

基于动态加权组合模型的 ATM 现金预测方法^①



杜 珊, 蔡为彬

(中国工商银行股份有限公司 软件开发中心, 珠海 519000)

摘 要: 本文提出了一种基于动态加权组合模型的智能现金预测方法, 可以对银行 ATM 设备的日常现金用量进行精准预测, 为日常现金调拨管理提供决策依据. 与以往使用的单算法预测不同, 本文对银行业务、交易流水与设备等特性进行分析, 据此组合 4 种单一机器学习模型, 提出并实现基于动态加权组合模型的智能算法. 该算法可以为银行现金用量管理提供更智能、更精准、更高效的预测手段, 有效压降现金库存总量与回钞率, 提升现金运用率. 此方法已在广东、重庆、江西、山西、北京等地区使用, 并取得良好效果.

关键词: 现金预测; 机器学习; 动态加权法

引用格式: 杜珊, 蔡为彬. 基于动态加权组合模型的 ATM 现金预测方法. 计算机系统应用, 2020, 29(8): 24-30. <http://www.c-s-a.org.cn/1003-3254/7569.html>

ATM Cash Forecasting Method Based on Dynamic Weighted Combination Model

DU Shan, CAI Wei-Bin

(Software Development Center, ICBC, Zhuhai 519000, China)

Abstract: A wise cash forecasting method based on a dynamic weighted combination model is proposed in this study, to precisely predict the daily cash consumption of ATM equipments so as to make a better decision for daily cash transfer management. Different from single-algorithm prediction used in the past, with analyzing characteristics of banking business, transaction flow, and equipment, etc., an intelligent algorithm based on a dynamic weighted combination model that combining 4 single machine learning models, is proposed and implemented in this study. This algorithm provides a more intelligent, more precise, and more efficient forecasting method for the management of bank cash consumption, effectively reduces the total amount of cash inventory and the rate of cash return, and improves the utilization rate of cash. This method has been used in Guangdong, Chongqing, Jiangxi, Shanxi, Beijing, and other areas with sound results.

Key words: cash forecasting; machine learning; dynamic weighted method

1 引言

商业银行 ATM 现金用量具有即时性、冗余性特征, 更加准确和精细化的预估现金需求量对于商业银行压降现金库存以及降低现金运营成本有着重要的意义^[1]. 其目标就是在 ATM 机中放置最佳数量的现金以满足不确定的需求. 但由于商业银行库存现金的来源和影响因素繁杂, 使得现金管理存在很大难度. 商业银行通常会采取保留过高库存现金总量的方式来满足客

户不确定的需求, 而过高的现金积压会增加银行经营成本和管理风险. 所以, 如何较为准确快速的计算出 ATM 设备未来一段时间的现金需求量, 已经成为商业银行 ATM 设备现金管理方面的热点和痛点问题.

国内外学者们对 ATM 设备现金预测进行了一定的研究, 国外学者 Simutis 等提出了支持向量机模型, 实验发现支持向量机的预测精度比神经网络的预测精度差一些^[2]. Acuna 等在人工神经网络和向量机的基础

① 收稿时间: 2020-01-20; 修改时间: 2020-02-25; 采用时间: 2020-03-11; csa 在线出版时间: 2020-07-29

上提出了 NARX 和 NARMAX 预测模型^[3]. Kumar 等基于 ANN 建立预测模型, 并以印度国家银行的数据作为训练集和测试集, 得到周预测和日预测值分别约为 95% 和 94%^[4]. Venkatesh 等基于神经网络和聚类建立模型, 预测的误差率最好约为 18%^[5]. 许琪使用 SVM 和 BP 算法预测 ATM 需求量, 结果显示使用 SVM 算法的误差标准偏差更小, 得到 SVM 算法比 BP 算法在 ATM 现金量预测中更优^[6]. 刘艳杰对支持向量回归机算法预测 ATM 现金量做了深入研究, 分别使用遗传算法的全局寻优技术改进支持向量回归算法和网格搜索法确定支持向量回归机算法计算出预测精度分别是 90.90% 和 71.38%, 得出改进的支持向量回归机具有更优的预测性能^[7]. 伍娜调整 BP 算法中网络隐含层结点数, 用以确定网络模型结构, 再优化 BP 算法的权值阈值, 用以解决原始 BP 算法预测 ATM 现金量过程中易陷入局部极值和收敛速度慢的缺点^[8]. 韦金香等组合改进 BP 网络算法和改进时间序列分析算法两种算法, 结合 SpringMVC 框架, 结合两种单算法的优势, 通过系统组合各单项算法权值, 从而得到一个精确度高, 适应性强的预测算法^[9].

目前, 对 ATM 的现金预测研究, 主要使用神经网络算法、支持向量机回归算法, 或者在此两种算法上的改进算法对 ATM 现金需求量进行建模预测, 而且一般只是在单台或者多台 ATM 设备上上进行小规模验证研究, 没有将模型效果推广到全国的 ATM 设备上上进行充分验证, 无法论证这些单一算法模型是否具备广泛实用性. 考虑到现金预测问题影响因素较多且各种因素之间关联复杂性比较高, 每台 ATM 设备情况都不相同, 本文提出一种基于动态加权组合模型的 ATM 现金预测方法, 该方法采用均匀权重向量集生成法, 通过动态组合匹配产生最优权重向量, 实现不同 ATM 设备、不同周期的多种算法最优组合模型, 能够提升模型预测精准度和模型效果稳定性, 该组合算法模型预测效果优于 LSTM 神经网络算法和遗传算法改进的 SVM 算法, 并具备大范围推广的特性.

2 构建动态加权组合模型

为了克服单一算法模型在不同 ATM 设备的建模效果, 即不同数据集上预测稳定性较差这一局限性, 本文借助排列组合的数学思想, 推出动态加权组合模型, 将提升模型预测效果的准确性和稳定性.

2.1 建立均匀权重向量集

ATM 现金预测系统整体流程图如图 1 所示, 其中最核心的模块是模型集成模块, 它确保了系统模型预测效果的稳定性和准确性.

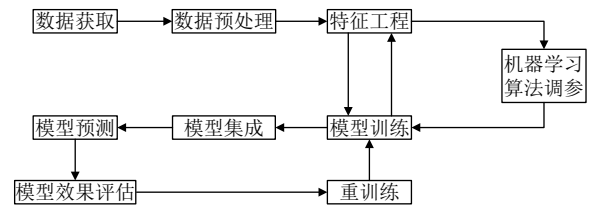


图 1 现金预测系统整体流程图

本文使用排列组合思想, 根据基模型个数动态生成均匀权重向量集合 I . 具体实现方法如下:

给定 n 和 H , n 代表基模型个数, $1/H$ 代表权重变化的最小粒度.

设定 $M = \{m_1, m_2, \dots, m_H\}$, 且 $m_1 \neq m_2 \neq \dots \neq m_H$; 使用插空法将集合 M 中的 1 分为 n 组, 可以得到所有组合的集合:

$$A = \left\{ \{a_{11}, a_{12}, a_{13}, \dots, a_{1n}\}, \{a_{21}, a_{22}, a_{23}, \dots, a_{2n}\}, \dots, \left\{ a_{A_{H+1}^{n-1}1}, a_{A_{H+1}^{n-1}2}, a_{A_{H+1}^{n-1}3}, \dots, a_{A_{H+1}^{n-1}n} \right\} \right\},$$

其中, A 中的每个子集合 $a_{i1} + a_{i2} + \dots + a_{ij} + \dots + a_{in} = H$; 且 $0 \leq a_{ij} \leq H$; 且 $a_{ij} \in N_+$.

得到第 i 个权重向量集合: $\frac{a_{i1}}{H} + \frac{a_{i2}}{H} + \dots + \frac{a_{ij}}{H} + \dots + \frac{a_{in}}{H}$.

得到权重向量集合 I 为:

$$I = \left\{ \left\{ \frac{a_{11}}{H}, \frac{a_{12}}{H}, \frac{a_{13}}{H}, \dots, \frac{a_{1n}}{H} \right\}, \left\{ \frac{a_{21}}{H}, \frac{a_{22}}{H}, \frac{a_{23}}{H}, \dots, \frac{a_{2n}}{H} \right\}, \dots, \left\{ \frac{a_{A_{H+1}^{n-1}1}}{H}, \frac{a_{A_{H+1}^{n-1}2}}{H}, \frac{a_{A_{H+1}^{n-1}3}}{H}, \dots, \frac{a_{A_{H+1}^{n-1}n}}{H} \right\} \right\}$$

获取每个单一模型在过去 D 天内的预测值和真实值, 为每个单一模型从权重向量集中 I 的每个权重向量, 并计算加权值, 算出该加权向量对应的 $RMSE$ 取值, $RMSE$ 取值最小的权重向量即为匹配的最优权重向量.

2.2 构建集成模型

假定有 n 个机器学习或者深度学习模型, 统称为基模型, 那么动态加权组合模型是指根据每个基模型在数据集上的预测效果, 动态生成每个基模型的权重, 再将每个基模型的预测结果按照权重进行加权的模型.

(1) 基模型选择

可用于 ATM 现金预测的模型总体上来说可以分为 3 类: 广义线性回归模型、基于树的回归模型和时间序列回归模型. 首先, 挑选出可用于时间序列问题预测的常用模型. 从广义线性回归模型中挑选出线性回归模型、K 近邻回归模型^[10]、SVR 回归模型、Elastic Net 回归模型, 基于树的回归模型中挑选出决策树回归模型、随机森林回归模型^[11]、GBDT 回归模型、LightGBM 回归模型和 XGBoost 回归模型^[12], 时间序列回归模型中挑选出 ARIMA 回归模型和 LSTM 神经网络模型^[13], 共计 11 个模型. 然后, 用挑选出的 11 个模型进行建模, 用于预测 ATM 设备未来现金需求量. 建模的数据来自全国不同地区的数百台试点 ATM 设备历史交易流水数据. 最后, 在测试集上得到 11 个模型在试点 ATM 设备上的现金需求量预测值, 并统计每个模型在所有试点 ATM 测试集上的预测值与真实值偏差在 30% 以内的百分比, 所占百分比越大, 代表模型预测效果接近真实值.

从试点设备模型的实验数据可以得到如下结论: 同一种算法不可能满足所有的 ATM 设备预测需求, 即同一种算法一般只是在一部分 ATM 设备上建模效果良好; 同一种算法在同一台 ATM 设备上不同时间段的模型预测效果是不一样的. 所以, 为了使得模型具备可推广性, 建模方式要满足 ATM 设备自动识别出其最合适的算法. 为了达到这一目的, 在本文中采取动态加权组合的集成模型方式在多种算法之间取得一个平衡, 以期在对 ATM 设备进行预测时, 可以加大预测效果较好的模型影响, 减少甚至消除预测效果较差的模型影响.

(2) 自适应的模型集成

本文采用动态加权的思想自适应地集成基础模型, 形成最终的组合模型. 集成模型算法思路如图 2 所示, 假设有基模型 $model_1$ 、 $model_2$ 、 $model_3$ 、 $model_4$, 则组合模型 $model = a \times model_1 + b \times model_2 + c \times model_3 + d \times model_4$, 其中, 权重 a 、 b 、 c 、 d 的取值是根据对应模型的预测效果系统自适应地从集合 I 中挑选的, 以期获取预测效果最优的组合模型, 这个过程即为集成模型 (组合模型).

模型权值确定方法如下: 从集合 I 中任意挑选一个权重向量 $w_k = \{w_{k1}, w_{k2}, \dots, w_{kn}\}$, $w_k \in I$, 表示模型 $model_1, model_2, \dots, model_n$ 的权重, 用 Y_1, Y_2, \dots, Y_n 表示模型 $model_1, model_2, \dots, model_n$ 在过去 T 天的预测值,

用 R 表示 ATM 设备在过去 T 天的客户真实需求值. 用 Y_{ij} 表示 $model_i$ 在过去第 j 天的预测值, R_j 表示过去第 j 天 ATM 客户真实需求金额, 定义的损失函数 $L(w_k)$: 当权重向量为 w_k 时, 求过去 T 天组合模型 $w_{k1}model_1 + w_{k2}model_2 + \dots + w_{kn}model_n$ 的预测结果和真实值之间的均方根误差, $L(w_k)$ 值越小, 代表组合模型预测结果越接近真实值, 模型准确度越高. 公式如下:

$$L(w_k) = \sum_{j=1}^T \sqrt{\left(R_j - \sum_{i=1}^n w_{ki}Y_{ij}\right)^2} / T$$

每次进行预测时, 都从集合 I 中挑选出 $L(w_k)$ 值最小的权重向量 w_k 进行对模型集成, 可以得到均方根误差最小的组合模型. 为了挑选出最合适的 w_k , 预测时间复杂度增加 $O(n)$.

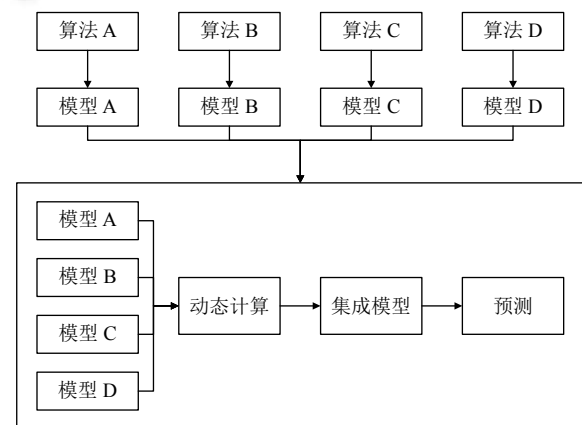


图 2 集成模型算法图

2.3 模型实现

ATM 现金预测模型实现的系统架构图如图 3 所示, 将一台 ATM 设备的试点开关打开后, 依次经过数据处理层、单模型训练层、组合模型预测层和个性化调优层, 即可获得 ATM 设备的预测结果.

系统架构中每一层的具体功能为:

数据处理层: 在这一层存储所有 ATM 设备的原始数据, 并在 Hadoop 大数据平台上进行数据清洗提高数据质量, 以及特征工程生成特征数据.

单一模型训练层: 使用网格搜索和贝叶斯优化对算法的超参数进行自动调优, 获得预测效果最好的算法模型.

组合模型预测层: 在这一层首先需要生成均匀权重向量集, 然后从向量集中动态的选择出使得多个模型加权效果最好的权重向量, 得到组合模型, 最后使用组合模型对 ATM 未来现金使用量进行预测.

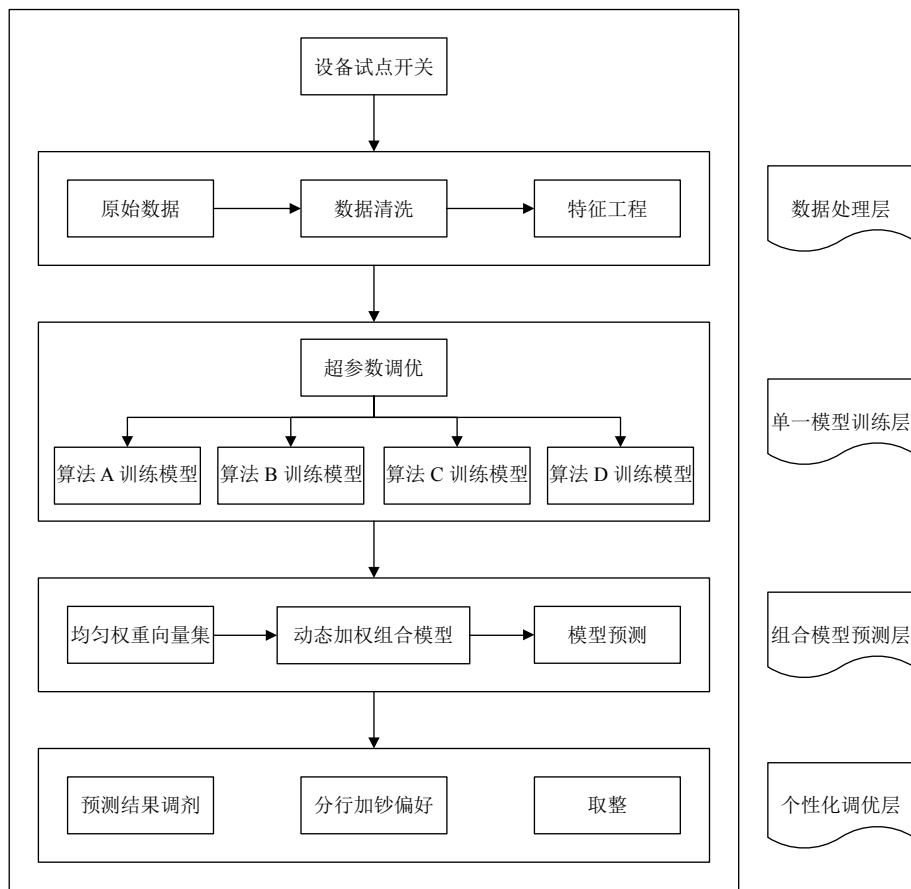


图3 ATM 现金预测系统架构图

个性化调优层: 首先, 根据每台 ATM 设备模型预测效果, 对组合模型的预测结果进行调剂, 然后, 根据分行加钞偏好 (加钞偏多还是偏少) 对调剂后的预测结果进行调整, 最后, 根据分行设定取整额度对调整后的预测结果进行取整, 取整后的值即为 ATM 现金预测系统提供的 ATM 设备未来现金需求量。

3 动态加权组合模型验证

3.1 数据来源

实验数据来源于商业银行 ATM 设备从 2017 年 2 月 1 日到 2019 年 2 月 20 日每天的真实日终存取款交易数据, 做了一定的移形处理。在数据不缺失的情况下, 每台设备都应该有数据记录。将原始数据进行数据预处理后, 需要把数据集划分为训练集、测试集和验证集, 进行模型训练和验证, 划分方式为: 2019 年 1 月 1 日到 2019 年 1 月 31 日的样本作为测试集, 2019 年 2 月 1 日到 2019 年 2 月 20 日的样本作为验证集, 剩余的样本作为训练集。

以一台 ATM 设备 (存取款一体机) 为例, 解析 ATM 现金预测系统是如何对原始数据进行预处理的。原始数据包含四列: 日期、每日取款总量、每日存款总量和每日轧差 (每日轧差=每日取款总量-每日存款总量), 其中每日轧差为标签。

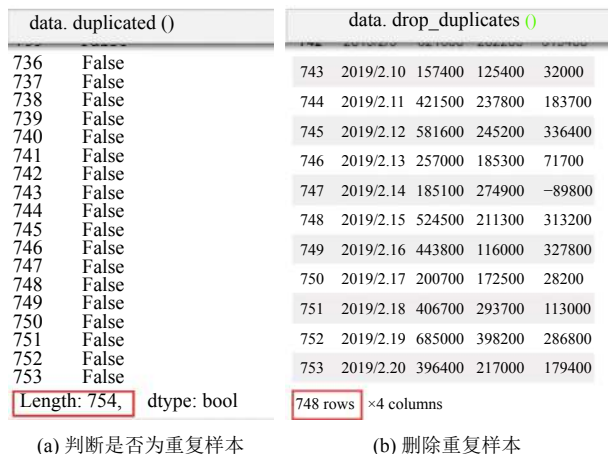
训练模型时, 采用网格搜索或者贝叶斯优化对每个算法的超参数实现自动调优, 寻找使得模型预测效果最优的超参数组合。

3.2 数据清洗及特征处理

原始数据中包含了大量的噪声数据, 直接用这样的原始数据生成特征数据, 会大大降低训练模型的效果, 数据清洗可以对原始数据中的重复样本、缺失样本、异常样本以及离群样本进行处理。考虑到现金预测问题为时间序列问题, 对缺失样本、异常样本和离群样本进行处理时, 采用修复样本的方式, 以确保数据集在时间上的连续性。数据清洗包括以下部分:

(1) 删除重复样本。使用 `duplicated()` 方法标记样本是否为重复样本, 得到图 4(a), 可以判断出设备原始数

据共有 754 个样本; 然后, 使用 drop_duplicates() 方法删除重复样本, 得到图 4(b), 还剩余 748 个样本。



(a) 判断是否为重复样本 (b) 删除重复样本

图 4 删除重复样本操作

(2) 修复缺失样本和异常样本. 使用 describe() 方法对删除重复样本后的数据进行一个描述分析, 查看属性的 count 值 (非 NA 值的数量)、mean 值 (均值)、std 值 (标准差)、min 值 (最小值)、25% 值 (第 25 百分位数)、50% 值 (第 50 百分位数即中位数)、75% 值 (第 75 百分位数) 和 max 值 (最大值), 通过这些值我们可以对数据做初步的了解, 并进一步分析出缺失样本和异常样本, 如图 5 所示。

	outcash	outcash	diff
count	7.480000e+02	748.000000	748.000000
mean	3.836789e+05	233273.930481	150404.946524
std	1.948108e+05	104671.054709	183241.967346
min	.000000e+00	0.000000	-588000.000000
25%	2.449750e+05	162250.000000	25250.000000
50%	3.410000e+05	219550.000000	112300.000000
75%	4.800750e+05	287025.000000	235600.000000
max	1.175400e+06	991700.000000	795800.000000

图 5 数据集的描述性分析

从图 5 中可以得到原始数据样本数量为 748, 且这些样本的每一列都不为空, 所以样本不存在部分属性为空的现象, 同时取款值 (outcash) 和存款值 (incash) 均为可被 100 整除的正整数, 所以数据中不存在异常样本。

(3) 修复离群样本. 找出可能的离群样本的方式为: ① 每日轧差为 0 的样本; ② 箱线图法, 具体判断方法为: 存取款轧差大于箱线图上边缘值或者小于下边缘值的样本很可能为离群样本, 然后对这些样本进行探索性分析来确定是否为离群样本. 采用箱线图法挑选出原始数据中可能的离群样本. 以“年-月”为最小单位绘制设备的箱线图, 如图 6 所示。

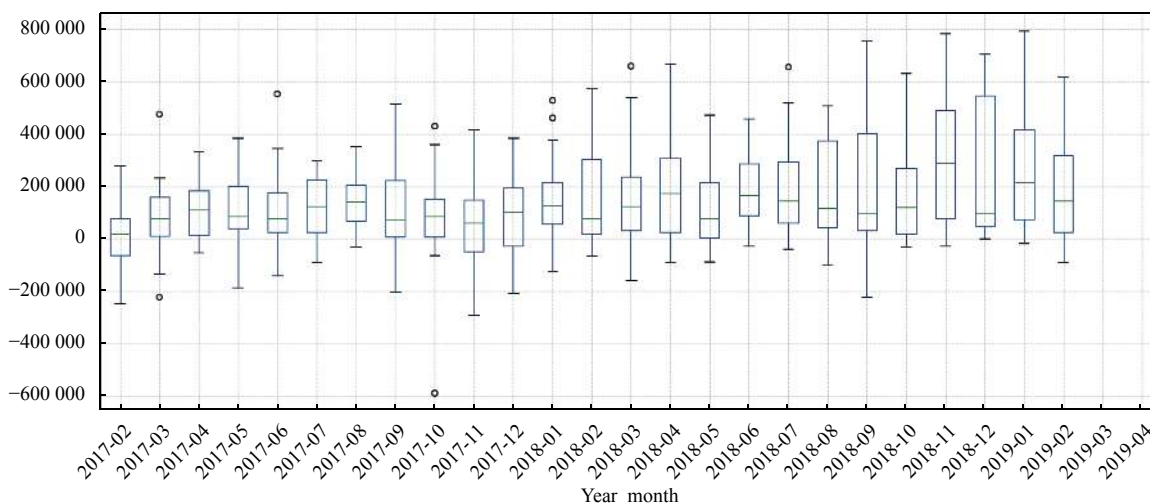


图 6 “年-月”维度的 ATM 现金需求量箱线图

以图 6 中 2017 年 10 月 1 日和 10 月 2 日这两条样本为例, 来进行离群点判断. 绘制该台设备 2017 年 9

月 20 日到 2017 年 10 月 30 日的现金需求量曲线图如图 7 所示, 2017 年 10 月 1 日样本对应的轧差为 -588 000,

这一天的样本现金需求量明显有异于其他样本, 直接判定为离群点. 2017年10月2日样本是对应的轧差为429 900, 虽然箱线图判定该样本对应的值超出了2017年10月份的上限值, 但是可以从图7中看出该设备在国庆节前两天就已经出现每日轧差大幅上涨情况, 所以样本2017年10月2日符合节假日的变化情况, 为正常样本.

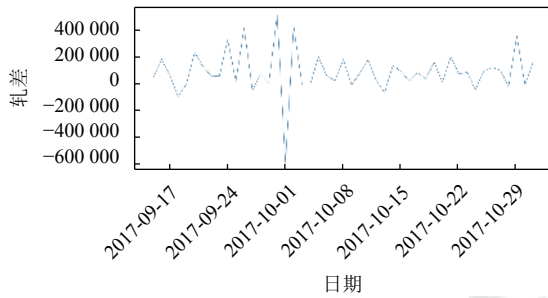


图7 现金需求量变化曲线

在数据特征处理中, 针对节假日、发薪日和还款日这些特殊日期进行单列操作, 可以衍生出47个新特征; 结合2017年至2019年该ATM设备所在区域天气状况数据以及地理位置数据, 进行单列操作衍生出18个新特征, ATM设备现金需求量变化具备周期性的特征, 对每日轧差这一列进行分组/聚合操作. 从同期、环比、往期、平均、统计这5个角度, 生成共计41个新特征. 所以, 通过特征构建, 特征维度变为106维.

3.3 动态加权组合模型效果

本文使用均方根误差(RMSE)来判断模型与实际数据的拟合度, 取值越小代表预测值和真实值误差越小.

第1步. 生成组合模型. 本文中的ATM现金预测系统使用K近邻回归模型、随机森林回归模型、XGBoost回归模型和LSTM神经网络模型4个模型作为基模型进行模型集成, 设定 $n=4$, $H=20$, 那么最小变化步长为0.05, 生成包含7980个权重向量的均匀权重向量集 I . 然后, 为了保证组合模型的预测效果, 在每次进行预测时, 系统都需要重新从均匀权重向量集 I 中挑选出一个权重向量, 用来加权4个基模型, 生成组合模型, 使得组合模型的近期ATM设备现金需求量变化趋势拟合最佳.

第2步, 分别使用4个基模型、遗传算法改进的SVM模型和组合模型在验证集上进行预测, 分别对比4个基模型和组合模型在验证集上的预测效果、遗传算法改进的SVM模型和组合模型在验证集上的预测效果.

4个基模型和组合模型在验证集上的预测效果如图8所示.

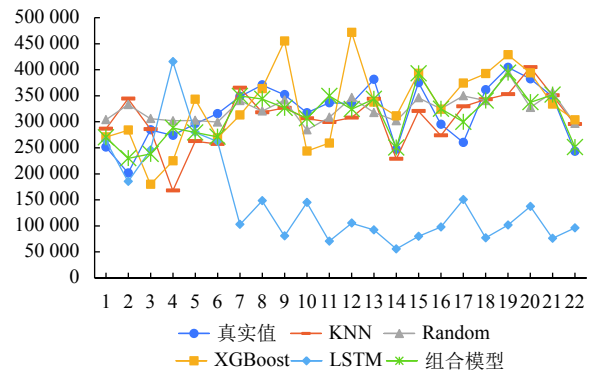


图8 4个基模型和组合模型预测效果对比图

考虑到ATM设备需要保证不缺钞的实际情况, 预测结果略大于真实轧差更符合实际需求. 在图8中, 圆点(•)折线代表ATM真实现金需求情况, 星号(*)折线代表组合模型预测情况, 其他折线代表4个基模型的预测情况, 可以看出组合模型预测效果整体上更加贴近真实情况, 且组合模型预测效果优于神经网络算法LSTM.

4个基模型和组合模型在验证集上的RMSE取值如表1所示.

表1 4个基模型和组合模型的RMSE值

效果	KNN	随机森林	XGBoost	LSTM	组合模型
RMSE值	22 152.657	20 232.189	27 630.287	88 529.637	10 873.069

从表1可以很直观地看出组合模型的RMSE取值最小, 证明组合模型预测效果优于4个基模型.

遗传算法改进的SVM模型与组合模型预测效果如图9所示, 同时遗传算法改进的SVM模型的RMSE取值为: 50 378.591, 组合模型RMSE取值为: 10 873.069, 所以在本论文数据集中组合模型预测效果优于遗传算法改进的SVM算法.

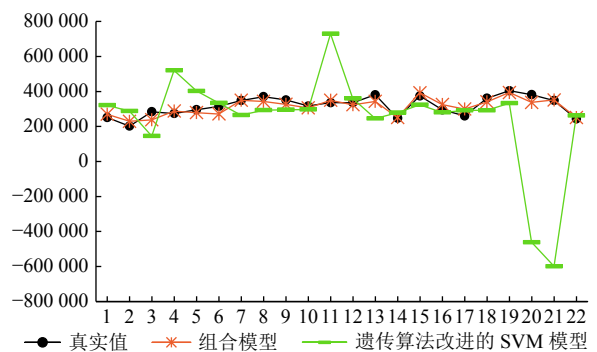


图9 遗传算法改进的SVM模型和组合模型预测效果对比图

所以,针对 ATM 现金预测这种具有大量不同数据集的业务场景,基于单一算法模型充分调参,并根据每个数据集变化趋势而动态调整多个单一模型权重,结合单一算法模型在该数据集上的不同时间段内的预测优势,形成动态加权组合模型.相比较于单一算法模型,组合模型在预测效果的准确性和稳定性上更具优势.

4 总结与展望

本文提出基于动态加权组合模型的 ATM 现金预测方法,在商业银行实现了方法运用,并推广 ATM 现金预测系统,完成了数万台 ATM 设备试点上线.该系统针对每台 ATM 设备进行单个模型训练,根据单个模型在 ATM 设备上的预测效果,在每次预测时进行动态加权,生成组合模型,提升了系统预测准确度和稳定性.自系统上线推广以来,效果显著,在 ATM 现金保障率不低于 99.3% 的情况下,ATM 设备回钞率平均降低 15~25%,现金运用率平均提升 10~22%.

后续,在提高组合模型的预测效果方面,有以下几种思路:(1)使用更适合该场景的单一算法替代原有单一算法来进行组合.(2)采用聚类分析等方法,分析出 ATM 的同类设备,它们具备相似的变化趋势和相近的现金需求量级别.对于同类设备的建模模型预测效果要优于单台设备的建模模型效果.而对于非同类设备联合建模效果比单台设备建模效果更差.

参考文献

- 1 刘艳杰. ATM 现金流预测研究现状与展望. 科技经济导刊, 2018, 26(24): 11-12.
- 2 Simutis R, Dilijonas D, Bastina L, *et al.* A flexible neural network for ATM cash demand forecasting. Proceedings of the 6th WSEAS International Conference on Computational Intelligence, Man-Machine Systems and Cybernetics. Stevens Point, WI, USA. 2007. 162-165.
- 3 Acuña G, Ramirez C, Curilem M. Comparing NARX and NARMAX models using ANN and SVM for cash demand forecasting for ATM. Proceedings of 2012 International Joint Conference on Neural Networks. Brisbane, QLD, Australia. 2012. 1-6.
- 4 Kumar PC, Walia E. Cash forecasting: An application of artificial neural networks in finance. International Journal of Computer Science & Applications, 2006, 3(1): 61-77.
- 5 Venkatesh K, Ravi V, Prinzie A, *et al.* Cash demand forecasting in ATMs by clustering and neural networks. European Journal of Operational Research, 2014, 232(2): 383-392. [doi: 10.1016/j.ejor.2013.07.027]
- 6 许琪. 基于支持向量机的 ATM 机现金需求预测研究 [硕士学位论文]. 杭州: 浙江工业大学, 2009.
- 7 刘艳杰. 基于 GA-SVR 的 ATM 现金需求量预测 [硕士学位论文]. 广州: 暨南大学, 2016.
- 8 伍娜. 基于改进遗传神经网络的 ATM 现金预测的研究 [硕士学位论文]. 广州: 暨南大学, 2016.
- 9 韦金香, 张建同. 银行 ATM 设备业务总量的时序特征分析及预测. 上海管理科学, 2017, 39(6): 25-28. [doi: 10.3969/j.issn.1005-9679.2017.06.005]
- 10 闭小梅, 闭瑞华. KNN 算法综述. 科技创新导报, 2009, (14): 37. [doi: 10.3969/j.issn.1674-098X.2009.14.024]
- 11 蔡永川, 李琦. 基于随机森林算法的 PM_{2.5} 浓度快速预测模型. 环境工程, 2018, 36(S1): 100-104, 129.
- 12 Chen TQ, He T, Benesty M, *et al.* xgboost: Extreme gradient boosting. <https://pbil.univ-lyon1.fr/CRAN/web/packages/xgboost/>, 2016.
- 13 Gers FA, Schmidhuber J, Cummins F. Learning to forget: Continual prediction with LSTM. Proceedings of the 9th International Conference on Artificial Neural Networks ICANN '99. Edinburgh, UK. 1999. 850-855.