

# 基于代价敏感 XGBoost 模型的企业 ESG 评级预测

鲍金敏

合肥工业大学 数学学院, 安徽 合肥 230009

DOI:10.61369/ASDS.2026060010

**摘 要 :** 针对企业环境、社会与治理 (ESG) 评级分布不平衡导致的预测偏差, 以及现有模型对不同行业 ESG 行为解释力不足的问题, 本文提出一种基于代价敏感极限梯度提升 (XGBoost) 与沙普利加法解释 (SHAP) 框架结合的改进预测模型。我们引入代价敏感学习策略, 并结合合成少数类过采样技术 (SMOTE), 强化模型对少数类等级样本的捕获能力。同时量化财务指标的边际增益并利用 SHAP 算法剖析行业异质性。研究结果表明代价敏感优化显著提升了极端评级的预测精度, “落后者”召回率91.72%, “领导者”召回率92.24%; 财务基本面和行业信息的引入有效修正了历史评级的依赖性, 宏平均受试者工作特征曲线下面积 (Macro AUC) 高达97.81%; 驱动机制存在行业分化, 重污染行业呈现出“资产密集型合规”逻辑, 而非重污染行业则展现出“效率溢价与市场敏感”特征。本文为投资者精准锚定优质 ESG 标的、排查尾部风险提供了高效的决策工具。

**关 键 词 :** ESG 评级; 代价敏感 XGBoost 模型; SHAP; 行业异质性

## Enterprise ESG Rating Prediction Based on Cost-Sensitive XGBoost Model

Bao Jinmin

School of Mathematics, Hefei University of Technology, Hefei, Anhui 230009

**Abstract :** Aiming at the prediction bias caused by the imbalance in the distribution of ESG ratings and the insufficient explanatory power of existing models for ESG behavior in different industries, this paper proposes an improved prediction model based on the combination of cost-sensitive limit gradient lifting (XGBoost) and Shapli addition explanation (SHAP) framework. We introduce a cost-sensitive learning strategy combined with synthetic minority oversampling (SMOTE) to enhance the model's ability to capture minority-class samples. At the same time, the marginal gain of financial indicators is quantified, and the SHAP algorithm is used to analyze the industry heterogeneity. The results show that cost-sensitive optimization significantly improves the prediction accuracy of extreme ratings, with a recall rate of 91.72% for "laggards" and 92.24% for "leaders". The introduction of financial fundamentals and industry information effectively corrects the dependence of historical ratings, and the macro average receiver operating characteristic area (Macro AUC) is as high as 97.81%. The driving mechanism is differentiated by industries, with heavy polluting industries showing the logic of "asset-intensive compliance", while non-heavy polluting industries showing the characteristics of "efficiency premium and market sensitivity". This article provides an efficient decision-making tool for investors to accurately anchor high-quality ESG targets and investigate tail risks.

**Keywords :** ESG rating; cost-sensitive XGBoost model; SHAP; Industry heterogeneity

## 引言

在全球经济逐步转向可持续发展、我国“双碳”目标持续推进的背景下, 环境、社会与治理 (ESG) 表现已不仅是监管部门推动企业绿色转型的重要抓手, 也在企业层面体现出较为明显的价值创造作用。从信号理论与代理理论视角看<sup>[1]</sup>, 管理层借助 ESG 评级向股东传递相关信息, ESG 表现较好的企业往往更容易获得投资者关注, 并对其融资方式和融资工具产生影响<sup>[2]</sup>。在价值创造方面, 研究表明, ESG 评价与资产回报率、股本回报率等财务绩效指标呈正相关关系, 是推动企业价值提升的重要因素<sup>[3, 4]</sup>。在具体应用中, 将 ESG 信息纳入制造业评估体系, 有助于增强模型对企业信用风险的识别与预警能力<sup>[5]</sup>。不过, 现有 ESG 评级体系仍存在第三方评价结果滞后、信息披露成本较高以及评级过程透明度不足等问题, 不同评级机构之间还存在较为明显的分歧现象<sup>[6]</sup>。在此背景下, 构建更为准确的 ESG 预测机制, 对引导企业绿色创新和促进高质量发展具有现实意义<sup>[7, 8]</sup>。

作者简介: 鲍金敏, 合肥工业大学数学学院, 硕士研究生, 研究方向为 ESG 数据分析与机器学习。

为提高 ESG 表现的预判准确性,学术界逐步将研究重点放在机器学习方法的应用上。DeLucia 等人<sup>[9]</sup>选用随机森林、支持向量机等模型,并基于汤森路透数据库实证分析,发现 ROA 与 ROE 对 ESG 指标具有较好的预测效果。Gupta 等人<sup>[10]</sup>则通过线性回归与随机森林回归方法,进一步分析利润率对 ESG 评分的影响。D'amato 等人<sup>[11]</sup>从财务报表角度出发,指出资产负债表项目对 ESG 评级具备显著解释力。随着算法不断发展,研究开始更多采用非线性模型以提升预测精度。CINI 等人<sup>[12]</sup>利用随机森林分类模型,对道琼斯欧洲斯托克 600 指数公司进行分析,结果表明资产负债变量及系统性风险对评级类别具有较强解释能力,模型预测效果较为理想。Kartal 等人<sup>[13]</sup>以土耳其上市公司 ESG 报告为样本,引入超级学习者算法,模型预测准确率达到 94.3%。TASKIND 等人<sup>[14]</sup>则发现,历史 ESG 评分与评级本身即可作为未来 ESG 表现的有效预测依据,为复杂数据驱动方法提供了简洁替代路径。Zhang 等人<sup>[15]</sup>结合 XGBoost 算法与 SHAP 解释方法,在提升评级预测准确性的同时,对各类指标的整体贡献进行了更为细致的分析。

尽管现有研究虽已取得一定进展,但仍存在若干不足。首先,在识别偏误方面,ESG 等级通常呈现明显的不平衡分布,传统模型多默认各类别地位相同,实际训练中往往降低了对少数类样本的识别能力,从而影响极端风险情形下的稳定性。其次,在评价维度上,尽管已有研究关注历史评级与财务表现之间的联系,但对其他基本面信息在评级预测中的边际作用缺乏清晰界定。再次,在解释层面,现有成果多停留在行业整体分析,对不同行业间决策逻辑差异的刻画仍不充分。

针对上述问题,本文构建基于改进 XGBoost 的企业 ESG 评级预测框架。具体而言:其一,在处理类别不平衡问题时,通过引入类别权重的代价敏感学习方法,在目标函数中设置差异化惩罚权重,从而增强模型对少数类别样本的识别能力,缓解由样本分布不均带来的偏置。其二,通过多维度消融实验,对其他基本面信息在 ESG 评级预测中的增量作用进行量化分析,为企业识别关键变量提供依据。其三,在结果解释方面,借助 SHAP 方法分析不同行业的驱动机制差异,进一步揭示重污染与非重污染企业在 ESG 表现上的不同逻辑。

本文围绕 ESG 评分的可解释预测问题,构建了较为系统的研究框架,并为后续相关研究提供参考依据。行文结构安排如下:第一部分介绍理论基础,第二部分说明模型与方法,第三部分呈现实证结果及模型表现,第四部分归纳主要结论并对未来研究进行展望。

## 一、基本方法

### (一) XGBoost 算法

XGBoost<sup>[16]</sup>是一种先进的集合学习算法,旨在提升梯度增强树的计算效率和预测性能,其通过迭代构建多棵决策树来拟合前一模型的残差。在 ESG 评级的预测背景下,XGBoost 算法旨在最小化由损失函数和正则化项两部分组成的目标函数 $L^{(t)}$ 。如下所示:

$$L^{(t)} = \sum_{i=1}^n l(y_i, \hat{y}_i^{(t-1)} + f_t(x_i)) + \Omega(f_t) \quad (1)$$

$$\Omega(f_t) = \gamma T + \frac{1}{2} \lambda \sum_{j=1}^T w_j^2 \quad (2)$$

$l(\cdot)$ 为度量 ESG 真实评级 $y_i$ 与预测评级 $\hat{y}_i$ 差异的损失函数; $\Omega(f_t)$ 为正则化项, $\gamma$ 和 $\lambda$ 是正则化项的系数,叶节点数量 $T$ 及权重 $w_j$ 用来控制模型复杂度。XGBoost 通过对目标函数进行二阶泰勒展开,利用一阶和二阶导数信息寻找最优分裂点,能够有效捕捉预测指标和评级之间的非线性关系。

### (二) 代价敏感学习

ESG 评级样本在等级分布上往往呈现不平衡特征,多数类别的样本在传统分类模型训练过程中会起到主导作用,导致模型对少数类别样本的判别能力不足,产生系统性的分类偏置问题。为缓解上述问题,本文引入基于类别权重调整的代价敏感学习策略<sup>[17]</sup>,对模型训练过程中的不同类别样本赋予差异化的重要性权重。

具体来说,代价敏感学习通过在模型损失函数中引入类别权重向量,使模型在优化过程中对少数类别样本的误判承担更高的

损失惩罚,引导模型更加关注少数类别的判别边界学习。本文在此基础上结合 SMOTE<sup>[18]</sup>过采样技术,构建“数据层重采样—算法层加权”的双重均衡机制。首先利用 SMOTE 方法对训练集中少数类别样本进行线性插值扩充,接着在 XGBoost 模型的目标函数中引入基于类别的权重向量,对不同等级样本在损失函数中的相对重要性进行调整。改进后的目标函数如(3)所示:

$$L_{cs}^{(t)} = \sum_{i=1}^n w_{y_i} \cdot l(y_i, \hat{y}_i^{(t-1)} + f_t(x_i)) + \gamma T + \frac{1}{2} \lambda \sum_{j=1}^T w_j^2 \quad (4)$$

$w_{y_i}$ 为类别权重向量。本文结合样本分布特征与实验调优结果对少数类等级赋予较高的惩罚权重,对多数类等级赋予基准权重。

### (三) SHAP 算法

SHAP 算法利用 Shapley 值来解释机器学习模型,旨在为每个特征为每个样本分配一个贡献值。通过计算其对应 SHAP 值的绝对平均值来量化特征的显著性。通过特征值与其 SHAP 值的图形表示,可以辨别每个特征对模型预测的影响<sup>[19]</sup>。该方法既可用于分析单一样本层面的预测机制,也能够整体层面揭示模型的特征依赖结构,有助于实现对模型预测行为的局部与全局理解。

设定样本数量为 $n$ ,特征集合为 $F$ ,特征维度为 $M$ 。对于第 $i$ 个样本,模型的预测输出可表示为各特征贡献值的加法和:

$$f(x_i) = \phi_0 + \sum_{j=1}^M \phi_{ij} \quad (5)$$

$f(x_i)$ 为解释模型, $\phi_0$ 为所有样本预测的平均值, $\phi_{ij}$ 代表第 $i$

个样本中第  $j$  个特征的贡献值。 $\phi_{ij}$  的计算公式如下：

$$\phi_{ij} = \sum_{S \subseteq F \setminus \{j\}} \frac{|S|!(M-|S|-1)!}{M!} [f_{S \cup \{j\}}(x_i) - f_S(x_i)] \quad (5)$$

$S$  是特征集合  $F$  中不包含特征  $j$  的子集， $f(S)$  表示特征子集  $S$  下模型的预测输出。本文采用 “One-vs-Rest” 策略<sup>[20]</sup>，分别计算每个目标类别对应的 SHAP 值， $I_j^{(k)}$  表示  $i$  个样本中第  $j$  个特征对类别  $k$  的贡献值，则第  $j$  个特征在类别  $k$  下的全局重要性定义为：

$$I_j^{(k)} = \frac{1}{N} \sum_{i=1}^N |\phi_{ij}^{(k)}| \quad (6)$$

同时，本文将各类别下的全局重要性进行等权累加用来刻画特征在整体多分类任务中的综合影响力，定义总重要性为：

$$I_j = \sum_{k=1}^K I_j^{(k)} \quad (7)$$

其中  $K=3$  表示评级类别数量。

二、模型与方法

（一）数据来源和变量说明

本文构建以历史 ESG 评级与财务指标为核心的预测指标体系。参考既有研究<sup>[2, 11]</sup>所用框架，本文筛选与 ESG 表现关联较强的关键财务变量，以保证评估结果具备完整性与可靠性。整个指标体系划分为五个维度。就盈利与营运情况而言，选取净资产收益率、总资产净利润率、息税前利润以及人均创利，用于衡量资源配置效率。围绕资本结构与风险水平，引入资产负债率、利息保障倍数以及自由现金流，以反映财务稳健程度。针对市场表现与成长性，采用年换手率、市盈率、营业收入增长率和资本密集度，对市场关注程度与扩张能力进行刻画。在行业划分方面，依据证监会行业分类标准，将样本区分为“重污染行业”和“非重污染行业”，并设置虚拟变量，用于识别行业差异。考虑到历史评级具备延续特征<sup>[14]</sup>，本文引入“上年综合得分”，以反映 ESG 表现的动态惯性。

本文数据来源于华证 ESG 评价体系与 CSMAR 数据库，相关指标定义见表1。

表1：指标定义

| 维度     | 变量名称    | 变量符号       | 变量定义与计算方法                         |
|--------|---------|------------|-----------------------------------|
| 历史评级特征 | 上年综合得分  | Score{t-1} | 企业在 t-1 年度的华证 ESG 综合得分            |
|        | 净资产收益率  | ROE        | 净利润 / 股东权益平均余额                    |
|        | 总资产净利润率 | ROA        | 净利润 / 总资产平均余额                     |
|        | 息税前利润   | EBIT       | 利润总额 + 利息支出                       |
|        | 人均创利    | PPL        | 净利润 / 员工总数                        |
|        | 营业收入增长率 | Growth     | (本期营收 - 上期营收) / 上期营收              |
| 财务盈利能力 | 资本密集度   | CI         | 总资产 / 营业收入                        |
|        | 年换手率    | Turnover   | 统计区间内股票累计成交量 / 流通总股数              |
|        | 资产负债率   | Lev        | 总负债 / 总资产                         |
|        | 利息保障倍数  | ICR        | 息税前利润 / 利息费用                      |
|        | 企业自由现金流 | FCFF       | 息税前利润 - 税收 + 折旧摊销 - 营运资本变动 - 资本开支 |
|        | 市盈率     | PE         | 每股市价 / 每股收益                       |
| 行业属性   | 重污染分组   | Pollute    | 虚拟变量：若企业属于重污染行业取1，否则取0            |

（二）样本划分与不平衡处理策略

本研究以华证 ESG 评级为主要数据来源，实现了严格的数据收集和预处理程序。我们选取2014—2023年沪深 A 股上市公司企业 ESG 表现、财务指标及行业属性为初始数据集。为确保数据有效性，本文实施如下预处理：首先，剔除样本期间内被标记为 ST、PT 或 ST\* 的异常经营企业；其次，针对部分缺失值采用 KNN 算法进行填补，并删除存在严重数据异常的公司。最终构建了一个包含610家公司、共6100个企业年度观测值的平衡面板数据集。

在评级划分方面，由于华证 ESG 评价分为九级，且样本高度集中于中位等级，为缓解类别失衡并提升模型预测精度，本文参照中证指数（CSI）官方标准，将原始等级整合为三类：将 AAA 至 BBB 级归为“领导者”，BB 至 CCC 级归为“平均者”，CC

及 C 级归为“落后者”。为便于模型计算，将三类标签编码为0、1、2。

此外，为进一步处理整合后仍存在的样本分布差异，本文采用 SMOTE 算法对数据集缓解不平衡处理，样本总量由初始的6100个扩充至10429个。样本均衡后的频率百分比见表2。

表2：样本均衡后频率的百分比

| 级别  | 频率    | 百分比（%） |
|-----|-------|--------|
| 领导者 | 3212  | 30.80  |
| 平均者 | 5589  | 53.59  |
| 落后者 | 1628  | 15.61  |
| 总计  | 10429 | 100.00 |

（三）代价敏感 XGBoost 模型参数设计

XGBoost 模型的性能依赖于超参数的调优<sup>[21]</sup>。本研究针对 ESG 三分类任务，对核心超参数进行了系统化配置，通过随机搜

索策略与交叉验证的结合，最终确定的最优超参数配置见表3。

表3：最优超参数配置超参数

| 超参数    | 取值                    | 说明                  |
|--------|-----------------------|---------------------|
| 最大树深度  | 12                    | 控制基学习器复杂度，用于刻画非线性关系 |
| 迭代轮数   | 1000                  | 表示基学习器数量，影响模型拟合能力   |
| 学习率    | 0.015                 | 控制新增树对模型的贡献权重       |
| 样本采样比例 | 0.8                   | 每轮随机抽取80%样本进行训练     |
| 特征采样比例 | 0.8                   | 每棵树随机选取80%特征建模      |
| 随机种子   | [42,100,2024,888,123] | 多随机种子重复实验并取平均值      |
| 代价敏感系数 | {0.3,1:1,2:2}         | 提高少数类别的惩罚权重         |

（四）模型评估指标

混淆矩阵用于评估多分类模型性能，通过统计实际类别与预测结果的对应关系，可以直观看出模型判别偏差。在本研究的ESG 评级三分类模型中，混淆矩阵C 是一个3×3 的方阵。横轴表示预测等级，纵轴表示实际等级，类别划分为“落后者”“平均者”“领导者”。矩阵中的元素  $C_{ij}$  表示实际等级为 i 但被分类为 j 的样本数量。

1. 准确性

准确率表示预测正确的样本在总体中的占比，其计算方式见公式（8）。

$$Acc = \frac{\sum_{i=0}^2 C_{ii}}{\sum_{i=0}^2 \sum_{j=0}^2 C_{ij}} \quad (8)$$

其中： $i, j \in \{0,1,2\}$  分别代表“落后者”“平均者”与“领导者”。

2. 精度、召回率和 F1 分数

精度用于衡量预测结果的可靠性，即被判为某一等级的样本中，真实属于该等级的比例。对应计算方式见公式（9）。

$$P_i = \frac{C_{ii}}{\sum_{j=0}^2 C_{ji}} \quad (9)$$

召回率反映识别能力，即在实际属于某一等级的样本中，被正确识别的比例，对于等级  $i$  其召回率公式如（10）所示。

$$R_i = \frac{C_{ii}}{\sum_{j=0}^2 C_{ij}} \quad (10)$$

F1 分数综合考虑精度与召回率，取二者调和平均值，用以衡量单一等级上的整体判别效果，计算方式见公式（11）。

$$F1_i = 2 \times \frac{P_i \times R_i}{P_i + R_i} \quad (11)$$

精度、召回率和 F1 分数的数值越高，表示模型在相应分类层面上表现更优越。考虑到“平均者”样本数量较多，容易对整体评价产生主导影响，本文引入宏平均方法。该方法对各类别指标

赋予相同权重后取平均值，从而更合理地反映模型在少数类别上的识别能力，其计算方式见公式（12）。

$$M_{macro} = \frac{1}{3} \sum_{i=0}^2 m_i \quad (12)$$

其中  $m_i$  表示第  $i$  类的对应性能指标。本文在未特指某一等级的性能指标情况下（如宏精度、宏召回率、宏 F1 值及 Macro AUC），均指代经宏平均法计算后的结果。

3. ROC 曲线与 AUC 值

ROC 曲线以图形方式展示了真向正类率与假向正类率之间的权衡关系。本研究分别绘制“落后者”“平均者”与“领导者”的 ROC 曲线。AUC 作为衡量模型分类性能的标量值，其数值越接近 1，代表模型的判别准确度越高。

三、模型预测结果与行业异质性分析

（一）预测性能与策略有效性分析

本文将平衡后的数据集按 8:2 的比例划分为训练集和测试集，并基于改进后的代价敏感 XGBoost 模型进行等级判别。为了实现精准的等级分类建模，本文将华证 ESG 评级整合为三个核心类别：第 0、1 和 2 类分别代表“落后者”“平均者”与“领导者”。为兼顾结果的可视化直观性与统计稳健性，本文在图示层面报告单一随机种子下的混淆矩阵结果，而在模型评价层面采用多随机种子重复实验并取均值作为最终结论依据。测试集样本的预测分布结果由图 1 的混淆矩阵呈现（以随机种子 123 的单次实验结果为例）。

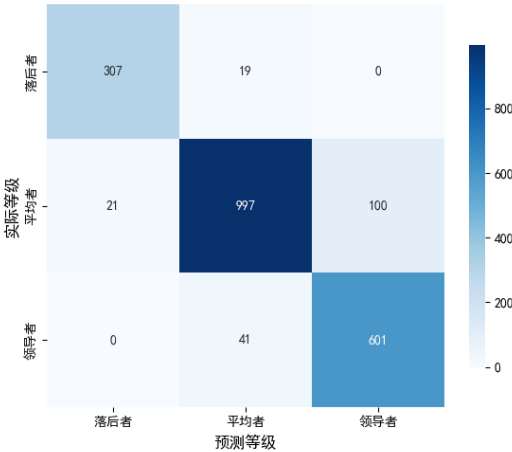


图1：代价敏感 XGBoost 模型的混淆矩阵

从图 1 可以看出，代价敏感 XGBoost 模型展现了显著的分类性能，样本高度集中在主对角线上。并且模型对于“落后者”与“领导者”的判别没有跨级误判，这意味着模型能够精准识别出表现最优与最差的企业。误判样本主要分布在相邻等级之间，如有 100 个样本所属于“平均者”被归类为“领导者”行列。为消除数据随机划分对模型评价的潜在干扰，本文在相同的超参数配置下，选取了 5 组不同的随机种子进行重复实验，并以各评估指标的算术平均值作为最终评价依据。实验结果如表 4 所示：

表4：基于代价敏感优化的模型预测性能对比检验

|                 | 准确率    | 宏精度    | 宏 F1 值 | 落后者召回率 | 领导者召回率 |
|-----------------|--------|--------|--------|--------|--------|
| XGboost 模型      | 89.95% | 89.91% | 89.70% | 87.85% | 90.53% |
| 代价敏感 XGboost 模型 | 90.22% | 89.62% | 90.18% | 91.72% | 92.24% |

通过表4，基准 XGboost 模型与代价敏感 XGBoost 模型在整体平均准确率和宏平均精度都能达到89%以上，体现了 XGBoost 框架在 ESG 评级预测中的效能。同时代价敏感学习发挥了重要的调节作用，能够在提升整体准确率的同时显著改善极端类别识别能力，使“落后者”召回率从87.85%提升至91.72%，“领导者”召回率从90.53%优化至92.24%。这种针对性优化，增强了模型对少数等级样本的捕捉能力，为投资者锚定尾部风险与优质标的提供了可靠的技术支撑。

本文进一步通过 ROC 曲线考察了代价敏感 XGBoost 模型对不同 ESG 等级的判别区分能力。根据图2，“落后者”“平均者”与“领导者”三类预测的 ROC 曲线均高度趋近于左上角，对应的单类 AUC 均值分别达到了0.9948、0.9685和0.9815。表明该模型在为投资者提供公司在 ESG 领域的相对位置方面具有高度准确性。

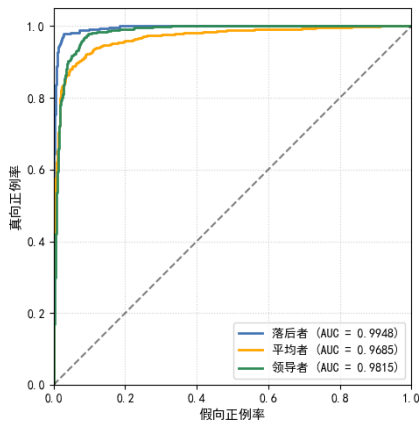


图2：代价敏感 XGboost 模型的 ROC 曲线

## （二）特征重要性分析

在完成代价敏感 XGBoost 模型构建后，本文引入 SHAP 方法分析特征贡献。为量化各变量影响，通过计算各特征在三个类别上的 SHAP 平均绝对值并进行堆叠排序。由此可以识别各特征在模型决策中的相对重要程度。

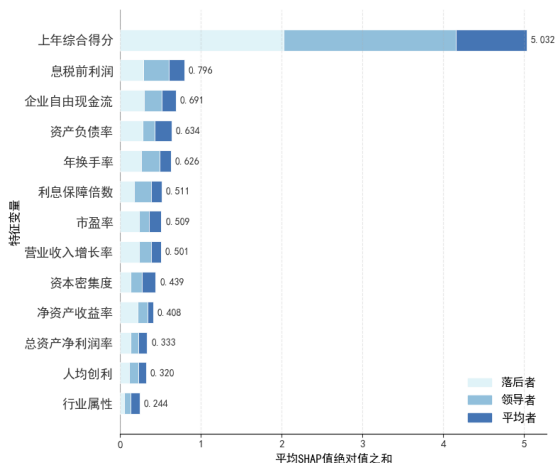


图3：特征全局贡献度量化排名

图3显示，“上年综合得分”对应 SHAP 均值之和达到5.032，

远高于其他变量。该结果表明，企业 ESG 表现存在明显路径依赖，历史评级对未来结果具有关键影响。

虽然“上年综合得分”占主导地位，但其他变量同样发挥作用。息税前利润、企业自由现金流以及资产负债率对应 SHAP 均值之和均超过0.6，对模型输出形成补充影响。其余指标也参与模型判别过程，并带来一定增量信息，相关结果见表5。

表5：财务变量对模型性能的整体增量分析

|                | 准确率    | 宏精度    | 宏平均召回率 | 宏 F1 值 | Macro AUC |
|----------------|--------|--------|--------|--------|-----------|
| 仅历史得分模型        | 74.18% | 74.07% | 68.32% | 70.35% | 87.48%    |
| 历史 + 财务 + 行业模型 | 90.22% | 89.62% | 90.86% | 90.18% | 97.81%    |

消融实验统一采用代价敏感超参数配置。表5结果表明，引入财务变量与行业属性后，单一维度预测局限得到缓解，模型整体性能明显改善。具体来看，准确率由74.18%提升至90.22%，宏平均精度由74.07%提高至89.62%，Macro AUC 达到97.81%。上述结果说明，历史评级具有较强解释力，但单一依赖会带来路径依赖偏差，财务与行业信息可作为补充进行修正。

## （三）行业异质性分析

鉴于各行业在环境规制、资源消耗及社会责任履行方面差异明显，本文将样本划分为重污染行业与非重污染行业两类。基于该划分，进一步检验代价敏感 XGBoost 模型在不同情境下的适用性，并分析其驱动机制差异。

### 1. 预测性能的稳健性检验

从表6可见，模型在不同行业样本中均表现出稳定的预测能力，并能够有效识别“落后者”与“领导者”。其中，不同行业属性下 Macro AUC 表现均高于全样本模型下的97.81%，表明当样本空间缩小至特定行业内部时，行业内部结构更为同质化，有利于模型学习稳定判别边界，从而展现出更优的判别灵敏度。其中非重污染行业的各项评价指标略高于重污染行业，其 F1 值达到93.29%，Macro AUC 值高达98.97%。这表明尽管重污染行业面临更为复杂的监管环境，其预测准确率与精度依然维持在较高水平，并与非重污染行业的差距极小。

表6：基于行业异质性的模型预测性能对比检验

| 模型组别   | 准确率    | 宏精度    | 宏平均召回率 | 宏 F1 值 | Macro AUC |
|--------|--------|--------|--------|--------|-----------|
| 重污染行业  | 92.06% | 92.49% | 92.05% | 91.87% | 98.57%    |
| 非重污染行业 | 93.43% | 93.70% | 93.43% | 93.29% | 98.97%    |

### 2. 驱动机制的差异化分析

本文引入 SHAP 方法，对重污染行业与非重污染行业的驱动差异进行分析。图4展示“领导者”类别预测结果中各特征的边际贡献分布，横轴为 SHAP 值，用于衡量变量对预测结果的影响方向与强度。结合特征取值分布可以进一步判断变量作用方向：红色表示取值较高，蓝色表示取值较低。若高取值样本主要集中在横轴正向区间，则说明该变量与评级结果呈正向关系。

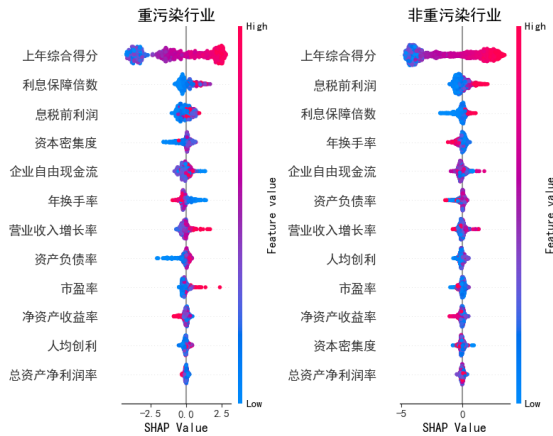


图4：“领导者”行业驱动机制对比分析

结果显示，不论企业是否属于重污染行业，“上年综合得分”始终处于核心位置，是能否进入“领导者”行列的关键因素。SHAP 分布进一步表明，该指标取值较高时，多集中在横轴正向区间，对预测结果产生明显正向作用。与此同时，其他财务变量在不同行业中呈现差异化特征。

就重污染行业而言，资本密集度重要性排名明显靠前，高于非重污染行业。结合 SHAP 分布可见，较高资本密集度通常对应更高评级结果，二者呈正向关系。这一现象反映出，在重资产行业中，固定资产投入规模往往与企业治理能力相关，进而影响 ESG 表现。

就非重污染行业而言，人均创利与年换手率的重要性更为突出，其排名均高于重污染行业。该结果表明，对于轻资产企业，模

型更依赖人力产出效率以及市场交易活跃程度来判断 ESG 等级。

四、结论

（一）企业 ESG 表现呈现明显路径依赖特征

各模型结果均显示，“上年综合得分”始终处于特征重要性首位。对应 SHAP 值集中于正向区间，对预测结果产生持续推动作用。该现象说明，ESG 评级具有延续性，既往建立的 ESG 治理体系与声誉积累会对后续表现产生长期影响。

（二）财务信息对评级预测具有补充作用

消融结果表明，引入财务指标后，模型准确率由 74.18% 提高至 90.22%，Macro AUC 达到 97.81%。这一变化表明，财务数据不仅反映经营状况，同时能够缓解历史评级带来的偏置，对预测结果形成有效补充。

（三）不同行业存在差异化驱动路径

分行业结果显示，重污染行业中资本密集度排名更高，而非重污染行业中，人均创利与年换手率更为关键。上述差异说明，重资产企业更依赖资本投入支撑表现，轻资产企业则更多受效率与市场活跃度影响。

（四）代价敏感方法提升极端类别识别能力

针对样本分布两端较少的情况，引入差异化权重后，模型对“落后者”与“领导者”的识别能力得到增强，对应召回率分别达到 91.72% 与 92.24%。该方法有助于提高极端类别识别效果，为风险识别与优质样本筛选提供参考。

参考文献

[1]JENSEN M C, MECKLING W H. Theory of the firm: managerial behavior, agency costs and ownership structure[J]. Journal of Financial Economics, 1976, 3(4): 305–360.

[2]D’ AMATO V, D’ ECCLESIA R, LEVANTESI S. ESG score prediction through random forest algorithm[J]. Computational Management Science, 2022, 19(2): 347–373.

[3]VELTE P. Does ESG performance have an impact on financial performance? Evidence from Germany[J]. Journal of Global Responsibility, 2017, 8(2): 169–178.

[4]SANDBERG H, ALNOOR A, TIBERIUS V. Environmental, social, and governance ratings and financial performance: evidence from the European food industry[J]. Business Strategy and the Environment, 2022, 32(4): 2471–2489.

[5]史瑜港. 融入 ESG 信息的制造业上市公司信用风险预警研究 [D]. 兰州: 西北师范大学, 2025.

[6]袁蓉丽, 江纳, 刘梦瑶. ESG 研究综述与展望 [J]. 财会月刊, 2022(17):128–134.

[7]梁四安, 李淑萍, 段军山, 等. ESG 表现如何影响企业绿色创新? —— 基于融资约束与分析师关注的中介效应 [J/OL].(2025–12–26)[2026–01–21].https://link.cnki.net/urlid/61.1294.N.20251226.1133.008.

[8]尚煊川, 蒋建洪. ESG 评级对企业高质量发展的影响研究 [J]. 商业观察, 2025, 11(34):106–111.

[9]DE LUCIA C, PAZIENZA P, BARTLETT M. Does good ESG lead to better financial performances by firms? Machine learning and logistic regression models of public enterprises in Europe [J]. Sustainability, 2020, 12(13): 5317.

[10]GARCÍA F, GONZÁLEZ-BUENO J, GUIJARRO F, et al. Forecasting the environmental, social, and governance rating of firms by using corporate financial performance variables: a rough set approach[J]. Sustainability, 2020, 12(8): 3324.

[11]D’ AMATO V, D’ ECCLESIA R, LEVANTESI S. Fundamental ratios as predictors of ESG scores: a machine learning approach[J]. Decisions in Economics and Finance, 2021, 44: 1087–1110.

[12]CINI F, FERRARI A. Towards the estimation of ESG ratings: a machine learning approach using balance sheet ratios [J]. Research in International Business and Finance, 2024, 73: 102653.

[13]Kartal M T, Taşkın D, Shahbaz M, Depren S K, Pata U K. Effects of environment, social, and governance (ESG) disclosures on ESG scores: investigating the role of corporate governance for publicly traded Turkish companies[J]. Journal of Environmental Management, 2024, 368: 122205.

[14]TASKIN D, SARIYER G, ACAR E, et al. Do past ESG scores efficiently predict future ESG performance?[J]. Research in International Business and Finance, 2025, 74: 102706.

[15]ZHANG J F, ZHAO Z X. Corporate ESG rating prediction based on XGBoost–SHAP interpretable machine learning model[J]. Expert Systems with Applications, 2026, 295: 128809.

[16]CHEN T, GUESTRIN C. XGBoost: a scalable tree boosting system[C]//Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. San Francisco: ACM, 2016: 785–794.

[17]张影. 基于代价敏感的 XGBoost 算法在 T2DM 患者并发糖尿病肾病预测中的应用 [D]. 湘潭: 湘潭大学, 2024.

[18]李爱华, 刘婉婷, 陈思帆, 等. 面向不平衡数据的 SMOTE–BO–XGBoost 集成信用评分模型研究 [J/OL].(2024–10–11)[2025–11–20].https://doi.org/10.16381/j.cnki.issn1003–207x.2023.0635.

[19]BEN YAN, ZIQI Y, YAN W. The measurement and early warning of daily financial stability index based on XGBoost and SHAP: Evidence from China[J]. Expert Systems with Applications, 2023, 227: 120245.

[20]HAND D J, TILL R J. A simple generalisation of the area under the ROC curve for multiple class classification problems[J]. Machine Learning, 2001, 45(2): 171–186.

[21]MEHDARY A, CHEHRI A, JAKIMI A, et al. Hyperparameter optimization with genetic algorithms and XGBoost: A step forward in smart grid fraud detection[J]. Sensors, 2024, 24(4): 1121.