

# 基于混合神经网络模型的卷烟投放预测方法

邹旺<sup>1</sup>,张吴波<sup>1,2</sup>,马勋政<sup>1,3</sup>

- 湖北汽车工业学院,电气与信息工程学院,湖北 十堰 442002;
- 武汉科技大学,计算机科学与技术学院,湖北 武汉 430081;
- 湖北省烟草公司十堰市公司,湖北 十堰 442000)

**摘要:**为科学制定卷烟产品的投放策略,提出1种基于混合神经网络模型的卷烟投放预测方法.通过结合 Boosting-GRU 多元预测模型、产品销售量至投放值计算以及 GRU 多分类模型3种方法来实现卷烟产品不同档位的投放策略的生成.以湖北省十堰市的卷烟产品历史销售数据为实验对象,对所提的方法进行验证.实验结果表明,该方法的产品销售量预测平均准确率达到97.64%,投放策略生成的平均绝对误差均低于20%.

**关键词:**多元预测模型;多分类模型;平均绝对误差

**中图分类号:**TP-9 **文献标志码:**A **文章编号:**1672-8513(2024)01-0117-06

烟草部门落实国家“总量控制,稍紧平衡”的方针,为了更好的消化社会库存,促进市场良性发展,在每轮投放中都会根据卷烟规格和不同档位的零售户来制定具体的投放策略.由于卷烟品种繁多,投放策略需要每轮重新制定,任务繁重,同时存在主观情感等局限性,难以实现大规模的准确计算.建立科学的卷烟投放预测模型有利于烟草专卖局对资源合理的利用和产品精确的投放,有利于推进精细化管理.

目前卷烟投放预测技术的研究主要包括传统方法和机器学习方法.传统方法是应用概率统计学和计量经济学的思想,建立概率模型或多组线性函数去拟合并预测卷烟历史销售数据.王伟民等<sup>[1]</sup>提出1种基于灰色马尔科夫模型的卷烟需求预测方法,并在10年的全国卷烟销量数据集上验证该模型的准确性.王诗豪等<sup>[2]</sup>集成差分自回归移动平均模型(ARIMA)、向量自回归模型(VAR)、支持向量回归(SVR)以及习惯消费下的卷烟需求模型来实现卷烟需求的预测.随着机器学习算法的广受瞩目,不少研究者逐渐将其引入到卷烟销量预测的任务中.武牧等<sup>[3]</sup>提出1种线性模型整合支持向量机(SVM)的卷烟销量预测方法,该模型相比ARIMA模型的误差率降低45.79%.为了进一步提升学习器的性能,研究者考虑将单个学习器融合向集成学习方向发展.目前,集成学习按照优化方向可以分为用于减少方差的装袋算法(Bagging)<sup>[4]</sup>、用于减少偏差的提升算法(Boosting)<sup>[5]</sup>、用于提升准确率的堆叠算法(Stacking)<sup>[6]</sup>3大类.韩伟民等<sup>[7]</sup>提出1种极端梯度提升(XGBoost)的卷烟创新产品的工商交易预测.机器学习方法中的深度学习算法的适用性更广、精度较优.在卷烟销量预测中广泛使用有门控神经网络特点的长短期记忆网络(LSTM)<sup>[8]</sup>和循环门控单元(GRU)<sup>[9]</sup>模型.邓超等<sup>[10]</sup>提出1种基于LSTM和BP神经网络模型的卷烟智能投放模型,该模型能实现产品销量的预测和卷烟产品投放策略的生成.

## 1 相关技术

### 1.1 GRU 神经网络

GRU 作为 LSTM 的改进优化模型,拥有更少的参数和简单的结构. GRU 只包含更新门、重置门两种门控结构,更新门负责对信息的筛选和存储;重置门负责对上一节点状态选择性遗忘. GRU 内部结构如图 1

收稿日期:2022-08-02.

基金项目:湖北省重点研究项目(TA02002);湖北省中央引导地方科技发展专项(2018ZYYD007).

作者简介:邹旺(1998-),男,硕士.主要从事机器学习、自然语言处理研究.

通信作者:张吴波(1977-),男,硕士,副教授.主要从事人工智能、机器学习、大数据挖掘研究.

所示。

其中,  $X^t$  表示当前节点的输入;  $h^{t-1}$  表示上一节点的隐藏状态;  $r$  为重置门;  $z$  为更新门;  $h'$  为候选隐藏状态;  $h^t$  表示传递给下一节点的隐藏状态. GRU 内部计算如式(1) - (4)所示。

$$r = \sigma(w^1[h^{t-1}, X^t] + b^1). \quad (1)$$

$$z = \sigma(w^2[h^{t-1}, X^t] + b^2). \quad (2)$$

$$h' = \tanh(w^3[r \otimes h^{t-1}, X^t] + b^3). \quad (3)$$

$$h^t = (1 - z) \otimes h^{t-1} + z \otimes h'. \quad (4)$$

其中:  $\sigma$  表示 Sigmoid 激活函数;  $\tanh$  为激活函数;  $[\ ]$  表示两个向量相连接;  $W^{1-3}$  为权重矩阵;  $b^{1-3}$  为偏置项。

### 1.2 Boosting - GRU 模型

Boosting 算法的思想是将多个基学习器串联组合,通过不断的调整样本权重来训练每个基学习器,然后计算每个基学习器的权重值,待所有基学习器训练完成后,按相应的权重值线性组合所有基学习器得到一个整体模型,以此来提升模型的性能.更新样本权重的目的是使本轮预测错误的样本在下一轮训练中得到更大的权重,受到更多的重视<sup>[11]</sup>.权重值用于评估每个基学习器的重要性,基学习器预测越准确在整体模型中的权重越大。

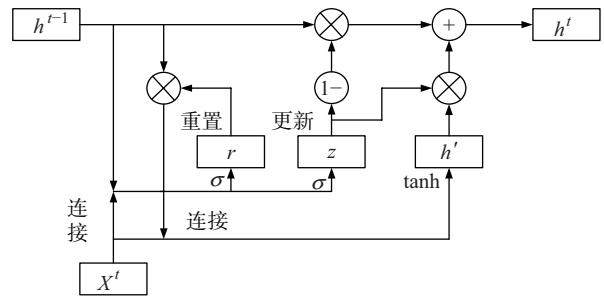


图 1 GRU 结构图

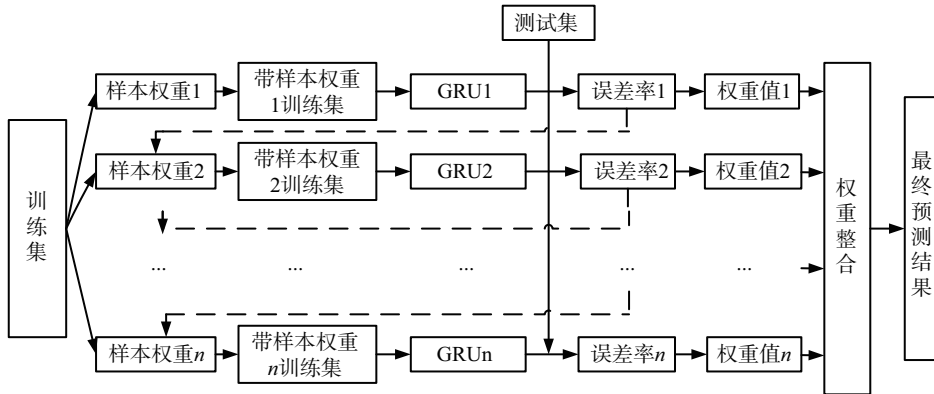


图 2 Boosting - GRU 模型结构

Boosting - GRU 模型的结构如图 2 所示,在 Boosting 模型中采用多个 GRU 模型作为基学习器.首先对训练集分配初始样本权重值,将带样本权重值 1 训练集输入到 GRU1 模型中进行训练.然后采用测试集进行模型的预测,计算预测的误差率并以此来更新样本权重 2 和计算当前基学习器的权重值 1.最后采用带样本权重值 2 训练集训练第二个基学习器,依次串联的训练所有的基学习器。

## 2 卷烟产品投放策略生成模型

产品投放策略生成模型如图 3 所示.为精确的预测投放值序列(即下一年 48 轮投放值组成的序列),卷烟产投放策略生成模型采用 3 种不同的方法来计算投放值.第 1 种方法首先采用多元预测模型得到产品销售量序列,然后通过销售量与投放值之间的关系来计算投放值序列 1;第 2 种方法直接采用多元预测模型来计算投放值序列 2;第 3 种方法采用多分类模型进行海量数据的训练并预测投放值序列 3,最后选择 3 种序列中误差值最小的序列为最终序列。

### 2.1 数据处理

数据处理将历史销售数据处理为模型所需的输入数据,历史销量数据为 A 卷烟产品的所有零售户的订购信息,其中包括时间、档位、投放值、订购量(销售量)、定足率、订单满足率共 6 条信息.首先将数据按不同的档位划分为 1 ~ 30 份数据,该数据可以用于训练 GRU 多分类模型.然后在档位划分的基础之上按时间进一步处理,得到数据包括总户数、订购户数、销售量(累加)、投放值、定足率(均值)、订单满足率(均值)、订

购率(计算). 其中定足率、订单满足率、订购率的计算方法如下:

$$\text{定足率} = (\text{订购量} \div \text{投放量}) \times 100\%$$

$$\text{订单满足率} = (\text{客户订购量} \div \text{客户需求量}) \times 100\%$$

$$\text{订购率} = (\text{订购户数} \div \text{总户数}) \times 100\%$$

最后将数据进行监督学习处理来用于训练 Boosting - GRU 多元预测模型.

### 2.2 多元预测模型原理

采用神经网络模型进行序列预测前需要对数据进行监督学习处理<sup>[12]</sup>,设定一定长度的“回溯”值. 例如,投放值序列为 15, 10, 10, 5, 5, 回溯值设置为 1, 则预测模型训练数据为  $x = [15, 10, 10, 5]$ 、 $y = [10, 10, 5, 5]$ ; 回溯值设置为 2, 则预测模型训练数据为  $x = [[15, 10], [10, 10], [10, 5]]$ 、 $y = [10, 5, 5]$ , 即采用前回溯个数据  $x$  来预测后一个数据  $y$ . 多元预测是将每一行的总户数、订购户数、销售量、投放值、定足率、订单满足率、订购率采用归一化处理至 0 ~ 1 之间, 然后将  $x$  作为数据  $y$  为标签输入 Boosting - GRU 模型进行训练, 通过 GRU 模型捕捉这 7 种数据之间的关系和各自数据随时间变化的前后联系, Boosting 算法能降低 GRU 模型预测的偏差, 模型输出的结果为产品销售量序列和投放值序列.

烟草行业常规的投放策略主要按“档位”来制定. 例如 A 产品卷烟在第 1 档位投放 5 条, 第 2 档位投放 6 条, ... 则属于 1 档位的零售户订购上限为 5 条, 属于 2 档位的零售户订购上限为 6 条, 依次类推. 由此来计算投放值的估值, 第  $N$  档位零售户下一轮关于 A 产品销售量  $S_N$  和投放值  $T_N$  之间的计算公式如下:

$$T_N = S_N \div (S_{N\text{总户数}} \times R_{N\text{订购率}}), 1 \leq N \leq 30. \tag{5}$$

其中, 销售量  $S_N$  通过模型预测得到;  $S_{N\text{总户数}}$  为第  $N$  档位零售户数量;  $R_{N\text{订购率}}$  为第  $N$  档位的订购率.

由于投放值只能为整数, 采用公式计算和模型预测的投放值会存在小数的问题, 模型采用四舍五入函数进行取整.

### 2.3 多分类模型原理

多分类模型采用有监督的学习方式, 数据集为同档位内的零售户所有信息, 将投放值作为标签  $y$ , 其它相关数据为  $x$ . 首先统计投放值有多少类值来确定模型属于几分类任务, 如果投放值设置为 1 ~ 15 那么模型设置为 15 分类. 然后将数据转化为计算机便于计算的形式, 即  $x$  数据归一化至 0 ~ 1 之间, 标签  $y$  数据采用独立热编码表示. 最后将  $(x, y)$  输入到 GRU 模型中进行训练, 模型会根据标签  $y$  来挖掘  $x$  数据之间的关联, 并通过全连接层和 Softmax 层计算投放值类别的概率. 由于模型在测试集上预测的不确定性, 同时还需要对预测的结果进行筛选来确定最适合该时间段的投放值, 筛选原则在某时间段预测结果类别数量最多则为该时间的投放值.

## 3 实验验证

### 3.1 数据集

实验的数据来源湖北省烟草公司十堰公司 2016 年 1 月 1 日至 2021 年 12 月 30 日的 2 630 万条历史销售数据, 为实现对具体产品不同档位零售户的投放预测, 选取 191 万条红金龙软装精品为实验对象. 对实验数据按照档位处理得到 30 份平均 6.4 万条的历史销售数据, 卷烟产品按周投放, 不包括假期每年共 48 轮投放, 按照时间段处理得到 288 条累计数据. 采用 2016 - 2020 年时间段内的数据为训练集, 2021 年时间段内的数据为测试集, 在 2021 年数据上验证投放策略生成模型的性能. 卷烟产品投放策略生成模型的各部分输入输出和数据集划分如表 1、表 2 所示.

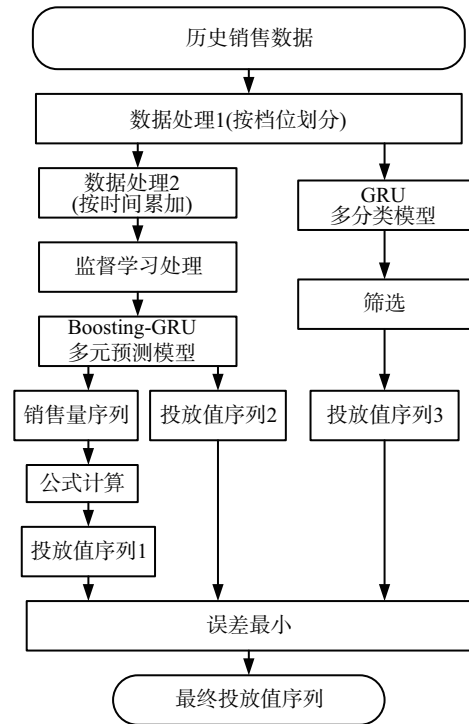


图 3 卷烟产品投放策略生成模型

### 3.2 实验设置及评价指标

实验在 Windows10 上进行,CPU 为 Intel(R) Core(TM) i9,3.10GHz,GPU 为 RTX3060 12G,编程语言为 Python3.7,深度学习框架为 Tensorflow 2.4.1 和 Keras 2.4.3.

实验中模型的参数设置:GRU 中隐藏单元数为 256;Boosting 算法中 GRU 模型的个数为 10;迭代次数为 20 次;批次大小设置为 32;学习率为  $1e-5$ ;优化器 Adam;为防止过拟合 dropout 设置为 0.2;预测模型的损失函数为均方根误差;分类模型的损失函数为交叉熵代价函数.

实验选用准确率和平均绝对误差(MAE)作为投放策略生成模型的衡量指标,其计算过程如下式(6)-(7)所示.

$$\text{准确率} = \left(1 - \frac{|\text{实际销售量} - \text{预测销售量}|}{\text{预测销售量}}\right) \times 100\%. \quad (6)$$

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^n |\hat{y}_i - y_i|. \quad (7)$$

其中,  $n$  为预测序列的长度 48;  $y_i$  表示实际序列值;  $\hat{y}_i$  表示预测的序列值.

### 3.3 销售量预测

销售量预测通过多元预测模型来实现,为了验证 Boosting-GRU 多元预测模型的准确性,实验同时采用支持向量回归(SVR)、逻辑回归(LR)、决策树(DT)、随机森林(RF)、季节性差分回归移动平均模型(SARIMA)、GRU 神经网络模型共 6 种回归预测方法作为对比.计算几种方法在 1~30 档位零售户的历史销售数据上预测销售量的准确率并取平均值,各种方法预测的平均准确率对比如下表 3 所示.为便于直观的观测模型预测的性能,列举多元预测模型在 1 档、15 档和 25 档零售户上红金龙软装精品卷烟 2021 年的销售量预测如下图 4 所示.

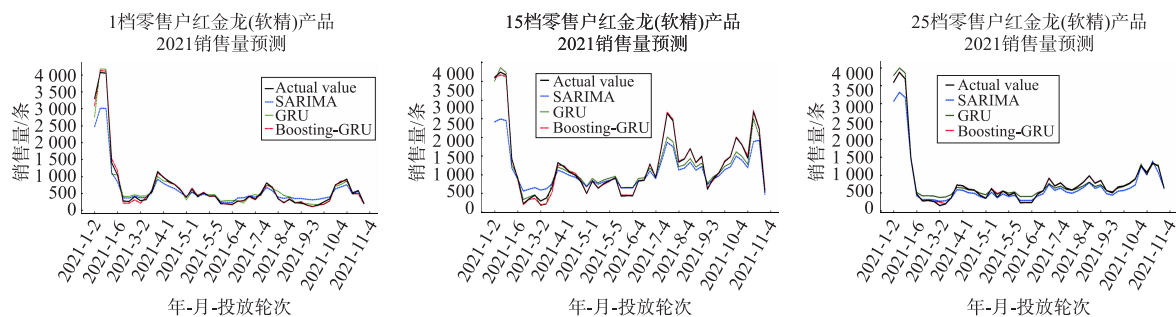


图 4 多元预测模型的预测对比

由表 3 可知,Boosting-GRU 模型在不同档位数据上的预测相比其它预测方法预测效果更优,平均预测准确率达到 97.64%,对比另外 6 种预测方法的平均准确率最低提升 1.67%,最高提升 5.44%.由图 4 观测可知,在列举的 1 档、15 档和 25 档零售户档位上预测的效果均为:Boosting-GRU > GRU > SARIMA,神经网络的预测结果要优于传统预测模型,而结合 Boosting 算法的神经网络模型能通过不断调整样本权重来降低预测结果的误差,从而提高整体模型的预测能力.

### 3.4 投放策略生成模型验证

卷烟产品投放策略生成模型在十堰市 2021 年红金龙软装精品的 1~30 档位零售户的历史销售数据上

表 1 模型各部分的输入输出

方法	输入	输出
多元预测模型	总户数、订购户数、销售量、投放值、定足率、订单满足率、订购率	销售量序列、投放值序列 2
多分类模型	订购量、定足率、订单满足率、投放值	投放值序列 3
销售量至投放值	销售量、总户数、订购户数	投放值序列 1

表 2 模型各部分的数据集划分

模型	训练集	测试集
多元预测模型	240 条	48 条
多分类模型	平均 5.3 万条	平均 1.1 万条

表 3 各种预测方法的平均准确率对比

预测方法	平均准确率/%
SVR	94.77
LR	92.38
DT	92.20
RF	95.97
SARIMA	94.50
GRU	96.75
Boosting-GRU	97.64

进行验证,通过观测模型销售量预测值以及投放值序列的平均绝对误差来验证模型的性能,实验结果如下表 4 所示. 为便于直观的观测投放策略生成模型性能,同时列举模型在 1 档、15 档和 25 档零售户上红金龙软装精品卷烟 2021 年的投放策略生成效果,如下图 5 所示.

表 4 投放策略生成模型验证

档位	实际销量/条	预测销量/条	平均绝对误差/%			
			投放值序列 1	投放值序列 2	投放值序列 3	最终序列
1 档	30 441	30 537	22.92	18.75	20.84	18.75
2 档	10 190	10 201	8.33	6.25	7.14	6.25
3 档	12 747	12 790	12.50	10.42	10.42	10.42
4 档	14 429	14 380	16.67	16.67	18.75	16.67
5 档	15 921	16 002	20.84	18.75	16.67	16.67
6 档	21 949	22 030	14.58	12.50	12.50	12.50
7 档	24 900	24 982	22.92	25.00	18.75	18.75
8 档	33 390	33 480	18.75	16.67	14.58	14.58
9 档	38 741	38 820	17.08	16.67	12.50	12.50
10 档	42 115	42 025	12.50	10.42	5.63	5.63
11 档	38 180	38 145	16.67	12.50	13.58	12.50
12 档	42 697	42 750	14.58	8.33	10.42	8.33
13 档	57 955	57 865	22.92	20.84	18.75	18.75
14 档	60 665	60 738	12.50	6.25	4.17	4.17
15 档	60 860	60 787	10.42	6.25	10.42	6.25
16 档	58 109	58 170	16.67	12.50	14.58	12.50
17 档	59 426	59 500	14.58	14.58	14.58	14.58
18 档	56 456	56 382	18.75	18.75	16.67	16.67
19 档	44 280	44 195	22.92	18.75	25.00	18.75
20 档	48 759	48 681	20.84	16.67	14.58	14.58
21 档	44 167	44 230	20.84	18.75	12.50	12.50
22 档	44 097	44 127	12.50	10.42	5.63	5.63
23 档	41 696	41 759	18.75	16.67	14.58	15.58
24 档	32 279	32 315	8.75	5.48	4.63	4.63
25 档	32 036	32 116	22.92	18.75	16.67	16.67
26 档	21 864	21 927	16.67	14.58	16.67	14.58
27 档	21 041	21 139	25.00	22.92	18.75	18.75
28 档	20 051	20 108	12.50	10.42	12.50	10.42
29 档	16 890	16 824	20.84	16.67	12.50	12.50
30 档	18 939	18 996	12.50	8.75	10.42	8.75

注:实际销售量、预测销售量为 2021 年该产品的总销售量;投放值序列为预测 2021 年 48 轮的序列.

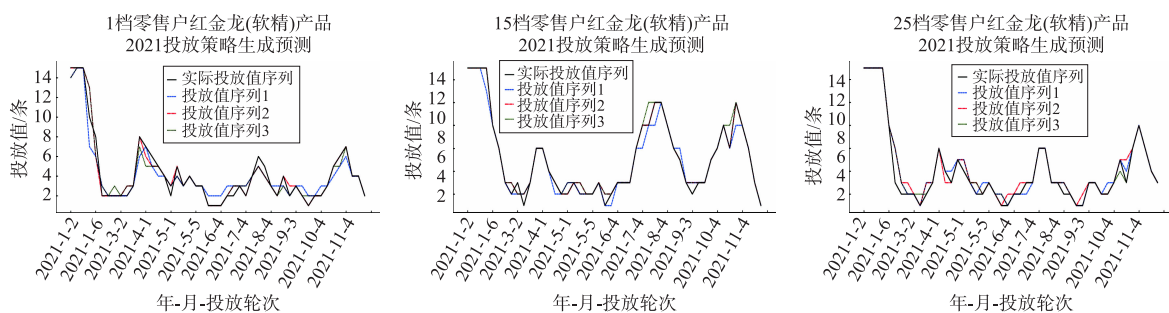


图 5 投放策略生成模型预测效果

由表 4 对比计算可知,实际销量和预测的销量的差值均在 100 条以内. 通过对比各投放值序列的误差,最终序列的平均绝对误差均低于 20%. 分析原因,最终序列是选择 3 种投放值序列误差最小的序列,因此最终序列更符合卷烟产品的投放策略. 由图 5 观测可知,投放策略生成模型分别采用 3 种不同的方法计算投放值序列,在 1 档和 15 档零售户数据上预测效果,投放值序列 2 要优于投放值序列 1 和投放值序列 3,模型最

终会选择投放值序列 2 作为最终投放策略;在 25 档零售户数据集上预测的效果,投放值序列 3 优于投放值序列 2 和投放值序列 1,即模型最终会选择投放值序列 3 作为最终投放策略。

#### 4 结语

为科学制定卷烟产品投放策略,结合 GRU 神经网络模型和 Boosting 集成学习算法各自的优点,提出一种混合神经网络模型的卷烟投放预测方法。该模型采用 Boosting - GRU 多元预测模型、产品销量至投放值计算以及 GRU 多分类模型 3 种方法来分别计算卷烟产品投放值序列,并选择误差最小为最终的投放值序列。在湖北省十堰市卷烟产品的历史销售数据上进行实验,实验结果表明,该模型能很好的完成卷烟投放策略生成任务。

#### 参考文献:

- [1] 王伟民,汪沅,张国安. 基于灰色马尔科夫模型的全国卷烟需求预测研究[J]. 中国烟草学报,2009,15(6):66-69.
- [2] 王诗豪,张晓妮,张云,等. 铜川市卷烟需求集成预测[J]. 中国烟草学报,2019,25(6):105-109.
- [3] 武牧,林慧苹,李素科,等. 一种基于支持向量机的卷烟销量预测方法[J]. 烟草科技,2016,49(2):87-91.
- [4] BREIMAN L. Bagging predictors[J]. Machine learning,1996,24(2):123-140.
- [5] FREUND Y, IYER R, SCHAPIRE R E, et al. An efficient boosting algorithm for combining preferences[J]. Journal of Machine Learning Research,2003,4(11):933-969.
- [6] PARK J H, HWANG J H, LEE C S, et al. Stacking sequence design of composite laminates for maximum strength using genetic algorithms[J]. Composite Structures,2001,52(2):217-231.
- [7] 韩伟民,杨应广. 卷烟创新产品工商交易预测研究[J]. 中国烟草学报,2022,28(1):115-122.
- [8] YU Y, SI X, HU C, et al. A review of recurrent neural networks: LSTM cells and network architectures[J]. Neural Computation, 2019,31(7):1235-1270.
- [9] CHO K, VAN M B, GULCEHRE C, et al. Learning phrase representations using RNN encoder - decoder for statistical machine translation[EB/OL]. (2014-6-3)[2014-9-3]. <http://arxiv.org/abs/1406.1078>.
- [10] 邓超,刘颂,王露笛,等. 基于深度神经网络的卷烟智能投放模型构建方法[J]. 烟草科技,2021,54(2):78-83.
- [11] 张文生,于廷照. Boosting 算法理论与应用研究[J]. 中国科学技术大学学报,2016,46(3):222-230.
- [12] 左亚尧,王铭锋,洪嘉伟,等. 一种基于集成学习的多元时间序列预测方法[J]. 小型微型计算机系统,2020,41(12):2475-247.

### Forecast method of cigarette release based on hybrid neural network model

ZOU Wang<sup>1</sup>, ZHANG Wu-bo<sup>1,2</sup>, MA Xun-zheng<sup>1,3</sup>

(1. School of Electrical and Information Engineering, Hubei University of Automotive Technology, Shiyan 442002, China;

2. School of Computer Science and Technology, Wuhan University of Science and Technology, Wuhan 430081, China;

3. Hubei Provincial Tobacco Company Shiyan City Corporation, Shiyan 442000, China)

**Abstract:** In order to scientifically formulate the cigarette product release strategy, a hybrid neural network model based cigarette release forecast method is proposed. By combining three methods: Boosting - GRU multivariate forecast model, product sales to release value calculation, and GRU multi - category model to achieve the generation of release strategies for different grades of cigarette products. The proposed method was validated by using historical sales data of cigarette products in Shiyan City, Hubei Province as the experimental object. The experimental results show that the method achieves an average accuracy of 97.64% in product sales volume forecast and the mean absolute error in release strategy generation were all below 20%.

**Key words:** multivariate prediction models; multiple classification models; mean absolute error

(责任编辑 段 鹏)