Finance 金融, 2025, 15(2), 415-422
Published Online March 2025 in Hans. https://www.hanspub.org/journal/fin
https://doi.org/10.12677/fin.2025.152043

Hans汉斯

基于XGBoost算法的上市公司财务风险预警研究

朱顺泉

广州华商学院数字金融学院，广东 广州

收稿日期：2025年2月6日；录用日期：2025年2月17日；发布日期：2025年3月11日

摘　要

在当今的大数据和人工智能时代，机器学习在企业财务预警领域应用是热门的话题，而XGBoost算法是最重要的机器学习方法之一。论文在回顾上市公司财务预警相关研究的基础上，由于传统统计方法、神经网络法、决策树和GBDT算法在财务预警方面存在运行效率低、计算速度慢、学习效率低、缺乏交叉验证、过拟合、无自动筛选、预测精度低等方面的问题，提出基于机器学习的XGBoost算法建立了中国上市公司的财务预警模型。结果发现：XGBoost算法对中国上市公司的财务预警有较好的识别，具有广泛的适用性和推广价值。论文丰富了传统的公司财务预警方法，提供了公司财务预警的新思路。

关键词

XGBoost算法，上市公司，财务预警，应用

Research on Early Warning of Listed Corporate Financial Risk Based on XGBoost Algorithm

Shunquan Zhu

School of Digital Finance, Guangzhou Huashang College, Guangzhou Guangdong

Received: Feb. 6th, 2025; accepted: Feb. 17th, 2025; published: Mar. 11th, 2025

Abstract

In today's era of big data and artificial intelligence, the application of machine learning in the field of enterprise financial early warning is a hot topic, and the XGBoost algorithm is one of the most important machine learning methods. Based on a review of relevant research on financial distress pre-

文章引用: 朱顺泉. 基于 XGBoost 算法的上市公司财务风险预警研究[J]. 金融, 2025, 15(2): 415-422.
DOI: 10.12677/fin.2025.152043

diction for listed companies, this paper proposes the establishment of a financial distress prediction model for Chinese listed companies using the XGBoost algorithm, which is rooted in machine learning. This proposal arises due to the shortcomings of traditional statistical methods, neural networks, decision trees, and the GBDT algorithm in financial distress prediction, such as low operational efficiency, slow computation speeds, poor learning efficiency, lack of cross-validation, overfitting, absence of automatic feature selection, and low prediction accuracy. The result shows that: XGBoost algorithm has a good identification of the financial warning of Chinese listed companies and has a wide applicability and promotion value. The paper enriches the traditional company financial warning method and provides the new idea of company financial warning.

Keywords

XGBoost Algorithm, Listed Companies, Financial Early Warning, Application

1. 研究背景与意义

近年来，公司财务预警受到了学术界、金融财务实业界和政府部门的广泛关注。作为金融市场的一个重要领域，资本市场在我国发展了近二十年，股票、债券、基金、衍生品等各种投资方式已经逐渐被社会大众所接受，但资本市场的不规范现象十分严重，上市公司财务造假、上市圈钱、损害金融机构和股民利益的情况屡禁不止，上市公司因财务状况异常而被特别处理的现象屡见不鲜。因此，如何应用先进、科学的方法对上市公司进行财务预警，营造公平竞争的市场环境，规范和加强企业财务分类监管，关系到我国资本市场的健康发展。在大数据的背景下，传统的财务预警法如统计与 BP 神经网络等分类方法难以解决运行效率低、计算速度慢、学习效率低、缺乏交叉验证、容易产生过拟合、无自动筛选、预测精度低等方面的问题，对于复杂的资本市场则需要探索新的财务预警方法——基于 XGBoost 算法的财务预警方法来解决上述诸方面的问题，以提高财务预警的速度和质量。因此，本文的研究不但可以丰富传统的财务预警理论与方法，而且对广大的投资者、企业经营者、政府管理部门和资本市场的健康发展有着重要的意义，财务预警的质量将直接影响着风险防范与成本费用控制等，对金融监管部门、商业银行、投资银行、基金公司、保险公司等金融机构、上市公司的生存发展产生重要的影响。

2. 文献综述

国外对于上市公司财务预警的研究，大致经历了以下几类：(1) 定性分析法；(2) 统计分析法；(3) 神经网络法；(4) 基于市场价值的 KMV 方法。(1) 定性分析法，主要是 5C 要素分析法和 LAPP 原则，5C 要素分析法主要从 Character 品格、Capacity 能力、Capital 资本、Collateral 担保、Condition 环境等 5 个方面来分析财务状况；LAPP 原则主要从 Liquidity 流动、Activity 活动、Profitability 盈利、Potentialities 潜力等 4 个方面分析财务状况。除此之外还有杜邦财务分析体系和沃尔比重分析法等，这些方法的缺点是：主观性较强，受人的主观影响大，为了克服定性分析能力差，缺乏整体概括、定量分析不足等问题，国外从 20 世纪 60 年代开始，普遍采用了统计分析法。(2) 统计分析法，Beaver (1966)以企业的 29 个财务指标建立指标体系，通过该指标体系分析企业违约状况，发现财务指标是影响企业违约的重要因素[1]。但是，统计法对数据要求严格，如：数据要服从多元正态分布、变量间不存在多重共线性、配对样本协

方差矩阵应相同等等，而现实中的数据难以满足这样的要求。因此，后续学者采用 Logistic 等方法建立模型，Ohlson (1980)利用 Logit 回归模型预测企业违约概率，发现在样本数量比较多时，财务指标对预测具有一定的可信度[2]。(3) 神经网络法，20 世纪 80 年代末和 90 年代初，随着信息技术的发展，神经网络方法引入了信用评级，并且它能解决非正态分布、非线性的信用分类问题，但它难以解决小样本数据、局部极小点、高维数、函数逼近与分类能力弱、学习速度慢等问题。(4) 市场价值 KMV 法，20 世纪 90 年代后期出现了许多新的财务分类模型，最具代表性的有：KMV 公司开发的 KMV 模型与 JPMorgen 银行 1997 年在 VaR 模型的基础上建立的 CreditMetrics 模型。但 CreditMetrics 模型有很多参数需要确定，如等级迁移矩阵、资产之间的相关系数、远期收益率等，这些参数来自长时期统计数据积累，目前中国大陆还少有类似的统计资料。我国学者对财务预警的研究，是从 20 世纪 90 年代中后期开始的。朱顺泉(2009)应用期权定价方法对上市公司财务分类问题进行了初步研究[3]。孙林(2022)基于一般均衡分析模型，论述发债企业经济类型、行业属性、业务开展情况等对债券违约的影响。然后基于 2014~2020 年我国上市公司财务数据构建三组面板数据模型，发现企业杠杆水平、多元化经营和行业差异等指标对企业债券违约率存在显著影响，高杠杆和多元化经营策略会不同程度削弱发债企业主营业务竞争力对债券违约的抑制效果[4]。在运用 Logistic 模型方面，程昊(2020)基于企业内外部风险影响因素，构建风险识别指标体系，使用 Logistic 模型识别债券违约现象，效果较好[5]。在 Z-score 模型运用方面，何鑫超(2022)通过紫光集团财务状况和 Z 值检测证明了紫光集团在发生违约前就具有很高的债券违约风险[6]。黄卿(2018)等人采用沪深 300 股指期货 1 分钟高频数据为研究对象，对比了 SVM、XGBoost 和神经网络模型的预测性能，研究结果表明，XGBoost 算法在预测准确率上表现最为出色[7]。唐一峰(2021)以贷款发放时间、债务收入比、信贷周转余额、分期付款金额等因素作为评价指标，并利用 AUC 值作为评价指标进行对比分析，比较分析 XGBoost 和 LightGBM 算法在贷款违约预测中的效果。研究发现，在预测准确性方面，XGBoost 算法略胜一筹[8]。黄颖和杨会杰(2021)聚焦于黄金价格的涨跌趋势预测，利用 XGBoost 算法进行特征优化，并与 LSTM 算法进行了对比[9]。王怡宁(2021)将高频数据作为研究对象，并得出结论：高频因子在 XGBoost 算法预测中展现出更高的准确度[10]。

综上所述，对财务预警的研究，基本上都是经典的、定性分析、统计分析和神经网络方法来完成的，这些方法难以解决运行效率低、计算速度慢、学习效率低、缺乏交叉验证、过拟合、无自动筛选、预测精度低等方面的问题，对于复杂的金融市场则需要探索新的财务预警方法来解决这些问题。而基于统计学习理论的 XGBoosts 算法来建立财务预计建模，能够有效解决运行效率低、计算速度慢、学习效率低、缺乏交叉验证、过拟合、无自动筛选、预测精度低等方面的问题，因此，本文在继承和综合国内外现有研究成果的基础上，试图以 Wind 上市公司财务指标数据库为数据样本，将 XGBoosts 算法应用于上市公司的财务预警建模，并进行应用研究。

3. XGBoosts 算法及优势分析

3.1. XGBoost 算法

XGBoost 是 eXtreme Gradient Boosting 的简称，即扩展梯度提升算法，是 GBDT (Gradient Boosting Decision Tree)的拓展和改进，由学者陈天奇在 2014 年首次提出。基础理论是 Gradient Boosting 模型，XGBoost 每一步迭代都产生一个弱预测模型，然后加权累加到总模型中，然后每一步弱预测模型生成的依据都是损失函数的负梯度方向，这样若干步以后就可以达到逼近损失函数局部最小值的目标。它的基学习器除了可以是决策树，也可以是线性分类器。这里主要以决策树为基学习器进行分析。

基本的决策树模型表达式为：$\hat{y}_i = \sum_{k=1}^{k} f_k(x_i), f_k \in U$，式中 $U = \left\{ f(x) = w_{q(x)} \right\}$，$f_k$ 代表着不同的决策

树，U 则为所有的决策树集合，w 为每一颗决策树对应的叶子权重，q 为对应的树结构。目标函数包括两部分：$obj=\sum_{i=1}^{n} l\left(y_i,\hat{y}_i\right)+\sum_{k=1}^{k}\Omega\left(f_k\right)$，式中，第一部分 l 是预测值 $\hat{y}_i$ 和目标真实值 y 之间的误差，第二部分是每颗决策树的复杂度的总和。该公式决策树模型中的目标函数无法用随机梯度下降 GSD 等传统方法优化，所以采用 boosting 方式训练，即每一次都在保留原模型上添加一个新函数 $f_t\left(x_i\right)$。选择在每一轮加入新函数是为了尽可能地让目标函数最大程度的减小。通过 f_t 可以优化这个目标函数，当误差函数 l 是平方误差时，可以直接改写目标函数；若是非平方误差的其他行数，则可以用泰勒展开获得近似的目标函数，移除常数项后，和是平方误差时的目标函数统一，即：$\tilde{L}(t)=\sum_{i=1}^{n}\left[g_i f_t\left(x_i\right)+1/2h_i f_t^2\left(x_i\right)\right]+\Omega\left(f_t\right)$。

3.2. XGBoost 算法优势分析

和传统的统计方法、神经网络法、决策树和 GBDT 算法等相比，当数据量比较大时，计算速度会很慢，XGBoost 改进了算法，并且调用 CPU 进行并行的多线程计算，引入了正则化项，在提高了计算速度的同时也可以提升精度。比较而言，XGBoost 有如下的优点：(1) 并行多线程计算：我们知道，传统决策树最耗时的步骤是对特征值进行排序。虽然 Boosting 算法都是顺序处理的，但 XGBoost 在迭代之前，先存为 block 结构进行预排序，然后在之后的每次迭代中重复使用该结构，降低了模型的计算总量，XGBoost 的并行处理相比 GBDT 模型大大提高了运算能力。(2) 开放性：用户可以在使用 XGBoost 模型时自行定义优化标准和目标，减少了使用模型时受到的限制。(3) 缺失值处理：当样本存在缺失值时，XGBoost 能自动学习分裂方向，同时在未来处理缺失值时可应用该处理方法。(4) 提高学习效率：XGBoost 在每次迭代之后，会为叶子结点分配学习速率，减少每棵树的影响，降低每棵树在模型中的权重，为后面深度算法学习提供更好的空间。(5) 正则化：在 GBDT 模型中，没有正则化，导致模型的过拟合度偏高，而 XGBoost 模型在处理时会引入正则项，提升算法结果的正则化，减少了算法的过拟合现象。(6) 交叉验证：方便选择最好的参数，比如你发现 30 棵树预测已经很好了，不用进一步学习残差了，那么停止建树；(7) 列抽样：XGBoost 借鉴随机森林的做法，支持列抽样，这样不仅能防止过拟合，还能降低计算。

4. 变量选择与建模样本

在 WIND 财务分析指标中，从盈利能力(30)、收益质量(5)、资本结构(17)、偿债能力(31)、营运能力(13)、现金流量(11)六个维度方面选取了 107 个财务指标变量。在使用传统的逻辑回归方法构建财务预警模型时，如果自变量较多，不需将全部的自变量加入到分析模型中进行拟合训练，可以先对自变量集合进行筛选，挑选一些合适的自变量放入模型。一般通过以下几个因素进行筛选：变量的预测能力、变量的鲁棒性(即变量健壮性 Robust，数值长时间保持相对稳定)、变量的简单性(易于获得和使用)、变量的易理解性等。XGBoost 算法具有自动筛选自变量的功能，只需把自变量放进模型里，会自动选出有用的自变量。通过使用 XGBoost 算法的训练结果发现：XGBoost 算法里使用了大于 5 次的特征有这些指标：主营业务比率、留存收益占总资产比率、EBITDA 占营业总收入比率、净资产收益率(加权)、CRGS 占营业收入比重、长期总策适合率、净资产收益率(扣除/平均)、营运资本占总资产比率、资本固定化比率、NOCF 占非流动负债比率、销售毛利率、有形净值债务率等，说明这些特征重要性比较高(注：这里 EBITDA 是息、税、折旧及摊销前利润，NOCF 是经营活动产生的现金流量净额，CRGS 是销售商品提供劳务收到的现金)。

将 WIND 财经数据库中未被 ST 的股票共 2028 家、经筛选过后的 ST 股票共 127 家，总计 2155 个样本随机打乱，然后按照 7:3 的比例划为训练集和测试集，其中训练集 1508 个，测试集 647 个。

如果将已经实际发生财务困境的年度定义为 T 年，那么 ST 公司的 T-3 的财务信息作为研究的样本，即若 2018 年被 ST 的公司，则选择 2015 年的财务信息数据作为 X 数据。非 ST 公司选择 2015 年的财务

数据作为 X 数据。Y 的定义，ST 为 1，非 ST 为 0。

5. XGBoost 模型的迭代与优化参数

通过不断迭代和优化，最终得到一个效果比较好和稳定的 XGBoost 模型，该模型的具体参数如下：

model = XGBClassifier (learning_rate = 0.05, n_estimators=42, max_depth = 3, min_child_weight = 1, gamma = 0, subsample = 0.6, colsample_bytree = 0.9, objective = 'binary: logistic', nthread = 4, scale_pos_weight = 1)

model.fit (train. iloc [:, 1:], train ['y'], eval_metric = "auc", eval_set = eval_set, verbose = True)

learning_rate：学习率，控制每次迭代更新权重时的步长，默认 0.3。值越小，训练越慢。典型值为 0.01~0.2；

n_estimators：总共迭代的次数，即决策树的个数，论文用了 42 棵决策树；

max_depth：树的深度，值越大，越容易过拟合；值越小，越容易欠拟合。这个参数的取值最好在 3~10 之间。本文选取值为 3；

min_child_weight：叶子节点最小权重，值越大，越容易欠拟合；值越小，越容易过拟合(值较大时，避免模型学习到局部的特殊样本)。本文选了一个比较小的值 1，因为这是一个极不平衡的分类问题。因此，某些叶子节点下的值会比较小；

gamma：惩罚项中叶子结点个数前的参数，起始值也可以选其它比较小的值，本文选 0；

subsample：0.6，随机选择 60%样本建立决策树；

colsample_bytree：0.9，随机选择 90%特征建立决策树；

objective：'binary: logistic'，定义学习任务及相应的学习目标，本文选取的目标函数二分类的逻辑回归问题，输出为概率；

nthread：这个参数用来进行多线程控制，应当输入系统的核数，如果希望使用 cpu 全部的核，就不要输入这个参数，算法会自动检测；

scale_pos_weight：在各类别样本十分不平衡时，把这个参数设定为一个正值，可以使算法更快收敛；

其它参数均使用默认参数。

6. 基于 XGBoost 模型的上市公司财务预警结论

6.1. AUC 评价角度的研究结论

机器学习实践中分类器常用的评价指标就是 AUC，AUC 值的含义是：如果一个分类器能输出得分 score，调整分类器的阈值，把对应的点画在图上，连成的这条线就是 roc，曲线下的面积就是 AUC。

首先 AUC 值是一个概率值，当随机挑选一个正样本以及一个负样本，当前的分类算法根据计算得到的 score 值将这个正样本排在负样本前面的概率就是 AUC 值。当 AUC 值越大，当前的分类算法越有可能将正样本排在负样本前面，即能够更好地分类。

计算出训练集和测试集按逻辑回归以及 XGBoost 模型分别的 AUC 值，结果如表 1。

Table 1. AUC values of the training and test sets of T-3 data via logistic regression and XGBoost
表 1. T-3 数据训练集和测试集经由逻辑回归和 XGBoost 的 AUC 值

AUC	逻辑回归	XGBoost
train	0.888	0.958
test	0.899	0.94

从表 1 可以看出：(1) XGBoost 的 AUC 值即便在结果较差的测试集也高达 94%，这说明利用 XGBoost 模型进行预警可以取得很好的效果；(2) XGBoost 模型的 AUC 无论是训练样本还是测试样本，都远远大于逻辑回归模型的 AUC 值；(3) XGBoost 的训练和测试样本的 AUC 相差不大，说明模型很稳定，没有过拟合情况。

6.2. 精确率和召回率评价角度的研究结论

通过混淆矩阵计算出来的不仅是 AUC，还有精确率和召回率等可用来评判模型的好坏。我们定义：TP：样本为正，预测结果为正；FP：样本为负，预测结果为正；TN：样本为负，预测结果为负；FN：样本为正，预测结果为负。则精确率(precision)：TP/(TP + FP)，正确预测为正占全部预测为正的比例；召回率(recall)：TP/(TP + FN)，正确预测为正占全部正样本的比例。

当评价 ST 时，精确率(precision)表示在预测为 ST 的样本中，真正是 ST 的样本所占的比例。召回率(recall)表示所有真正是 ST 的样本中，预测为 ST 所占的比例。

当评价非 ST 时，精确率(precision)表示在预测为非 ST 的样本中，真正是非 ST 的样本所占的比例。召回率(recall)表示所有真正是非 ST 的样本中，预测为非 ST 所占的比例。

以预测概率 0.4 为阈值，超过 0.4 预测为 ST 公司，其他的预测为非 ST 公司。

逻辑回归和 XGBoost 的精确率与召回率的结果分别如表 2 和表 3 所示。

Table 2. Accuracy and recall of logistic regression results for T-3 data
表 2. T-3 数据逻辑回归结果的精确率和召回率

预测真实	非 ST	ST	recall
非 ST	2016	30	0.99
ST	84	43	0.34
precision	0.96	0.59	

Table 3. Accuracy and recall of the XGBoost model results for the T-3 data
表 3. T-3 数据 XGBoost 模型结果的精确率和召回率

预测真实	非 ST	ST	recall
非 ST	2016	12	0.99
ST	69	58	0.46
precision	0.97	0.83	

从表 2 和表 3 可见：(1) 从模型的评价指标来看，XGBoost 模型对于样本数据在 AUC、精确率、召回率三个评价指标都要优越于逻辑回归的结果，模型预测的结果也更为精确；(2) XGBoost 模型预测 ST 的精确率 0.83 远大于逻辑回归的精确率 0.59；(3) XGBoost 模型预测 ST 的召回率 0.46 远大于逻辑回归的召回率 0.34。

6.3. 数据时间维度对于结果影响的研究结论

从逻辑上，越接近企业财务异常年度，企业经营往往已经开始出现问题，数据通常越糟糕，也越能反映企业的经营状况趋于恶化，而偏离企业出现问题的年份越久远，企业往往还没意识到经营状况或者外部环境的恶化，预警指标也不能显著揭露风险，所以可以理解 T-3 的数据预测显著性要好于 T-4 或者更早的数据。

以 T-4 数据为例，其分别采用逻辑回归模型和 XGBoost 模型结果的评价指标情况如表 4~6。

Table 4. AUC values for the T-4 data using logistic regression and XGBoost-model results
表 4. T-4 数据用逻辑回归和 XGBoost 模型结果的 AUC 值

AUC	逻辑回归	XGBoost
train	0.822	0.915
test	0.789	0.864

Table 5. Accuracy and recall of logistic regression results for T-4 data
表 5. T-4 数据逻辑回归结果的精确率和召回率

预测真实	非 ST	ST	recall
非 ST	2018	10	0.995
ST	100	27	0.21
precision	0.95	0.73	

Table 6. Accuracy and recall of the XGBoost model results for the T-4 data
表 6. T-4 数据 XGBoost 模型结果的精确率和召回率

预测真实	非 ST	ST	recall
非 ST	2014	14	0.99
ST	79	48	0.38
precision	0.96	0.77	

从上面的数据分析结果来看，在 AUC、精确率和召回率三个评价指标方面，T-4 的数据的有效性和显著性要远落后于 T-3 数据，T-4 数据相比 T-3 数据预测结果不足以推断企业 T 年的财务风险状况，继而也无法去可靠地分析企业 T 年被出现财务危机的可能性。

基金项目

本研究为广东省重点建设学科科研能力提升项目(项目编号 2024ZDJS113)；广州华商学院应用型示范专业－金融科技专业建设项目 HS2024SFZY08；广州华商学院金融科技专业核心课程教研室建设项目 HS2024ZLGC43 等阶段性成果。

参考文献

[1] Beaver, W.H. (1966) Financial Ratios as Predictors of Failure. *Journal of Accounting Research*, **4**, 71-111. https://doi.org/10.2307/2490171

[2] Ohlson, J.A. (1980) Financial Ratios and the Probabilistic Prediction of Bankruptcy. *Journal of Accounting Research*, **18**, 109-131. https://doi.org/10.2307/2490395

[3] 朱顺泉. 基于期权定价理论的上市公司信用分类建模及应用研究[J]. 统计与信息论坛, 2009, 24(7): 23-38.

[4] 孙林, 孙健. 上市公司债券违约的影响因素分析——基于企业财务杠杆和多元化经营的研究[J]. 上海金融, 2022(11): 2-11

[5] 程昊, 朱芳草, 等. 我国债券违约风险评估模型的构建与应用[J]. 开发性金融研究, 2020(6): 76-89.

[6] 何鑫超. 企业债券违约成因及风险防范[D]: [硕士学位论文]. 昆明: 云南财经大学, 2022

[7] 黄卿, 谢合亮. 机器学习方法在股指期货预测中的应用研究-基于 BP 神经网络、SVM 和 XGBoost 的比较分析[J].

数学的实践与认识, 2018, 48(8): 297-307

[8] 唐一峰. 基于 XGBoost 算法和 LightGBM 算法的贷款违约预测模型研究[J]. 现代计算机, 2021, 27(32): 33-37

[9] 黄颖, 杨会杰. 基于 XGBoost 和 LSTM 模型的金融时间序列预测[J]. 科技和产业, 2021, 21(8): 158-162

[10] 王怡宁. 基于 XGBoost 的高频交易选股研究[D]: [硕士学位论文]. 上海: 东华大学, 2021.