

# 基于自适应增量集成学习的非平稳金融时间序列预测

于慧慧, 戴群

(南京航空航天大学计算机科学与技术学院/人工智能学院, 南京 211106)

**摘要:** 金融市场对于社会经济的发展非常重要, 因此金融时间序列预测 (Financial time series prediction, FTSP) 一直是人们研究的焦点。至今, 许多基于统计分析和软计算的方法被提出以解决 FTSP 问题, 其中大多数方法将金融时间序列 (Financial time series, FTS) 视为或转化为平稳序列进行处理。但是, 由于绝大部分 FTS 是非平稳的, 因此这些方法通常存在伪回归或预测性能不佳等问题。本文提出了一种自适应增量集成学习 (Self-adaptive incremental ensemble learning, SIEL) 算法, 用于解决非平稳金融时间序列预测 (Non-stationary FTSP, NS-FTSP) 问题。SIEL 算法的主要思想是为每个非平稳金融时间序列 (Non-stationary FTS, NS-FTS) 子集增量地训练一个基模型, 然后使用自适应加权规则将各基模型组合起来。SIEL 算法的重点在于数据权重和基模型权重的更新: 数据权重基于当前集成模型在最新数据集上的性能进行更新, 其目的不是为了数据采样, 而是为了权衡误差; 基模型权重基于其所处环境进行自适应更新, 且基模型在越新环境下的性能应具有越高的权重。此外, 针对 NS-FTS 的特征, SIEL 算法提出了一种能协调新旧知识以及应对环境重演的策略。最后, 给出了 SIEL 算法在 3 个 NS-FTS 数据集上的实验结果, 并将其与已有算法进行了对比。实验结果表明, SIEL 算法能很好地解决 NS-FTSP 问题。

**关键词:** 非平稳金融时间序列预测; 自适应增量集成学习; 数据权重; 基模型权重

**中图分类号:** TP391      **文献标志码:** A

## Non-stationary Financial Time Series Prediction Based on Self-adaptive Incremental Ensemble Learning

YU Huihui, DAI Qun

(College of Computer Science and Technology / College of Artificial Intelligence, Nanjing University of Aeronautics & Astronautics, Nanjing 211106, China)

**Abstract:** The financial market is very important to the development of social economy, so financial time series prediction (FTSP) has always been the research focus. So far, many methods based on statistical analysis and soft computing have been proposed to solve FTSP problems, most of which treat financial time series (FTS) as or convert them into stationary time series. However, since most FTSs are non-stationary, these methods usually have problems such as false regression or poor prediction performance. Therefore, this paper proposes a novel self-adaptive incremental ensemble learning (SIEL) method to solve the problem of non-stationary FTSP (NS-FTSP). The main idea of the SIEL algorithm is

**基金项目:** 国家自然科学基金 (61473150) 资助项目。

**收稿日期:** 2020-09-24; **修订日期:** 2020-10-24

to incrementally train a base model for each non-stationary financial time series (NS-FTS) subset, and then ensemble the base models using the adaptive weighting rule. The focus of the SIEL algorithm is the update of data weight and base model weight. The weight of data is updated based on the performance of the current ensemble model on the latest dataset, and its purpose is not to sample the data, but to weigh the error; the weight of the base model is adaptively updated based on its environment, and the performance of the base model in the newer environment should have a higher weight. In addition, in view of the characteristics of NS-FTS, the SIEL algorithm proposes a strategy to coordinate new and old knowledge and cope with the recurrence of the environment. Finally, the paper gives the experimental results of the SIEL algorithm on three NS-FTS datasets and compares them with the existing algorithms. Experimental results show that the SIEL algorithm can solve the NS-FTSP problem well.

**Key words:** non-stationary financial time series prediction (NS-FTSP); self-adaptive incremental ensemble learning (SIEL); data weight; base model weight

## 引 言

金融市场在现代社会的经济发展中起着非常重要的作用。在过去的几十年中,对金融市场动向的分析已在金融、工程和数学等领域得到了广泛的研究,这些研究可以分为基础分析和技术分析两类。基础分析研究了影响市场走势的所有经济因素,它对于支持长期决策非常有用。技术分析认为,价格已经包含了影响市场走势的所有基本面,它通常将历史价格和一些技术指标建模为历史时间序列,并认为历史环境总会重演。

时间序列是在固定采样间隔内收集到的一系列观测结果<sup>[1]</sup>。现实世界中的许多动态过程都可以建模为时间序列,例如股票价格变动<sup>[2]</sup>、城市天气<sup>[3]</sup>、黑子数和疾病发生率等。如果要预测这些过程来帮助生产、生活,则需要了解时间序列预测(Time series prediction, TSP)。TSP是指基于历史时间序列生成模型来预测未来趋势的过程。与金融产品相关的时间序列称为金融时间序列(Financial time series, FTS),诸多因素的影响导致其极为复杂。从20世纪80年代研究者将机器学习(Machine learning, ML)应用于金融市场开始,ML在金融市场预测方面取得了诸多进展,研究者也逐渐认识到现实生活中的FTS大多为非平稳金融时间序列(Non-stationary FTS, NS-FTS)。但是,ML在NS-FTS上的应用仍有诸多挑战,主要包括:(1) NS-FTS被认为是确定性混沌的时间序列。它对初始条件敏感,因此利用历史数据对其进行长期预测是不现实的。(2) NS-FTS具有非线性、非平稳性的特征,这导致其数据分布会随着时间发生改变,如何跟踪变化并及时反应变化是一个难题。(3) NS-FTS具有高噪声的特征,这意味着无论建立何种模型,都无法获得序列的完整信息,需要消除噪声的影响;(4) 一些NS-FTS具有周期性特征,如何在历史环境重现时准确快速地做出反应,是值得思考的问题。

现有的非平稳金融时间序列预测(Non-stationary FTSP, NS-FTSP)方法大致可以分为3类:(1)传统的统计学方法。它是最早应用于NS-FTS的一类方法,其中的代表方法有差分整合滑动平均自回归(Auto-regressive integrated moving average, ARIMA)模型<sup>[4]</sup>和广义自回归条件异方差(Generalized auto-regressive conditional heteroscedasticity, GARCH)模型<sup>[5]</sup>。ARIMA模型曾被广泛地应用于NS-FTSP任务,但是由于它需要通过差分过程对NS-FTS进行平稳化处理,所以必然会影响模型的预测精度和效率。GARCH模型是一种使用过去变化和方差来预测将来变化的方法,它可以很好地拟合具有波动特征的序列。但是同ARIMA一样,它也需要对NS-FTS进行平稳化处理。可见传统统计学各种方法由于自身线性结构的约束,在实际应用中并不能达到很好的效果。(2)计算智能方法。此类方

法有很多,其中的代表方法有人工神经网络(Artificial neural network, ANN)<sup>[6]</sup>、模糊逻辑(Fuzzy logic, FL)<sup>[7]</sup>和支持向量机(Support vector machine, SVM)<sup>[8]</sup>。ANN作为非线性预测模型具有良好的自学能力、函数逼近能力,在NS-FTSP领域得到了广泛的应用。但是ANN往往需要大量的训练数据,网络结构和网络参数难以确定,收敛速度慢,容易陷入局部最优。FL通过产生模糊的NS-FTS,模拟人类思维得出不确定的知识,然后将其变换为精确的预测值来完成预测任务。但如何生成合适的模糊规则是一大局限。SVM结构简单、泛化能力强,在NS-FTSP领域上得到了较好的应用,但只适用于小样本学习。(3)混合预测方法。实验表明,单一模型很难全面反映预测对象的整体变化规律,特别是在NS-FTSP这种具有高度不确定特征的任务上,所以混合预测方法应运而生,并成为现阶段最流行的方法。混合预测方法通过组合多种方法,包括上述提到的传统统计学方法和计算智能方法,来得到综合信息,从而提高最终模型的预测性能和稳定性。

实现混合预测方法的方式有两种:一是通过增量学习(IL)<sup>[9]</sup>和集成学习(Ensemble learning, EL)<sup>[10]</sup>,比如文献[11-12];二是通过分阶段思想来实现,比如文献[13-14],基本思想是先将NS-FTS进行分解,然后分别用模型进行预测,最后再将预测结果组合起来。IL指一个模型在保存大部分旧知识的同时,不断从新数据中学习新知识的过程。应用于非平稳环境中的IL算法需要满足以下条件<sup>[15-16]</sup>:(1)任何训练数据集仅被学习一次,所学知识将以增量方式存储在模型参数中;(2)因为最新的数据集可以最好地表示当前环境,所以应根据知识与当前环境之间的关系对知识进行分类并动态更新;(3)模型应具有协调旧知识与新知识之间冲突的策略,即应有一种机制来监视模型在新旧数据上的性能;(4)模型应及时遗忘无关的知识,但在历史环境发生重现时,又能重新回忆起这些知识。EL指构建并结合多个基模型来完成学习任务的过程,它有3个要素:数据采样、训练基模型和组合基模型。至今EL已广泛应用于各种机器学习任务,比如特征选择、置信度估计、不平衡数据分类、以及非平稳环境学习等方面。主流的EL方法有以随机森林<sup>[17]</sup>为代表的装袋(Bagging)方法<sup>[18]</sup>和以AdaBoost<sup>[19]</sup>为代表的提升(Boosting)方法<sup>[20]</sup>两大类。Bagging是一种基于自助采样法并行生成不存在强依赖关系的基模型的方法。Boosting是一种串行生成存在强依赖关系的基模型的方法。AdaBoost作为Boosting的代表算法,其主要思想是:每次通过最新的数据分布迭代地训练一个基模型,再根据该基模型的性能调整样本分布,减少擅长数据的权重,增加不擅长数据的权重,以使先前不擅长的数据在后续训练过程中得到更多的关注。最后减少误差大的基模型的权重,增加误差小的基模型的权重,并将其加权组合起来。通过文献[21]知道,应用于非平稳环境学习的EL方法有3个思路:(1)更改组合规则以使已存在的基模型适应非平稳环境,例如加权多数投票算法<sup>[22]</sup>;(2)利用新数据集在线更新模型,如在线提升算法<sup>[23]</sup>;(3)添加新的基模型或利用新基模型替换集合中贡献最小或年龄最小的成员<sup>[24]</sup>。

结合以上背景,本文提出了一种自适应增量集成学习(Self-adaptive incremental ensemble learning, SIEL)算法,用于解决NS-FTSP问题。其思想是为每个NS-FTS子集增量地训练一个基模型,然后使用自适应加权规则将所有基模型组合起来。SIEL算法的具体步骤如下:(1)初始化数据分布,此时所有数据的权重是相同的。(2)计算当前集成模型在最新数据集上的相对误差,并根据该误差更新数据权重。(3)在最新数据集上训练一个基模型,将其作为新的集成成员以增量地构建基模型集合。(4)评估所有基模型在最新数据集上的性能,并对那些表现较差的基模型进行处理。值得注意的是,对于每个基模型来说,由于其生成时间不同,获得的性能评估数量也不同。(5)计算所有基模型的加权平均误差以自适应地更新基模型权重并消除噪声对基模型的影响。每个基模型在越新数据集下的性能具有越大的权重(因为越新的数据集越能表示下一个新环境,尤其是在NS-FTS中)。(6)使用加权规则组合每个基模型。此外,考虑到极限学习机(Extreme learning machine, ELM)<sup>[25]</sup>作为特殊的神经网络模型具有较高的可扩展性和较低的计算复杂度,但其隐含层参数的随机性又会影响集成模型的稳定性和鲁棒

性。所以,本文选择它的改进算法核极限学习机(Extreme learning machine with kernels, ELMK)<sup>[26]</sup>作为 SIEL 算法的基模型。

SIEL算法的主要创新和贡献如下:(1)提出了一种新规则来自适应地更新数据权重。与AdaBoost在提升阶段中基于单个基模型更新数据权重的过程不同,SIEL算法基于当前集成模型的性能来更新数据权重,且更新后的数据权重不用于采样,而是用于评估误差。该规则有助于提高最终模型的泛化性和鲁棒性。此外,该规则为预测误差较小的数据赋予了较高的权重,从而加重了对此类数据表现不佳的模型的惩罚。(2)提出了一种新规则来自适应地构建基模型。与AdaBoost在提升阶段中基于总体数据训练基模型的过程不同,SEIL算法根据数据块训练基模型,提高了基模型的多样性。(3)提出了一种新规则来自适应地组合基模型。与AdaBoost在集成阶段中单纯地根据基模型误差决定基模型权重的过程不同,该规则可以使表现较差的基模型的权重变小甚至为零,从而达到暂时遗忘无关知识的目的。但当历史训练环境重演时,该基模型又会被分配较高的权重从而使忘掉的知识重新被记起,提高了算法的稳定性。并且SIEL算法基于时间对基模型误差进行了加权,为基模型在越新环境中的性能分配越高的权重,从而能有效应对下一个新环境。这使SIEL算法在保持稳定性的同时,提高了本身的可塑性。(4)为解决NS-FTSP问题提供了一个通用的框架,并取得了优异的性能。

## 1 算法描述

### 1.1 ELM 算法

ELM是由黄光斌等提出的一种特定的单隐藏层前馈网络(Single-hidden layer feedforward networks, SLFNs)模型。与传统的SLFNs相比,其不同之处在于:(1)输入层与隐藏层之间的连接权重,以及隐藏层的偏置可以随机产生,也可以手动设置;(2)隐藏层和输出层之间的连接权重不需要通过迭代来反复调整,而是直接通过最小二乘法求解方程来确定。因此,在大部分情况下,ELM能在更短的时间内获得更优的泛化性能,且避免了局部最优问题。ELM算法结构如图1所示。

假设有  $N$  个训练样本  $(x_i, t_i), i = 1, 2, \dots, N$ , 其中  $x_i \in \mathbb{R}^d, t_i \in \mathbb{R}^m$ 。若设置隐藏层节点数为  $L$ , 那么 ELM 可以被表示为

$$y_i = f_L(x) = \sum_{j=1}^L \beta_j g(w_j \cdot x_i + b_j) \quad (1)$$

式中:  $y_i$  代表第  $i$  个样本的期望值,  $w_j$  代表输入层和第  $j$  个隐藏层节点之间的连接权重,  $\beta_j$  代表第  $j$  个隐藏层节点和输出层之间的连接权重,  $g(\cdot)$  和  $b_j$  分别代表隐藏层的激活函数以及第  $j$  个隐藏层节点的偏置。

那么对于这  $N$  个样本,可以得到

$$H\beta = Y \quad (2)$$

式中隐藏层输出矩阵为

$$H = \begin{bmatrix} g(w_1 \cdot x_1 + b_1) & \dots & g(w_L \cdot x_1 + b_L) \\ \vdots & & \vdots \\ g(w_1 \cdot x_N + b_1) & \dots & g(w_L \cdot x_N + b_L) \end{bmatrix}_{N \times L} \quad (3)$$

理想情况下的 ELM 应满足

$$\|H\beta - T\| = 0 \quad (4)$$

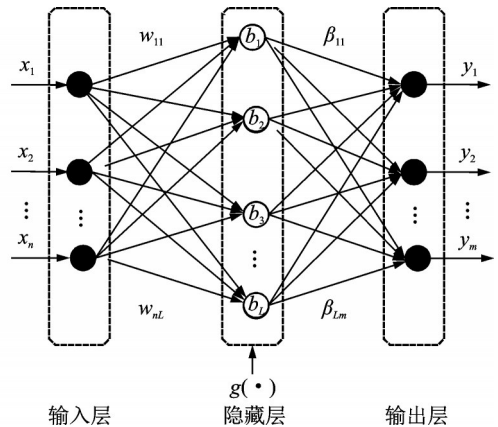


图1 ELM算法的结构图

Fig.1 Block diagram of ELM algorithm

根据相关知识,可以得到

$$\beta = H^+ T \quad (5)$$

式中  $H^+$  代表  $H$  的广义逆矩阵。

### 1.2 ELMK 算法

作为 ELM 的改进算法,ELMK 使用核映射代替了 ELM 的随机特征映射,在保证泛化性能的同时,避免了随机性,加强了鲁棒性。ELMK 在多分类、图像分类和回归预测等方面都有着广泛的应用。

ELMK 定义如下

$$K_{i,j} = \varphi(\mathbf{x}_i) \cdot \varphi(\mathbf{x}_j) = K(\mathbf{x}_i, \mathbf{x}_j) \quad (6)$$

式中  $\varphi(\cdot)$  代表未知的特征映射,能将原空间样本映射为另一个特征空间的向量。

ELMK 的目标是求该特征空间内的最优输出权重  $\beta$ , 即求解

$$\begin{aligned} \text{Min } L_p &= \frac{1}{2} \|\beta\|^2 + \frac{C}{2} \xi_i^2 \\ \text{s.t. } \varphi(\mathbf{x}_i) \beta &= t_i - \xi_i \end{aligned} \quad (7)$$

式中  $\xi_i$  表示训练误差,  $C$  表示正则化系数。基于拉格朗日乘子法,最终 ELMK 的目标可以转换成

$$L_D = \frac{\|\beta\|^2 + C \sum_{i=1}^N \xi_i^2}{2} - \sum_{i=1}^N \alpha_i (\varphi(\mathbf{x}_i) \beta - t_i + \xi_i) \quad (8)$$

式中  $\alpha_i$  表示第  $i$  个拉格朗日乘子。

### 1.3 SIEL 算法

SIEL 算法作为一种基于自适应增量集成学习的通用框架,可以有效解决 NS-FTSP 问题。SIEL 算法的主要思想是为每个 NS-FTS 子集增量地训练一个基模型,然后使用自适应加权规则将所有基模型组合起来。SIEL 算法的重点在于数据权重和基模型权重的更新。图 2 展示了 SIEL 算法的基本流程。

SIEL 算法的具体步骤如下:

(1) 数据预处理。首先,使用式(9)对非平稳金融时间序列  $\text{NSFTS} = \{d_1, d_2, \dots, d_N\}$  进行归一化处理,得到序列 DATA。

$$d'_i = \frac{d_i - d_{\min}}{d_{\max} - d_{\min}} \quad (9)$$

式中:  $d_i$  表示原始数据,  $d_{\min}$  表示原序列上的最小值,  $d_{\max}$  表示原序列上的最大值,  $d'_i$  表示归一化后的数据。然后,将 DATA 转换为预测任务所需的训练集  $\text{Train} = \{(\mathbf{x}_1, y_1), (\mathbf{x}_2, y_2), \dots, (\mathbf{x}_n, y_n)\}$  和测试集  $\text{Test} = \{(\mathbf{x}_{n+1}, y_{n+1}), (\mathbf{x}_{n+2}, y_{n+2}), \dots, (\mathbf{x}_{n+ts}, y_{n+ts})\}$ , 其中  $n = N - tw - ts$ ,  $tw$  表示时间窗大小,  $ts$  表示测试集大小。然后把 Train 分成  $T$  部分,得到  $D^t = \{(\mathbf{x}_1, y_1), (\mathbf{x}_2, y_2), \dots, (\mathbf{x}_{m^t}, y_{m^t}), t = 1, 2, \dots, T\}$ ,

其中  $\sum_{t=1}^T m^t = n$ 。

(2) 当新的数据子集  $D^t$  到来时,初始化数据分布。此时,所有数据的权重是相同的,即

$$D^t(i) = w^t(i) = \frac{1}{m^t} \quad i = 1, 2, \dots, m^t \quad (10)$$

(3) 更新并归一化数据权重。首先,计算当前集成模型  $H^{t-1}$  在  $D^t$  上的相对误差

$$E^t = \sum_{i=1}^{m^t} D^t(i) * \left( \frac{\text{abs}(H^{t-1}(\mathbf{x}_i) - y_i^t)}{\max(\text{abs}(H^{t-1} - y^t))} \right)^2 \quad (11)$$

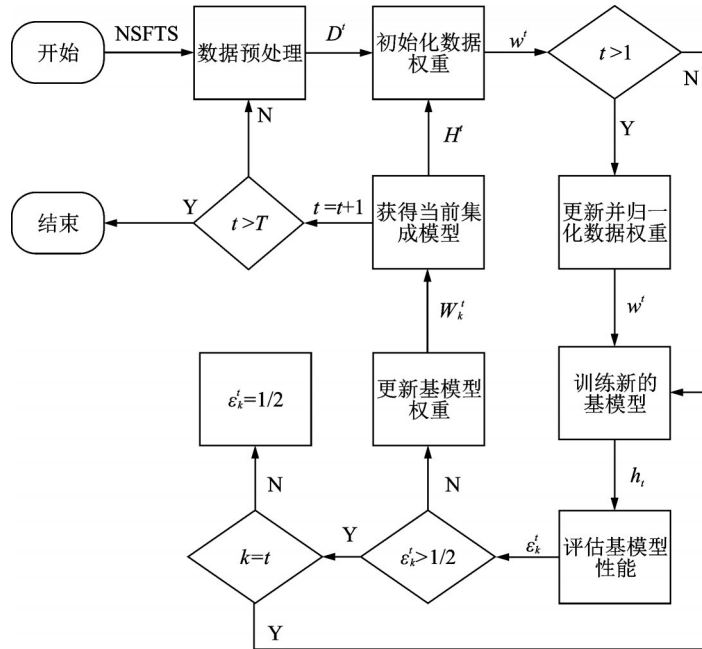


图2 SIEL算法的结构图

Fig.2 Block diagram of SIEL algorithm

然后,根据  $E^t$  更新数据权重

$$w^t(i) = \frac{1}{m^t} * E^t \left( 1 - \frac{\text{abs}(H^{t-1}(x_i) - y_i^t)}{\max(\text{abs}(H^{t-1} - y^t))} \right)^2 \quad (12)$$

更新的数据权重不是为了数据采样,而是为了在式(14)中评估和均衡误差。最后,对  $w^t(i)$  进行归一化处理,即

$$D^t = w^t / \sum_{i=1}^{m^t} w^t(i) \quad (13)$$

(4) 在  $D^t$  上使用 ELMK 训练一个新的基模型  $h_t$ , 将其作为新的集成成员来增量地构建基模型集合。

(5) 评估此时所有基模型  $h_k (k = 1, 2, \dots, t)$  在  $D^t$  上的性能

$$\epsilon_k^t = \sum_{i=1}^{m^t} D^t(i) * \left( \frac{\text{abs}(h_k(x_i) - y_i^t)}{\max(\text{abs}(h_k - y^t))} \right)^2 \quad (14)$$

对表现较差的模型进行处理: 当  $\epsilon_{k=t}^t > \frac{1}{2}$  时, 重新训练一个  $h_t$ ; 当  $\epsilon_{k<t}^t > \frac{1}{2}$ , 令  $\epsilon_{k<t}^t = \frac{1}{2}$ 。这样处理的目的是保证集成模型中新成员的性能; 二是使集成模型具有协调新旧知识以及应对重现环境的能力。

(6) 计算所有基模型的加权平均误差以自适应地更新基模型权重并消除噪声对基模型的影响。首先根据式(15~17)对基模型的误差进行时间加权, 以便为每个基模型在更新数据集下的性能分配到更大的权重值, 然后根据式(18)计算各个基模型的权重。

$$\omega_k^t = 1 / (1 + e^{-(t-k)}) \quad (15)$$

$$\omega_k^t = \omega_k^t / \sum_{j=0}^{t-k} \omega_k^{t-j} \quad (16)$$

$$\bar{\beta}_k^t = \sum_{j=0}^{t-k} \omega_k^{t-j} (\epsilon_k^{t-j} / (1 - \epsilon_k^{t-j})) \quad (17)$$

$$W_k^t = \lg(1/\bar{\beta}_k^t) / \sum_{k=1}^t \lg(1/\bar{\beta}_k^t) \quad (18)$$

(7) 使用加权规则组合每个基模型, 得到一个新的集成模型, 即

$$H^t = \sum_{k=1}^t h_k \cdot W_k^t \quad (19)$$

这时, 生成的集成模型  $H^t$  将通过影响步骤(3)中数据权重的更新来影响基模型权重的更新。理想情况下, 只要数据输入不中断, SIEL 算法将会始终生成新的集成模型来适应下一个新环境。

## 2 实验设置和结果分析

### 2.1 实验数据集

本文使用 3 个 FTS 数据集来评估 SIEL 算法的性能: 美元兑人民币汇率(USD / CNY)、上证综合指数(SSE)和日经指数(N225)。3 个 FTS 数据集均从 Yahoo finance<sup>[27]</sup> 获得。

由于并不是所有的 FTS 都具有非平稳特征, 因此出于严谨性考虑, 在实验之前, 使用增广迪基-福勒(Augmented Dickey-Fuller, ADF)<sup>[28]</sup> 检验方法对这 3 个 FTS 数据集进行了非平稳检验, 检验结果如表 1 所示。从表 1 可以看出, 这 3 个 FTS 数据集都是非平稳的, 均为 NS-FTS 数据集。

表 1 各数据集的 ADF 检验结果

Table 1 Results of ADF test on each dataset				
数据集	$p$	$c$	$A/\%$	$h$
USD/CNY	0.941 8	-1.941 5	5	0
SSE	0.604 8	-1.941 8	5	0
N225	0.245 5	-1.941 6	5	0

### 2.2 实验设置

#### 2.2.1 算法参数设置

SIEL 算法主要涉及 3 个参数: 时间窗大小  $tw$ 、子数据集数目  $T$  和基模型 ELMK 的核函数 Kernel。 $tw$  和  $T$  过大过小都会对模型的泛化性能产生负面影响, 因此需要对其进行合理设置。通过实验发现, 对于所选数据集,  $tw$  在 6 和 16 之间,  $T$  在 2 和 6 之间, SIEL 算法可以获得比较好的预测效果。此外, 关于 Kernel, 经过交叉验证<sup>[29]</sup> 可证明高斯核函数是比较好的选择。

#### 2.2.2 对比算法设置

为进一步证明 SIEL 算法的性能, 选取了 5 个对比算法: 在线极限学习机(Online sequential ELM, OS-ELM)<sup>[30]</sup>、改进的误差最小化在线极限学习机(Online sequential improved error minimized ELM, OSIEM-ELM)<sup>[31]</sup>、双增量学习算法(Double incremental learning, DIL)<sup>[32]</sup>、核极限学习机(ELMK)和以 ELM 为基模型的 SIEL 算法(SIEL based on ELM, SIEL-ELM)。

OS-ELM 和 OSIEM-ELM 都属于在线极限学习机, 可以逐段或逐块地学习数据。DIL 算法类似于本文提出的 SIEL 算法, 它们都属于 IL 和 EL 算法。与 DIL 算法进行比较, 可以验证 SIEL 算法能否更好地解决 NS-FTSP 问题。与 SIEL-ELM 算法和 ELMK 算法相比, 可以验证 SIEL 中提出的基于自适应增量集成学习的通用框架对于算法性能的改善是否具有显著效果。

#### 2.2.3 性能指标设置

为了评估 SIEL 算法和各个对比算法的性能, 选取了两个性能指标: 均方根误差(RMSE)和平均绝

对误差(MAE)。

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (x_i - \widehat{x}_i)^2} \quad (20)$$

$$\text{MAE} = \frac{\sum_{i=1}^n |x_i - \widehat{x}_i|}{n} \quad (21)$$

式中: $x_i$ 代表数据的真实值; $\widehat{x}_i$ 代表数据的预测值; $n$ 表示数据集的大小。

为减少偶然因素对实验结果的影响,进行了20次独立重复实验。

### 2.3 实验结果

表2和表3展示了SIEL算法在其参数 $tw$ 、 $T$ 取不同值时在各数据集上的RMSE结果,其中最小值以粗体显示。从结果可以看出, $tw$ 和 $T$ 过大或过小都会对模型的泛化性能产生负面影响。针对本文的3个NS-FTS数据集, $tw$ 的合适范围为 $[6, 16]$ , $T$ 的合适范围为 $[2, 6]$ 。

表2 取不同 $tw$ 值时SIEL算法在各数据集上的RMSE结果

Table 2 RMSEs obtained by SIEL algorithm at different values of  $tw$  on each dataset

数据集	$tw$							
	4	6	8	10	12	14	16	18
USD/CNY	2.36e-02	2.14e-02	<b>2.08e-02</b>	2.13e-02	2.18e-02	2.21e-02	2.27e-02	2.35e-02
SSE	2.90e-02	2.83e-02	2.79e-02	2.69e-02	2.71e-02	2.62e-02	<b>2.56e-02</b>	2.64e-02
N225	1.98e-02	<b>1.93e-02</b>	2.05e-02	2.07e-02	2.20e-02	2.23e-02	2.29e-02	2.32e-02

表3 取不同 $T$ 值时SIEL算法在各数据集上的RMSE结果

Table 3 RMSEs obtained by SIEL algorithm at different values of  $T$  on each dataset

数据集	$T$				
	2	3	4	5	6
USD/CNY	<b>2.08e-02</b>	2.16e-02	2.16e-02	2.30e-02	3.48e-02
SSE	<b>2.56e-02</b>	2.62e-02	2.68e-02	2.72e-02	2.98e-02
N225	2.23e-02	2.13e-02	1.95e-02	<b>1.93e-02</b>	1.99e-02

表4和表5分别展示了不同算法在各数据集上的RMSE和MAE结果。从表中可以看出,SIEL算法在每个数据集上的预测性能均优于其他算法,证明SIEL算法能更好地解决NS-FTSP问题。此外,SIEL、SIEL-ELM和ELMK这3种算法之间的比较,值得注意。从实验结果可以看出,SIEL算法和SIEL-ELM算法之间的RMSE比ELMK的RMSE更接近且更小。因此可以得出结论,SIEL算法中提出的基于自适应增量集成学习的通用框架对于算法性能的改善具有显著效果。

表6展示了SIEL算法和对比算法之间关于RMSE和MAE的 $t$ -测试结果。从表中可以看出,与ELMK、DIL、OSELM和OSIEM-ELM算法相比,在5%显著性水平下,SIEL算法在实验数据集上的预测性能得到了显著改善。但是SIEL算法与其变种算法SIEL-ELM之间的 $t$ -测试结果存在 $H=0$ ,这是因为与基模型的选择相比,SIEL中提出的基于自适应增量集成学习的通用框架更有利于算法性能的提高,基模型的选择相对不太重要。



表 4 不同算法在各数据集上的 RMSE 结果

Table 4 RMSEs obtained by different algorithms on each dataset

数据集	算法					
	ELMK	DIL	OS-ELM	OSIEM-ELM	SIEL-ELM	SIEL
USD/CNY	2.23e-02	2.18e-02	2.12e-02	2.22e-02	2.10e-02	2.08e-02
SSE	2.78e-02	2.87e-02	2.81e-02	3.15e-02	2.72e-02	2.55e-02
N225	2.06e-02	1.99e-02	1.99e-02	2.04e-02	1.95e-02	1.93e-02

表 5 不同算法在各数据集上的 MAE 结果

Table 5 MAEs obtained by different algorithms on each dataset

数据集	算法					
	ELMK	DIL	OS-ELM	OSIEM-ELM	SIEL-ELM	SIEL
USD/CNY	1.46e-02	1.29e-02	1.29e-02	1.34e-02	1.32e-02	1.28e-02
SSE	2.11e-02	2.18e-02	2.06e-02	2.40e-02	2.09e-02	1.90e-02
N225	1.62e-02	1.58e-02	1.55e-02	1.57e-02	1.51e-02	1.49e-02

表 6 SIEL 和其他对比算法的  $t$ -测试结果

Table 6  $t$ -test results between SIEL and other comparative algorithms

数据集	ELMK		DIL		OS-ELM		OSIEM-ELM		SIEL-ELM	
	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE
USD/CNY	1	1	1	1	1	1	1	1	0	0
SSE	1	1	1	1	1	1	1	1	1	1
N225	1	1	1	1	1	1	1	1	0	0

注： $H=1$ 表示拒绝零假设 $H_0$ ，即以RMSE或MAE为标准，在5%显著性水平下，SIEL算法与其他算法相比，泛化性能有显著提高。 $H=0$ 表示不能拒绝零假设 $H_0$ ，即以RMSE或MAE为标准，在5%显著性水平下，SIEL算法与其他算法相比，泛化性能没有显著提高。

为了更直观地展示实验结果，分别在3个NS-FTS数据集上绘制了各算法的预测结果图。图3~5展示了各算法在USD/CNY、SSE和N225数据集上的预测值和实际值，其结果作了归一化处理。

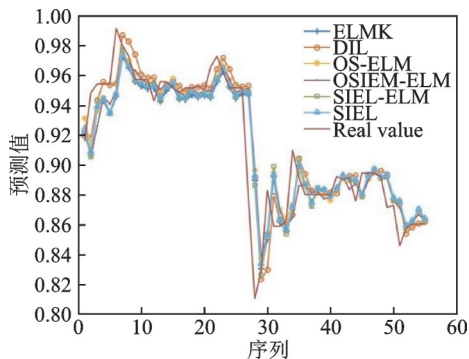


图3 不同算法在USD/CNY数据集上的预测结果

Fig.3 Prediction results of different algorithms on USD/CNY dataset

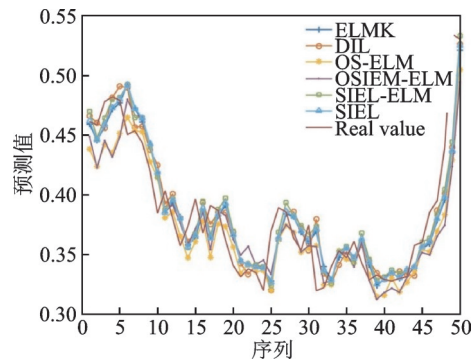


图4 不同算法在SSE数据集上的预测结果

Fig.4 Prediction results of different algorithms on SSE dataset

从图3~5可以看出,各算法的预测趋势与实际趋势基本一致,但SIEL算法的预测趋势与实际趋势最接近。这也意味着SIEL在某种程度上缓解了“稳定性-可塑性”困境,从而能有效解决NS-FTSP问题。

### 3 结束语

本文提出了一种基于自适应增量集成学习的方法SIEL来解决NS-FTSP问题。SIEL算法为每个NS-FTS子集增量地训练一个基模型,并使用自适应加权规则将它们组合在一起。实验证明,SIEL算法在解决NS-FTSP问题上取得了令人满意的结果。

但是SIEL算法仍存在一些不足。例如,每个基模型的训练过程都是独立的,这与在IL中充分利用历史知识的想法相矛盾。下一步的计划是使用“迁移学习”来加强基模型之间的关系并引入“剪枝”技术来更新基模型数量,以进一步改善算法性能。

#### 参考文献:

- [1] METCALFE A V, COWPERTWAIT P S P. Introductory time series with R[M]. New York:Springer-Verlag, 2009.
- [2] SCHÖNEBURG E. Stock price prediction using neural networks: A project report[J]. Neurocomputing, 1990, 2(1): 17-27.
- [3] SHI X J, CHEN Z, WANG H, et al. Convolutional LSTM network: A machine learning approach for precipitation nowcasting [C]//Proceedings of the 28th International Conference on Neural Information Processing Systems. [S.l.]:[s.n], 2015, 1: 802-810.
- [4] BOX G E P, JENKINS G M, REINSEL G C, et al. Time series analysis: Forecasting and control[M]. [S.l.]: John Wiley & Sons, 2015.
- [5] FRANCQ C, ZAKOIAN J M. GARCH models: Structure, statistical inference and financial applications[M]. [S.l.]: John Wiley & Sons, 2019.
- [6] CAO J, LI Z, LI J. Financial time series forecasting model based on CEEMDAN and LSTM[J]. Physica A: Statistical Mechanics and its Applications, 2019, 519: 127-139.
- [7] ADEMOLA POPOOLA S A, AHMAD K. A fuzzy-wavelet method for analyzing non-stationary time series[C]//Proceedings of the 5th International Conference on Recent Advances in Soft Computing. Nottingham, UK: [s.n.], 2004: 231-236.
- [8] CAO L, GU Q. Dynamic support vector machines for non-stationary time series forecasting[J]. Intelligent Data Analysis, 2002, 6(1): 67-83.
- [9] SYED N A, HUAN S, KAH L, et al. Incremental learning with support vector machines[C]//Proceedings of International Joint Conference on Artificial Intelligence Sweden: Morgan Kaufmann Publishers, 1999: 352-356.
- [10] DIETTERICH T G. Ensemble learning in the Handbook of Brain Theory and Neural Networks [M]. Cambridge: MIT Press, 2002.
- [11] VAN HEESWIJK M, MICHE Y, LINDH-KNUUTILA T, et al. Adaptive ensemble models of extreme learning machines for time series prediction[C]//Proceedings of International Conference on Artificial Neural Networks. [S.l.]: Springer, 2009: 305-314.
- [12] CHACÓN H D, KESICI E, NAJAFIRAD P. Improving financial time series prediction accuracy using ensemble empirical mode decomposition and recurrent neural networks[J]. IEEE Access, 2020, 8: 117133-117145.
- [13] ZHANG Y A, YAN B, AASMA M. A novel deep learning framework: Prediction and analysis of financial time series using CEEMD and LSTM[J]. Expert Systems with Applications, 2020, 159: 113609.
- [14] 许锐. 基于 EEMD 的金融时间序列多尺度分析[D]. 合肥:中国科学技术大学, 2016.

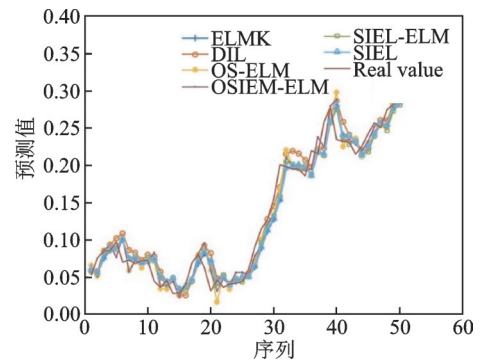


图5 不同算法在N225数据集上的预测结果  
Fig.5 Prediction results of different algorithms on N225 dataset

- XU Rui. Multiscale analysis of financial time series based on EEMD[D]. Hefei: University of Science and Technology of China, 2016.
- [15] GIRAUD-CARRIER C. A note on the utility of incremental learning[J]. *Ai Communications*, 2000, 13(4): 215-223.
- [16] DRUCKER H. Improving regressors using boosting techniques[C]//*Proceedings of the 14th International Conference on Machine Learning*. [S.l.]: Morgan Kaufmann, 1997, 97: 107-115.
- [17] BREIMAN L. Random forests[J]. *Machine Learning*, 2001, 45(1): 5-32.
- [18] BREIMAN L. Bagging predictors[J]. *Machine Learning*, 1996, 24(2): 123-140.
- [19] KUNCHEVA L I. Combining pattern classifiers: Methods and algorithms[M]. [S.l.]: John Wiley & Sons, 2014.
- [20] FREUND Y, SCHAPIRE R E. A decision-theoretic generalization of on-line learning and an application to boosting[J]. *Journal of Computer and System Sciences*, 1997, 55(1): 119-139.
- [21] KUNCHEVA L I. Classifier ensembles for changing environments[C]//*Proceedings of International Workshop on Multiple Classifier Systems*. [S.l.]: Springer, 2004: 1-15.
- [22] BLUM A. Empirical support for winnow and weighted-majority algorithms: Results on a calendar scheduling domain[J]. *Machine Learning*, 1997, 26(1): 5-23.
- [23] OZA N C, RUSSELL S. Online ensemble learning[M]. Berkeley: University of California, 2001.
- [24] NISHIDA K, YAMAUCHI K, OMORI T. ACE: Adaptive classifiers-ensemble system for concept-drifting environments [C]//*Proceedings of International Workshop on Multiple Classifier Systems*. [S.l.]: Springer, 2005: 176-185.
- [25] HUANG G B, ZHU Q Y, SIEW C K. Extreme learning machine: A new learning scheme of feedforward neural networks[C]//*Proceedings of 2004 IEEE International Joint Conference on Neural Networks*. [S.l.]: IEEE, 2004, 2: 985-990.
- [26] HUANG G B, ZHOU H, DING X, et al. Extreme learning machine for regression and multiclass classification[J]. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, 2011, 42(2): 513-529.
- [27] Yahoo. Yahoo finance[EB/OL]. (2019-11-30)[2020-8-20]. <https://finance.yahoo.com/>.
- [28] CHEUNG Y W, LAI K S. Lag order and critical values of the augmented Dickey-Fuller test[J]. *Journal of Business & Economic Statistics*, 1995, 13(3): 277-280.
- [29] MINOWA Y. Verification for generalizability and accuracy of a thinning-trees selection model with the ensemble learning algorithm and the cross-validation method[J]. *Journal of Forest Research*, 2008, 13(5): 275.
- [30] LIANG N Y, HUANG G B, SARATCHANDRAN P, et al. A fast and accurate online sequential learning algorithm for feedforward networks[J]. *IEEE Transactions on Neural Networks*, 2006, 17(6): 1411-1423.
- [31] XUE J, LIU Z, GONG Y, et al. Time series prediction based on online sequential improved error minimized extreme learning machine[C]//*Proceedings of ELM-2015*. [S.l.]: Springer, 2016, 1: 193-209.
- [32] LI J, DAI Q, YE R. A novel double incremental learning algorithm for time series prediction[J]. *Neural Computing and Applications*, 2019, 31(10): 6055-6077.

#### 作者简介:



于慧慧(1994-),女,博士研究生,研究方向:模式识别与数据挖掘, E-mail: huihuiyu11@163.com。



戴群(1974-),通信作者,女,博士,教授,博士生导师,研究方向:模式识别与数据挖掘, E-mail: daiqun@nuaa.edu.cn。

(编辑:王静)