

doi: 10.19920/j.cnki.jmsc.2026.06.005

# 基于多样性正则化 Adaboost 的动态选择集成预测方法<sup>①</sup>

成 晟<sup>1,2</sup>, 冯 晗<sup>1,2</sup>, 张逸飞<sup>1,2</sup>, 张 晓<sup>1,2</sup>, 王 珏<sup>1,2\*</sup>, 汪寿阳<sup>1,2</sup>

(1. 中国科学院数学与系统科学研究院, 北京 100190; 2. 中国科学院大学经济与管理学院, 北京 100190)

**摘要:** 集成预测中模型多样性与模型预测精度的权衡问题是当前的研究热点之一。针对集成预测中存在的过拟合、预测精度低等问题, 本研究提出了一种基于多样性正则化的两阶段动态选择集成策略。第一阶段, 通过定义一种权衡多样性与预测精度的损失函数, 构建了全新的集成预测多样性正则化策略; 第二阶段, 提出一种动态选择集成方法, 能够有效筛选兼顾多样性与预测精度的候选预测模型。实验结果表明, 所提出的两阶段选择集成方法在有效增加模型多样性的情况下显著提高预测精度, 从而较好地提升集成预测模型的泛化性能。

**关键词:** 多样性正则化; 动态选择; Adaboost; 集成预测

**中图分类号:** C934 **文献标识码:** A **文章编号:** 1007-9807(2026)06-0063-12

## 0 引 言

集成学习作为机器学习领域中的热点研究方向, 其核心目标是通过合理组合多个基模型来获得更好的泛化性能与预测可靠性, 目前已经在图像分类<sup>[1]</sup>、预测、模式识别等多个领域中广泛应用。在集成学习中, 多样化基模型生成方法与最优组合策略始终是两大核心研究重点<sup>[2]</sup>。

模型的预测精度与模型多样性, 是决定集成模型整体预测性能<sup>[3-5]</sup>的两个关键因素, 二者相辅相成、缺一不可。其中, 多样性作为评估基模型之间差异程度的重要分歧度量指标<sup>[6]</sup>, 其核心价值在于为集成性能提升提供基础。根据 Hu 等人<sup>[7]</sup>研究, 多样化的基模型能够产生不相关的误差, 有利于提升预测精度。集成学习可以有效地利用这种多样性, 进而在不显著增加偏差<sup>[8,9]</sup>的情况下有效降低预测方差。尽管 Bagging<sup>[10]</sup>和 Adaboost<sup>[11]</sup>等经典集成方法在实际应用中能够在一定程度上增加基模型的多样性, 但这些方法均未将多样性作为显式的优化目标, 也未主动寻求多

样性的最大化, 因此在复杂数据场景下的性能提升存在局限。

需要明确的是, 多样性虽为提高集成预测性能的必要条件, 但其最大化未必能实现最优泛化性能<sup>[4,12]</sup>。已有大量研究和实验表明<sup>[6]</sup>, 集成学习的准确性与多样性之间并不存在严格的强相关性, 过度追求多样性反而会降低模型的精度<sup>[13]</sup>。这也意味着, 集成学习的核心竞争力仍依赖于基模型自身的准确性<sup>[14]</sup>, 多样性的提升必须建立在保证基模型精度的基础之上。因此, 如何在准确性与多样性之间建立科学合理的权衡机制, 实现二者的协同优化, 成为进一步提升集成学习性能的关键突破口。目前国内外学者已围绕这一核心问题展开了大量探索<sup>[15-17]</sup>, 包括 Mao 等人<sup>[18]</sup>提出的端到端加权集成学习框架, 该框架通过构建一个联合优化模型, 利用个体分类器之间的相关性及其相应的权重, 以实现多样性与精度之间的平衡。Giacinto 和 Roli<sup>[19]</sup>面向图像分类任务, 锚定基模型间的误差独立性(即分类多样性), 围绕精度与多样性的协同优化提出一套自动化的神经网络

① 收稿日期: 2022-09-29; 修订日期: 2024-02-25。

基金项目: 国家自然科学基金资助项目(72271229; 71771208; 71988101)。

通讯作者: 王 珏(1978—), 女, 山东青岛人, 博士, 研究员, 博士生导师。Email: wjue@amss.ac.cn

集成方法。

上述研究均以在预测精度与多样性之间实现平衡为核心目标,进而缓解集成模型的过拟合问题。然而,这些研究通常仅在模型生成过程完成后,才将多样性作为优化准则之一,未能在模型迭代生成阶段实现精度与多样性的动态协同优化。本研究则在模型生成阶段即将多样性纳入考量,以权衡多样性与预测精度为核心准则,依次生成多样性更优的模型,并在后续选择集成阶段进一步提升集成多样性。简言之,本研究提出一种基于预测精度与多样性权衡的两阶段集成学习方法。第一阶段为基模型生成阶段,针对回归任务定义了双误度量测度,并将其用于多样性正则化。该阶段的优势在于,迭代生成基模型时能够关注到多个模型均难以准确学习的样本,从而在 Adaboost 算法中同时兼顾模型精度与多样性,生成具有适当多样性的基模型组。此阶段在保证基模型具备一定预测精度的同时,最大化模型之间的多样性,为实现最优的集成学习奠定基础。第二阶段为模型筛选阶段,一方面剔除 Adaboost 算法中因过度关注难以准确预测样本而产生较大偏差的基模型;另一方面筛选出因提升多样性而导致预测偏差较大的基模型。在此基础上动态保留高多样性基模型,最终筛选出兼具高多样性与优良预测精度的模型子集。

本研究的边际贡献主要体现在以下四个方面。

1) 本研究拓展了多样性在回归任务中的应用,定义了适用于回归任务的双误度量测度。该测度可依据预测精度与预测偏差方向两项准则,衡量回归任务中难以准确预测的样本占比,为后续模型多样性评估奠定基础;

2) 本研究构建了一种全新的兼顾多样性与预测精度的目标损失函数。通过设置多样性正则化因子,实现对模型多样性与损失的差异化权重配置,有效改善了传统 Adaboost 算法生成基模型易同质化的问题;

3) 本研究设计了一种新型动态选择集成策略,实现对基模型的筛选与集成权重的重新分配。通过权衡基模型多样性与集成模型的预测精度,动态选取基模型以进一步提升集成性能;

4) 在多个数据集上开展实验,分别采用权衡

多样性与预测精度的损失函数生成候选模型,并进行动态选择集成预测。实验结果均表明,所提出的损失函数与选择集成策略能显著提高集成模型的多样性与预测精度,验证了该方法的有效性与鲁棒性。

## 1 模型框架介绍

### 1.1 Adaboost 回归方法

Adaboost 算法是一种经典的集成学习方法。该方法通过迭代生成多个基学习器,每个基学习器会更加关注上一个基学习器预测偏差较大的样本<sup>[20]</sup>,之后将多个基学习器进行组合,从而构建具有较高精度的强学习器。在回归任务中,第  $t$  个模型在第  $j$  个样本点的预测偏差定义如式(1)所示

$$Bias_t(j) = \frac{|y_t(j) - \hat{y}_t(j)|}{\max_j |y_t(j) - \hat{y}_t(j)|} \times 100\% \quad (1)$$

其中  $y_t(j)$  表示第  $t$  个模型在第  $j$  个样本点的预测值,  $\hat{y}_t(j)$  表示集成模型在第  $j$  个样本点的预测值。

Adaboost 算法认为弱学习器与强学习器之间存在等价关系,即一个强学习器可以由多个弱学习器线性表示。当数据呈现高度非线性特征而难以直接预测,导致单个强学习器难以有效拟合时,可通过训练多个基学习器并将其集成,构建出性能更优的强学习器以完成预测任务。面向回归任务的 Adaboost 算法具体实现步骤如下。

1) 输入: 训练数据集  $X$ , 基学习器算法  $h(\cdot)$ , 训练次数  $T$ 。

2) 权重初始化: 设置初始训练集权重为平均权重  $D_1(x) = \frac{1}{m}$ 。

3) 循环: 对于每个训练轮次  $t = 1, 2, \dots, T$ 。

① 使用权重  $D_t$  训练第  $t$  个基学习器  $h_t = h(D_t(x))$ 。

② 计算第  $t$  个模型偏差  $Bias_t(j) = \frac{|y_t(j) - \hat{y}_t(j)|}{\max_j |y_t(j) - \hat{y}_t(j)|} \times 100\%$ 。

$$\textcircled{3} \text{ 计算误差 } e_t = \sum_{j=1}^N w_{t-1}(j) \times \text{Bias}_t(j) .$$

$$\textcircled{4} \text{ 计算 } \beta_t = \frac{e_t}{1 - e_t} .$$

$$\textcircled{5} \text{ 更新训练集权重 } D_{t+1}(j) = D_t(j) \times \beta_t \times \text{Bias}_t(j) .$$

$$\textcircled{6} \text{ 计算第 } t \text{ 个模型的权重 } \alpha_t = \ln \frac{1}{\beta_t} .$$

4) 输出: 完成所有轮次后, 返回集成预测

$$H(x) = \sum_{t=1}^T \alpha_t h_t .$$

## 1.2 两阶段 Adaboost 选择集成策略

### 1.2.1 基于多样性正则化的 Adaboost 方法

在采用 Adaboost 算法进行回归类集成学习时, 往往存在一定的局限性. 算法在迭代过程中会驱使基模型倾向于关注上一轮迭代中预测误差较大的样本, 但此类样本误差偏高并非完全源于其本身难以预测, 也可能由基学习器过拟合或欠拟合等因素导致. 因此, 在回归任务中, Adaboost 易使当前基模型过度关注前序基学习器的预测残差, 进而降低模型整体稳定性. 回归任务可选用决策树、逻辑回归、神经网络等作为基学习器. 以神经网络为例, 由于其训练过程存在随机性, 单次训练往往难以充分捕捉高误差样本的有效信息, 同时对预测精度较高的样本也易出现学习不充分的问题, 致使以神经网络为基模型的 Adaboost 回归算法常面临算法收敛困难的状况.

为应对上述问题, 本研究旨在进一步提升基模型多样性, 同时使各基模型聚焦于多轮迭代中均难以拟合的样本, 而非仅关注上一轮拟合效果较差的样本. 为此, 借鉴分类任务中双误度量<sup>[20]</sup>思想, 从上述两方面优化模型性能, 提升基模型间的多样性. 双误度量是衡量两个分类模型共同分类错误占比的指标. 该指标值越大, 表明两个模型共同错分样本占比越高, 模型间多样性越低; 反之, 双误度量值越小, 共同错分样本占比越低, 模型多样性则越优. 在 Adaboost 算法迭代过程中引入双误度量, 能够在最大化模型多样性的同时, 精准捕捉多基学习器难以预测的样本, 从而提升预测的泛化能力. 双误度量的计算如下

$$DF = \frac{d}{m} \quad (2)$$

其中  $d$  表示两个学习器共同预测错误的样本个数,  $m$  表示样本个数. 但在回归任务中, 每个样本点的预测结果均与真实值之间存在偏差, 且预测偏差同时包含幅值与方向两方面信息, 难以直接采用传统双误度量进行刻画. 为此, 本研究对回归任务重新定义双误度量测度, 用以衡量同方向难以准确预测的样本占比. 为便于表述, 将该多样性测度指标称为 SDDP (same direction difficult to predict measure), 具体形式为

$$SDDP = \frac{dp}{m} \quad (3)$$

其中  $dp$  衡量回归任务难以准确预测的样本个数. 首先定义第  $t$  个模型对第  $i$  个样本的难以准确预测度量  $dp_{t,i}$  为

$$dp_{t,i} = \frac{y_i - \hat{y}_{t,i}}{y_i} \times 100\% \quad (4)$$

其中  $y_i$  表示第  $i$  个样本的真实值,  $\hat{y}_{t,i}$  表示第  $t$  个模型对第  $i$  个样本的预测值. 进而定义

$$dp_t = \sum_{i=1}^m dp_{t,i} \quad (5)$$

其中  $dp_t$  用以衡量第  $t$  个模型是否被准确预测<sup>[21]</sup>

$$dp_i = \begin{cases} 1, & dp_{t-1,i} dp_{t,i} > \delta \text{ 或 } dp_{t-1,i} dp_{t,i} < -\delta \\ 0, & \text{其它} \end{cases} \quad (6)$$

其中  $\delta$  无固定取值, 可根据数据集特性与任务需求调整, 本研究中统一设置为 0.8.

基模型多样性是集成学习成功的关键, 但不同基模型之间多样性的增加必然会导致整个学习器的平均精度降低, 反之亦然<sup>[22, 23]</sup>. 当找到二者的最优平衡点时<sup>[24]</sup>, 可显著提升集成模型的预测性能. 基于此, 本研究构造了多样性正则化损失函数, 用于权衡预测精度与多样性. 该损失函数不仅包含了本轮训练的损失项, 同时考虑降低前两轮模型计算得到的 SDDP 值, 即前两轮模型均难以准确预测且预测误差方向相同的样本信息, 从而使后续迭代过程能够聚焦于多轮训练中均难以准确预测的样本. 上述多样性正则化损失函数表示为

$$\text{DivLoss}_k = w \times SDDP_{k-2} + (1-w) \times MAE_k \quad (7)$$

其中  $SDDP_{k-2}$  表示第  $k-2$  次迭代与第  $k-1$  次迭代生成模型之间的多样性,  $w$  是权衡训练误差和多样性正则项的正则化参数, 其取值为前  $k-2$  轮迭代生成 SDDP 值的平均值.

基于上述定义的多样性正则化损失函数,本研究设计了基于多样性的 Adaboost 改进算法.为便于后文表述,将该算法称为 D-Adaboost 算法 (Adaboost based on Diversity),其主要实现流程如下.

1) 输入: 训练数据集  $X$ , 基学习器算法  $h(\cdot)$ , 预设训练次数  $T$ .

2) 权重初始化: 设置初始训练集权重为平均权重  $D_1(x) = \frac{1}{m}$ , 即初始权重为平均分配.

3) 循环: 对于每个训练轮次  $t = 1, 2, \dots, T$ .

① 基于当前样本权重  $D_t$ , 训练第  $t$  个基学习器  $h_t = h(D_t(x))$ .

② 计算第  $t$  个模型在当前样本权重下的预测偏差  $Bias_t(j) = \frac{|y_t(j) - \hat{y}_t(j)|}{\max_j |y_t(j) - \hat{y}_t(j)|} \times 100\%$ .

③ 计算误差  $e_t = \sum_{j=1}^N w_{t-1}(j) \times Bias_t(j)$ .

④ 计算  $\beta_t = \frac{e_t}{1 - e_t}$ .

⑤ 更新训练集权重  $D_{t+1}(j) = D_t(j) \times \beta_t \times Bias_t(j)$ .

⑥ 计算第  $t$  个模型的权重  $\alpha_t = \ln \frac{1}{\beta_t}$ .

⑦ 当  $t > 3$  时, 计算  $SDDP_{t-1}$ .

⑧ 更新第  $t + 1$  个模型的损失函数为  $DivLoss_{t+1}$ .

4) 输出: 完成所有轮次训练后, 返回集成预测  $H(x) = \sum_{i=1}^T \alpha_i h_i$ .

### 1.2.2 动态选择集成策略

采用 Adaboost 进行预测时, 基模型个数的确定依赖于迭代次数, 而迭代次数的合理选择往往存在困难. 基模型个数过少容易导致精度不足、多样性单一等问题; 而过多则会造成训练复杂度攀升、模型过拟合等问题.

因此, 需综合考量迭代次数对多样性与预测精度的影响: 既要保证生成基模型个数充足, 使集成学习能够充分挖掘样本信息、获得较好的预测精度, 也要确保各基模型之间具有较好的多样性, 避免模型同质化. 然而, 寻找这两者的最优权衡将带来巨大的计算成本, 包括模型训练成本、剪枝成

本和预测过程中的计算成本<sup>[25]</sup>.

选择集成策略可以有效降低上述成本, 且相较于使用全部模型进行集成, 合理的选择集成策略往往能获得最优的预测效果<sup>[26]</sup>, 即通过有效剔除偏差过大的基模型, 可进一步提升集成模型的预测精度. 为此, 本研究设计了基于动态选择集成策略的预测方法, 简称为 DS-Adaboost 算法 (Adaboost based on Dynamic Selection), 将提出的两阶段选择集成方法简称为 DDS-Adaboost 算法 (Adaboost based on Diversity and Dynamic Selection). 算法的具体流程如下所示.

1) 输入: 模型选择阈值  $Diff$ ; 基模型组  $h_k, k = 1, \dots, T$ ; 基模型组集成系数  $\alpha_k, k = 1, \dots, T$ .

2) 初始化: 选择平均绝对百分比误差 (MAPE) 最小的模型  $f_1(x) = h_{\min MAPE}(x)$  作为集成模型组中的首个模型;  $SDDP_1 = 0$ .

3) 循环: 当  $Diff - \sum_{k=1}^K SDDP_k > 0$  时

① 分别计算集成模型组  $f(x)$  与未选入模型组的模型  $h(x)$  的  $SDDP$ , 并记  $SDDP_k = SDDP_{\max}(f_{k-1}(x), h_k(x))$ .

② 从未选入动态选择模型组的模型中, 选取与模型组多样性最大的模型加入,  $f_{k+1}(x) = f_k(x) + f_{\max SDDP}(x)$ .

③ 当  $MAPE_{f_{k+1}} > MAPE_{f_k}$  时, 从模型组中剔除一个能够使得剔除后精度提升最大的模型,  $f_{k+1}(x) = f_{k+1}(x) - h_j(x)$ .

④ 归一化集成模型系数  $\alpha'$ .

4) 输出: 迭代完成后, 输出最终集成预测

$$F(x) = \sum_{i=1}^T \alpha'_i F_i.$$

上述模型选择阈值  $Diff$  定义为如式(8)所示

$$Diff = \frac{|y_t - \hat{y}_t|}{y_t} \times rate \times 100\% \quad (8)$$

其中  $rate$  表示模型选择比率, 为取值范围在  $[0, 1]$  之间的常数, 用于调控  $Diff$  的贡献权重.

综上, DDS-Adaboost 方法的整体框架如图 1 所示. 首先, 输入预测样本与初始权重, 利用 D-Adaboost 算法迭代生成具有较好多样性的基模型; 然后, 将得到的基模型纳入模型池, 通过动态选择集成策略筛选出兼顾多样性与预测精度的基

模型;最后,将筛选后的基模型进行加权组合,得到模型的最终预测输出。

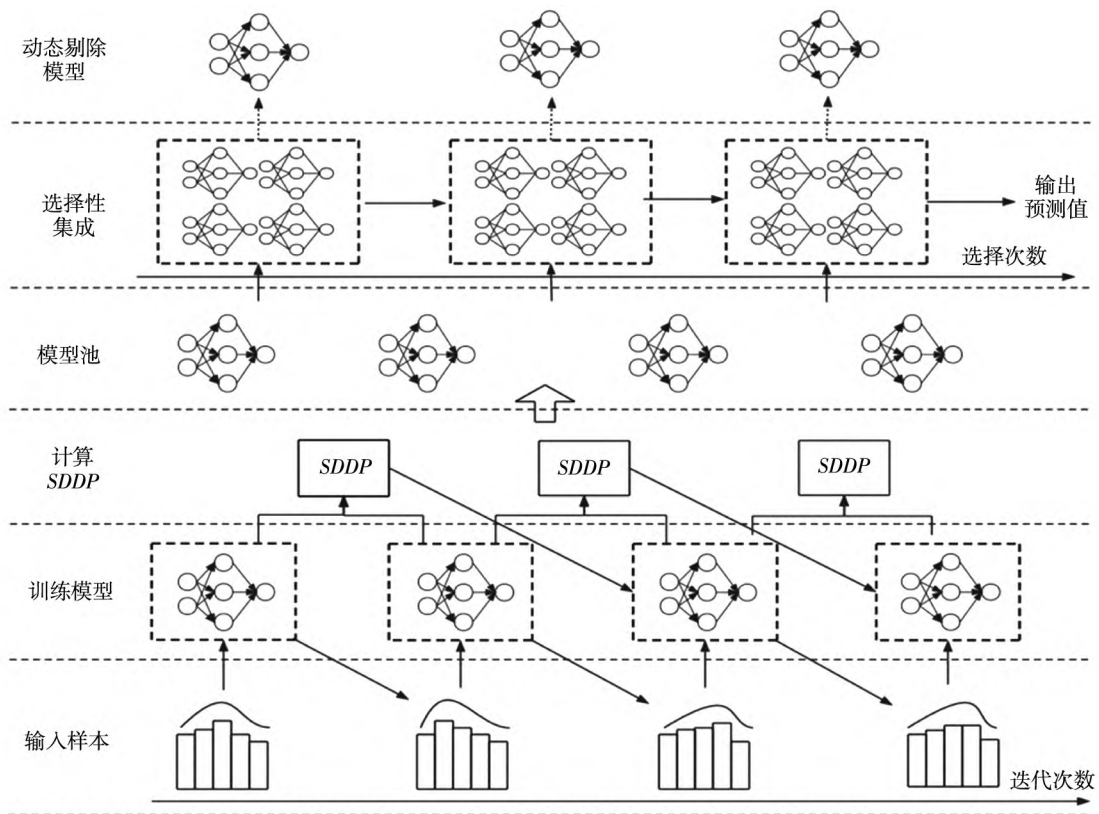


图 1 DDS-Adaboost 集成方法框架  
Fig.1 Framework of DDS-Adaboost

表 1 数据集信息

Table 1 Dataset information

序号	数据集	样本量	变量数
1	Airfoil	1 503	5
2	BostonHousing	506	13
3	ENB	768	8
4	Valuation	414	6
5	Machine	209	7
6	No2	500	7
7	Slump	103	9
8	Toxicity	908	6
9	WTI	363	198

## 2 实验分析

本实验主要分为三部分. 第一部分为基模型多样性分析,通过在 Adaboost 迭代过程中引入 SDDP 策略并考察其对基模型损失函数的影响,验证该策略能否有效提升模型多样性;第二部分聚焦选择集成的有效性分析,探讨本研究所提选择集成方法对提升集成预测精度的实际作用;第三部分以 WTI 原油价格预测数据集为例,详细阐述模型参数设定与统计性检验流程.

为验证所提出的 DDS-Adaboost 选择集成算法在数据集上的适用性,本研究选取 8 组来自 University of California Irvine (UCI) 的公开数据集进行实验.此外,还选择了 WTI 原油价格预测数据集,以评估该方法在高维数据场景下的预测性能,具体数据集信息如表 1 所示.

为保证实验结果具有可比性,以上所有数据集均使用 BP 神经网络(BPNN)作为基模型,网络结构均为单隐含层,包含 20 个隐藏单元.各数据集均按 8:2 划分为训练与测试集,前 80% 的样本作为训练集,后 20% 的样本用于模型效果的验证

与评估. 实验记录各数据集上  $T$  次重复训练的训  
练误差和测试误差均值.

### 2.1 基模型多样性分析

本实验将改进的 Adaboost 方法用于各数据  
集, 验证在基模型的损失函数中加入  $SDDP$  后, 采  
用 Adaboost 算法进行集成预测能否有效增加基  
模型多样性, 防止过拟合现象. 针对回归数据集,  
用  $dp$  来度量基模型与集成模型之间的多样性. 实  
验结果如表 2 所示.

表2 Adaboost 基模型多样性

Table 2 Diversity of Adaboost base models

数据集	Adaboost	D- Adaboost	DS- Adaboost	DDS- Adaboost
Airfoil	0.64	0.88	0.36	0.43
Boston Housing	10.23	10.90	3.67	3.79
ENB	2.76	36.10	0.73	5.20
Valuation	6.43	7.15	2.23	3.68
Machine	14.31	6.92	3.03	7.53
No2	1.70	3.45	0.50	1.47
Slump	4.24	1.84	2.06	2.13
Toxicity	2.82	3.98	0.71	1.80
WTI	1.39	1.46	0.97	1.02
Average	4.95	8.08	1.58	3.01

通过对比 Adaboost 与 D-Adaboost、DS-Ada-  
boost 与 DDS-Adaboost 可知, 在基模型的损失函数  
中引入  $SDDP$  后, 模型平均多样性得到显著提升.  
其中 D-Adaboost 较 Adaboost 的平均多样性提高  
了 63.23%; DDS-Adaboost 较 DS-Adaboost 平均多  
样性上升了 90.51%. 分数据集来看, 与 Adaboost  
相比, D-Adaboost 仅在 Machine 和 Slump 数据  
集上基模型多样性略有下降, 其余七个数据集的  
多样性均有提升. 与 DS-Adaboost 相比, DDS-Ada-  
boost 在所有数据集上的基模型多样性均实现  
提升.

选择集成策略通过剔除偏差较大的基模型实  
现性能优化, 该过程会不可避免地降低基模型之  
间的多样性, 这一规律在两组对照实验中均得到  
验证: 一组为基准 Adaboost 算法与引入选择集成

的 DS-Adaboost 算法, 另一组为 D-Adaboost 算  
法与引入选择集成的 DDS-Adaboost 算法. 从各数  
据集平均结果来看, DS-Adaboost 较 Adaboost 的模型  
平均多样性下降 68.08%; DDS-Adaboost 较  
D-Adaboost 的模型平均多样性下降 62.75%.

综上, 虽然在引入选择集成策略后, DDS-Ad-  
aboost 较 D-Adaboost 多样性有所下降, 但与 DS-  
Adaboost 相比, 在提升预测精度的同时, 仍保持显  
著的整体多样性水平. 这说明通过引入  $SDDP$  之  
后再进行选择集成, 可以更好地保留模型的整体  
多样性, 为权衡预测精度与模型多样性的关系提  
供一种有效的平衡策略.

### 2.2 集成方法有效性分析

表 3 给出了不同集成方法在均方根误差  
(RMSE)、平均绝对误差(MAE)和平均绝对百分  
比误差(MAPE)的对比结果. 从表中可以看出, 在  
本研究采用的 9 个实验数据集上, Adaboost 算  
法的预测精度均优于单一 BPNN 模型, MAPE 平  
均降低 20.78%. 实验结果表明, 相较于单一预测  
模型, 基于 Adaboost 的集成预测方法具有明显  
优势.

与传统 Adaboost 算法相比, D-Adaboost 通  
过强化模型多样性策略进一步提升了预测精度, 在  
全部 9 个数据集上均实现性能改善, MAPE 平  
均降低 18.41%. 同时, DS-Adaboost 采用选择集  
成策略亦能有效降低预测误差, 除 ENB 数据集外,  
在其余 8 个数据集上预测精度均有提升, MAPE  
平均降低 3.40%. 将 DDS-Adaboost 与传统的 Ad-  
aboost 方法对比可知, DDS-Adaboost 在三种评  
价指标上均表现最优: RMSE 平均降低 14.76%,  
MAE 平均降低 18.54%, MAPE 平均降低  
25.44%. 由此说明, 在 Adaboost 集成框架中引入  
 $SDDP$  并结合选择性集成, 能够进一步提升预测  
性能.

此外, 对比选择集成与多样性增强两类策略  
可以发现, D-Adaboost 在 7 个数据集上的表现均  
优于 DS-Adaboost. 这表明, 在本实验条件下, 提升  
模型多样性比选择策略对改善集成预测性能的作  
用更为关键和显著.

表3 各集成策略预测表现

Table 3 Prediction performance of each ensemble strategy

模型	评价指标	Airfoil	BH	ENB	Valuation	Machine	No2	Slump	Toxicity	WTI
BPNN	RMSE	5.64	4.25	2.55	8.06	51.36	0.55	3.84	0.94	9.85
	MAE	4.48	3.29	2.07	5.89	29.24	0.42	2.90	0.72	8.20
	MAPE	3.58	24.47	9.78	15.12	28.71	13.91	7.65	21.28	10.30
Adaboost	RMSE	4.80	3.51	2.21	7.10	37.51	0.52	2.43	0.82	8.84
	MAE	3.79	2.72	1.63	5.36	17.83	0.39	1.93	0.61	7.19
	MAPE	2.97	18.51	6.85	14.65	14.99	12.81	5.25	17.94	9.29
DS-Adaboost	RMSE	4.70	3.40	2.49	7.11	37.47	0.52	2.53	0.81	8.32
	MAE	3.78	2.75	1.76	5.15	17.56	0.39	2.01	0.60	6.84
	MAPE	2.96	15.26	7.37	13.67	14.25	12.78	5.22	17.77	8.63
D-Adaboost	RMSE	4.63	3.36	1.76	7.12	36.79	0.51	2.35	0.82	7.86
	MAE	3.63	2.70	0.92	5.17	14.57	0.39	1.70	0.59	6.07
	MAPE	2.84	14.65	3.76	13.86	7.38	12.61	4.14	17.11	8.19
Median	RMSE	4.38	3.39	1.40	6.60	29.69	0.52	1.74	0.82	7.99
	MAE	3.39	2.48	1.00	4.90	15.29	0.40	1.39	0.59	6.44
	MAPE	2.65	16.39	3.71	13.45	7.39	11.92	4.17	17.19	8.49
Average	RMSE	4.37	3.43	1.46	6.97	29.61	0.51	1.68	0.82	8.35
	MAE	3.35	2.60	0.98	4.97	14.88	0.40	1.30	0.59	6.74
	MAPE	2.63	19.75	3.81	13.44	7.04	11.87	3.88	17.17	8.57
DDS-Adaboost	RMSE	4.37	3.35	1.61	6.99	28.49	0.51	1.47	0.78	7.05
	MAE	3.34	2.36	0.88	5.19	13.25	0.38	1.29	0.57	5.42
	MAPE	2.61	13.03	3.64	13.37	6.76	11.71	3.33	16.52	7.09

此外,通过对比 DDS-Adaboost 与中位数(Median)权重、平均(Average)权重两种集成方法可知,本研究提出的方法在大部分评估指标和数据集上优于这两种集成策略,进一步说明基于多样性与预测精度权衡的新方法具有更优的预测性能和更稳健的泛化能力。

### 2.3 参数设定及检验——以 WTI 原油价格数据集为例

本节以 WTI 原油现货价格数据集为例,介绍模型参数设定与统计性检验过程。本研究所用 WTI 原油价格为跨度 30 年的月度数据,共 363 个样本点,涵盖 198 个与原油市场相关的变量,包括原油库存、流量与成本、金融市场指数、宏观经济指数等,相关变量选取参考 Wang、Zhao<sup>[27,28]</sup> 等人构建的数据集。

选取上述变量模拟油价系列主要基于两点考量:其一,各变量与油价密切相关,分别代表油价形成的不同驱动因素;其二,油价系列与这些因素之间呈现出噪声显著、关系不稳定和非线性的特

征<sup>[29]</sup>。尽管如此,单一变量仍可能在特定时点为油价走势提供有效信息<sup>[30]</sup>。因此,通过纳入尽可能丰富的变量以充分挖掘信息,可更好地体现所提方法在高维数据建模中的优势。

为缓解 WTI 原油现货价格数据集中高维特征可能引发的过拟合问题,首先通过特征选择方法对变量进行筛选。本研究采用 TABU 特征选择方法<sup>[31]</sup>筛选得到与原油价格最相关的变量,包括:欧佩克国家进口原油的平均上岸成本、安哥拉进口原油的平均上岸成本、利比亚原油产量、工业部门消耗的化石燃料总量、运输部门消耗的石油总量、废弃物能源消耗、风能消耗产量、木材能源消耗、钢铁产品产能、石油和天然气开采产能、城市居民消费价格指数、道琼斯指数、有效联邦基金利率、伦敦黄金市场黄金定价等。

利用上述提出的 DDS-Adaboost 方法,在基预测器个数为 50 的条件下进行选择集成。图 2 直观展示了所提出方法的预测效果,基于 DDS-Adaboost 集成预测模型在原油价格预测任务中具有

显著有效性.

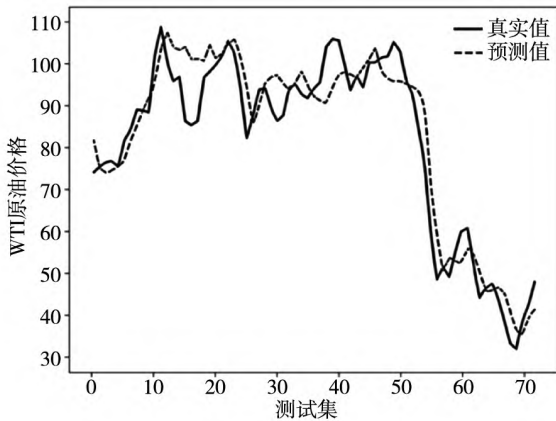


图2 WTI原油价格预测结果

Fig.2 Prediction results of WTI crude oil prices

从图2还可以看出,WTI原油价格在实验期间出现了一次大幅下跌,半年内跌幅高达50%,地缘政治冲突也未能遏制这轮油价暴跌.此轮暴

跌主要受美国退出量化宽松货币政策、美元走强和全球石油需求不足等因素共同驱动.对比预测结果与真实油价走势可见,本研究提出的集成预测方法能较好地捕捉到油价波动的拐点与趋势变化.

### 2.3.1 选择集成阈值确定

对于选择集成策略,阈值 Diff 作为关键超参数,会对模型预测精度产生一定影响,因此探究 Diff 取值与预测精度之间的关系具有重要意义.如图3所示,随着训练集内 Diff 值的增大,集成模型的预测误差呈现先减小后增大的趋势.这表明在初始阶段选择模型可以提高预测精度,但当所选模型到达最优子集后,若继续增加模型数量,集成模型开始出现过拟合现象,预测精度也随之下降.综上分析,在后续预测中选择 Diff 值为 2.05 进行进一步的分析.

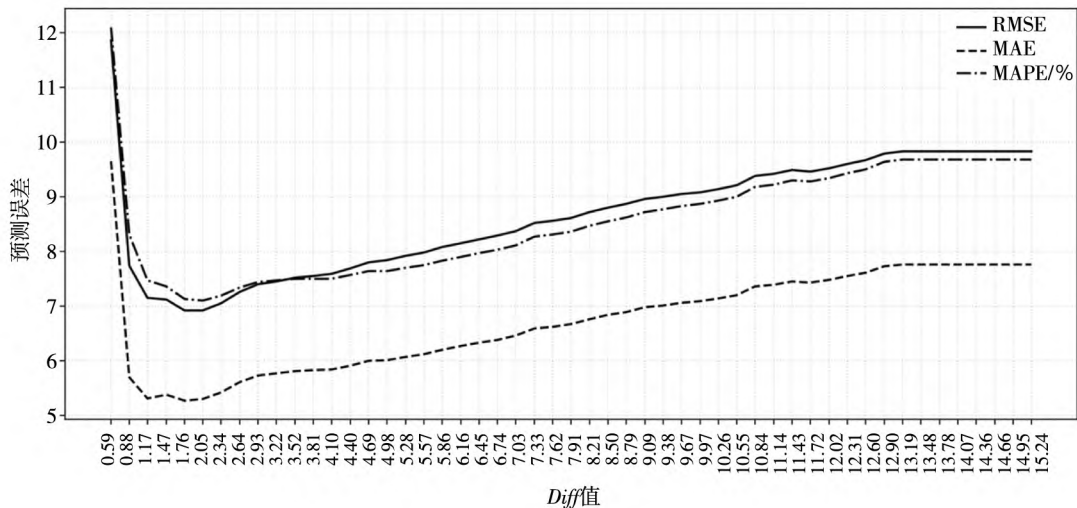


图3 选择阈值确定

Fig.3 Determination of selection threshold

### 2.3.2 基模型数量确定

对于 Adaboost 算法,如何确定基学习器的个数同样是需要解决的关键问题.若基学习器数目过少,模型学习不充分;基学习器数目过多,模型会过度关注难预测的样本,容易出现过拟合及计算复杂度过高的问题.因此,本研究结合多样性与精度两方面来确定 Adaboost 算法生成基预测器的最优个数,结果如图4所示.

由图4(a)可知,使用 Adaboost 进行集成预测时,随着基学习器的个数增加,集成预测精度提升较小,整体基本保持稳定;而使用 D-Adaboost 进

行集成预测时,测试误差呈现显著降低趋势,当基预测器个数达到50个之后,集成预测精度趋于平稳,不再出现明显变化.

由图4(b)可知,无论使用 Adaboost 还是 D-Adaboost 进行集成预测,随着基学习器的个数增加,预测组合的多样性均呈提升态势,但提升幅度逐渐放缓,最终趋于稳定.其中,D-Adaboost 算法的优势尤为明显,在基预测器较少的情况下就能显著提升基预测器的多样性,从而提高集成预测精度.当迭代次数达到50次后,随着基学习器个数的增多,D-Adaboost 的多样性保持稳

定. 这表明, 针对 WTI 原油数据集, 使用 Adaboost 对 50 个基学习器进行集成学习, 已能较为全面地覆盖难以预测样本, 基学习器之间形成

了良好的多样性; 若继续增加基学习器个数, 不仅无法显著提升多样性, 还可能引发过拟合风险.

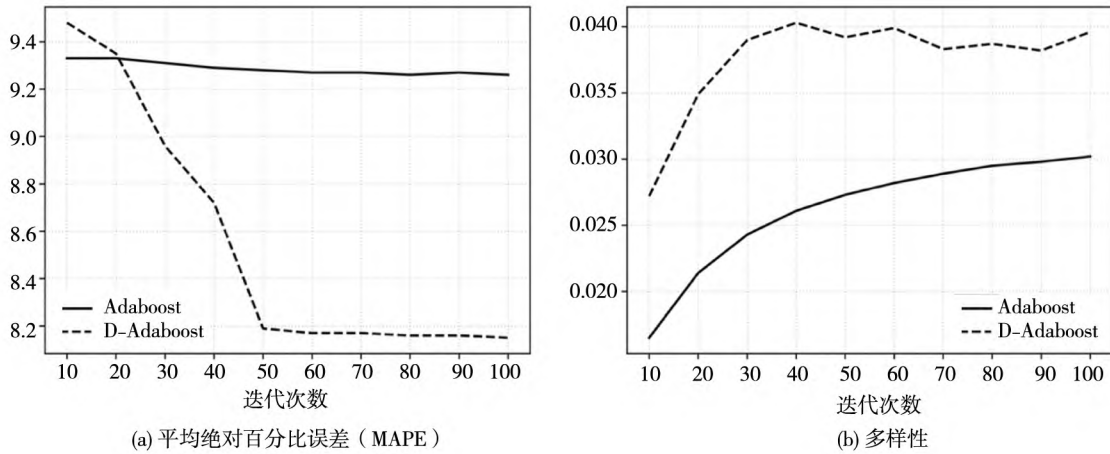


图 4 基模型数量与预测精度、多样性的关系

Fig. 4 Relationship between the number of base models and prediction accuracy and diversity

### 2.3.3 MCS 检验

本节采用 MCS 统计检验方法对 Adaboost 与 D-Adaboost 两种算法生成的 50 个基模型集开展预测性能检验, 以筛选出最优的模型子集. 通过对比

MCS 检验所得最优模型子集与本研究选择集成策略得到的模型子集之间的对应关系, 判断选择集成策略筛选出的模型是否均具备优良的预测能力. 图 5 给出了  $p$  值分别为 0.05 和 0.1 时的检验结果.

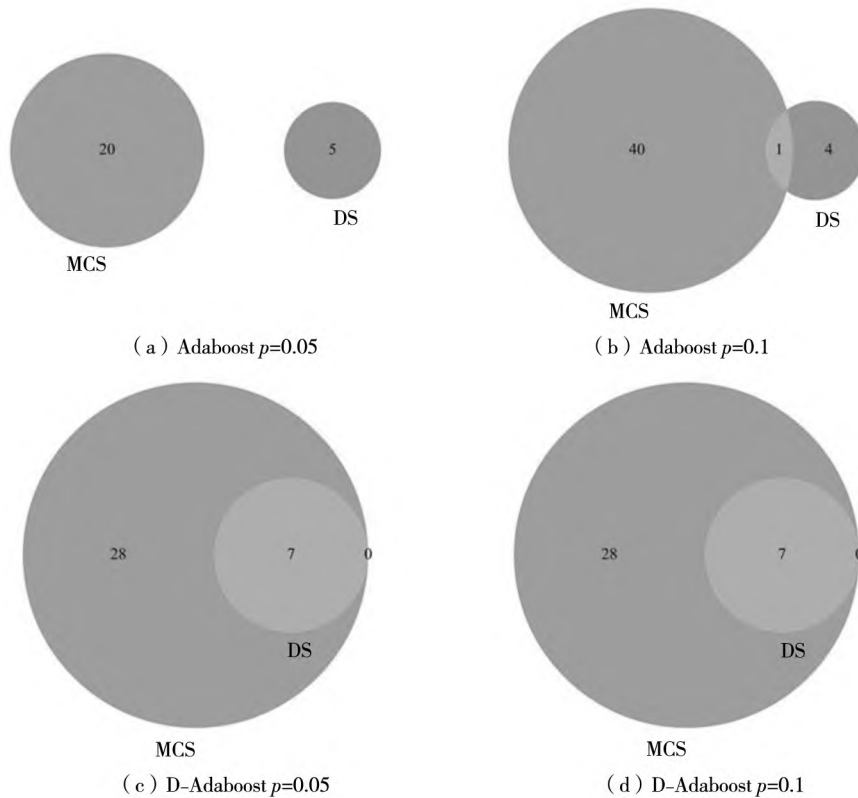


图 5 MCS 检验结果

Fig. 5 MCS test results

由图5可知,在 $p$ 值为0.05的显著性水平下,Adaboost生成的子模型通过统计检验的有20个,选择集成得到的子模型有5个,但二者无交集;在 $p$ 值为0.1的显著性水平下,Adaboost生成的子模型通过统计检验的有41个,选择集成得到的子模型有5个,但同时满足两项条件的只有1个.这表明利用DS-Adaboost进行集成预测,在 $p$ 值为0.05和0.1的显著性水平下均不能通过MCS检验.

在 $p$ 值为0.05和0.1的显著性水平下,DDS-Adaboost方法选择的所有子模型,均包含于MCS检验得到的最优模型子集中,即该策略选出的所有子模型在 $p$ 值为0.05和0.1的显著性水平下,均通过了对应显著性水平下的MCS检验.

综上,本研究所提两阶段Adaboost选择集成算法所选择的模型,在预测精度上无显著性差异,且均属于最优候选模型范畴.同时,该结果进一步说明了在损失函数中引入多样性指标,能够有效提升集成模型的预测精度与泛化性能.

#### 参考文献:

- [1]Rokach L. Ensemble-based classifiers[J]. Artificial Intelligence Review, 2010, 33(1): 1-39.
- [2]Geman S, Bienenstock E, Doursat R. Neural networks and the Bias/Variance dilemma[J]. Neural Computation, 1992, 4(1): 1-58.
- [3]Zhou Z H. Ensemble Methods: Foundations and Algorithms[M]. Boca Raton: CRC Press, 2012.
- [4]Sagi O, Rokach L. Ensemble learning: A survey[J]. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 2018, 8(4): e1249.
- [5]Dai Q, Ye R, Liu Z. Considering diversity and accuracy simultaneously for ensemble pruning[J]. Applied Soft Computing, 2017, 58: 75-91.
- [6]Kuncheva L I, Whitaker C J. Measures of diversity in classifier ensembles and their relationship with the ensemble accuracy[J]. Machine Learning, 2003, 51(2): 181-207.
- [7]Hu X. Using rough sets theory and database operations to construct a good ensemble of classifiers for data mining applications[C]//Proceedings of 2001 IEEE International Conference on Data Mining. San Jose: IEEE, 2001: 233-240. <https://doi.org/10.1109/ICDM.2001.989524>.
- [8]Aiolfi M, Timmermann A. Persistence in forecasting performance and conditional combination strategies[J]. Journal of Econometrics, 2006, 135(1): 31-53.
- [9]Merentitis A, Debes C, Heremans R. Ensemble learning in hyperspectral image classification: Toward selecting a favorable bias-variance tradeoff[J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2014, 7

### 3 结束语

多样性增强是提升集成学习泛化能力的关键之一.适度增加模型多样性能够有效改善集成模型的泛化能力,但若过度追求多样性则会导致模型精度下降.为更好地权衡多样性与预测精度的关系,本研究提出一种基于多样性增强的两阶段选择集成方法.首先,构建了融合多样性与预测精度的损失函数,并基于该损失利用Adaboost生成基模型,改善了传统Adaboost在迭代生成基模型时易陷入局部最优的问题;其次,设计了动态选择集成策略,能够筛选出兼顾多样性与预测精度的基模型以完成集成预测.将提出的集成方法在9个数据集上进行验证,结果表明,该方法能够有效提升基模型间的多样性,并显著增强集成模型的预测精度.

考虑到目前集成学习中多样性度量尚未形成统一标准,未来研究将进一步探讨不同多样性度量形式对集成性能的影响.此外,集成权重分配也是影响集成精度的重要原因之一.因此,如何在权重优化中也引入多样性度量,也将是后续研究的重点方向.

- (4): 1089 – 1102.
- [10] Breiman L. Bagging predictors [J]. *Machine Learning*, 1996, 24(2): 123 – 140.
- [11] Freund Y, Schapire R E. Experiments with a new boosting algorithm [C]//*Proceedings of the 13th International Conference on Machine Learning*. Bari: Morgan Kaufmann, 1996: 148 – 156.
- [12] Zhang Y, Burer S, Street W N. Ensemble pruning via semi-definite programming [J]. *Journal of Machine Learning Research*, 2006, 7: 1315 – 1338.
- [13] Bi Y. The impact of diversity on the accuracy of evidential classifier ensembles [J]. *International Journal of Approximate Reasoning*, 2012, 53(4): 584 – 607.
- [14] Kuncheva L I. A bound on kappa-error diagrams for analysis of classifier ensembles [J]. *IEEE Transactions on Knowledge and Data Engineering*, 2011, 25(3): 494 – 501.
- [15] Zhang H, He H, Zhang W. Classifier selection and clustering with fuzzy assignment in ensemble model for credit scoring [J]. *Neurocomputing*, 2018, 316: 210 – 221.
- [16] Goldberg D E. *Genetic Algorithms in Search, Optimization, and Machine Learning* [M]. Reading: Addison-Wesley, 1989.
- [17] Mao S, Chen J W, Jiao L, et al. Maximizing diversity by transformed ensemble learning [J]. *Applied Soft Computing*, 2019, 82: 105580.
- [18] Mao S, Lin W, Jiao L, et al. End-to-end ensemble learning by exploiting the correlation between individuals and weights [J]. *IEEE Transactions on Cybernetics*, 2021, 51(5): 2835 – 2846.
- [19] Giacinto G, Roli F. Design of effective neural network ensembles for image classification purposes [J]. *Image and Vision Computing*, 2001, 19(9 – 10): 699 – 707.
- [20] 李 航. 统计学习方法 [M]. 北京: 清华大学出版社, 2012.  
Li Hang. *Statistical Learning Methods* [M]. Beijing: Tsinghua University Press, 2012. (in Chinese)
- [21] Solomatine D P, Shrestha D L. AdaBoost. RT: A boosting algorithm for regression problems [C]//*2004 IEEE International Joint Conference on Neural Networks*. Budapest: IEEE, 2004: 1163 – 1168. <https://doi.org/10.1109/IJCNN.2004.1380102>.
- [22] Tang E K, Suganthan P N, Yao X. An analysis of diversity measures [J]. *Machine Learning*, 2006, 65(1): 247 – 271.
- [23] Krogh A, Vedelsby J. Neural network ensembles, cross validation, and active learning [C]//*Advances in Neural Information Processing Systems*, Denver: MIT Press, 1995, 7: 231 – 238.
- [24] Yang C, Yin X C, Hao H W, et al. Classifier ensemble with diversity: Effectiveness analysis and ensemble optimization [J]. *Acta Automatica Sinica*, 2014, 40(4): 660 – 674.
- [25] Dai Q, Zhang T, Liu N. A new reverse reduce-error ensemble pruning algorithm [J]. *Applied Soft Computing*, 2015, 28: 237 – 249.
- [26] Zhou Z H, Wu J, Tang W. Ensembling neural networks: Many could be better than all [J]. *Artificial Intelligence*, 2002, 137(1 – 2): 239 – 263.
- [27] Wang J, Zhou H, Hong T, et al. A multi-granularity heterogeneous combination approach to crude oil price forecasting [J]. *Energy Economics*, 2020, 91: 104790.
- [28] Zhao Y, Li J, Yu L. A deep learning ensemble approach for crude oil price forecasting [J]. *Energy Economics*, 2017, 66: 9 – 16.
- [29] Tang Y, Xiao X, Wahab M I M, et al. The role of oil futures intraday information on predicting US stock market volatility [J]. *Journal of Management Science and Engineering*, 2021, 6(1): 64 – 74.
- [30] 董 坤, 谢海滨, 汪寿阳. 中国股票市场的石油效应之谜 [J]. *管理科学学报*, 2012, 15(11): 45 – 53.  
Dong Kun, Xie Haibin, Wang Shouyang. The mystery of oil effect in China's Stock market [J]. *Journal of Management Sciences in China*, 2012, 15(11): 45 – 53. (in Chinese)

[31]Hedar A R , Wang J , Fukushima M. Tabu search for attribute reduction in rough set theory [J]. *Soft Computing* ,2008 ,12 ( 9) : 909 –918.

## **A dynamic selection ensemble forecasting method based on diversity–regularized Adaboost**

*CHENG Sheng*<sup>1,2</sup> , *FENG Han*<sup>1,2</sup> , *ZHANG Yi-fei*<sup>1,2</sup> , *ZHANG Xiao*<sup>1,2</sup> , *WANG Jue*<sup>1,2\*</sup> , *WANG Shou-yang*<sup>1,2</sup>

1. Academy of Mathematics and Systems Science , Chinese Academy of Sciences , Beijing 100190 , China;
2. School of Economics and Management , University of Chinese Academy of Sciences , Beijing 100190 , China

**Abstract:** The interplay between model diversity and predictive accuracy in ensemble forecasting represents a pivotal area of current research focus. Addressing the prevalent challenges of overfitting and suboptimal predictive accuracy in ensemble forecasting , this study proposes a two-stage dynamic selection ensemble strategy predicated on diversity regularization. In the first stage , a novel ensemble forecasting diversity regularization strategy is devised by instituting a loss function that judiciously balances diversity against predictive accuracy. The second stage introduces a dynamic selection ensemble methodology capable of proficiently identifying candidate predictive models that adeptly reconcile diversity with predictive accuracy. Experimental outcomes derived from publicly available datasets attest to the proposed two-stage selection ensemble strategy’s efficacy in notably enhancing predictive accuracy while concurrently bolstering model diversity. This dual advancement substantially elevates the ensemble forecasting model’s generalization capacity and reduces predictive errors.

**Key words:** diversity regularization; dynamic selection; Adaboost; ensemble forecasting