用于分类客户响应度与信用风险的多层方法

Ayomide Afolabi、Ebere Ogburu 和 Symon Kimitei 美国肯尼索州立大学数据科学学院，肯尼索，
电子邮箱：aafolab4@students.kennesaw.edu；eogburu@students.kennesaw.edu；skimitei@kennesaw.edu

摘要

本研究评估了三种不同模型（响应模型、风险模型和响应-风险模型）中各类分类器的性能，研究主题涉及信用卡邮件营销活动和违约预测。在响应模型中，极端随机树分类器的召回率最高（79.1%），凸显了其识别信用卡优惠潜在响应者的效能。相较之下，在风险模型中，随机森林分类器表现出显著的特异度（84.1%），这对于识别违约可能性最低的客户至关重要。此外，在多类别响应风险模型中，随机森林分类器达到了最高准确率（83.2%），表明其能够有效识别信用卡邮件营销活动的潜在响应者和低风险信用卡用户。本研究通过优化多种性能指标，解决了一个具体的信用风险和邮件响应度业务问题。

关键词 —逾期、信用卡优惠、邮件响应、信用风险、响应、风险、响应-风险、极端随机树分类器、随机森林分类器、二分类、多分类等。

I. 引言

信用卡用户还款逾期或违约会加剧信用卡债务和核销，这构成了财务风险，可能对银行和其他信用卡发行公司的盈利能力造成不利影响。截至2023年12月，美国

近年来，由于机器学习和统计模型能够实现较高的优惠作出回应且不会拖欠还款的潜在客户。

本研究旨在探索不同的机器学习模型，以识别更有

II. 文献综述

正如 Siddiqi 所指出的 [13], , 评估客户的信用 worthiness 通常包括：根据其历史数据建立统计模型，将模型应用于潜在客户以评估其信用worthiness，并持续使用既定业务指标监测模型的有效性。Abdou 和 Pointon 指出，采用逻辑回归等分类模型 [14], 有助于汇总客户数据，从而准确预测还款行为或盈利能力。为预测信用卡违约者，Venkatesh 等人 [3] 采用了包括朴素贝叶斯和随机森林在内的传统机器学习算法，而 Fan [7] 则比较了支持向量机与决策树的预测性能。

为进一步提升这些传统机器学习模型的性能，Ebiaredoh-Mienye 等人 [6] 将从堆叠稀疏自编码器中提取的特征作为这些模型的输入；Almajid [12] 则结合独热编码与合成少数类过采样技术（SMOTE），解决信用卡数据通常存在的数据不平衡问题。在探索深度学习模型的应用时，Sun 等人 [4] 研究了深度学习算法相较于传统机器学习方法的效果，旨在开发一种预测系统，帮助信贷发行机构建模信用卡逾期风险。此外，Kim 等人 [5] 提出了基于深度神经网络的变分序列深度嵌入，其预测客户未来序列逾期的表现优于其他基线模型。

集成技术已广泛应用于信用卡用户分类。Lee 等人 [8]着重探讨了使用 LightGBM、XGBoost 和 CatBoost 等算法进行分类。类似地，Lawi 等人 [9]提出了一种最小二乘支持向量机集成方法，而 Zhao 等人 [10] 和 Gao [11] 则分别深入研究了 CatBoost 和 XGBoost-LSTM 的有效性。

III. 材料与方法

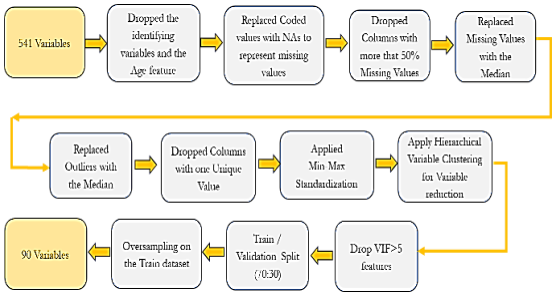

图1：数据预处理流程图

Atlanticus 数据集包含若干代表缺失数字值的编码

关于标签变量，缺失的goodbad类别定义为未开设

y_i = { 0 如果个人开通了信用卡且未逾期 1 如果个人开通了信用卡且逾期 2 如果个人未开通信用卡 ..... (1)

在该数据集中，所有观测值中有15.6%属于类别0，

此外，为减轻异常值的影响，偏离均值超过 3 个标

为从归一化后的变量中选出合适的变量，我们进行了聚类。采用层次聚类技术识别相似特征组，并通过树状图将其可视化。根据某个变量在所属聚类中的高 R平方值，将其选为聚类代表变量，从而进一步缩减变量集。

从筛选后的变量中，为了检测并解决可能存在的多重共线性问题，我们采用了方差膨胀因子法。多重共线性是指预测变量之间存在高度相关性，导致标准误增大，并使系数估计值不可靠。我们从数据集中剔除了方差膨胀因子大于 5 的预测变量，因为这些变量可能带来问题。经过这一处理，我们得到了建模过程所需的变量。

我们采用三个模型来解决业务问题：响应模型、风险模型和响应-风险模型。为构建响应模型，我们将逾期客户与未逾期客户合并，生成一个新的类别 1，代表对邮件作出响应的客户；未开卡的客户则被赋予新的类别 0，代表未对邮件作出响应的客户。响应-风险模型用于进一步评估另外两个二分类模型的效能。

在构建这些机器学习模型之前，先将数据集划分为训练集和验证集，其中 70% 用于训练，30% 用于验证，以优化模型性能。

对数据集的分析发现存在类别不平衡问题：15.56% 的观测值对应未逾期且开通过卡的个人，9.24% 对应开通过卡但已逾期的个人，75.2% 对应从未开通过卡的个人。从分布情况可以看出，类别比例存在偏斜。未开通过卡的客户在数据集中占很大比例。出现此类问题时，标准分类器往往会偏向数量较多的类别，而非数量较少的类别，这通常会导致虚假性能指标。为解决这一问题，我们使用了不平衡学习自适应合成采样方法（ADASYN）Python 库，通过生成难以分类的少数类样本的合成样本，对训练集进行过采样。必须对训练集进行过采样，以避免数据泄漏和过拟合。

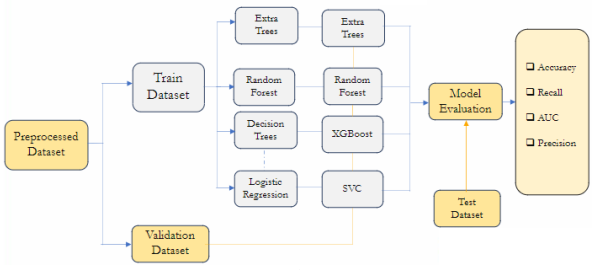

图 2：模型构建流程图

随后，利用训练数据集分别对响应、风险和响应-风险这三个模型中的多种机器学习算法进行拟合。通过这一过程，选出表现最佳的模型，并将这些模型应用于验证数据集。为确保模型针对我们的目标进行了优化，我们采用了多项性能指标。准确率衡量预测总数中正确预测所占的比例。精确率表示所有正例预测中真正例预测所占的比例，而召回率则计算

真正例预测占实际正例总数的比例。ROC曲线下面积（曲线下面积）也是一项重要指标，用于衡量模型在不同阈值设置下区分正类和负类的能力。此外，我们还使用特异度来评估风险模型正确识别负例（未违约）的能力。

为获得这些稳定且稳健的结果，我们采用了一项重要流程：调优超参数。超参数调优是优化机器学习模型的关键步骤，因为它可以调整模型参数，使模型在特定指标上取得更好的表现。通过系统地探索超参数空间，我们得以确定最优配置，在准确率、精确率、召回率、曲线下面积等各项性能指标之间实现最佳权衡。

IV. 结果与讨论

i.) 响应模型

下方的表1展示了我们为构建响应模型而考察的三种最佳分类器的性能，其中极端随机树分类器的召回率最高，达到79.1%。值得注意的是，在识别可能响应定向信用卡邮寄优惠的潜在响应者时，召回率是一项关键指标。这也体现了可能会积极响应邮件优惠的客户所占比例，因此是衡量邮件营销活动成效的重要指标。

Response Model								
	Accuracy		Precision		Recall		Auc	
Models	Train	Validation	Train	Validation	Train	Validation	Train	Validation
Random Forest Classifier	0.663	0.588	0.635	0.347	0.788	0.745	0.718	0.686
XGBoost	0.982	0.814	0.998	0.693	0.965	0.455	0.999	0.784
Extra Trees Classifier	0.641	0.561	0.612	0.337	0.794	0.791	0.694	0.693

表1：三种最佳响应分类器的性能指标

分类器

Response Model Tuning Hyperparameter		
Models	Maximum Depth	Number of Estimators
Random Forest Classifier	1	60
XGBoost	16	265
Extra Trees Classifier	1	280

表2：响应模型超参数

根据图1中的数据，我们可以深入了解模型在预测客户对定向邮件营销活动的响应方面的表现。首先，图中显示模型成功识别了 58,297 个真正例（TP），即正确预测出会响应邮件营销活动的客户。这一结果非常理想，因为这代表了潜在的互动和转化机会。同样，模型还正确识别了 38,842 个真负例（TN），准确找出了不会响应此次营销活动的客户。这一结果也很可取，因为它确保资源不会浪费在不感兴趣的人群上。

然而，有 114,859 个假正例（FP），即模型将实际上不会响应的客户错误地分类为响应者。这些情况可视为误报，也意味着错失了更有效地分配资源的机会。最令人担忧的是，模型将 15,440 个假负例（FN）预测为会响应邮件营销活动的客户，而实际上他们应被标记为非响应者。这是最不理想的结果，因为它意味着失去了促成互动和转化的机会。

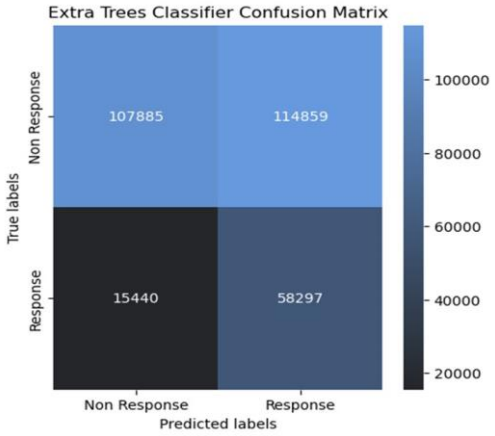

图1：响应变量极端随机树分类器的混淆矩阵

下方的图 2 中，ROC 图展示了不同决策阈值下敏感

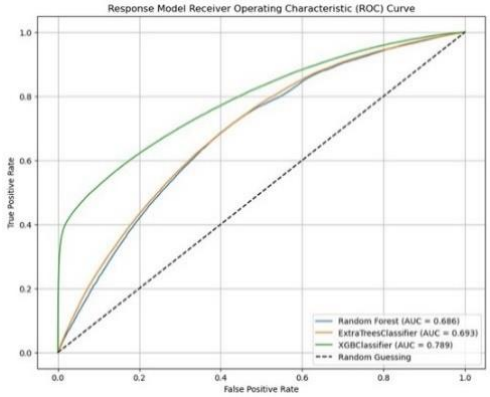

图 2：响应变量极端随机树分类器的 ROC 曲线

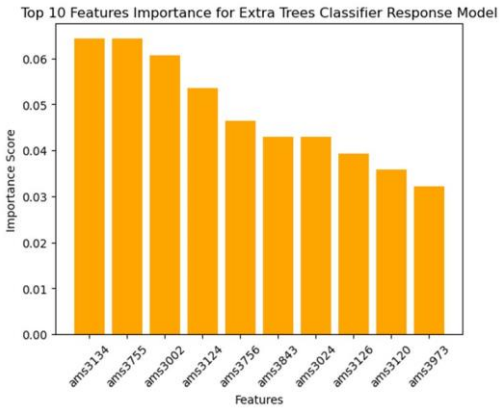

图 3：响应模型中最重要的 10 个特征

ii.) 风险模型

表3展示了我们为风险模型考察的三个最佳分类器的表现，其中随机森林分类器的特异度最高，为 0.841。值得注意的是，特异度是识别在信贷使用期间最有可能不违约的客户的一项关键指标。该结果基于经过优化的模型得出，超参数调优的目标是最大化特异度指标。如表4所示，随机森林分类器的最佳超参数为最大深度 1 和 280 个估计器。

Risk Model								
	Accuracy		Specificity		Recall		Auc	
Models	Train	Validation	Train	Validation	Train	Validation	Train	Validation
Random Forest Classifier	0.995	0.735	0.994	0.841	0.995	0.557	1.000	0.780
XGBoost	0.995	0.731	0.994	0.836	0.995	0.555	1.000	0.779
Extra Trees Classifier	0.991	0.732	0.989	0.826	0.994	0.575	1.000	0.779

表 3：排名前三的风险分类器的性能指标

Response Model Tuning Hyperparameter		
Models	Maximum Depth	Number of Estimators
Random Forest Classifier	1	60
XGBoost	16	265
Extra Trees Classifier	1	280

表 4：风险模型超参数

图 4 显示，共有 15,275 个真正例（TP），即模型预测为会违约的客户数量。相反，图中还显示有 38,842 个真负例（TN），表示模型正确识别出未违约的客户。然而，模型产生了 12214 个假负例（FN），因此将实际并未违约的客户归类为违约客户。这些情况属于误报，因为这些客户实际上并未违约。

此外，还有 7406 个假正例（FP），这表示被预测为会违约、但被随机森林模型错误地分类为非违约客户的数量。这种结果尤其不理想，因为这意味着未能准确识别高风险客户，错失了机会。

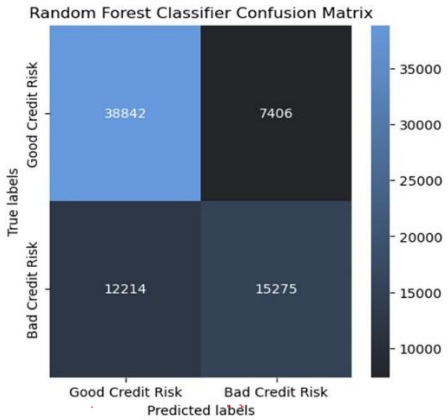

图 4：风险模型随机森林的混淆矩阵

如图 5 所示，三个模型的曲线下面积均约为 0.779。例如，特异度为 0.84 意味着，在 100 个实际未违约者中，

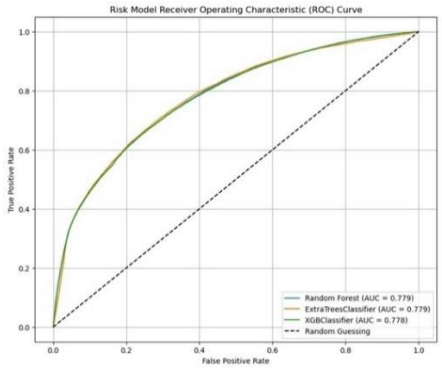

图 5：风险模型随机森林 ROC 曲线

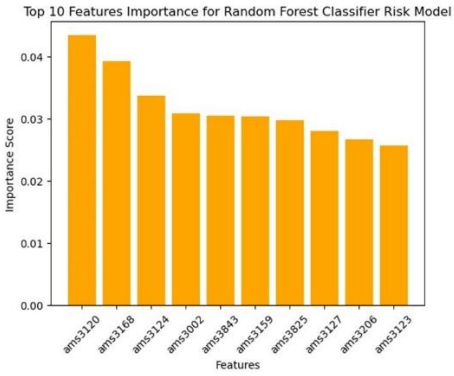

图 6：风险模型中最重要的 10 个特征

最终，通过量化风险模型所关联的信用风险来确定盈利能力。当客户被分类为信用风险较高时，将记为 -$600；而选中信用状况良好的客户时，则记为 $200。机器学习模型的盈利能力通过计算盈利评分矩阵与混淆矩阵的点积得出。如表 5 所示，使用 XGBoost、极端随机树分类器和随机森林分类器所获得的利润分别为 $5,122,400、$4,709,400 和 $5,309,400。由此可见，随机森林分类器带来的利润最高。

Risk Model Profitability			
	XGBoost	Extra Trees Classifier	Random Forest Classifier
Profit	$5,122,400	$4,709,400	$5,309,400

表 5：风险模型实现的最高利润

iii.) 多类别响应风险模型

下表 6 展示了我们在构建响应-风险模型时探索的三种最佳分类器的性能。随机森林分类器的准确率最高，为 0.823，曲线下面积为 0.786。对于多类别问题，准确率是一项有用的指标，因为它能全面衡量模型在所有类别上的表现。我们将这些指标选为最佳指标，用于识别最有可能响应定向营销邮件活动的客户比例，以及最有可能不发生信用卡违约的客户比例。我们对响应-风险模型的超参数进行了调优，重点是最大化

准确率。对于最佳模型——随机森林分类器，其最优超参数为最大深度 42 和 337 个估计器。

Response-Risk Model								
	Accuracy		Precision		Recall		Auc	
Models	Train	Validation	Train	Validation	Train	Validation	Train	Validation
Random Forest Classifier	0.996	0.823	0.996	0.763	0.996	0.559	1.000	0.786
XGBoost	0.996	0.823	0.994	0.766	0.996	0.557	1.000	0.783
Extra Trees Classifier	0.993	0.815	0.993	0.720	0.993	0.563	1.000	0.780

表 6：排名前三的响应-风险

分类器的性能指标

Response-Risk Model Tuning Hyperparameter		
Models	Maximum Depth	Number of Estimators
Random Forest Classifier	42	337
XGBoost	24	33
Extra Trees Classifier	47	310

表 7：模型风险模型超参数

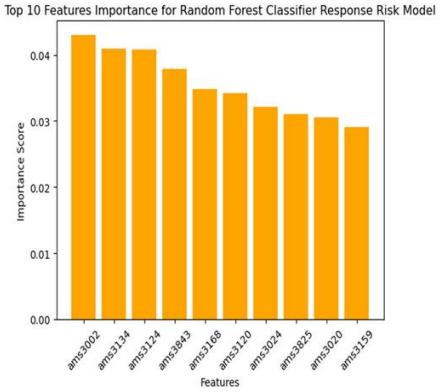

图 7：排名前 10 的重要风险模型特征

图 3、图 6 和图 7 展示了最重要的前 10 个特征。有趣的是，有两个特征在这三个模型中的排名都很高。这两个特征是 ams3002 和 ams3124，分别表示过去 24 个月和 6 个月内的查询次数。值得探究的是，这些特征在集成模型中的排名是否也会很高。这些特征在针对该业务问题的三种建模方法中均展现出较强的预测能力和稳健性。

V. 结论

总之，极端随机树分类器是准确预测哪些客户可能响应定向邮件营销活动的最佳选择。在响应风险多分类情境中，随机森林分类器表现最佳。

我们的结果表明，在识别最有可能不会违约的已开户客户时，针对真负例进行优化的机器学习模型效果最佳。特异度指标显示，随机森林分类器擅长识别最有可能不会违约（信用风险较低）的客户。因此，该分类器的盈利能力最高。

本项目凸显了机器学习方法在精准识别可能对信用卡优惠作出响应的潜在信用卡用户方面的优势，同时在目标营销策略中降低逾期风险。

V1. 参考文献

1. 美国联邦储备委员会。（2023年。）。消费者信贷 - G.19。检索自https://www.federalreserve.gov/releases/g19/current/

2. 美国联邦储备委员会。（2023年）。商业银行贷款和租赁的核销率和逾期率。检索自 https://www.federalreserve.gov/releases/chargeoff/chgallsa.htm

3. A., Venkatesh, A., & Gracia, S.（2016）。信用卡违约者预测：分类器性能的比较研究。《国际计算机应用杂志》，*145*(7), 36–41。https://doi.org/10.5120/ijca2016910702

4. Sun, T., & Vasarhelyi, M. A. (2018). 预测信用卡逾期：深度神经网络的应用。《会计、金融与管理智能系统》，25(4), 174–189. https://doi.org/10.1002/isaf.1437 5. Kim, Y., Cho, Y., Lee, H., & Moon, I.-C. (2021). 使用 Vade- Seq2Seq 预测信用卡连续逾期。2021年 IEEE国际系统、人类与控制论大会（SMC）。https://doi.org/10.1109/smc52423.2021.9659039 6. Ebiaredoh-Mienye, S. A., Esenogho, E., & Swart, T. G. (2021). ~~一种用于提高信用卡违约预测能力的人工~~神经网络技术：堆叠稀疏自编码器方法。《国际电气与计算机工程杂志》（IJECE），11(5), 4392. https://doi.org/10.11591/ijece.v11i5.pp4392-4402 7. Fan, J. (2023). 基于信用卡数据的支持向量机和决策树模型进行信用违约预测。BCP Business & ~~Management, 38, 28–33.~~ https://doi.org/10.54691/bcpbm.v38i.3666 8. Lee, W., Lee, S., & Seok, J. (2023). 使用异构集成进行信用卡违约预测。2023年第十四届国际泛在与未来网络大会（ICU~~FN~~）。https://doi.org/10.1109/icufn57995.2023.10199756 9. Lawi, A., & Aziz, F. (2018). 使用最小二乘支持向量机集成对信用卡违约客户进行分类。2018年第三届国际信息学与计算大会（ICIC）。https://doi.org/10.1109/iac.2018.8780427 10. Zhao, Y. (2023). 一种基于 CatBoost 的信用卡违约预测方法。Atlantis智能系统精选，178–184. https://doi.org/10.2991/978-94- 6463-222-4 17_11. Gao, J., Sun, W., & Sui, X. (2021). 基于 ~~XGBoost-LSTM 模型的信用卡用户违约预测~~研究。《自然界离散动力学》与

学会，*2021*年，1–13。
https://doi.org/10.1155/2021/5080472

12. Almajid, A. S. (2022). 使用 SVM-smote 和独热编码优化多层感知机，以处理不平衡数据并进行信用卡违约预测。信息系统与技术进展期刊，*3*(2), 67–74。
https://doi.org/10.15294/jaist.v3i2.57061

13. Siddiqi, N.（2005）。《信用风险评分卡：开发与实施智能信

14. Abdou, Hussein & Pointon, John。（2011）。H. A. ABDOU AND J. POINTON。《国际智能会计与财务管理系统期刊》。18。59-88。