

基于随机森林算法的 ESG 评价模型及其应用（附录 2）

外审意见

意见 1、对目标变量的定义存在概念错误

文章 4.2.1 定义“超额收益率 = 个股收益率 - wind 全 A 同期收益率”作为目标标量，形成后续建模基础。

但是，该种超额收益率计算方法存在概念性错误，原因如下：

- （1）评估投资组合的业绩，仅计算平均收益并不足够，因为没有考虑风险调整，因此，直接相减的方法计算超额收益率是不能接受的。
- （2）经风险调整的业绩测度指标通常包括夏普测度、特雷纳测度等。

意见 2、模型关键假设中，存在概念错误

文章“3.2 模型关键假设”所述的逻辑如下：

具体逻辑可通过现金流贴现（Discounted Cash Flow，简称 DCF）模型进行解释，DCF 模型由分子现金流以及分母贴现率组成，从分子端来看，ESG 水平高的公司有较强的同业竞争优势，其盈利能力和股利水平也会更高，进而估值较高；从分母端来看，由于贴现率（资本成本）与股票的风险程度相关，ESG 水平高的公司，其系统性风险和特质性风险相对较低，因而贴现率也会处于较低水平，进而估值较高。

现金流贴现模型：
$$P = \sum_{t=1}^n \frac{CF_t}{(1+r)^t}$$

前段论述存在以下问题：

- （1）认为“ESG 水平高的公司，其系统性风险和特质性风险相对较低”。
- “系统性风险”又称市场风险，是不可分散风险。是指由于某种因素的影响和变化，导致股市上所有股价的下跌，从而给股票持有人带来损失的可能性。系统风险的诱因发生在企业外部，上市公司本身无法控制它。因此，假设 ESG 水平高的公司，其系统性风险相对较低，有悖于普遍接受的理论。

意见 3、模型评分规则与结果数据之间关系存在矛盾

文章“4.3.2 ESG 评分实证分析”

图表 16：ESG 评分主要统计量

	ESG 评分	E 评分	S 评分	G 评分
--	--------	------	------	------

样本量	4,637	4,637	4,637	4,637
均值	71	21	22	29
标准差	9	6	6	5
最小值	41	4	4	13
25%分位数	65	17	18	25
50%分位数	71	21	22	29
75%分位数	77	24	26	32
最大值	100	38	41	45

按照作者在“4.3.1 打分方法”中，对分数的计算规则如下：

- 1) 对三级议题进行指标方向和评价范围划分，依据指标反映公司状况的正负向，在指标方向上分为正向、负向、中性和类别指标，在评价范围上分为全市场和行业内指标；
- 2) 针对不同方向的指标，在评价范围内进行排序，以排序值作为得分基础；
- 3) 将指标得分与对应权重相乘，作为该项指标的最终得分；
- 4) 逐级向上加总各级议题得分，并得出 ESG 总分。

同时，“4.3.2 ESG 评分实证分析”中明确“评分采用 0-100 分制”

综合以上两点，图表 16：中 ESG 评分最大值为 100，即为满分，意味着
E 评分、S 评分、G 评分都应该为满分，即，最大值。但目前 E 评分、S 评分、G 评分的最大值相加（=38+41+45=134），该值大于满分 100。存在逻辑矛盾。

作者回复

意见 1 的反馈说明

本文建模的目的是寻找 ESG 的 alpha。在机器学习监督算法的建模中，需要考虑的是样本的区分度，即在 ESG 因素驱动下个股的收益。并非为了计算个股或者组合的超额收益率。

实际建模过程中，我们分别考虑了个股年度收益率是否为正数、个股年度收益率是否超过所在申万一级行业指数收益率、个股年度收益率是否超过 wind 全 A 指数收益率等，主要考虑的是样本的区分度情况。

在实证阶段，即 5. ESG 评分返回测试中，我们采用了组合年化收益率、夏普比率、波动率和最大回撤等风险绩效指标对投资组合进行回测，回测结果较好。

意见 2 的反馈说明

个股 ESG 评分之所以受到国内外投资机构的广泛关注，是因为 ESG 因子在追求 alpha 的同时降低了风险暴露。ESG 管理水平高的公司，通常其面临的 β 风

险较低，其抵御系统性风险的能力较强。结合近年来中特估、国企改革等系列价值创造行动，我们可以看到 ESG 关键议题以健全的公司治理为基础，追求更具绿色和社会责任的高质量可持续发展，与新一轮国企改革目标相一致，ESG 管理是国企改革的重要抓手。

从另一个角度，我们结合 CAPM 模型（即资本资产定价模型）来看，模型中 β 值是衡量资产系统性风险的指标，它反映了该资产相对于市场整体的波动情况，对于每一家上市公司，其 β 值是不同的。

基于上述分析我们假设 ESG 水平高的公司对系统性风险的抵御能力更强，面临的系统性风险更低。

意见 3 的反馈说明

我们在 4.3.1 打分方法中表述了指标最终得分基于指标得分乘以权重，并且就文章中不同的行业分组其权重并不一样。

在实际计算中，我们对三级议题的 48 个指标均计算了其百分制下的得分，乘以对应的指标权重（权重之和为 1），所有个股的 ESG 得分落在 0-100 之间。

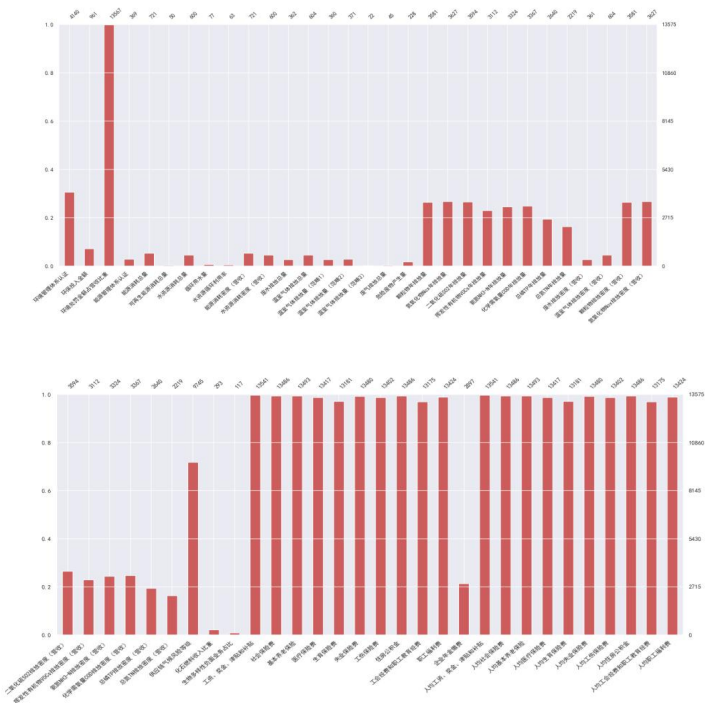
基于随机森林算法的 ESG 评价模型及其应用（附录 1）

图 2：环境排放指标各行业缺失率¹



注：上图为 5 种环境排放类指标缺失率的堆积柱形图

图 3：数据缺失分布样例



¹ 环境排放指标各行业缺失率基于建模原始数据计算

图 4：异常值分布样例

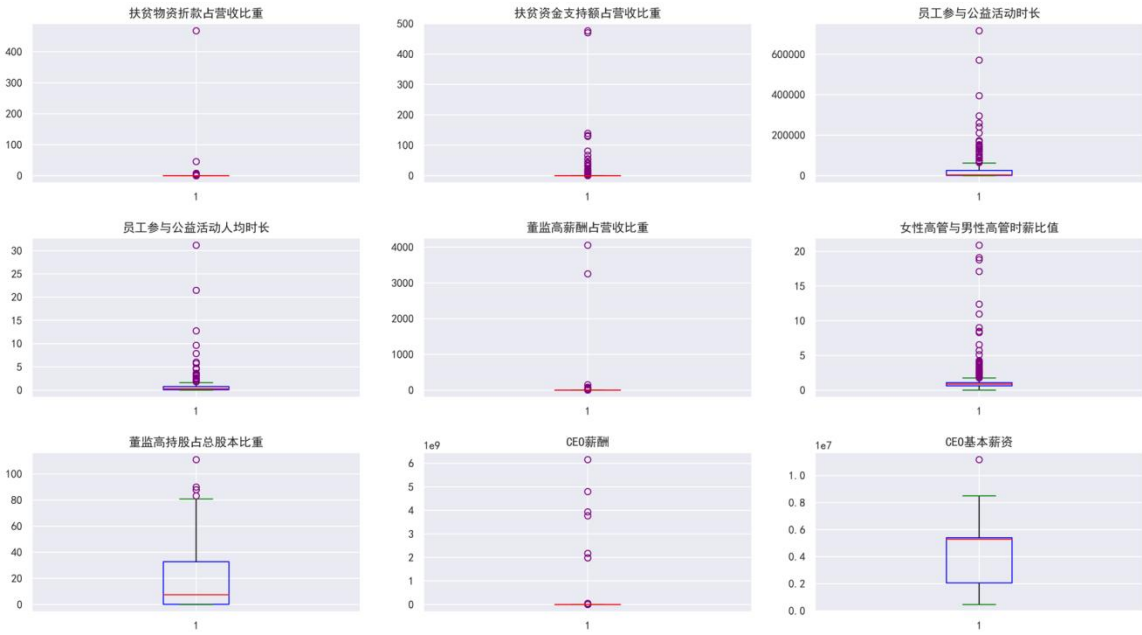
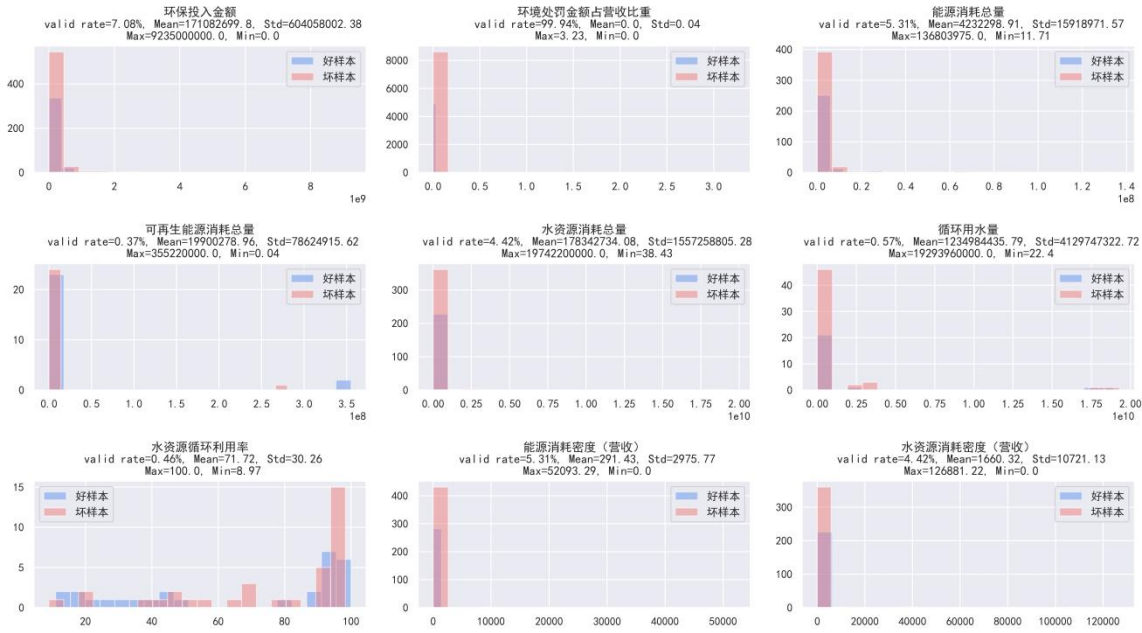


图 5：各指标类别分布样例



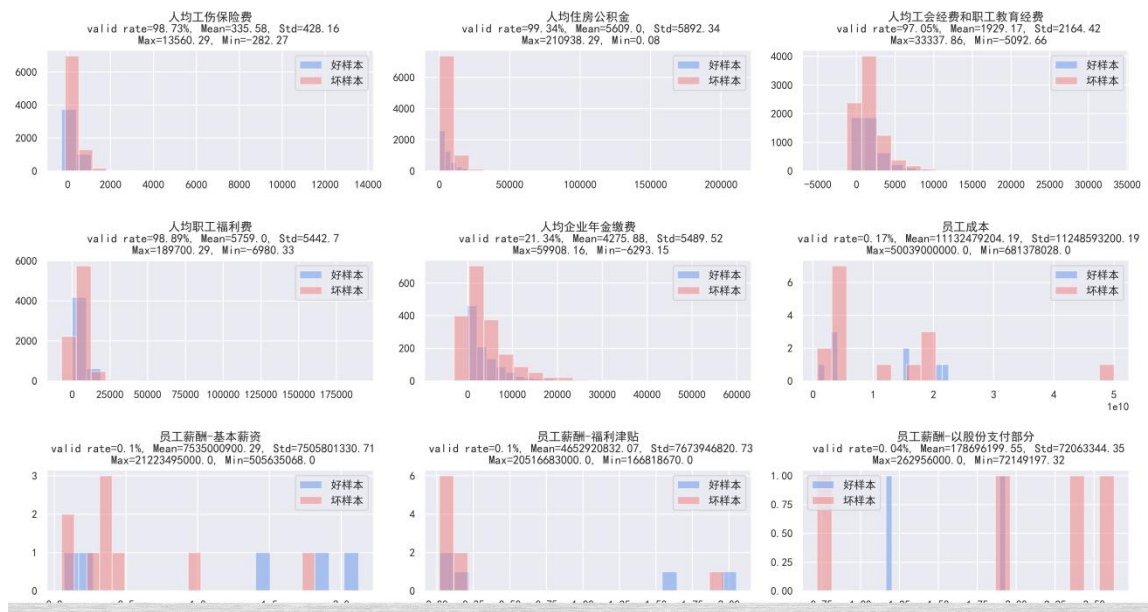


图 6：特征相关性分析样例

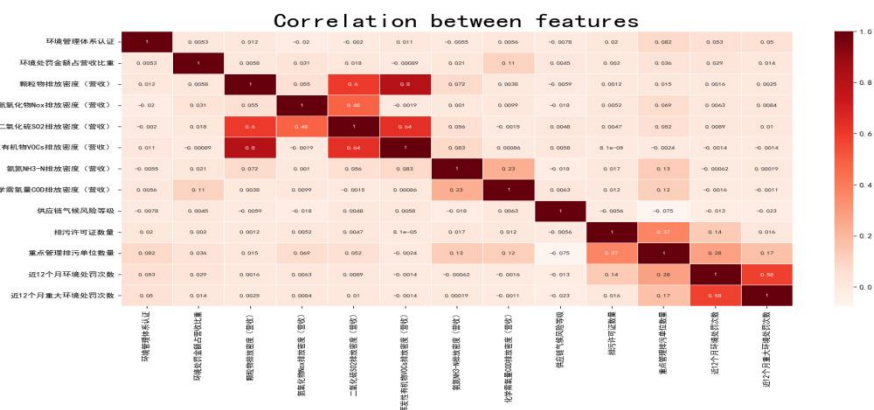


图 7: 随机森林模型示意图

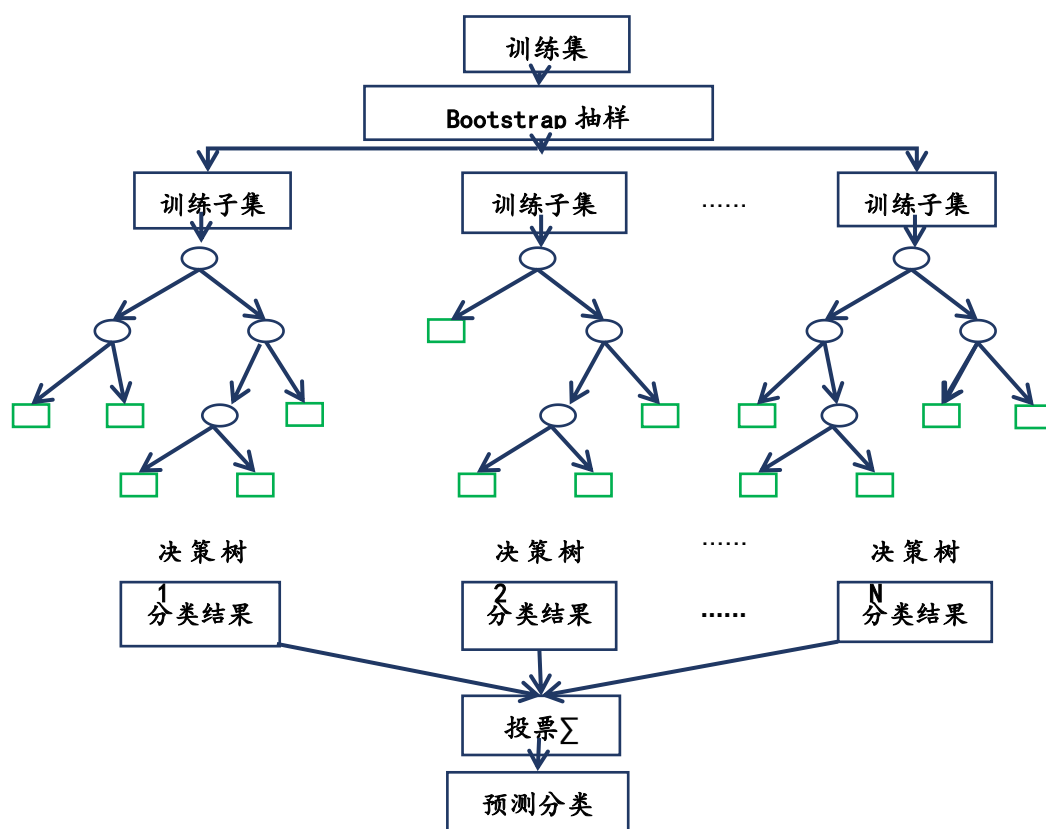


图 8: 随机森林中决策树样例

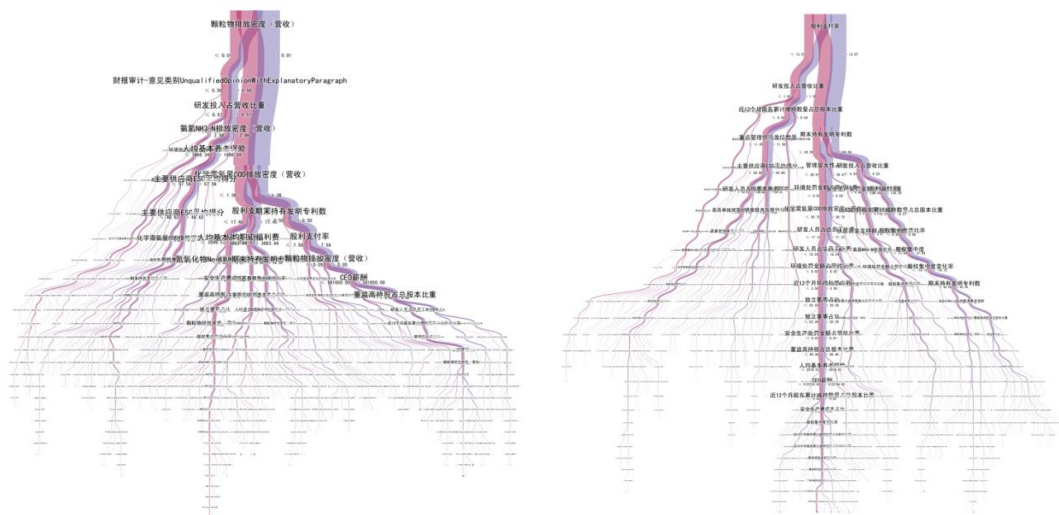


图 9：子模型 1（左图）和 2（右图）ROC 曲线

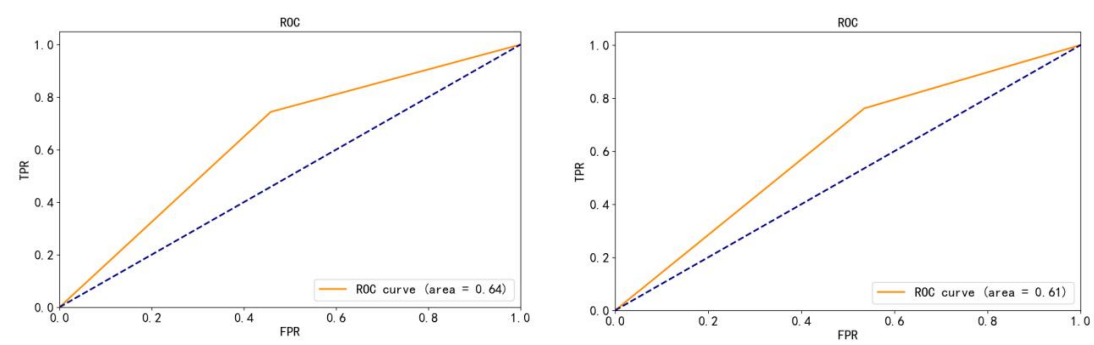


图 10：子模型 1 各算法比较

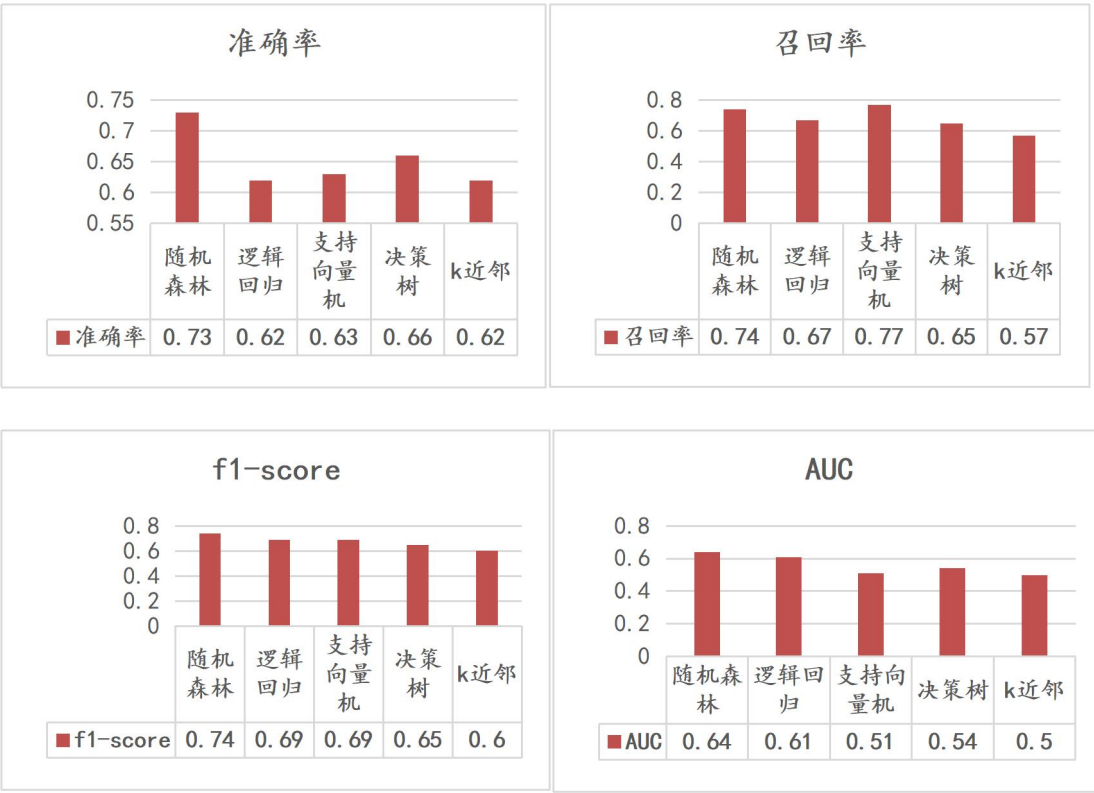
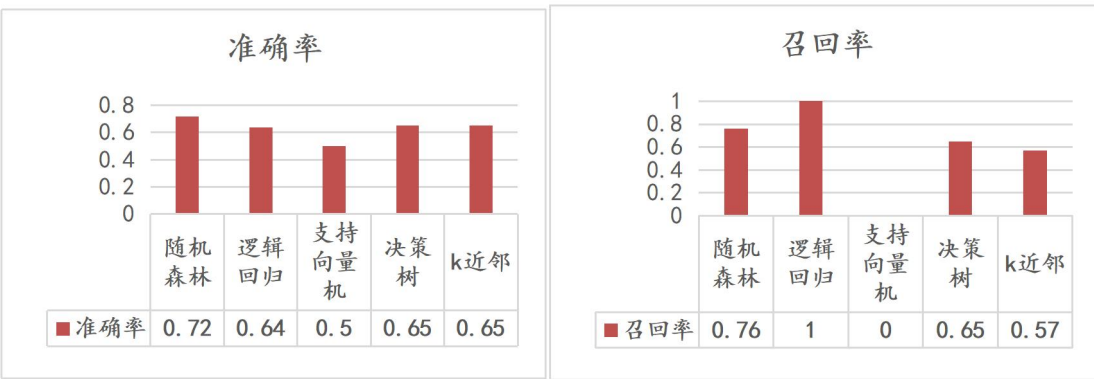


图 11：子模型 2 算法比较



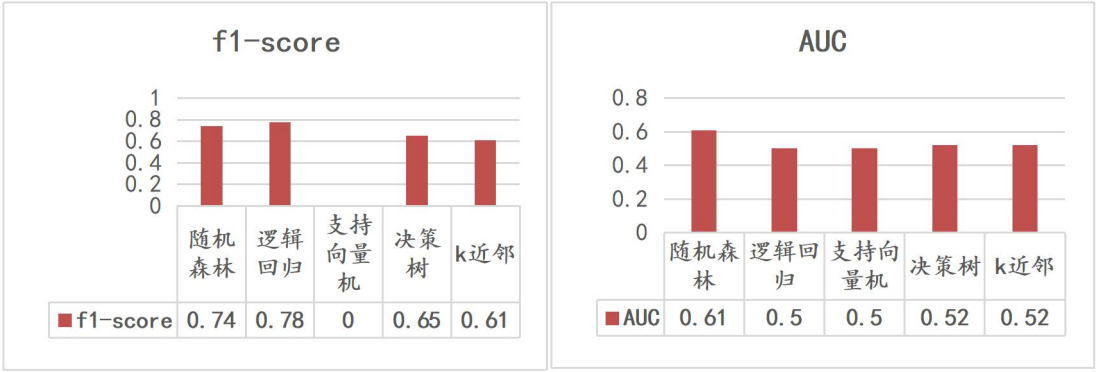


表 2: One-hot 编码样例

类别取值	One-hot 编码
（标准）无保留意见	10000
带强调事项的无保留意见	01000
保留意见	00100
无法表示意见	00010
否定意见	00001

表 4: 子模型 2 算法比较

算法	准确率	召回率	f1-score	AUC
随机森林	0.72	0.76	0.74	0.61
逻辑回归	0.64	1.00	0.78	0.50
支持向量机	0.50	0.00	0.00	0.50
决策树	0.65	0.65	0.65	0.52
k 近邻	0.65	0.57	0.61	0.52

表 6: ESG 评分主要统计量

	ESG 评分	E 评分	S 评分	G 评分
样本量	4,637	4,637	4,637	4,637
均值	71	21	22	29
标准差	9	6	6	5
最小值	41	4	4	13
25%分位数	65	17	18	25
50%分位数	71	21	22	29
75%分位数	77	24	26	32
最大值	100	38	41	45

