面临财务困境,风险等级 0, 需采取防控措施。结果显示,70 家企业风险等级 1, 21 家风险等级 0, 后者应关注财务异常,进行问题分析和处理。

2 模型预测对比

2.1 逻辑回归模型

将 X1, X2, X3, X4 项作为自变量,而将观测值作为因变量,训练集比例设置为 0.8,采用 lbfsgs 优化算法,并且使用 L2 正则化,且设置截距,进行逻辑回归建模。

最终模型在测试集上得到的准确率为 88.89%,精确率(综合)为 88.89%,召回率(综合)为 88.89%,f1-score(综合)为 0.89。模型效果良好。

模型代码如下:

```
model = LogisticRegression(solver='lbfsgs',
penalty='l2', fit_intercept=True, max_iter=100,
tol=0.001)

model.fit(x_train, y_train)
```

2.2 BP 神经网络

研究结果表明:选取 X1, X2, X3, X4 作为自变量,以风险分类作为因变量,设定训练集比例为 0.8,各隐层网络节点数为(100),并采用 relu 函数作为激活函数,构建了神经网络模型。

最终模型在测试集上的准确率达到 83.33%,综合精确率为 81.25%,综合召回率为 83.33%,综合 f1-score 为 0.82。由此可见,模型表现出较为良好的预测性能。

模型代码如下:

```
model=MLPClassifier(activation='relu', alpha=1.0E-4, hidden_layer_sizes=(100), learning_rate='constant', learning_rate_init=0.001, batch_size='auto', max_iter=200, solver='adam', tol=1.0E-4)
```

```
model.fit(x_train, y_train)
```

2.3 XGBOOST

以 X1, X2, X3, X4 为自变量,观测值为因变量,训练集比例为 0.8,使用 gbtree 提升器,学习器数量 100,样本和特征采样率均为 1.0,分裂收益阈值、L1 正则化和 L2 正则化分别设为 0.0、0.0 和 1.0 进行 XGBoost 建模。

模型在测试集上准确率 94.44%,精确率 96.30%,召回率 94.44%,f1-score 为 0.95,显示了良好的性能。

模型代码如下:

```
model=xgb.XGBClassifier(booster='gbtree', n_estimators=100, eta=0.1, max_depth=6, subsample=1.0, colsample_bytree=1.0, min_child_weight=1.0, gamma=0.0, lambda=1.0, alpha=0.0')
```

```
model.fit(x_train, y_train)
```

2.4 预测结果对比

表 4 模型性能对比(测试集)

评估指标	逻辑回归模型	BP 神经网络模型	XGBoost 模型
准确率	88.89%	83.33%	94.44%
综合精确率	88.89%	81.25%	96.30%
综合召回率	88.89%	83.33%	94.44%
综合 F1-score	0.89	0.82	0.95
过拟合风险	低	中等	低
模型解释性	高	低	中等

总体来看，三种模型在预测结果上都达到了较高的准确率，逻辑回归模型的解释性较强；BP 神经网络模型能够捕捉非线性关系；XGBoost 模型特征重要性明确，三者都可以为房地产上市公司的财务预警提供一定的参考。

从实际情况上来看，XGBoost 模型在风险类别 0（高风险企业）的召回率为 67%，高于逻辑回归模型和 BP 神经网络模型，表明了 XGBoost 对小样本风险捕捉更优，并且训练速度较快（并行计算），所以优先选择 XGBoost 模型，结合特征重要性优化企业风控策略，其在预测性能上全面领先，适合高精准度要求的风险预警。

3 结语

本研究融合因子分析和机器学习技术，构建了基于因子分析和 BP 神经网络的财务风险预警模型，实证分析显示其准确率和 F1-score 分别为 83.33% 和 0.82，有效识别房地产企业财务风险，模型通过提取四个核心因子，降低数据维度，揭示房地产行业关键风险因素。引入现金流量和行业特定指标，提高预警精准性和适用性，实现了财务风险预警从静态评价到动态预测的转变。

研究显示房地产企业财务风险具有政策敏感性和现金流脆弱性。模型通过因子分析纳入非稳定性指标，避免了传统单一阈值模型的误判风险。在 92 家样本企业中，21 家（22.8%）被划分为高风险企业。这些企业普遍面临短期偿债能力不足、存货积压严重和现金流断

裂风险等问题。

企业需建立全周期风险管理体系，包括动态监控和预警机制，如财务驾驶舱系统跟踪核心指标并设定预警阈值。现金流优化策略包括实施“以销定产”模式和拓展融资渠道。政策合规管理要求设立小组解读监管动态，调整土地储备。监管层应构建行业风险联防机制，包括数据共享、预警平台建设，差异化政策调控，

以及风险处置和市场稳定措施，如设立行业纾困基金和参考 TARP 计划稳定市场。

参考文献

- [1] 张荣艳, 廖萌. Z 计分模型在房地产上市公司财务预警中适用性检验 [J]. 财会通讯, 2012.
- [2] 潘丹, 万芷寒. 基于功效系数法的新能源汽车企业财务风险预警 [J]. 生产力研究, 2023.
- [3] 贾玉婷. 煤炭上市公司财务风险预警研究 [D]. 河北地质大学, 2016.
- [4] 欧国良, 吴刚, 朱祥波. 基于因子分析法的房地产企业财务风险预警研究 [J]. 社会科学家, 2018.

基金项目：2024 年校级大学生创新创业训练计划项目
(项目编号：2024cxcy105)

作者简介：刘佳鑫（2002 年），女，汉，浙江省杭州市，本科在读，研究方向：金融工程。