DOI:10.14188/j.1671-8844.2025.0157 文章编号:1671-8844(XXXX)XX-0001-14

# 融合多维特征的受贿风险预测模型及分析

朱金媛,刘丹

(中国刑事警察学院公安信息技术与情报学院,辽宁 沈阳 110035)

**摘要:**针对当前受贿案件多发和犯罪手段较为隐蔽,以及案发风险和受贿诱因难以量化分析的问题。提出基于中国裁判文书网公示的已判决生效文书提取特征变量,首次将随机森林算法融合夏普利加性解释(SHapley Additive exPlanations,SHAP)应用于受贿案件的预测与分析研究领域。在宏观层面,模型能够输出受贿风险的预测概率,其中,基于递归特征消除(Recursive Feature Elimination,RFE)进行特征选择的模型在Accuracy、Macro $F_1$和Macro AUC指标上的结果分别为:0.8929、0.8792和0.9662;在微观层面,利用SHAP定量探讨具体因素对风险分类的贡献程度。将为各级纪委监委精准识别受贿风险、高效开展监督执纪和科学制定反腐策略提供强有力的技术支撑与科学依据。

**关键词:**随机森林;受贿案;SHAP;风险分析

**中图分类号:**TP391 **文献标志码:**A

# Explainable bribery risk prediction model based on multidimensional feature integration

ZHU Jinyuan, LIU Dan

(School of Public Security Information Technology and Intelligence, Criminal Investigation Police University of China, Shenyang 110035, China)

**Abstract**: Regarding the current issues of frequent bribery cases with relatively concealed means, as well as the difficulties in quantitative analysis of case occurrence risks and bribery inducements. To effectively advance anti-corruption initiatives and achieve early warning of bribery cases, this study constructs predictive features based on legally effective judgment documents disclosed on China Judgments Online. For the first time in this research domain, we integrate the random forest algorithm with SHAP for bribery case prediction and analysis. At the macro level, the model generates probabilistic risk predictions, with the RFE-optimized version achieving performance metrics of 0.8929 (Accuracy), 0.8792 (Macro $F_1$), and 0.9662 (Macro AUC). At the micro level, SHAP analysis quantitatively delineates feature-specific contributions to risk classification. This dual-level analytical framework will provide disciplinary inspection and supervision commissions with robust technical support and scientific evidence for precisely identifying bribery risks, enhancing supervisory efficiency, and formulating data-driven anti-corruption strategies.

**Key words**: random forest; bribery case; SHAP; risk analysis

**收稿日期:**XXXX-XX-XX

**作者简介:**朱金媛(1998-),女,硕士研究生,主要从事计算机应用技术研究,E-mail: zhujinyuan_cipuc@qq.com

**通信作者:**刘丹(1969-),女,教授,主要从事计算机应用技术、公安图像技术和数据挖掘等研究,E-mail:liudan_sxzljyx@cipuc.edu.cn

**本文引用格式:**朱金媛,刘丹.融合多维特征的受贿风险预测模型及分析[J].武汉大学学报(工学版),XXXX,XX(XX):1-14.

**Citation:**ZHU Jinyuan, LIU Dan. Explainable bribery risk prediction model based on multidimensional feature integration[J]. Engineering Journal of Wuhan University, XXXX, XX(XX):1-14.

# 引言

腐败如同毒瘤,侵蚀着国家肌体,损害着社会公平正义,阻碍着经济发展。根据十四届全国人大二次会议上的最高人民法院工作报告,2023年审结贪污贿赂等职务犯罪案件2.4万件2.7万人,同比增长19.9%。这一数据不仅揭示了腐败问题的严重性,也反映出我国反腐败斗争的持续高压态势。在这些职务犯罪案件中,受贿案件是腐败的典型表现形式之一,涉案人员的身份使得此类案件的社会危害性尤为突出。

犯罪预测是预防和打击犯罪的重要手段,随着大数据和人工智能技术的发展,预测模型在公安工作中的应用日益广泛。围绕盗窃犯罪预测研究,颜靖华等[1]基于长短期记忆网络(Long Short-Term Memory,LSTM)开展日盗窃犯罪数量时间序列预测;张齐等[2]基于词频-逆文档频率(Term Frequency-Inverse Document Frequency,TF-IDF)特征提取与极端梯度提升(eXtreme Gradient Boosting,XGBoost)算法,实现了盗窃犯罪类型的多标签识别;王娟等[3]采用CNN - LSTM - Attention模型,结合时空与环境特征,对五个城市的盗窃犯罪案发数量进行预测,显著提升了预测精度,验证了模型有效性及多特征融合优势。针对人员特征的研究,张强[4]与李荣岗[5]聚焦支持向量机(Support Vector Machine,SVM)的应用优化,前者通过SVM构建治安高危人员预警模型,结合静态与动态数据实现风险评估,后者引入粒子群优化(Particle Swarm Optimization,PSO)算法优化参数,结果显示模型预测的特征与真实特征的平均相似度达到了78.8%;王佳[6]提出了基于贝叶斯网络的犯罪人特征刻画模型和基于模糊多准则决策方法的目标关联模型,分别应用于确定犯罪人特征和嫌疑人的排查分析。

SHAP框架是一种可解释性工具,能够揭开黑箱模型的决策逻辑,已经在多个领域应用。在金融风控领域,祝起祯等[7]利用SHAP分析居民投资风险偏好的影响因素,为金融产品的风险匹配和投资者适当性管理提供了科学依据。在医疗诊断领域[8-10],机器学习融合SHAP分析,使得各个特征对模型所预测目标的影响程度和交互作用关系可视化,提升了模型在医疗决策中的可信度。在生态环境领域,刘瑞雪等[11]结合SHAP分析确定消落带植物群落物种丰富度、物种多样性和谱系多样性的关键环境因素;王书苏等[12]通过SHAP量化$PM_{10}$、排放源、温度等特征因子对兰州市主城区$PM_{2.5}$浓度的影响程度。

在腐败治理领域,研究人员[13-15]提出了技术赋能腐败治理的机理、路径和挑战等,但是大多停留在理论层面的机理阐释与技术功能的宏观设想,目前缺乏具体算法模型的构建。本文基于中国裁判文书网公示的已判决生效的文书,创新性地提取受贿人特征、受贿行为特征及行贿人特征等多维数据。基于这些特征,采用随机森林算法构建了受贿风险预测模型,并对比三种特征选择方法,以优化模型性能和提升预测准确度。在宏观层次上,通过对各类变量的系统挖掘,模型的输出变量为受贿风险预测概率量化值。在微观层次上,进一步利用模型的特征重要性排名及SHAP可解释性分析,定量探讨了具体因素对风险分类的贡献程度,从而揭示关键变量的作用机制。

# 1 算法原理

## 1.1 随机森林算法

随机森林算法是由Leo Breiman[16]等人于2001年提出的一种集成学习算法,其分类过程的原理如图1所示。算法的核心思想是通过构建多棵决策树形成“森林”,并采用集成学习机制进行协同预测:对于分类任务采用投票法,回归任务则采用平均法。具体而言,算法实现包含双重随机化过程:首先,通过Bootstrap抽样对原始训练集进行有放回采样,生成多个差异化子训练集;其次,在每棵决策树的节点分裂时,从全部特征中随机选取特征子集,基于基尼系数或信息增益准则选择最优分裂方式。这种双重随机机制有效保证了基学习器间的多样性。

相较于单一决策树,随机森林算法展现出以下的显著优势:(1)预测准确性方面,通过集成多棵低相关性的决策树预测结果,显著降低模型方差,提升泛化性能;(2)特征处理能力方面,可自适应处理高维数据,无需特征缩放等预处理,对连续变量与离散变量具有天然兼容性;(3)鲁棒性方面,Bagging策略与随机子空间方法共同作用,有效抑制过拟合现象,同时对缺失值和异常值保持较强容错能力。基于这些优势,随机森林算法在医学、经济学、计算机视觉、工业制造和自然语言处理等众多领域广泛应用,为各领域数据分析与决策提供有力技术支撑[17-18]。

本文中的模型使用的受贿案件数据涵盖受贿人、受贿行为和行贿人特征,数据维度高且类型不同。随机森林算法不仅能够有效处理高维和多类

型数据,还能够通过随机抽样和随机子特征选择降低过拟合风险,提升模型的泛化能力。此外,该算法的集成机制使得每棵决策树均能够捕捉不同的局部特征模式,通过对大量决策树结果进行投票,生成稳定而准确的预测。同时,随机森林算法能够输出特征重要性排序,为后续利用SHAP进行模型解释性分析提供基础,从而揭示影响受贿风险分类的关键因素。

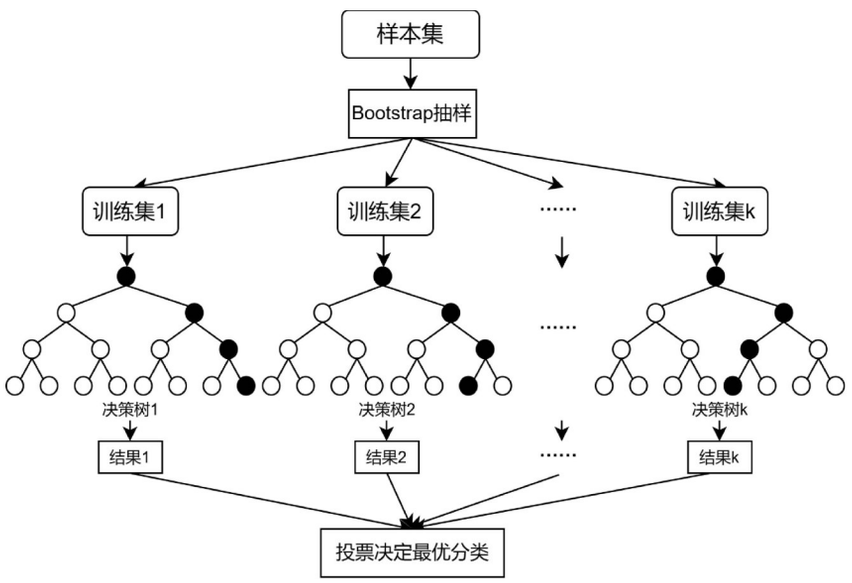



**图1 随机森林分类算法流程图**

**Fig.1 Flow chart of the random forest classification algorithm**

### 1.2 SHAP分析

SHAP是一种基于博弈论中Shapley值理论[19]的解释性方法,能够为机器学习模型提供一致性和局部准确性的解释[20]。其核心思想是将模型预测结果视为合作博弈中玩家共同带来的收益,并通过计算每个玩家(即特征)对最终预测结果的边际贡献,量化每个特征的重要性。

Shapley值的计算公式为:

$$\varnothing_i=\sum_{S\times(F\{i\})}\frac{|S|!(|F|-|S|-1)!}{|F|!}\times\left[f_{S\cup\{i\}}\left(x_{S\cup\{i\}}\right)-f_s\left(x_s\right)\right]\#(1)$$

式中:$\varnothing_i$为特征$i$的Shapley值;$|F|$为特征总数;$F\{i\}$为除特征$i$外所有可能的特征组合的集合;S为$F\{i\}$的特征集合;$f_S(x_S)$为S中特征的模型预测;$f_{S\cup\{i\}}(x_{S\cup\{i\}})$为S中特征加特征$i$的模型预测。

SHAP分为多种类型,包括适用于树模型的Tree SHAP,与模型无关的Kernel SHAP以及用于深度学习的Deep SHAP等。本文引入SHAP具有重要意义,一方面,它可以从全局层面识别影响受贿风险分类的关键特征;另一方面,可利用瀑布图在局部层面剖析各个特征如何共同驱动模型输出。通过SHAP分析的可视化结果,相关职能部门可明确识别高风险人群与行为特征,从而有针对性地开展风险预警和资源配置,实现精准打击与系统治理。

## 2 模型构建

本文基于已判决生效的受贿案件,运用随机森林建模方法,从受贿人、受贿行为及行贿人特征中挖掘隐藏的且有价值的信息,模型构建的流程如图2所示。本节将依次介绍构建数据集、变量定义、样本分析、数据预处理及特征选择过程。

### 2.1 数据集

本文共自行收集了140例案件,均来源于中国裁判文书网公示的2023-2024年度受贿罪生效判决书。判决书包含了受贿人信息、行贿人信息、犯罪事实经过、证据列举与法律适用等,完全满足本研究的实证分析需求。基于算法设计对变量控制的要求,所选择的案件均为单独受贿案件[21]。

### 2.2 变量

案件的变量旨在刻画案件全貌,也为构建风险预测模型和风险因素分析奠定坚实的数据基础。变量的选取遵循科学、客观和实用的标准:首先,所选变量必须为直观且不易被误解的指标,确保信息来源的客观性;其次,变量应真实反映行为特征,如受贿次数以及受贿款物等,从而有效描述受贿行为

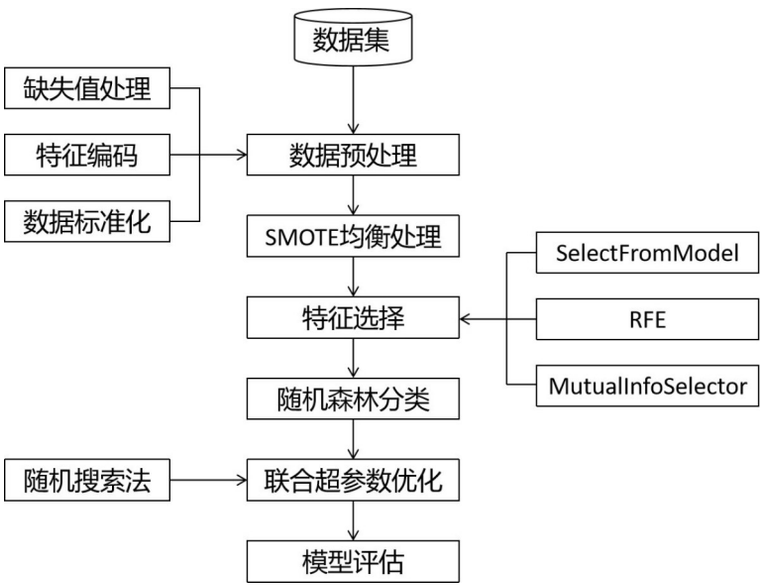



**图2 受贿风险预测模型流程图**

**Fig.2 Flow chart of the bribery risk prediction model**

的实际情况;最后,变量需能够揭示受贿人与行贿人之间的互动关系,为后续模型的解释和决策提供全面依据。基于对50起已判决生效的受贿案件中的相关内容进行深入挖掘,共提取出17个受贿人变量,47个受贿行为变量以及52个行贿人变量。

对受贿人特征的系统研究有助于定位腐败高发人群,本研究共收集17个受贿人特征,包括性别、年龄、教育情况、政治面貌、工作履历、退休状况、所在单位类型、职务级别等。

对受贿行为特征的多维度解析可为精准打击、资金追踪、制度漏洞修复提供依据,助力构建全链条反腐体系[22-23]。本研究共收集47个有关受贿行为的特征,包括受贿数额、受贿次数、受贿款物的类别及去向、受贿时间地点等。

对行贿人特征的系统研究有助于识别行贿主要群体及剖析其行为模式。本研究共收集52个行贿人相关特征,包括行贿人数、单个行贿人行贿次数、与受贿人结识方式、行贿人职业、行贿人所涉及企业的性质和行贿时经营年限、行贿名义、行贿人涉及行业等。

### 2.3 样本分析

样本分析旨在探究数据特性、研判数据质量,审视样本及其特征呈现出的概率分布信息,进而为构建预测模型提供有效指引。

在本研究的140例样本中,受贿人特征方面,从性别来看,男性受贿人占比93.6%,女性受贿人占比6.4%;从案发时年龄来看,40周岁以下占比11.4%,40周岁及以上至50周岁占比32.9%,50周岁及以上至60周岁占比45.7%,60周岁以上占比10.0%;从受教育情况来看,专科及以下占比28.6%,本科占比61.4%,研究生及以上占比10.0%。

受贿行为特征方面,从收受款物来看,93.6%的案件中存在现金收受行为,41.4%的案件中存在银行转账行为,8.5%的案件中存在房产收受行为,6.4%的案件中存在车辆收受行为;从收受方式来看,22.1%的受贿人通过其亲属收受,20.7%的受贿人以他人银行账户收受;从受贿地点来看,50.7%的案件中地点有办公场所,38.5%的案件中地点有私人住宅,27.1%的案件中地点有车内。

行贿人特征方面,从性别来看,20.0%案件中存在女性行贿人;从行贿人职业来看,6.4%的案件中存在政府人员,35.0%的案件中存在工程承包商,83.5%的案件中存在私营企业主;从行贿名义来看,41.1%的案件存在行贿人出于节日缘由拜访受贿人,36.0%案件中存在行贿人以借款名义行贿,7.8%的案件存在行为人以分红、利润名义行贿。

### 2.4 数据预处理

在依据《最高人民法院、最高人民检察院关于办理贪污贿赂刑事案件适用法律若干问题的解释》确立的数额分层基础上,结合实证样本统计分析以及强化高风险预警灵敏度的需求,将受贿行为特征中的受贿数额划分成20万元以下、20万元(含)至150万元(不含)、150万元及以上三个区间,依次代表低、中、高三个风险等级并分别编码为0、1、2,作为目标变量。

除去“受贿数额”这一目标变量后,其余的115

个特征则作为输入变量。输入变量中,按照数据类型划分,有11个数值类变量,6个多分类变量,98个二分类变量,部分特征变量的名称及类型如表1所示。不同类型的变量进行了差异化处理:针对数值型变量,使用中位数填充缺失值并进行标准化;针对二分类变量,进行二值化处理,1代表该特征在样本中出现,0代表未出现;针对多分类特征,使用众数填充缺失值并进行独热编码。

### 2.5 特征选择

机器学习领域中特征选择主要有过滤器(filter)[24]、封装法(wrapper)[25]和嵌入法(embedded)[26]三种模式,模型的性能表现很大程度上依赖于特征,恰当的特征选择不仅能够减少冗余信息和噪声,还能提高模型的泛化能力和训练效率。本研究共140个样本,每个样本包含116个特征,样本较少而变量种类多、特征维度相对高,易产生冗余信息和噪声,导致模型训练过程中出现过拟合现象,因此,筛选出对预测最有贡献的特征尤为重要,同时也便于后续在案件微观层次上的分析。

在特征选择阶段,选取了三种不同的特征选择方法:SelectFromModel方法、递归特征消除(RFE)方法[27]以及互信息选择器(MutualInfoSelector)[28],分别属于嵌入法、封装法和过滤器三种模式。首先,SelectFromModel方法基于内置模型计算各个特征的重要性评分,再依据预设阈值自动筛选出贡献较大的变量。此方法操作简单、计算高效,能够迅速剔除冗余特征,适合作为初步特征筛选手段;其次,递归特征消除法通过不断构建模型并迭代删除对预测贡献较小的特征,利用交叉验证寻找最优特征组合。该方法不仅能充分考虑特征之间的交互效应,还能逐步精炼出一组最具代表性的特征,适合在复杂数据情境下进行深入挖掘,但是,计算过程相对耗时;最后,互信息选择器依据信息论中的互信息概念,评估每个特征与目标变量之间的相关性,既能捕捉线性关系,也能揭示非线性依赖。该方法在高维数据中表现出色,能够发现传统方法容易忽略的统计依赖关系。

**表1 部分特征变量名称及类型**

**Table 1 Names and types of some feature variables**

| 编号 | 变量名称 | 变量类型 |
|---|---|---|
| 1 | 受贿人_性别 | 二分类 |
| 2 | 受贿人_案发时年龄分类 | 多分类 |
| 3 | 受贿人_教育情况 | 多分类 |
| 4 | 受贿人_政治面貌为党员 | 二分类 |
| 5 | 受贿人_案发时已退休 | 二分类 |
| 6 | 受贿人_所在单位类型 | 多分类 |
| 7 | 受贿人_同一岗位任职最长年限 | 数值类 |
| 8 | 受贿人_曾在出生地任职 | 二分类 |
| 9 | 受贿人_开始受贿时职务级别 | 多分类 |
| ⋮ | ⋮ | ⋮ |
| 18 | 受贿行为_首次受贿数目(万元) | 数值类 |
| 19 | 受贿行为_受贿数额(万元) | 数值类 |
| 20 | 受贿行为_受贿次数 | 数值类 |
| 21 | 受贿行为_单次收受大于10万 | 二分类 |
| 22 | 受贿行为_受贿款物:现金收受 | 二分类 |
| 23 | 受贿行为_受贿款物:银行转账 | 二分类 |
| 24 | 受贿行为_受贿款物:房产收受 | 二分类 |
| 25 | 受贿行为_受贿款物:车辆收受 | 二分类 |
| 26 | 受贿行为_以他人银行账户收受 | 二分类 |
| 27 | 受贿行为_受贿地点:私人住宅 | 二分类 |
| 28 | 受贿行为_受贿地点:酒店/宾馆 | 二分类 |
| 29 | 受贿行为_受贿地点:车内 | 二分类 |
| 30 | 受贿行为_受贿时间:有在节日 | 二分类 |
| 31 | 受贿行为_款物去向:用于购置房产 | 二分类 |
| 32 | 受贿行为_款物去向:用于购车 | 二分类 |
| 33 | 受贿行为_受贿过程有中间人参与 | 二分类 |
| 34 | 受贿行为_存在主动索贿 | 二分类 |
| 35 | 受贿行为_中途存在向行贿人退还行为 | 二分类 |
| ⋮ | ⋮ | ⋮ |
| 65 | 行贿人_行贿人数 | 数值类 |
| 66 | 行贿人_有行贿人单次行贿 | 二分类 |
| 67 | 行贿人_有行贿人多次行贿 | 二分类 |
| 68 | 行贿人_与受贿人结识方式:工作关系 | 二分类 |
| 69 | 行贿人_职业:私营企业主 | 二分类 |
| 70 | 行贿人_当面贿赂 | 二分类 |
| 71 | 行贿人_将贿款、物放置于特定地点 | 二分类 |
| 72 | 行贿人_名义:节日拜访 | 二分类 |
| 73 | 行贿人_名义:借款 | 二分类 |
| 74 | 行贿人_行贿时间跨度:5年内 | 二分类 |
| 75 | 行贿人_行业:建筑/房地产 | 二分类 |
| ⋮ | ⋮ | ⋮ |

## 3 算法测试与评估

### 3.1 算法测试

本研究基于VS Code开发环境、Python 3.11版本编写程序,调用scikit-learn、imbalanced-learn及其它常用数据处理库开展实验。首先,对“受贿数额”这一目标变量进行分箱离散化处理,划分为前述的低、中、高三个类别。采用分层抽样策略[29]将数据按8:2划分为训练集与测试集,确保类别分布一致性。在特征预处理阶段,对不同类型变量分别进行处理,应用SMOTE过采样技术均衡样本。为提高模型性能,选择三种特征选择方法(SelectFromModel、RFE和MutualInfoSelector)进行特征选择,

并将其嵌入包含预处理和随机森林分类器的管道中。采用5折分层交叉验证,将训练集划分为5个子集,每次用4个子集训练,1个子集验证,每折计算宏平均$F_1$分数(Macro $F_1$),最终分数为5折平均值,以Macro $F_1$分数为优化目标,对随机森林的树数量、最大深度、叶节点最小样本数等超参数进行30组随机搜索调优,保存最优模型。

### 3.2 模型性能评估

模型在随机搜索阶段采用5折交叉验证,用Macro $F_1$分数评估各超参数组合的表现。在此过程中,对于每种特征选择方法(SelectFromModel、RFE、MutualInfo),均计算其交叉验证的Macro $F_1$分数,从而比较不同特征选择方案的效果,选择交叉验证得分最高的方法作为最佳方案。选定最佳特征选择方法后,利用最终模型在测试集上进行预测,并使用准确率(Accuracy)、精确率(Precision)、召回率(Recall)、$F_1$分数($F_1$-Score)、宏平均$F_1$分数(Macro $F_1$)、宏平均AUC(Macro AUC)等指标进行度量,从而确保模型在平衡性和泛化能力上的整体优异效果。表2为三种特征选择方法下模型的相关指标结果。

本研究中存在类别不平衡问题,高、低风险类样本占比较少,而中风险类样本占比大于50%,Macro $F_1$分数同时考虑精确率和召回率,避免模型仅对中风险类有较高预测准确率而牺牲对高、低风险类的预测质量,提升模型整体泛化性能,因此在超参数调优过程中以其为优化目标。

对于三分类问题,Macro $F_1$分数的计算过程[30]如下:

(1)对每个类别($i$)分别计算其$F_1$分数:

$$F_{1\,\mathrm{class}_i}=2\times\frac{\mathrm{Precision}_i\times\mathrm{Recall}_i}{\mathrm{Precision}_i+\mathrm{Recall}_i}\#(2)$$

$$\mathrm{Precision}_i=\frac{\mathrm{TP}_i}{\mathrm{TP}_i+\mathrm{FP}_i}\#(3)$$

$$\mathrm{Recall}_i=\frac{\mathrm{TP}_i}{\mathrm{TP}_i+FN_i}\#(4)$$

(2)计算Macro $F_1$分数:

$$\mathrm{Macro}\,F_1=\frac{F_{1\,\mathrm{class}_0}+F_{1\,\mathrm{class}_1}+F_{1\,\mathrm{class}_2}}{3}\#(5)$$

式中:$F_{1\,\mathrm{class}_i}$为类别$i$的F1分数;$\mathrm{Precision}_i$为类别$i$的精确率;$\mathrm{Recall}_i$为类别$i$的召回率;$\mathrm{TP}_i$(True Positive)为正确预测为第$i$类的样本数量;$\mathrm{FP}_i$(False Positive)为错误预测为第$i$类的样本数量;$\mathrm{FN}_i$(False Negative)为错误预测为其它类别的样本数量。

**表2 三种特征选择方法作用下模型性能比较**

**Table 2 Comparison of model performance under three feature selection methods**

| 特征选择方法 | 交叉验证分数 | 选择特征数量 | Accuracy | Macro $F_1$ | Macro AUC |
|---|---|---|---|---|---|
| SelectFromModel | 0.7544 | 33 | 0.8571 | 0.8658 | 0.9531 |
| RFE | 0.7718 | 15 | 0.8929 | 0.8792 | 0.9662 |
| MutualInfo | 0.7687 | 15 | 0.7857 | 0.8002 | 0.9281 |

对比表2的实验结果可以看出,三种特征选择方法下的模型性能表现较为接近。RFE方法在交叉验证分数上比另外两种方法高0.4%-2.3%;在选择特征数量上,RFE方法与MutualInfo方法均为15个,便于特征重要性的分析;在Accuracy、Macro $F_1$和Macro AUC指标上,RFE方法的表现均强于SelectFromModel方法和MutualInfo方法。

综上所述,本文选择RFE方法下的随机森林模型进行后续研究,表3为该模型最终选择的参数。表4为模型性能指标结果,各个风险等级的预测结果均较为稳定,特别是高风险样本能得到有效识别。图3为概率分布图,直观地显示了模型对每个测试集样本在低、中、高风险分类下的预测概率,反应了样本在风险等级间的区分程度。75%的样本有一个突出的概率柱(某类别概率>0.6),表明模型区分风险等级的置信度较高,而柱高接近的则有助于识别预测结果接近边界的样本。如图4所示,基于RFE方法作用下随机森林内置特征重要性评估,前三个特征变量为“行贿人_几人单次行贿”、“行贿人_几人多次行贿”和“受贿人_同一岗位任职最短年限”,表明行贿人的规模和受贿人的职务履历在预测受贿风险时起关键作用。

### 3.3 基于SHAP的解释性分析

本节基于SHAP可解释性技术,系统解析了受贿风险预测模型中各个特征变量的作用机制。主要包括基于训练集样本的SHAP特征重要性排序图和蜂群图,用于分析特征变量对风险预测的贡献强度和正负方向。同时,针对测试集中具体样本绘制SHAP瀑布图,将预测风险类别的过程可视化。

图5、6和图7分别为低、中、高风险类别的SHAP特征重要性排序图,通过计算平均绝对SHAP值,按照其大小对特征进行排序,反映各个特征对预测结果的贡献强度,不考虑正负方向的影响。由图可知,不同风险类别下,各个特征对模型的贡献强度有所不同;此外,中风险类别的平均绝对SHAP值低于其它两类,表明其决策时依赖更分散的特征组合,而非少数主导因素。具体而言,低

**表 3 参数及数值**

**Table 3 Parameters and values**

| 参数名称 | 参数含义 | 参数值 |
|---|---|---|
| feature_selection__step | 每次迭代淘汰的特征数量 | 2 |
| feature_selection__n_features_to_select | 最终保留的特征数量 | 15 |
| classifier__n_estimators | 决策树的数量 | 100 |
| classifier__min_samples_split | 节点分裂所需最小样本数 | 5 |
| classifier__min_samples_leaf | 叶子节点最小样本数 | 3 |
| classifier__max_features | 寻找最佳分割时的特征子集大小 | 'log2' |
| classifier__max_depth | 树的最大深度 | 6 |

**表 4 RFE 特征选择方法作用下模型性能**

**Table 4 The performance of the model under the RFE**

| 指标<br>类别 | Precision | Recall | $F_1$−Score | Support |
|---|---|---|---|---|
| 0 | 0.80 | 0.80 | 0.80 | 5 |
| 1 | 0.93 | 0.87 | 0.90 | 15 |
| 2 | 0.89 | 1.00 | 0.94 | 8 |

风险案件中,“受贿人_同一岗位任职最短年限”贡献强度最大,其次是“受贿人_首次受贿时年龄”,说明受贿人的岗位任职稳定性以及首次受贿年龄等因素对低风险案件影响显著;中风险案件中,“受贿人_首次受贿时年龄”与“受贿人_同一岗位任职最短年限”仍居前列,“受贿人_政治面貌为党员”的重要性凸显,表明除个人任职特征外,政治身份在中风险评估中权重增加;高风险案件中,“行贿人_几人多次行贿”和“行贿人_几人单次行贿”等代表行贿规模和持续性的特征成为高风险案件的关键驱动特征。

图 8、9 和图 10 分别为低、中、高风险类别的 SHAP 蜂群图,它们进一步展示了不同风险等级中各个特征变量影响的方向性(SHAP 值为正表示增加当前类别预测概率,为负则降低当前类别预测概率),揭示了风险影响因素的动态差异。低风险类别的蜂群图中,大多数特征呈现“左红右蓝”的分布,例如,“受贿人_同一岗位任职最短年限”和“受贿人_首次受贿时年龄”等特征处于低值区间(蓝色)时,会提高低风险预测概率;中风险类别的蜂群图中,总体上呈现红蓝均匀分布,反映了该类别作为风险过渡状态的特点,模型决策依赖受贿人职业背景及行贿规模等多因素的组合;高风险类别的蜂群图中,普遍特征呈现“左蓝右红”的分布,“行贿人_几人多次行贿”和“行贿人_几人单次行贿”等这些与行贿人次和频次相关的特征处于高值区间时,会显著提升高风险预测概率。

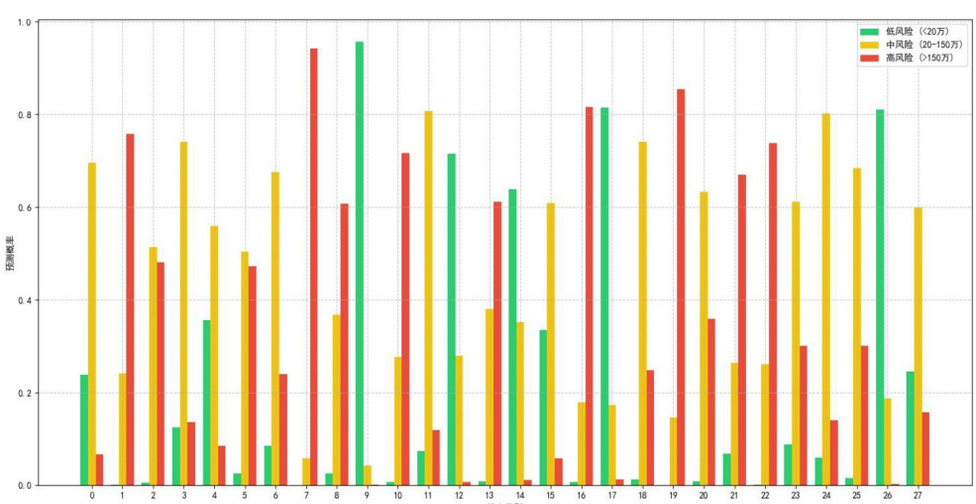



**图 3 测试集样本三分类预测概率分布**

**Fig.3 Probability distribution of three-class prediction for samples in the test set**

图 11、12 和图 13 为 SHAP 瀑布图,分别展示测试集中具体样本(图中为第 17 号样本)的各个特征对预测低、中、高风险概率的贡献路径,从基础值(全体样本平均预测概率)开始,通过叠加各个特征

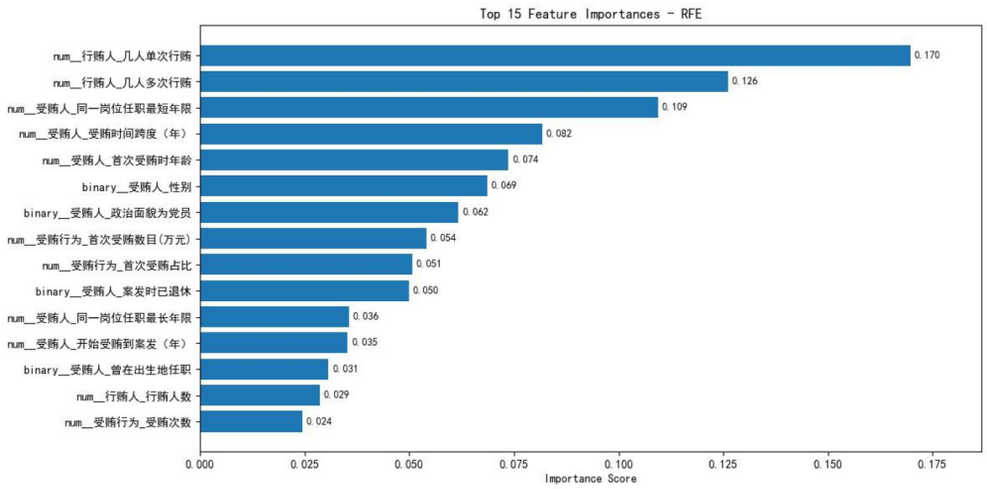



**图4　RFE方法作用下模型内置特征重要性排序**

**Fig.4　The ranking of built-in feature importance of the model under the RFE**

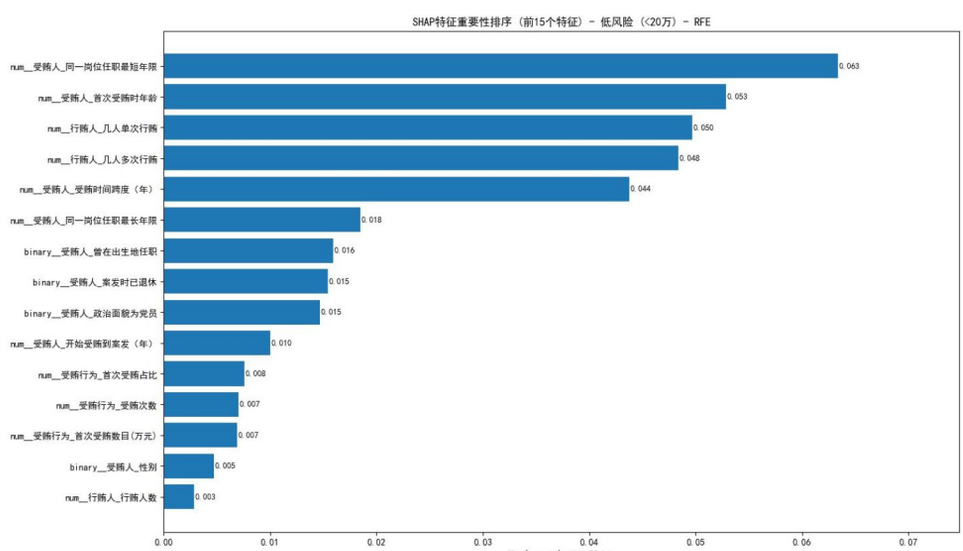



**图5　低风险类别SHAP特征重要性排序图**

**Fig.5　Ranking chart of SHAP feature importance for the low-risk category**

SHAP值形成该样本的预测概率值。

综上所述,受贿风险驱动因素呈现差异性:随着风险的增加,关键驱动因素从受贿人职业背景逐步转向行贿人规模与受贿行为的持续性。据此,提出以下反腐败建议:第一,重点监测年轻公职人员的行为轨迹,建立动态预警机制,合理优化岗位任职周期,降低因长期固定岗位可能产生的风险累积;第二,强化对公职人员的常态化监督与分层分类廉政教育,尤其针对党员干部群体;第三,构建公职人员社会关系与资金流动监测网络,对异常的多人高频次交往、非必要资金聚集等潜在风险信号进行筛查,及时捕捉利益输送的苗头,实现腐败风险的精准干预。

## 4　结论

本文构建了基于三种特征选择方法(SelectFromModel、递归特征消除RFE和互信息选择器MutualInfoSelector)的受贿风险预测模型,并对比了各种方法作用下模型表现的性能。实验结果表明,RFE方法取得的效果最佳,其在测试集上的Accuracy、Macro $F_1$和Macro AUC分别达到0.8929、0.8792和0.9662。因此,用该模型进行后续深入研究,通过SHAP可解释性分析,揭示了受贿风险演化的内在逻辑。其现实意义在于,通过识别不同风险等级的核心影响因素,对潜在腐败行为进行前瞻性识别与干预,为纪检监察部门提供科学的决策依据,实现从“被动查处”向“主动预防”的战略转型,

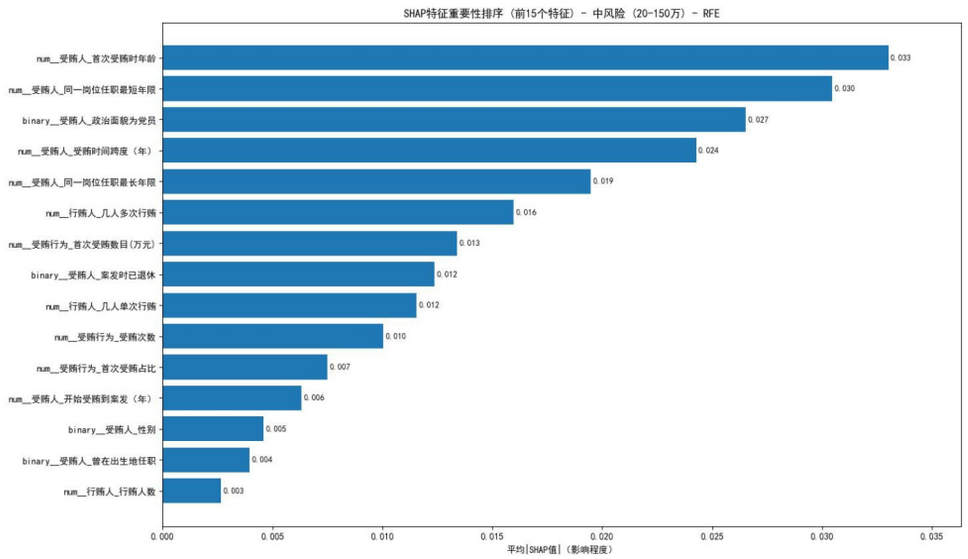



**图6　中风险类别SHAP特征重要性排序图**

**Fig.6　Ranking chart of SHAP feature importance for the medium-risk category**

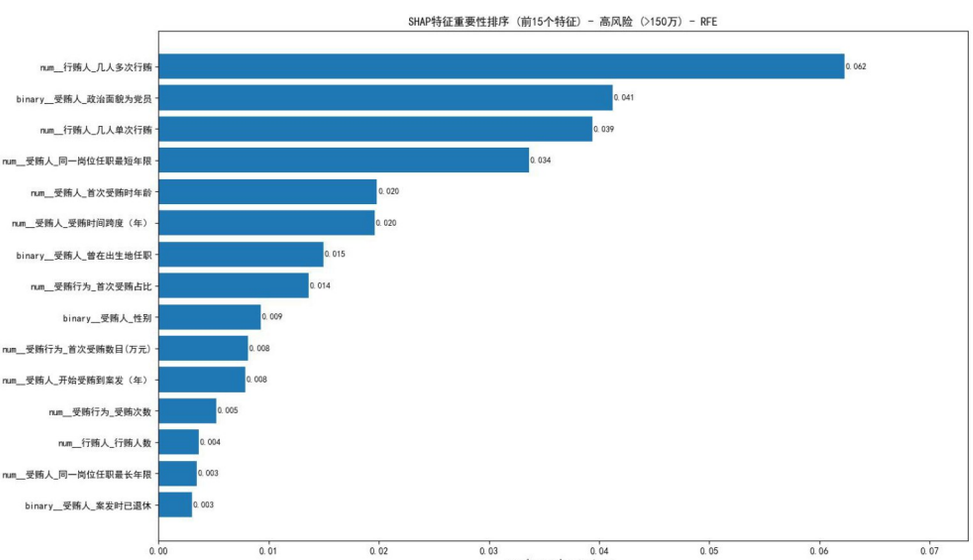



**图7　高风险类别SHAP特征重要性排序图**

**Fig.7　ranking chart of SHAP feature importance for the high-risk category**

减少国家和社会的公信力流失与财产损失。

尽管本文为受贿案件定量分析提供了新的视角,但仍存在一定局限性。第一,收集样本数量相对有限,可能影响模型泛化能力;第二,受贿行为并非孤立发生,而是嵌入于复杂的宏观经济中,本研究尚未充分纳入经济环境数据,如经济产业结构、财政与公共投资情况等;第三,数字智能时代背景下,滋生的新型腐败呈现主体复杂化、形式隐蔽化和危害广泛化的特点[31]。为进一步完善模型与应用效果,未来研究可沿以下方向推进:一是扩大样本规模并优化数据结构,增强模型的普适性和鲁棒性;二是融合金融交易记录、社交网络关系等信息,借助自然语言处理、图神经网络和时序异常检测等技术与方法构建更复杂的模型,实现对受贿行为的实时动态监测和预警。

**参考文献:**

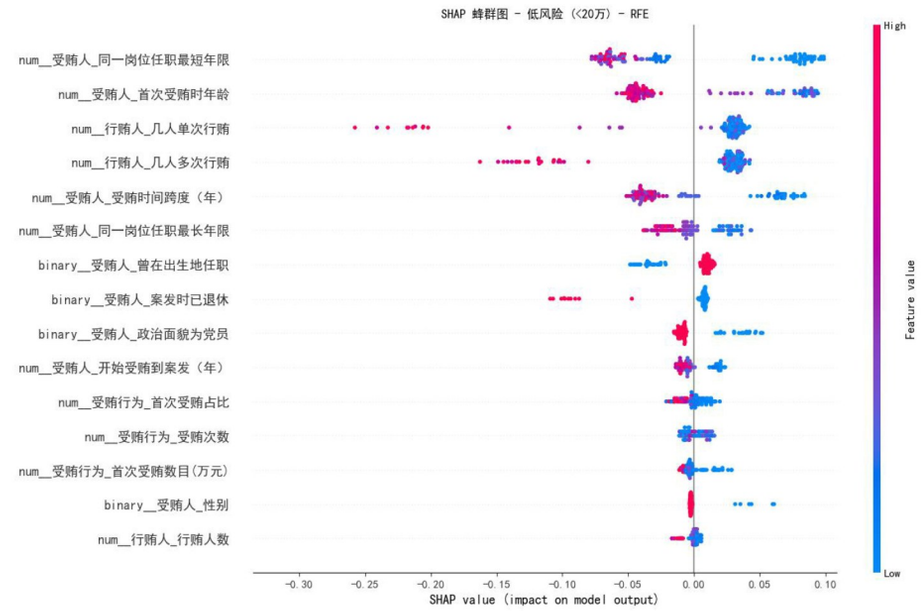



**图 8　低风险类别 SHAP 蜂群图**

**Fig.8　SHAP swarm plot for the low-risk category**

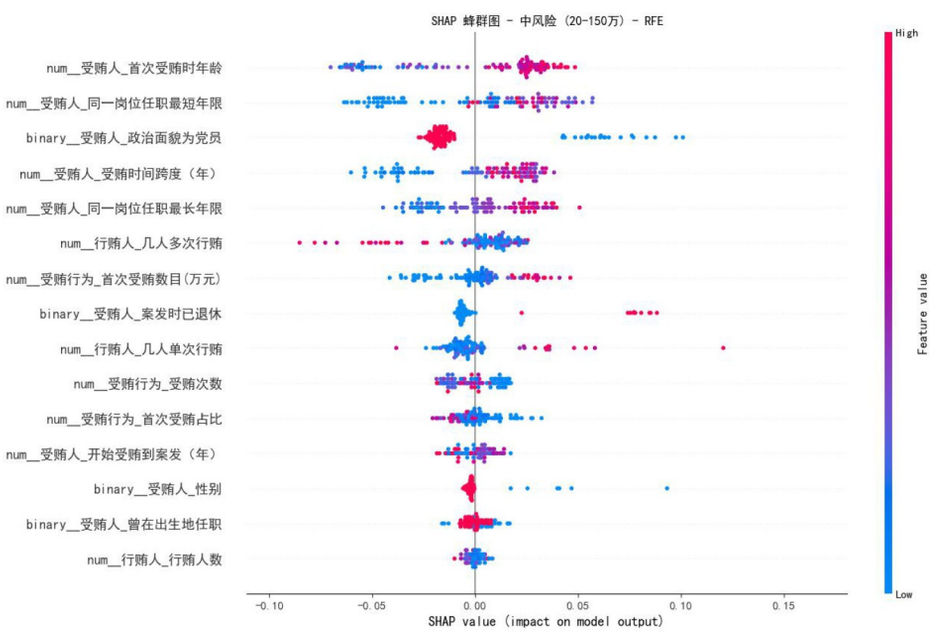



**图 9　中风险类别 SHAP 蜂群图**

**Fig.9　SHAP swarm plot for the medium-risk category**

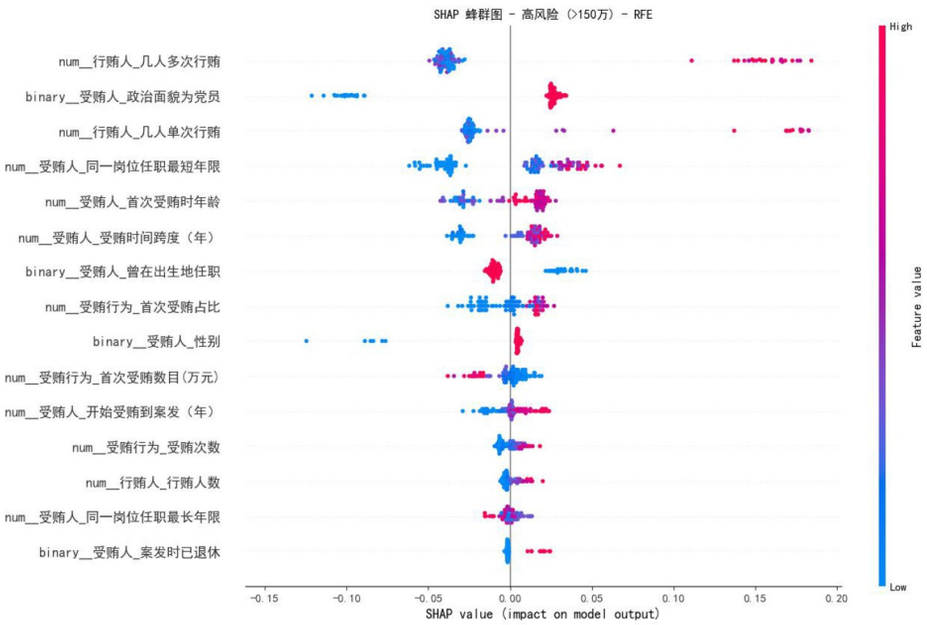



**图 10　高风险类别 SHAP 蜂群图**

**Fig.10　SHAP swarm plot for the high-risk category**

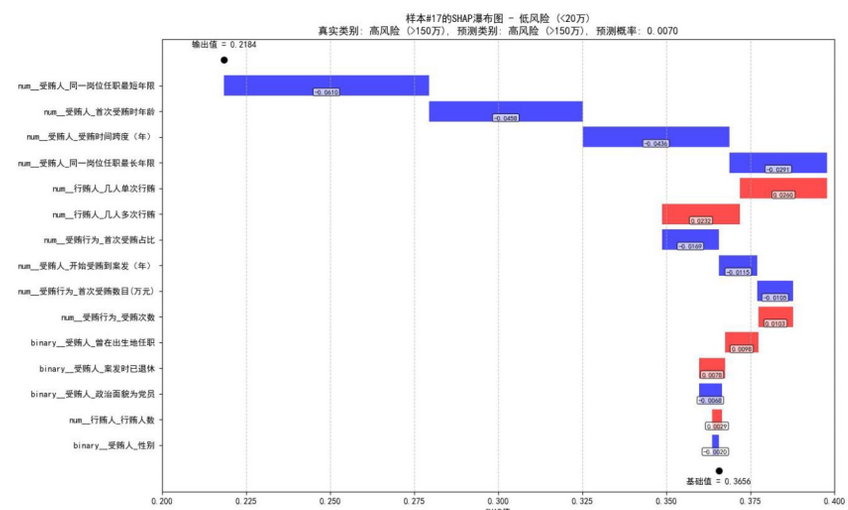



**图 11　样本#17SHAP 瀑布图(低风险)**

**Fig.11　SHAP waterfall plot of sample #17 (low-risk)**

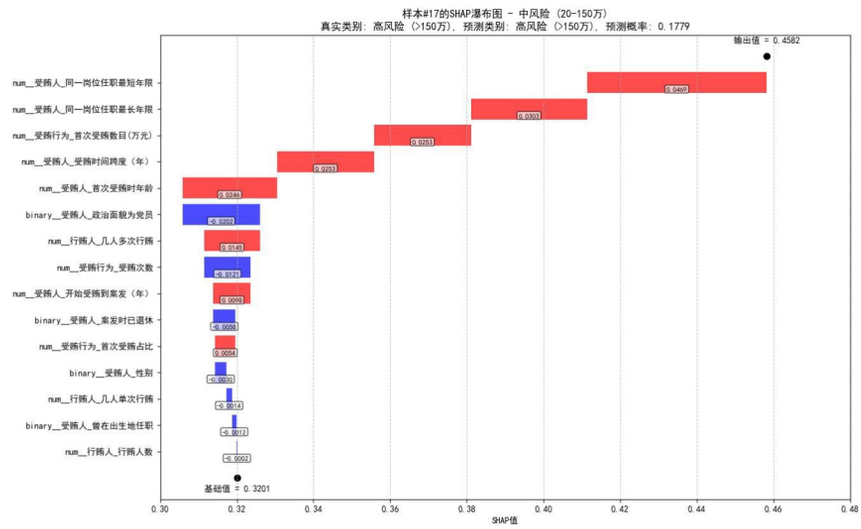



**图 12　样本#17SHAP瀑布图(中风险)**

**Fig.12　SHAP waterfall plot of sample #17 (medium-risk)**

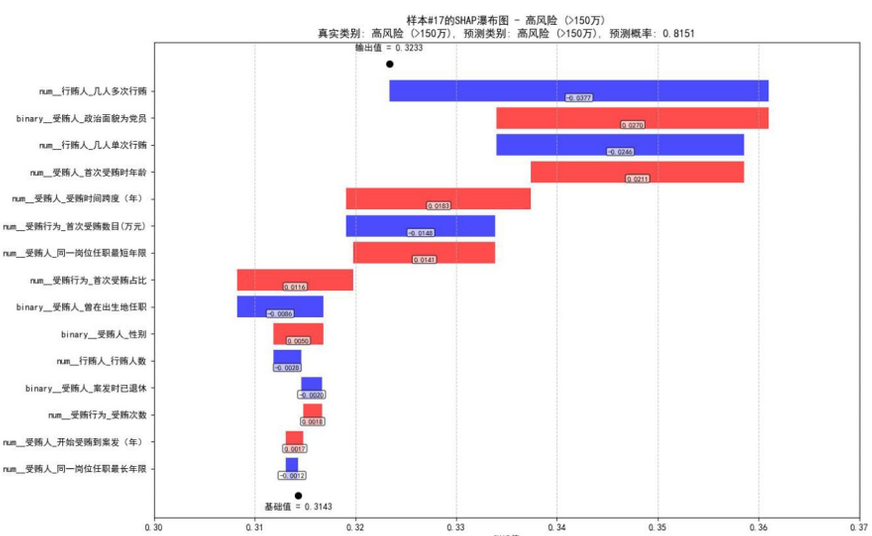



**图 13　样本#17SHAP瀑布图(高风险)**

**Fig.13　SHAP waterfall plot of sample #17 (high-risk)**

[1] 颜靖华,侯苗苗.基于LSTM网络的盗窃犯罪时间序列预测研究[J].数据分析与知识发现,2020,4(11):84-91.
YAN Jinghua, HOU Miaomiao. Predicting Time Series of Theft Crimes Based on LSTMNetwork[J]. Data Analysis and Knowledge Discovery, 2020, 4(11): 84-91.

[2] 张齐,李雪琛.基于机器学习的多标签盗窃犯罪类型识别方法研究[J].中国人民公安大学学报(自然科学版),2023,29(1):88-93.
ZHANG Qi, LI Xuechen. Research on multi-label theft crime type identification method based on machine learning[J]. Journal of People's Public Security University of China (Natural Science Edition), 2023, 29(1), 88-93.

[3] 王娟,樊畅,管雨翔.基于CNN-LSTM-Attention模型的盗窃犯罪分析与预测[J].大数据时代,2024,(10):27-40.
WANG Juan, FAN Chang, GUAN Yuxiang. Theft crime analysis and prediction based on CNN-LSTM-Attention model[J]. Big Data Era,2024, (10), 27-40.

[4] 张强,何乐平.基于支持向量机的治安高危人员风险预警方法[J].管理评论,2016, 28 (11): 245-251.
ZHANG Qiang, HE Leping. Risk early-warning method for high-risk public security individuals based on support vector machine[J]. Management Review, 2016, 28 (11): 245 - 251.

[5] 李荣岗,孙春华,姬建睿.基于支持向量机的嫌疑人特

征预测[J].计算机工程,2017,43(11):198-203.
LI Ronggang, SUN Chunhua, JI Jianrui. Suspect feature prediction based on support vector machine[J]. Computer Engineering, 2017, 43(11): 198 - 203.

[6] 王佳.杀人案件犯罪嫌疑人研判模型研究[D].北京:清华大学,2020.
WANG Jia. Research on the suspect profiling model of homicide cases[D]. Beijing:Tsinghua University, 2020.

[7] 祝起祯,苟小菊.基于LightGBM和SHAP方法的居民投资风险偏好的实证分析[J].金融科技时代,2024,32(04):79-84.
ZHU Qiling, GOU Xiaoju. Empirical analysis of residents′ investment risk preferences based on LightGBM and SHAP methods[J]. Financial Technology Times, 2024, 32(04): 79 - 84.

[8] 程祉元,张博良,蔡雨晨,等.融合随机森林与SHAP的心脏病预测及其特征分析研究[J].智能计算机与应用,2023,13(11):172-179.
CHENG Zhiyuan, ZHANG Boliang, CAI Yuchen, et al. Research on heart disease prediction and feature analysis by combining random forest and SHAP[J]. Intelligent Computer and Applications, 2023, 13(11): 172 - 179.

[9] 金东镇,郭城楠,彭芳,等.从SHAP到概率——可解释性机器学习在糖尿病视网膜病变靶向脂质组学研究中的应用[J].中国卫生统计,2023,40(4):511-515.
JIN Dongzhen, GUO Chengnan, PENG Fang, et al. From SHAP to probability: Application of interpretable machine learning in targeted lipidomics for diabetic retinopathy[J]. Chinese Journal of Health Statistics, 2023, 40(4): 511 - 515.

[10] 王驰.基于XGBoost与SHAP的"通督调神"不同针刺组方治疗失眠症的疗效预测研究[D].南京:南京中医药大学,2024.
WANG Chi. Efficacy prediction of different acupuncture formulas for treating insomnia based on XGBoost and SHAP[D]. Nanjing: Nanjing University of Chinese Medicine, 2024.

[11] 刘瑞雪,李佳轩,李云.水库消落带植物多样性空间格局预测模型及环境解释——基于XGBoost-SHAP模型框架[J].生态学报,2024,44(21):9652-9669.
LIU Ruixue, LI Jiaxuan, LI Yun. Prediction model and environmental interpretation of plant diversity spatial patterns in reservoir drawdown zones: Based on the XGBoost - SHAP model framework[J]. Acta Ecologica Sinica, 2024, 44(21): 9652 - 9669.

[12] 王书苏,万雨勤,仝纪龙,等.基于PMF和XGBoost-SHAP的兰州市主城区PM2.5来源解析及特征因子量化研究[J].环境科学学报,2025,45(04):313-321.
WANG Shusu, WAN Yuqin, TONG Jilong, et al. PM2.5 source apportionment and quantification of characteristic factors in Lanzhou's main urban area based on PMF and XGBoost - SHAP[J]. Acta Scientiae Circumstantiae, 2025, 45(04): 313 - 321.

[13] 杜治洲,赵雅星.大数据赋能隐性腐败治理的机理与路径[J].廉政文化研究,2023,14(4):16-22.
DU Zhizhou, ZHAO Yaxing. Mechanism and path of implicit corruption governance empowered by big data [J]. Research on Integrity Culture, 2023, 14(4): 16 - 22.

[14] 杜晓燕,郭吴昊,谢怡阳.人工智能赋能新型腐败治理:机理、挑战与路径[J].陕西行政学院学报,2024,38(4):11-17.
DU Xiaoyan, GUO Wuhao, XIE Yiyang. AI-empowered governance of new forms of corruption: Mechanisms, challenges and paths[J]. Journal of Shaanxi Administration Institute, 2024, 38(4): 11 - 17.

[15] 齐敏,任建明,牛朝辉.大数据环境下反腐机构的动态能力:框架、阻碍与对策[J].湖湘论坛,2023,36(3):55-67.
QI Min, REN Jianming, NIU Zhaohui. Dynamic capabilities of anti-corruption institutions in the era of big data: Framework, obstacles and countermeasures[J]. Huxiang Luntan, 2023, 36(3): 55 - 67.

[16] Breiman L. Random forests[J]. Machine Learning, 2001, 45(1): 5-32.

[17] Hastie T, Tibshirani R, Friedman J. The Elements of Statistical Learning: Data Mining, Inference, and Prediction (2nd ed.)[M]. New York: Springer, 2009.

[18] Géron A. Hands - on Machine Learning with Scikit - Learn, Keras, and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems (2nd ed.)[M]. Sebastopol: O′Reilly Media, 2019.

[19] Shapley L S. A value for n-person games[C]//Kuhn H W, Tucker A W. Contributions to the Theory of Games II. Princeton: Princeton University Press, 1953: 307 - 317.

[20] Lundberg S M, Lee S I. A unified approach to interpreting model predictions[C]//Proc of the 31st International Conference on Neural Information Processing Systems, 2017: 4768 - 4777.

[21] 张明楷.刑法学(第六版)[M].北京:法律出版社,2021.
ZHANG Mingkai. Criminal Law (6th Edition)[M]. Beijing: Law Press, 2021.

[22] 程文浩.新型腐败的识别要点和惩治策略[J].人民论坛,2021,(17):47-49.
CHENG Wenhao. Identification essentials and punish-

ment strategies of new types of corruption[J]. People's Tribune, 2021, (17): 47 - 49.

[23] 苏江涛,舒绍福.新型腐败:特征、诱因及惩治[J].学术探索, 2024, (3): 46-56.
SU Jiangtao, SHU Shaofu. New types of corruption: Characteristics, inducements, and countermeasures[J]. Academic Exploration, 2024, (3): 46 - 56.

[24] Huan Liu, Lei Yu. Toward integrating feature selection algorithms for classification and clustering[J]. IEEE Trans on Knowledge and Data Engineering, 2005, 17 (4):491-502.

[25] Yu Lei, Liu Huan. Efficient feature selection via analysis of relevance and redundancy[J]. The Journal of Machine Learning Research, 2004, 5:1205-1224.

[26] Oh I S, Lee J S, Moon B R. Hybrid genetic algorithms for feature selection[J]. IEEE Trans on Pattern Analysis and Machine Intelligence, 2004, 26(11): 1424-1437.

[27] Guyon I, Weston J, Barnhill S, et al. Gene selection for cancer classification using support vector machines [J]. Machine learning, 2002, 46: 389-422.

[28] Peng H, Long F, Ding C. Feature selection based on mutual information criteria of max-dependency, max-relevance, and min-redundancy[J]. IEEE Transactions on pattern analysis and machine intelligence, 2005, 27 (8): 1226-1238.

[29] Kohavi R. A study of cross-validation and bootstrap for accuracy estimation and model selection[C]// Proceedings of the 14th International Joint Conference on Artificial Intelligence (IJCAI). Montreal, Canada: Morgan Kaufmann, 1995: 1137-1143.

[30] Sokolova M, Lapalme G. A systematic analysis of performance measures for classification tasks[J]. Information processing & management, 2009, 45(4): 427-437.

[31] 毛昭晖,朱星宇.新型腐败的特征与类型——警惕传统型腐败向新型腐败的嬗变[J].理论与改革, 2022, (04): 84-98.
MAO Zhaohui, ZHU Xingyu. Characteristics and types of new forms of corruption: Beware of the transformation from traditional to new corruption[J]. Theory and Reform, 2022, (04): 84 - 98.