



DOI:10.12404/j.issn.1671-1815.2308252

引用格式:王威娜,李明莉.基于Shapelets的多元时间序列分类方法[J].科学技术与工程,2025,25(1):252-261.

Wang Weina, Li Mingli. Multivariate time series classification method based on Shapelets[J]. Science Technology and Engineering, 2025, 25(1): 252-261.

# 基于 Shapelets 的多元时间序列分类方法

王威娜,李明莉

(吉林化工学院信息与控制工程学院,吉林 132022)

**摘要** 多元时间序列分类是众多领域的关键问题,但是当前多元时序分类研究面临着原始数据高维、精度不足、可解释性缺乏等问题,这使得模型性能提升受限,准确率难以满足实际需求。针对上述问题,提出基于 Shapelets 的多元时间序列分类方法。首先,利用自适应邻居的无监督 Shapelet 学习将 Shapelet 变换与自适应权重结合,用于自动学习显著多元 Shapelets;然后,将该方法与 Shapelet 相似性和类标约束项结合,增强模型可解释性和分类准确性;最后,提出模型的优化策略,用以获取最优的 Shapelets,进一步提高模型的分类精度。与 3 种不同类型 11 个算法在 11 个公开数据集上进行比较,实验结果表明提出算法具有较高的分类精度。

**关键词** 多元时间序列;多元时间序列分类;Shapelets 学习;优化策略

**中图分类号** TP391; **文献标志码** A

## Multivariate Time Series Classification Method Based on Shapelets

WANG Wei-na, LI Ming-li

(School of Information and Control Engineering, Jilin Institute of Chemical Technology, Jilin 132022, China)

**[Abstract]** Multivariate time series classification is a key problem in many fields, but the current research on multivariate time series classification is faced with some problems, such as high dimensionality of original data, low accuracy, and lack of interpretability, which limits the performance improvement of models and makes it difficult to meet the actual requirements. Aiming at above problem, a multivariate time series classification method based on Shapelets was proposed. Firstly, unsupervised Shapelet learning of adaptive neighbors was used to automatically learn significant multivariate Shapelets by combining Shapelets transform and adaptive weights. Then, the method was combined with Shapelet similarity and class label constraint to enhance the interpretability and classification accuracy of the model. Finally, the optimization strategy of the model was proposed to obtain the best Shapelets to further improve the classification accuracy of the model. Three different types of 11 algorithms were compared on 11 public data sets, and the experimental results show that the proposed algorithm has high classification accuracy.

**[Keywords]** multivariate time series; multivariate time series classification; Shapelets learning; optimization strategy

随着人工智能的发展,时间序列数据的数量呈直线上升趋势,对其分析与研究已引起学术界和工业界的广泛关注<sup>[1]</sup>。早期研究者主要关注一元时间序列分类问题,但由于描述对象单一,导致其在众多领域的应用存在局限性。因此,越来越多的科研工作者开始深入研究多元时间序列<sup>[2]</sup>。多元时间序列具有广泛的实际应用,例如生物医学诊断<sup>[3]</sup>、股票金融市场<sup>[4]</sup>、交通流量<sup>[5]</sup>、工业制造<sup>[6]</sup>、人类活动识别<sup>[7]</sup>等。多元时间序列可以被看作是多个一元时间序列的集合,在整个学习过程中不仅要在一元时间序列进行分析,还需要充分考虑各变量之间的相关性。以自动驾驶为例,为做出更精准

的决策,需分析多种传感器信息,包括摄像头和激光雷达等,并将其综合起来进行全面决策<sup>[8]</sup>。多元时间序列分类的研究主要包括基于距离的方法<sup>[9]</sup>、基于特征的方法<sup>[10]</sup>、基于深度学习的方法<sup>[11]</sup>和基于模型的方法<sup>[12]</sup>。基于距离的方法主要通过计算不同时间序列之间的距离进行分类,同类间距离相近而不同类间距离较远,常用方法包括欧几里得距离、基于概率的距离、动态时间规整(dynamic time warping, DTW)距离<sup>[13]</sup>。动态时间规整距离已在多元时间序列分类领域展现出卓越性能,但是该方法的计算过程十分烦琐。基于特征的方法通过从原始时间序列数据中提取相关特征,并将这些特征作

收稿日期:2023-10-23 修订日期:2024-10-08

基金项目:国家自然科学基金(62266046);吉林省自然科学基金(YDZJ202201ZYTS603);吉林省教育厅科研项目(JJKH20230281KJ)

第一作者:王威娜(1981—),女,汉族,吉林吉林人,博士,教授。研究方向:时间序列分析、数据挖掘。E-mail:wangweina@jlicet.edu.cn。

投稿网址:www.stae.com.cn

为输入进行分类,主流方法包括二维奇异值分解、时间序列符号表示<sup>[14]</sup>、Shapelets 方法<sup>[15-17]</sup>。在基于深度学习的方法中,循环神经网络<sup>[18]</sup>、卷积神经网络<sup>[19]</sup>等方法已成为多元时序分类领域备受瞩目的研究对象,这些方法通过构建模型学习时间序列数据的内部关系和结构,提取特征实现分类,在分类准确度方面具有明显优势,但是可解释性较弱<sup>[20]</sup>。基于模型的方法将原始多元时间序列实例转换为模型参数,建立数学模型获取时间序列数据特征,并利用该模型进行分类。与其他方法相比,基于 Shapelets 的分类方法越来越受到科学学者的重视<sup>[21]</sup>,这是因为此类方法可以高效地获取信息,并具有较高的可解释性和准确度。

Shapelets 是指时间序列中最具辨识性的子序列,这一概念由 Keogh 等<sup>[22]</sup>于 2009 年首次提出。近年来,Shapelets 已成为备受关注的研究热点。Shapelets 具有很好的可解释性,能够有效地区分时间序列数据中的类间差异,另外 Shapelets 是从时间序列中选取的最具有代表性的特征,同时也达到了对时序数据降维的效果,但 Shapelets 的选取过程具有较高的时间复杂度,限制了模型的发展与应用<sup>[23]</sup>。大量快速发现 Shapelets 的算法被提出,此类方法通过信息增益度量的熵修剪方法替换传统欧几里得距离的计算,后续也有研究者提出搜索空间的修剪方法,用以提升 Shapelets 的发现速度<sup>[24]</sup>。Lines 等<sup>[25]</sup>利用最优的 Shapelets 与时间序列之间的距离作为时间序列的新特征,并结合其他分类器实施最终分类,为分类算法的发展提供了新思路。Grabocka 等<sup>[26]</sup>提出学习时间序列 Shapelets 算法(learning time-series Shapelets, LTS),该算法基于特定目标函数进行优化 Shapelets,结果表明与其他 Shapelets 的分类器相比,LTS 算法的准确率得到显著提高。最近,Guillaume 等<sup>[27]</sup>引入扩张 Shapelet 变换(dilated Shapelet transform, DST),获得包含膨胀概念的 Shapelet,并利用 Shapelet 特征增强分类能力,有效扩展了 Shapelets 的应用领域。Chen 等<sup>[28]</sup>提出可解释时序分类的小波选择方法,通过定义位置度量和距离度量评价每个候选 Shapelet 的判别能力,从而在分类结果中提供更具可解释性的洞察力,但在筛选 Shapelet 的过程中,此方法的计算复杂度较高。Liu 等<sup>[29]</sup>提出基于具有规范时序特征 Shapelet 的时序分类算法,利用提取的 Shapelet 建立新的特征表示,构建随机森林分类器,保证算法的泛化能力。虽然将 Shapelet 应用于时序分类的研究已取得长足发展,但是仍面临着 Shapelet 选取过程烦琐和显著性欠缺的挑战。王威娜等<sup>[30]</sup>提出基于优化 Shapelet 的时序分类方法,虽然选取 Shapelet

过程简便且具有可解释性,但是只考虑了单变量时间序列,对多变量问题未深入研究。李晨等<sup>[31]</sup>提出基于优化和两阶段筛选的 Shapelets 提取方法,虽然能够获取优异的 Shapelets 集合,但是后续筛选过程的复杂度较高。杨骏等<sup>[32]</sup>提出面向时序有序分类的 Shapelet 抽取算法,结合 SAX 与 Shapelet,并通过布隆过滤器对候选 Shapelet 进行剪枝,但该过程也相对烦琐。

针对上述问题,现提出基于 Shapelets 的多元时间序列分类方法(classification method based on Shapelets on multivariate time series, SMTS)。首先,利用自适应邻居的无监督 Shapelet 学习方法,将 Shapelet 变换与自适应权重结合用于自动学习显著多元 Shapelets,并根据自适应权重的计算,获取 Shapelets 的显著性。然后,将 Shapelet 学习方法与 Shapelets 相似性和类标约束项结合,计算 Shapelets 之间的相似程度,此过程与类标约束项结合,能够更好地减小类标之间的差异,缩小模型误差。最后,提出优化策略获取最优且不等长的多元 Shapelets,进而以此为基础实现多元时间序列分类。该算法通过自适应权重的无监督学习方法突出 Shapelets 的显著性,并融合 Shapelets 的多样性、相似性和类标约束,优化 Shapelets 的选取过程,减少分类复杂度的同时提高分类的准确性。

## 1 预备知识

列出的定义分别为时间序列、Shapelet、时间序列和 Shapelet 之间的距离度量以及 Shapelet 变换技术。

**定义 1 (时间序列)** 时间序列是一组序列数据,它由  $n$  个数值型数据组成,一条完整的时间序列可表示为  $t = \{x_1, x_2, \dots, x_n\}$ ,其中  $x_1, x_2, \dots, x_n$  按照时间顺序排列。多条时间序列可以组成一个时间序列集,即  $T = \{t_1, t_2, \dots, t_m\}$ 。多元时间序列数据中包含  $l$  个时间序列,即  $MTS = \{MTS_1, MTS_2, \dots, MTS_l\}$ ,每个时间序列为  $MTS_i \in \mathbf{R}^{P \times Q}$ , $P$  为每个样本包含的一元时间序列的个数, $Q$  为一维样本的长度。

**定义 2 (Shapelet)** Shapelet 是时间序列中特征最显著的一段子序列。在一元时间序列分类中,选出  $K$  个最优的 Shapelet,并记为  $S \in \mathbf{R}^{K \times M}$ , $M$  为时间序列的长度,即  $S = \{S_1, S_2, \dots, S_K\}$ 。

**定义 3 (时间序列和 Shapelet 之间的距离)** 一条时间序列  $t$  与第  $k$  个 Shapelet  $S_k$  之间的距离如式 (1) 所示。它通过滑动窗口,依次计算  $S_k$  与时间序列所有子序列之间的欧式距离,并选取最短距离作为最终结果。

$$\text{dist} = \min_{q=1,2,\dots,\bar{q}} \left[ \frac{1}{l_k} \sum_{k=1}^{l_k} (t_{q+k-1} - s_k)^2 \right] \quad (1)$$

**定义 4 (Shapelet 变换技术)** 利用定义 3 可将时间序列和 Shapelet 之间的距离作为新的特征,把时间序列映射到新特征空间中的过程称为 Shapelet 变换技术<sup>[29]</sup>,经 Shapelet 变换之后数据集转换为  $D$ ,记为  $D = \{D_1, D_2, \dots, D_n\}$ ,其中  $D_i = \{D_{i,1}, D_{i,2}, \dots, D_{i,K}\}$ ,由于  $K < Q$ ,所以在保留时间序列关键信息的同时实现了数据的降维处理。

## 2 多元时间序列分类方法

### 2.1 模型介绍

提出基于 Shapelets 的多元时间序列分类模型,该模型由自适应邻居的无监督 Shapelet 学习及 Shapelet 最小相似性和类标约束项共同构建。首先,自适应邻居无监督 Shapelet 学习方法可自动学习时序数据显著信息,进而获取多元 Shapelets;然后,结合 Shapelet 最小相似性和类标约束项构建候选 Shapelets 模型;最后,优化更新 Shapelets、分类器和稀疏邻居矩阵直到其收敛,从而获得最优的 Shapelets,并以此为基础实施多元时序数据分类。模型的具体流程如图 1 所示。

#### 2.1.1 自适应邻居的无监督 Shapelet 学习

Shapelet 学习是解决时间序列分类问题的关键技术之一<sup>[33]</sup>。自适应邻居的无监督 Shapelet 学习方法可以有效解决不同序列的重要度计算和多元 Shapelets 的更新策略<sup>[34]</sup>中存在的问题,该方法可以自动学习多元 Shapelets 信息,当给定候选多元 Shapelets 时,可以自动确定每个信息变量的显著性。

给定候选的多元 Shapelets,可将多元时间序列转换为基于多元 Shapelets 的矩阵  $\mathbf{X} = [X_{ij}]$ ,表达式为

$$\begin{aligned} X_{ij} &= \text{dist}[t^{(i)}, s^{(j)}] \\ &= \min_{q=1,2,\dots,\bar{q}} \sum_{f=1}^F \beta_k^f \left\{ \frac{1}{l_k} \sum_{h=1}^{l_k} \left[ t_{q+h-1}^{f(i)} - s_h^{f(j)} \right]^2 \right\} \\ &= \min_{q=1,2,\dots,\bar{q}} \sum_{f=1}^F \beta_k^f d_{ijq}^f \end{aligned} \quad (2)$$

式(2)中: $X_{ij}$ 为第  $i$  个多元时间序列与第  $j$  个多元 Shapelets 之间的距离; $\bar{q} = q_i - l_j + 1$  为时间序列  $t_i$  与 Shapelet  $S_j$ 之间的总段数; $q_i$ 为多元时间序列的长度; $l_j$ 为多元 Shapelets 的长度; $F$ 为多元时间序列中变量的个数; $\beta$ 为不同变量的重要性权重; $t_{q+k-1}^{f(i)}$ 为在第  $f$  个一元时间序列中长度为  $(q+k-1)$  的第  $i$  条时间序列; $s_k^{f(j)}$ 为在第  $f$  个 Shapelet 中长度为  $k$  的第  $j$  行 Shapelet。利用最小函数近似表示  $\mathbf{X} = [X_{ij}]$  为

$$\begin{aligned} X_{ij} &= \min_{q=1,2,\dots,\bar{q}} \sum_{f=1}^F \beta_k^f d_{ijq}^f \\ &\approx \frac{\sum_{q=1}^{\bar{q}} \left[ \sum_f \beta_k^f d_{ijq}^f \exp\left(\alpha \sum_f \beta_k^f d_{ijq}^f\right) \right]}{\sum_{q=1}^{\bar{q}} \exp\left(\alpha \sum_f \beta_k^f d_{ijq}^f\right)} \\ &= \frac{\sum_{f=1}^F \beta_k^f \left[ \sum_{q=1}^{\bar{q}} d_{ijq}^f \exp\left(\alpha \sum_f \beta_k^f d_{ijq}^f\right) \right]}{\sum_{q=1}^{\bar{q}} \exp\left(\alpha \sum_f \beta_k^f d_{ijq}^f\right)} \\ &= \sum_{f=1}^F \beta_k^f FX_{ij}^f \end{aligned} \quad (3)$$

式(3)中:

$$FX_{ij}^f = \frac{\sum_{q=1}^{\bar{q}} d_{ijq}^f \exp\left(\alpha \sum_f \beta_k^f d_{ijq}^f\right)}{\sum_{q=1}^{\bar{q}} \exp\left(\alpha \sum_f \beta_k^f d_{ijq}^f\right)} \quad (4)$$

式(4)中: $\alpha$ 用以控制函数的精度,根据前人的工作经验<sup>[35]</sup>,将  $\alpha$  设置为  $-100$ ;令  $\mathbf{X}_k$  为  $X_{ij}$  构成的矩阵,  $FX_k^f = \{FX_{ij}^f\}_{j=1}^J$  表示第  $f$  个多元时间序列与第  $i$  个多元 Shapelets 之间的距离。

针对不同变量重要性的差异,学习方法可自适应确定不同变量的重要性权重。权重  $\beta_k$  包括测量  $\beta_k^l$  与  $FX_k$  之间一致性的可学习权重  $\beta_k^f$  和多个变量之间相等的固定权重  $\beta_k^f$ ,其中  $\beta_k$  的学习过程可以看作是一种平行注意机制。另外,  $\beta_k$  的计算可以表示为

$$\begin{aligned} \beta_k &= (1 - \mu) \beta_k^l + \mu \beta_k^f \\ &= (1 - \mu) \frac{\exp(b_k)}{\sum_f \exp(b_k^f)} + \mu \frac{1}{F} \end{aligned} \quad (5)$$

$$b_k = b_k + \hat{\mathbf{X}}_k^T FX_k \quad (6)$$

$$\hat{\mathbf{X}}_k = \frac{\|\hat{\mathbf{X}}_k\|}{1 + \|\hat{\mathbf{X}}_k\|^2} \mathbf{X}_k \quad (7)$$

式中: $\mu$ 为权重参数,取值为  $\mu \in [0, 1]$ ;  $b_k$  通过度量  $\mathbf{X}_k$  与  $FX_k$  的一致性来迭代细化。

#### 2.1.2 拉普拉斯映射

拉普拉斯映射在无监督学习中被广泛采用<sup>[30]</sup>,其作用是使得相似时间序列尽可能具有相同的伪类标签。 $\mathbf{G} \in \mathbf{R}^{n \times n}$  表示经过 Shapelet 变换的时间序列相似矩阵,即

$$\mathbf{G}_{(ij)} = \exp\left[-\frac{\|\mathbf{X}_{(:,i)} - \mathbf{X}_{(:,j)}\|^2}{\sigma^2}\right] \quad (8)$$

式(8)中: $\sigma$ 为 RBF 内核参数; $\mathbf{X}_{(:,i)}$ 和  $\mathbf{X}_{(:,j)}$ 分别为  $\mathbf{X}$  矩阵中第  $i$  列的子矩阵和第  $j$  列的子矩阵。根据获得的  $\mathbf{G}$ ,相应的拉普拉斯正则化项可表示为

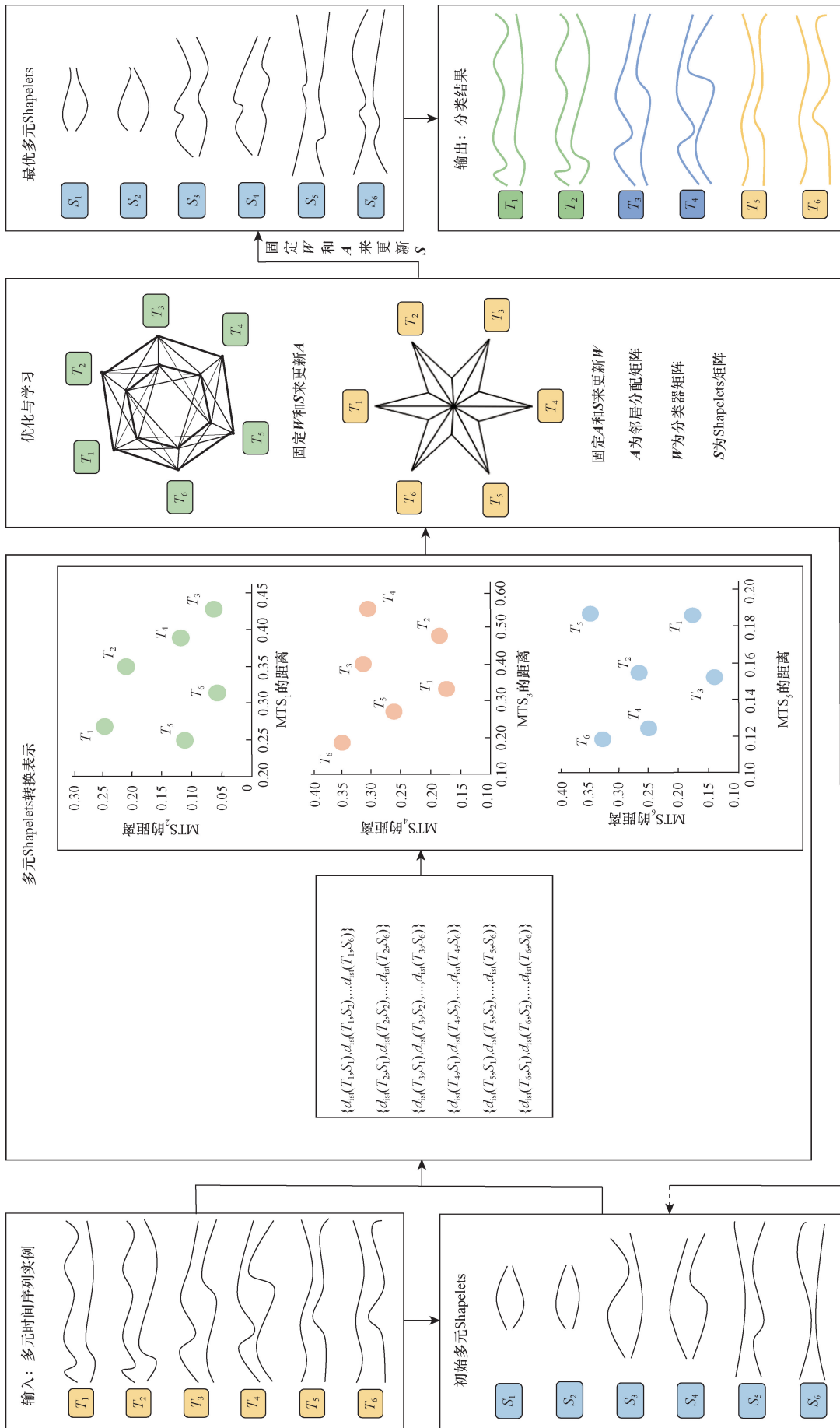


图 1 模型流程图示意图  
Fig. 1 The flow of the proposed model

$$\begin{aligned} & \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \mathbf{G}_{(ij)} \left\| \mathbf{Y}_{(:,i)} - \mathbf{Y}_{(:,j)} \right\|_2^2 \\ &= \frac{1}{2} \sum_{k=1}^c \sum_{i=1}^n \sum_{j=1}^n \mathbf{G}_{(ij)} \left[ \mathbf{Y}_{(ki)} - \mathbf{Y}_{(kj)} \right]^2 \\ &= \sum_{k=1}^c \mathbf{Y}_{(k,:)} (\mathbf{D}_G - \mathbf{G}) \mathbf{Y}_{(k,:)}^T \\ &= \text{tr}(\mathbf{Y} \mathbf{L}_G \mathbf{Y}^T) \end{aligned} \quad (9)$$

式(9)中: $\mathbf{D}_G(i, i) = \sum_{j=1}^n \mathbf{G}_{(ij)}$ 为对角矩阵; $\mathbf{L}_G = \mathbf{D}_G - \mathbf{G}$ 。

### 2.1.3 Shapelets 相似性最小化

为保障 Shapelets 的多样性,尽量剔除相似特性的 Shapelets,从而增加 Shapelets 相似性模块,使得输出的 Shapelets 具有鲜明特色。假设 Shapelets 相似矩阵为  $\mathbf{H} \in \mathbf{R}^{k \times k}$ , Shapelets  $S_m$  和  $S_n$  之间的相似性矩阵元素  $\mathbf{H}_{(mn)}$  为

$$\mathbf{H}_{(mn)} = e^{-\frac{\|d_{mn}\|^2}{\sigma^2}} \quad (10)$$

式(10)中: $d_{mn}$ 为 Shapelet 的  $S_m$  与  $S_n$  之间的距离,计算方法与式(2)相似。

### 2.1.4 类标约束项

模型中设立两项类标约束项,其目的是减少生成类标和实例类标之间的差异性,增强类标分配的准确性。

第一项是通过  $F$  范数实现计算实例类标和生成类标之间误差的平方和,获得规范化模型参数的同时也能使误差最小化,利用式(11)所示的函数实现误差最小化,表达式为

$$\min_{\mathbf{W}} \left\| \mathbf{W}^T \mathbf{X} - \mathbf{Y} \right\|_F^2 \quad (11)$$

第二项是计算生成类标与变换矩阵转置的乘积以及实际类标和变换矩阵转置的乘积,将得到的前后两个结果相减,通过  $F$  范数获得误差最小值,即

$$\min_{\mathbf{Y}} \left\| \mathbf{Y} \mathbf{D}^T - (\mathbf{Y} - \mathbf{I}) \mathbf{D}^T \right\|_F^2 = \min_{\mathbf{Y}} \left\| (2\mathbf{Y} - \mathbf{I}) \mathbf{D}^T \right\|_F^2 \quad (12)$$

式(12)中: $\mathbf{D} \in \mathbf{R}^{k \times n}$ 为 Shapelet 变换矩阵; $\mathbf{Y} \in \mathbf{R}^{c \times n}$ 为类标约束矩阵; $\mathbf{I}$ 为单位矩阵; $\mathbf{W} \in \mathbf{R}^{k \times c}$ 为分类器矩阵。

### 2.1.5 模型目标函数

当给定候选的多元 Shapelets  $\mathbf{S}$  后, SMTS 可以获得基于多元 Shapelets 的表示  $\mathbf{X}(\mathbf{S})$ 。为了学习信息丰富的多元 Shapelets, SMTS 同时学习邻居分配矩阵  $\mathbf{A}$ 、分类器  $\mathbf{W}$  和候选多元 Shapelets  $\mathbf{S}$ 。为保证获取的多元 Shapelets 能够代表多元时间序列的不同特征,构造目标函数包含 6 个主要成分。

第一部分  $\left[ \sum_{i,j=1}^n \left\| \mathbf{X}(\mathbf{S})_{.i} - \mathbf{X}(\mathbf{S})_{.j} \right\|_F^2 \mathbf{A}_{ij} \right]$  为

邻居一致性项,它能更准确地表示时间序列之间的相似性,并且采用的最优邻居可提高算法的效率。

第二部分 ( $\| \mathbf{A} \|^2$ ) 为邻居正则化项,优化模型的损失和复杂度,减少过拟合,提高模型的泛化能力。

第三部分  $\{ \text{tr}[\mathbf{Y} \mathbf{L}(\mathbf{A}) \mathbf{Y}^T] \}$  为拉普拉斯正则化项,确保多元时间序列数据之间的局部相似性,并提高模型的可解释性。

第四部分 [ $\| \mathbf{H}(\mathbf{S}) \|^2$ ] 为多元 Shapelet 相似性项,用以保证 Shapelet 的多样化,同时也会使得分类器能够更好地区分不同类别的时间序列,提高模型分类精度。

第五部分 ( $\| \mathbf{W} \|_F^2$ ) 为最优分类器,因为它可以更精确地匹配不同类别的时间序列样本,用以提高分类模型的准确性;

第六部分 [ $\| \mathbf{W}^T \mathbf{D}(\mathbf{S}) - \mathbf{Y} \|_F^2$  和  $\| (2\mathbf{Y} - \mathbf{I}) \times \mathbf{D}(\mathbf{S})^T \|_F^2$ ] 为两项类标约束项,对分类器进行正则化和约束,使其更加符合数据的分布,这样有助于减小过拟合的风险,使得数据之间的差异最小化,确保分类任务的性能。

因此, SMTS 的目标函数可以表示为

$$\begin{aligned} \min_{\mathbf{A}, \mathbf{Y}, \mathbf{S}} & \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \left\| \mathbf{X}(\mathbf{S})_{.i} - \mathbf{X}(\mathbf{S})_{.j} \right\|_F^2 \mathbf{A}_{ij} + \\ & \frac{\gamma}{2} \left\| \mathbf{A} \right\|^2 + \lambda_0 \text{tr}[\mathbf{Y} \mathbf{L}(\mathbf{A}) \mathbf{Y}^T] + \frac{\lambda_1}{2} \left\| \mathbf{H}(\mathbf{S}) \right\|_F^2 + \\ & \frac{\lambda_2}{2} \left\| \mathbf{W} \right\|_F^2 + \frac{\lambda_3}{2} \left\| \mathbf{W}^T \mathbf{D}(\mathbf{S}) - \mathbf{Y} \right\|_F^2 + \\ & \frac{\lambda_4}{2} \left\| (2\mathbf{Y} - \mathbf{I}) \mathbf{D}(\mathbf{S})^T \right\|_F^2 \\ \text{s. t. } & \rightarrow \mathbf{A}_i^T \mathbf{1} = 1, 0 \leq \mathbf{A}_{ij} \leq 1, \mathbf{Y} \mathbf{Y}^T = \mathbf{I} \end{aligned} \quad (13)$$

## 2.2 优化与学习

采用坐标下降法求解 SMTS 目标函数,具体方法如下。

### 2.2.1 固定 $\mathbf{W}$ 和 $\mathbf{S}$ 更新 $\mathbf{A}$

当固定  $\mathbf{W}$  和  $\mathbf{S}$  时, SMTS 的目标函数退化为

$$\begin{aligned} \min_{\mathbf{A}} F(\mathbf{A}) &= \frac{1}{2} \sum_{i,j=1}^n \left\| \mathbf{X}(\mathbf{S})_{.i} - \mathbf{X}(\mathbf{S})_{.j} \right\|_F^2 \mathbf{A}_{ij} + \\ & \frac{\gamma}{2} \left\| \mathbf{A} \right\|^2 + \lambda_0 \text{tr}[\mathbf{Y} \mathbf{L}(\mathbf{A}) \mathbf{Y}^T] \\ \text{s. t. } & \rightarrow \mathbf{A}_i^T \mathbf{1} = 1, 0 \leq \mathbf{A}_{ij} \leq 1, \mathbf{Y} \mathbf{Y}^T = \mathbf{I} \end{aligned} \quad (14)$$

利用谱分析中的基本方程得

$$\text{tr}(\mathbf{Y} \mathbf{L}_A \mathbf{Y}^T) = \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \left\| \mathbf{Y}_{.i} - \mathbf{Y}_{.j} \right\|_F^2 \mathbf{A}_{ij} \quad (15)$$

则上述函数  $F(\mathbf{A})$  可表示为

$$\begin{aligned}
& \min_{A^T \mathbf{I} = 1, 0 \leq A_{ij} \leq 1} \sum_{j=1}^n \left\| \mathbf{X}(\mathbf{S})_{\cdot i} - \mathbf{X}(\mathbf{S})_{\cdot j} \right\|_F^2 \mathbf{A}_{ij} + \\
& \gamma \left\| \mathbf{A} \right\|^2 + \lambda_0 \sum_{j=1}^n \left\| \mathbf{Y}_{\cdot i} - \mathbf{Y}_{\cdot j} \right\|_F^2 \mathbf{A}_{ij} \\
& \Leftrightarrow \min_{A^T \mathbf{I} = 1, 0 \leq A_{ij} \leq 1} \sum_{j=1}^n \left\| \frac{1}{2\gamma} \mathbf{d}_i^A + \mathbf{A}_{ij} \right\|^2 \\
& \Leftrightarrow \min_{A_i, \eta, \beta_i} F(\mathbf{A}_i, \xi, \beta_i) = \\
& \frac{1}{2} \left\| \mathbf{A}_i + \frac{\mathbf{d}_i}{2\gamma} \right\|_F^2 - \xi (\mathbf{A}_i^T \mathbf{I} - 1) - \beta_i^T \mathbf{A}_i
\end{aligned} \quad (16)$$

式(16)中:  $\eta > 0, \beta_i > 0$  为拉格朗日乘数;  $\mathbf{I}$  为全为 1 的列向量。由于  $\mathbf{A}$  是稀疏的, 因此只有与  $\mathbf{X}(\mathbf{S})_{\cdot i}$  最近的  $k$  个邻居才可能连接到  $\mathbf{X}(\mathbf{S})_{\cdot i}$ 。然后, 确定  $F(\mathbf{A}_i, \xi, \beta_i)$  的最优解为

$$\mathbf{A}_i = \max \left( \xi - \frac{\mathbf{d}_i}{2\gamma}, 0 \right) \quad (17)$$

$\mathbf{A}_i$  有  $\bar{k}$  个非零元素, 其中

$$\xi = \frac{1}{k} + \frac{1}{2k\gamma} \sum_{j=1}^{\bar{k}} d_{ij} \quad (18)$$

$$\gamma = \frac{1}{N} \sum_{i=1}^N \left( \frac{\bar{k}}{2} d_{i, \bar{k}+1} - \frac{1}{2} \sum_{j=1}^{\bar{k}} d_{ij} \right) \quad (19)$$

### 2.2.2 固定 $\mathbf{A}$ 和 $\mathbf{W}$ 更新 $\mathbf{S}$

当固定  $\mathbf{A}$  和  $\mathbf{W}$  时, SMTS 的目标函数退化为

$$\begin{aligned}
\min_{\mathbf{S}} F(\mathbf{S}) &= \sum_{i,j=1}^n \left\| \mathbf{X}(\mathbf{S})_{\cdot i} - \mathbf{X}(\mathbf{S})_{\cdot j} \right\|_F^2 \mathbf{A}_{ij} + \\
& \frac{\lambda_1}{2} \left\| \mathbf{H}(\mathbf{S}) \right\|_F^2 + \frac{\lambda_3}{2} \left\| \mathbf{W}^T \mathbf{D}(\mathbf{S}) - \mathbf{Y} \right\|_F^2 + \\
& \frac{\lambda_4}{2} \left\| (\mathbf{2Y} - \mathbf{I}) \mathbf{D}(\mathbf{S})^T \right\|_F^2
\end{aligned} \quad (20)$$

采用迭代算法更新  $\mathbf{S}$ , 即  $\mathbf{S}_{i+1} = \mathbf{S}_i - \eta \nabla \mathbf{S}_i$ , 其中  $\eta$  为学习率。  $F(\mathbf{S})$  对  $\mathbf{S}$  的导数为

$$\begin{aligned}
\frac{\partial F(\mathbf{S})}{\partial \mathbf{s}_{kp}^f} &= \lambda_1 \mathbf{H}(\mathbf{S}) \frac{\partial \mathbf{H}(\mathbf{S})}{\partial \mathbf{s}_{kp}^f} + \\
& \sum_{i,j} \left\{ \mathbf{A}_{ij} [\mathbf{X}(\mathbf{S})_{ki} - \mathbf{X}(\mathbf{S})_{kj}] \left[ \frac{\partial \mathbf{X}(\mathbf{S})_{ki}}{\partial \mathbf{s}_{kp}^f} - \right. \right. \\
& \left. \left. \frac{\partial \mathbf{X}(\mathbf{S})_{kj}}{\partial \mathbf{s}_{kp}^f} \right] \right\} + \lambda_3 (\mathbf{W}^T \mathbf{D} - \mathbf{Y}) \frac{\partial \mathbf{D}(\mathbf{S})}{\partial \mathbf{s}_{kp}^f} + \\
& \lambda_4 (\mathbf{2Y} - \mathbf{I}) \frac{\partial \mathbf{D}(\mathbf{S})}{\partial \mathbf{s}_{kp}^f}
\end{aligned} \quad (21)$$

此外, 式(21)第一项中的  $\frac{\partial \mathbf{H}(\mathbf{S})}{\partial \mathbf{s}_{kp}^f}$  变为

$$\frac{\partial \mathbf{H}(\mathbf{S})}{\partial \mathbf{s}_{kp}^f} = -\frac{2}{\sigma^2} \mathbf{H}(\mathbf{S})_{ij} \hat{d}_{ij}^S \frac{\partial \hat{d}_{ij}^S}{\partial \mathbf{s}_{kp}^f} \quad (22)$$

式(22)中:  $\hat{d}_{ij}^S$  为  $S_i$  和  $S_j$  之间的距离;  $\hat{d}_{ij}^S$  和  $\frac{\partial \hat{d}_{ij}^S}{\partial \mathbf{s}_{kp}^f}$  的计算方法与  $\mathbf{X}(\mathbf{S})_{ij}$  和  $\frac{\partial \mathbf{X}(\mathbf{S})_{ij}}{\partial \mathbf{s}_{kp}^f}$  相同。式(21)的第二

项  $\frac{\partial \mathbf{X}(\mathbf{S})_{ij}}{\partial \mathbf{s}_{kp}^f}$  变为

$$\begin{aligned}
\frac{\partial \mathbf{X}(\mathbf{S})_{ij}}{\partial \mathbf{s}_{kp}^f} &= \frac{1}{E_2} \sum_{q=1}^{\bar{q}} \exp \left( \alpha \sum_f \beta_k^f d_{ijq}^f \right) \times \\
& [1 + \alpha \beta_k^f d_{knq}^f] E_1 - \alpha E_2 \beta_k^f \frac{\partial d_{ijq}^f}{\partial \mathbf{s}_{kp}^f}
\end{aligned} \quad (23)$$

$$E_1 = \sum_{q=1}^{\bar{q}} \beta_k^f d_{ijq}^f \exp \left( \alpha \sum_f \beta_k^f d_{ijq}^f \right) \quad (24)$$

$$E_2 = \sum_{q=1}^{\bar{q}} \exp \left( \alpha \sum_v \beta_k^f d_{ijq}^f \right) \quad (25)$$

式中:

$$\frac{\partial d_{ijq}^f}{\partial \mathbf{s}_{kp}^f} = \frac{2}{l_k} [\mathbf{S}_p^{f(i)} - m t_{q+p+1}^{f(j)}], \bar{q} = q_i - l_j + 1.$$

综上,  $\nabla \mathbf{S} = \frac{\partial F(\mathbf{S})}{\partial \mathbf{S}}$  可由式(21)~式(25)共同

计算得出。

### 2.2.3 固定 $\mathbf{A}$ 和 $\mathbf{S}$ 更新 $\mathbf{W}$

当固定  $\mathbf{A}$  和  $\mathbf{S}$  时, SMTA 的目标函数退化为

$$\min_{\mathbf{W}} F(\mathbf{W}) = \min_{\mathbf{W}} \left[ \frac{\lambda_2}{2} \left\| \mathbf{W} \right\|_F^2 + \frac{\lambda_3}{2} \left\| \mathbf{W}^T \mathbf{D}(\mathbf{S}) - \mathbf{Y} \right\|_F^2 \right] \quad (26)$$

然后求解  $\mathbf{W}$  的最小值, 对  $F(\mathbf{W})$  关于  $\mathbf{W}$  进行求导计算, 公式为

$$\begin{aligned}
\frac{\partial F(\mathbf{W})}{\partial \mathbf{W}} &= \lambda_2 \mathbf{W} + \lambda_3 \mathbf{D}(\mathbf{W}^T \mathbf{D} - \mathbf{Y})^T \\
&= (\lambda_2 \mathbf{I} + \lambda_3 \mathbf{D} \mathbf{D}^T) \mathbf{W} - \lambda_3 \mathbf{D} \mathbf{Y}^T
\end{aligned} \quad (27)$$

令其等于 0, 得到  $\mathbf{W}$  的最终结果为

$$\mathbf{W}_{i+1} = (\lambda_2 \mathbf{I} + \lambda_3 \mathbf{D}_i \mathbf{D}_i^T)^{-1} (\lambda_3 \mathbf{D}_i \mathbf{Y}_i^T) \quad (28)$$

## 2.3 算法整体步骤

提出的多元时序分类算法首先给出自适应邻居的 Shapelet 学习方法, 计算时间序列和 Shapelet 之间的距离, 并将其映射到 Shapelet 空间中, 使得获取的 Shapelet 更精确, 保证了可解释性的同时也提高了算法的效率; 然后, 结合 Shapelet 相似性和类标约束项, 从而获得信息更丰富的 Shapelets, 同时减少了数据之间的差异, 使其更加符合数据的分布, 这样有助于减小过拟合的风险, 提高模型的稳定性和泛化能力; 最后, 利用坐标下降法进行优化迭代更新, 从而得到最优的 Shapelets, 并以此为基础实施分类。提出的分类算法的伪代码如算法 1 所示。

**算法 1** 基于 Shapelet 的多元时序数据分类算法

输入: 多元时间序列  $T$ , 不同长度的多元 Shapelets, 迭代次数  $i_{\max}$ , 学习比率  $\eta$ , 参数  $\lambda_0, \lambda_1, \lambda_2, \lambda_3, \lambda_4, \alpha, \mu$

输出: 多元时间序列的分类结果  $T^l$

1. 初始化  $S, Y, W$ ;
2. 根据式(3) ~ 式(7), 计算  $T$  和  $S$  之间的距离, 并映射到 Shapelet 空间中得到  $X(S)$ ;
3. 根据  $S$  和式(10)计算  $H(S)$ ;
4. 通过式(16) ~ 式(17)更新  $A$ ;
5. 通过式(28)更新  $W$ ;
6. For  $i = 1, 2, \dots, i_{\max}$  do
7. 通过式(21)计算  $\nabla S_i$ ;
8. 通过公式  $S_{i+1} = S_i - \eta \nabla S_i$  来更新  $S_i$ ;
9. End for
10.  $S^{\text{best}} = S_{i_{\max}}$ ; ( $S^{\text{best}}$  为获得的最佳 Shapelets)
11. 计算  $S^{\text{best}}$  与时间序列的距离并进行排序, 确定  $S^{\text{best}}$  的类标, 获得带类标的 Shapelets  $S^l$ ;
12. 将得到的  $S^{\text{best}}$  与时间序列的距离按照类别数进行排序, 并根据  $S^l$  确定最终时间序列的类标, 获得带有分类标签的时间序列  $T^l$ ;

Return  $T^l$

**2.4 复杂度分析**

如算法 1 所述, 每次迭代计算矩阵  $X(S)$  和  $H(S)$  的计算复杂度分别为  $O(MQNKL)$  和  $O(M^2L^2K^2)$ , 其中  $M$  为多元时间序列中的变量数,  $Q$  为每个时间序列的最大长度,  $L$  为每个 Shapelet 的最大长度,  $N$  为指多元时间序列的大小,  $K$  为  $S$  的最大值。针对 SMTS 的学习过程中后续  $A, W, S$  的依次迭代更新, 通过固定  $S$  迭代更新  $A, W$  直到其收敛, 更新  $A$  和  $W$  的计算复杂性分别是  $O(N^2K + N^2C)$  和  $O(N^2K + K^3 + NKC + K^2C)$ , 其中  $C$  为类数。当更新  $A$  和  $W$  直到收敛时, 复杂度为  $O\{J[N(NC + K^2 + NK + KC) + K^2(K + C)]\}$ ,  $J$  为迭代的最大次数。当更新  $S$  直到收敛时, 复杂度为  $O[i_{\max}(M^2Q^2N^2 + M^2L^2K^2 + N^2K)]$ ,  $i_{\max}$  为最大的迭代次数。综上所述, 总复杂度为  $O(G\{MQNKL + M^2L^2K^2 + J[N(NC + K^2 + NK + KC) + K^2(K + C)] + i_{\max}(M^2Q^2N + M^2L^2K^2 + N^2K)\})$ , 其中  $G$  是 SMTS 的最大迭代次数。因为  $K, C \leq N$ , 所以最终 SMTS 的时间复杂度可简化成  $O[G(M^2Q^2N^2i_{\max})]$ 。

**3 实验与分析**

实验采用 3 GHz 和 16 GB 内存的 CPU, Window 10 操作系统, 编程环境为 MATLAB2021。提出算法在 11 个数据集上分别与 3 种不同类型(11 个)分类算法进行对比, 即基于距离的分类算法<sup>[32]</sup>、基于 Shapelet 的分类算法<sup>[33-36]</sup>、基于网络的

分类算法<sup>[37-40]</sup>, 以此验证本文算法的有效性。

**3.1 实验数据集**

利用 11 个真实的多元时间序列数据集<sup>[41]</sup> 评估所提出的模型, 数据集分别为 Atrial Fibrillation (AF)、Basic Motions (BM)、Epilepsy (Ep)、Ethanol Concentration (EC)、Finger Movements (FM)、Hand Movement Direction (HMD)、Racket Sports (RS)、Heartbeat (Hb)、Self Regulation SCP1 (SRS1)、Stand Walk Jump (SWJ)、Self Regulation SCP2 (SRS2), 具体信息如表 1 所示。数据的维度为 2 ~ 61, 长度为 30 ~ 2 500。对部分数据集进行了预处理, 并利用训练集调整参数, 利用测试集进行评估。

表 1 UEA 数据集

Table 1 UEA data set

| 数据集  | 训练集 | 测试集 | 序列数 | 序列长度  | 类别数 |
|------|-----|-----|-----|-------|-----|
| AF   | 15  | 15  | 2   | 640   | 3   |
| BM   | 40  | 40  | 6   | 100   | 4   |
| Ep   | 137 | 138 | 3   | 206   | 4   |
| EC   | 261 | 263 | 3   | 1 751 | 4   |
| FM   | 316 | 100 | 28  | 50    | 2   |
| HMD  | 160 | 74  | 10  | 400   | 4   |
| Hb   | 204 | 205 | 61  | 405   | 2   |
| RS   | 151 | 152 | 6   | 30    | 4   |
| SRS1 | 268 | 293 | 6   | 896   | 2   |
| SRS2 | 200 | 180 | 7   | 1 152 | 2   |
| SWJ  | 12  | 15  | 4   | 2 500 | 3   |

**3.2 参数设置**

有两种不同类型的多元 Shapelets: 相同长度的多元 Shapelets 和不同长度的多元 Shapelets。在实验中, 正则化参数为  $\lambda_0, \lambda_1, \lambda_2, \lambda_3$  和  $\lambda_4$ , 从  $\{10^{-4}, 10^{-2}, 10^0, 10^2, 10^4\}$  中确定, 参数  $\mu$  从  $\{0, 0.5, 1\}$  中选取, 内部迭代次数  $i_{\max}$  设为 50, 学习率设为 0.01。

**3.3 准确率比较**

3.3.1 基于距离的分类算法准确率对比

提出算法与 3 种基于距离的算法  $ED_1, DTW_1$  和  $DTW_D$ <sup>[36]</sup>, 在 UEA 的 11 个多元时间序列数据集上进行比较, 实验对比结果如表 2 所示。表 2 表明提出算法在除 Ep 以外的 10 个数据集上得到最优的分类结果, 且在 AF、BM 等数据集上均得到 100% 的准确率。

提出算法仅在 Ep 上稍有欠缺, 因为该数据集的规律性不强, 每个类别的个数不一致。但是整体的实验结果表明, 在基于距离的分类算法中提出算法在准确度方面具有明显优势。

表2 基于距离的分类算法准确率对比

Table 2 Comparison on accuracy of classification algorithms based on distance

| 数据集  | 准确率/%           |                  |                  |              |
|------|-----------------|------------------|------------------|--------------|
|      | ED <sub>1</sub> | DTW <sub>1</sub> | DTW <sub>D</sub> | SMTS         |
| AF   | 0.167           | 0.267            | 0.200            | <b>1.000</b> |
| BM   | 1.000           | 1.000            | 0.975            | <b>1.000</b> |
| Ep   | 0.564           | <b>1.000</b>     | 0.964            | 0.920        |
| EC   | —               | 0.361            | 0.323            | <b>0.844</b> |
| FM   | —               | 0.489            | 0.530            | <b>0.730</b> |
| HMD  | —               | 0.210            | 0.206            | <b>0.716</b> |
| Hb   | 0.683           | 0.500            | 0.604            | <b>0.776</b> |
| RS   | 0.869           | 0.891            | 0.818            | <b>0.921</b> |
| SRS1 | 0.841           | 0.806            | 0.775            | <b>0.997</b> |
| SRS2 | 0.447           | 0.489            | 0.539            | <b>1.000</b> |
| SWJ  | 0.333           | 0.333            | 0.200            | <b>1.000</b> |

注:加粗字体表示分类最高准确率。

3.3.2 基于 Shapelet 的分类算法准确率对比

提出算法与基于 Shapelet 的算法 ShapeNet<sup>[37]</sup>, gRFS<sup>[38]</sup>, Shapelet\_D-S<sup>[39]</sup>, Multi-shapelet<sup>[40]</sup> 进行比较的结果如表3所示。可以看出,提出算法在大部分数据集均取得最好的结果。在 Hb 和 RS 数据集上,Shapelet\_D-S 算法仅以微小的优势胜过本文算法,但是从整体上看,利用本文方法获得的 Shapelet 集合具有显著性,使得时间序列基于 Shapelet 分类算法的结果在这 11 个数据集上具有明显优势。

3.3.3 基于网络的分类算法准确率对比

提出算法与基于网络的分类算法 TapNet<sup>[42]</sup>, SMATE<sup>[43]</sup>, DA-Net<sup>[44]</sup>, MF-Net<sup>[45]</sup> 准确率对比的结果如表4所示,可以看出,本文算法的准确率在6个数据集上均为最高,在4个数据集上达到100%的准确率,在与最先进的 MF-Net 算法进行对比时,提出算法也以明显优势获胜,这不仅表明提出算法的优越性,同时进一步显示出本文算法有效地提高了时间序列分类精度。

根据不同类型算法准确率列出统计图如图2所示。条形柱代表在同种类型中不同算法在11个数据集上胜出的个数。可清楚看出,本文算法在准确率上具有优势,在3种类型中均取得了突出成果。

进一步,对算法进行了 Nemenyi 非参数统计检验,并绘制了关键差异图,结果如图3和图4所示。首先,将所有算法在6个公共数据集上进行 Nemenyi 非参数统计检验并绘制关键差异图,结果如图3所示。可以看出,本文算法的性能与其他分类器相比达到了最高的准确率。然后,将基于距离的分类算法和基于 Shapelet 的分类算法在11个数据集上进行统计检验,得到的关键差异图如图4所示。图4表明提出算法的分类结果最优,体现了提出算法的先进性和优越性,在与 Shapelet-D\_S 的对比中也以0.1的优势获胜。

表3 基于 Shapelet 的分类算法准确率对比

Table 3 Comparison on accuracy of classification algorithms based on Shapelet

| 数据集  | 准确率/%        |              |              |                |              |
|------|--------------|--------------|--------------|----------------|--------------|
|      | ShapeNet     | gRFS         | Shapelet_D-S | Multi-Shapelet | SMTS         |
| AF   | 0.167        | 0.267        | 0.550        | 0.500          | <b>1.000</b> |
| BM   | <b>1.000</b> | <b>1.000</b> | <b>1.000</b> | <b>1.000</b>   | <b>1.000</b> |
| Ep   | 0.982        | 0.979        | <b>1.000</b> | 0.982          | 0.920        |
| EC   | —            | 0.346        | 0.725        | —              | <b>0.844</b> |
| FM   | —            | 0.582        | 0.642        | —              | <b>0.730</b> |
| HMD  | —            | 0.431        | 0.504        | —              | <b>0.716</b> |
| Hb   | 0.756        | 0.640        | <b>0.802</b> | 0.781          | 0.776        |
| RS   | 0.875        | 0.891        | <b>0.935</b> | 0.918          | 0.921        |
| SRS1 | 0.867        | 0.823        | 0.900        | 0.884          | <b>0.997</b> |
| SRS2 | 0.789        | 0.517        | 0.643        | 0.868          | <b>1.000</b> |
| SWJ  | 0.400        | 0.333        | 0.450        | 0.833          | <b>1.000</b> |

注:加粗字体表示分类最高准确率。

表4 基于网络的分类算法准确率对比

Table 4 Comparison of accuracy of classification algorithms based on network

| 数据集  | 准确率/%        |              |        |        |              |
|------|--------------|--------------|--------|--------|--------------|
|      | TapNet       | SMATE        | DA-Net | MF-Net | SMTS         |
| AF   | 0.333        | 0.133        | 0.414  | 0.466  | <b>1.000</b> |
| BM   | <b>1.000</b> | <b>1.000</b> | 0.925  | 0.950  | <b>1.000</b> |
| HMD  | 0.365        | 0.527        | 0.347  | 0.445  | <b>0.716</b> |
| Hb   | 0.727        | 0.727        | 0.626  | 0.692  | <b>0.776</b> |
| SRS2 | 0.550        | 0.556        | 0.561  | 0.533  | <b>1.000</b> |
| SWJ  | 0.400        | 0.200        | 0.400  | 0.400  | <b>1.000</b> |

注:加粗字体表示分类最高准确率。

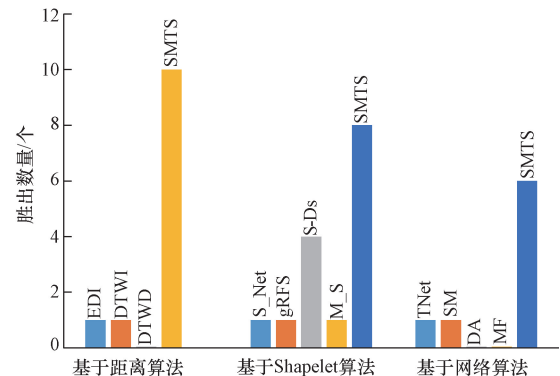
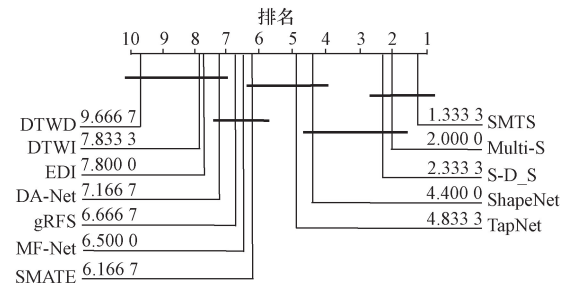
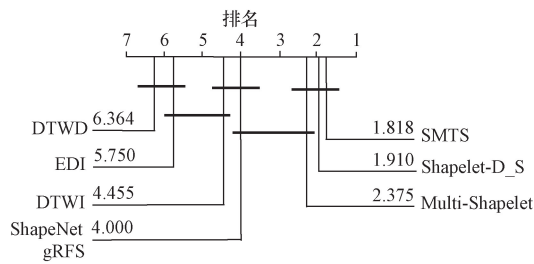


图2 不同类型算法准确率对比  
Fig. 2 Accuracy comparison of different types of algorithms



水平连线表示同一分组内分类器之间没有显著差异  
图3 基于算法的关键差异图  
Fig. 3 Key difference graph based on algorithm



水平连线表示同一分组内分类器之间没有显著差异

图4 基于数据集的关键差异图

Fig. 4 Key difference graph based on data set

## 4 结论

提出基于 Shapelets 的多元时间序列分类方法。首先,利用自适应邻居的无监督 Shapelet 学习方法将 Shapelet 变换和自适应权重结合,并用于自动学习显著多元 Shapelets。通过自适应权重的计算,得到不同 Shapelet 的重要性,进而提高模型的分类精度。然后,该方法与 Shapelet 相似性和类标约束项相结合,通过 Shapelet 相似性可以增强模型可解释性,类标约束项使得实际类标和生成类标的差异最小化,提高分类的准确性。最后,提出模型的优化策略来获取最优的且不同长度的多元 Shapelets。

将提出算法在 UEA 的 11 个数据集上进行验证,并且分别针对 3 种不同类型 11 个算法进行对比,得到以下结论。

(1) 在与基于距离的 3 个算法进行比较时,提出算法的优势十分明显。基于 Shapelet 分类也会对数据进行降维,减少了分类的复杂度,同时也提高了分类的精度。

(2) 在与基于 Shapelet 的 4 个算法进行对比时,提出算法在大部分数据集上的结果均为最佳,整体上优于其他先进算法的分类效果,这体现了本文算法的先进性和有效性。

(3) 在与基于网络的 4 个算法进行对比时,提出算法在 6 个数据集上均为最佳结果,甚至在 4 个数据集上均取得 100% 的准确率,展现了本文算法的显著优势。

准确率对比图和关键差异图进一步表明本文算法的优越性。本文算法相比于其他 3 种类型的算法,在分类精度上取得了明显的优势,同时也表明在获取 Shapelet 集合的可解释性和分类准确性的效果十分显著。未来将探索更高层次的 Shapelet 集合,使得模型获得更好的效率和分类精度。

### 参考文献

[1] Feremans L, Cule B, Goethals B. PETSC: pattern-based embedding for time series classification[J]. Data Mining and Knowledge

Discovery, 2022, 36(3): 1015-1061.

[2] 张旭, 张亮, 金博, 等. 基于不确定性的多元时间序列分类算法研究[J]. 自动化学报, 2023, 49(4): 790-804.

Zhang Xu, Zhang Liang, Jin Bo, et al. Research on classification algorithm of multivariate time series based on uncertainty[J]. Acta Automatica Sinica, 2023, 49(4): 790-804.

[3] Velasco I, Sipols A, De Blas C S, et al. Motor imagery EEG signal classification with a multivariate time series approach[J]. BioMedical Engineering OnLine, 2023, 22(1): 1-24.

[4] Lee J, Youn H L, Poon J, et al. StockEmotions: discover investor emotions for financial sentiment analysis and multivariate time series[J]. ArXiv Preprint ArXiv, 2023: 2301.09279.

[5] Yan Y, Zhang S, Tang J, et al. Understanding characteristics in multivariate traffic flow time series from complex network structure[J]. Physica A: Statistical Mechanics and its Applications, 2017, 477: 149-160.

[6] Rahman M M, Farahani M A, Wuest T. Multivariate time-series classification of critical events from industrial drying hopper operations: a deep learning approach[J]. Journal of Manufacturing and Materials Processing, 2023, 7(5): 164.

[7] Gupta A, Gupta H P, Biswas B, et al. A fault-tolerant early classification approach for human activities using multivariate time series[J]. IEEE Transactions on Mobile Computing, 2020, 20(5): 1747-1760.

[8] Zhang N, Sun S. Multiview unsupervised shapelet learning for multivariate time series clustering[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2022, 45(4): 4981-4996.

[9] Orsenigo C, Vercellis C. Combining discrete SVM and fixed cardinality warping distances for multivariate time series classification[J]. Pattern Recognition, 2010, 43(11): 3787-3794.

[10] Cheng Z, Yang Y, Wang W, et al. Time2graph: revisiting time series modeling with dynamic Shapelets[C]//Proceedings of the AAAI Conference on Artificial Intelligence. Menlo Park: AAAI, 2020: 3617-3624.

[11] Cho S, Chang W, Lee G, et al. Interpreting internal activation patterns in deep temporal neural networks by finding prototypes[C]//Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery & Data Mining. New York: Association for Computing Machinery, 2021: 158-166.

[12] Wu G, Zhang H, He Y, et al. Learning Kullback-Leibler divergence-based Gaussian model for multivariate time series classification[J]. IEEE Access, 2019, 7: 139580-139591.

[13] Berndt D J, Clifford J. Using dynamic time warping to find patterns in time series[C]//Proceedings of the 3rd International Conference on Knowledge Discovery and Data Mining. Palo Alto: AAAI, 1994: 359-370.

[14] Zhao B, Xing H, Wang X, et al. Classification-oriented distributed semantic communication for multivariate time series[J]. IEEE Signal Processing Letters, 2023, 30: 369-373.

[15] Le X M, Tran M T, Huynh V N. Learning perceptual position-aware shapelets for time series classification[C]//Joint European Conference on Machine Learning and Knowledge Discovery in Databases. Cham: Springer Nature Switzerland, 2022: 53-69.

[16] 李可, 燕晗, 顾杰斐, 等. 基于 Shapelets 时间序列的多源迁移学习滚动轴承故障诊断方法[J]. 中国机械工程, 2022, 33(24): 2990-2996.

Li Ke, Yan Han, Gu Jiefei, et al. Multi-source transfer learning

- rolling bearing fault diagnosis method based on Shapelets time series [J]. *China Mechanical Engineering*, 2022, 33 (24): 2990-2996.
- [17] Li G, Yan W, Wu Z. Discovering Shapelets with key points in time series classification [J]. *Expert Systems with Applications*, 2019, 132: 76-86.
- [18] Hssayni E H, Joudar N E, Ettaouil M. A deep learning framework for time series classification using normal cloud representation and convolutional neural network optimization [J]. *Computational Intelligence*, 2022, 38(6): 2056-2074.
- [19] Liu C L, Hsaio W H, Tu Y C. Time series classification with multivariate convolutional neural network [J]. *IEEE Transactions on Industrial Electronics*, 2018, 66(6): 4788-4797.
- [20] Liu M, Ren S, Ma S, et al. Gated transformer networks for multivariate time series classification [J]. *ArXiv Preprint ArXiv*, 2021: 2103. 14438.
- [21] 孙其法. 基于多样化 Shapelets 的不平衡时间序列分类方法 [J]. *信息与电脑(理论版)*, 2021, 33(18): 45-47.  
Sun Qifa. Unbalanced time series classification method based on diversified Shapelets [J]. *Information and Computers (Theory)*, 2021, 33(18): 45-47.
- [22] Ye L, Keogh E. Time series Shapelets: a new primitive for data mining [C]//*International Conference on Knowledge Discovery and Data Mining*. New York: ACM, 2009: 947-956.
- [23] Liang Z, Wang H. FedST: federated Shapelet transformation for interpretable time series classification [J]. *ArXiv Preprint ArXiv*, 2023: 2302. 10631.
- [24] Yang J, Jing S, Huang G. Accurate and fast time series classification based on compressed random shapelet forest [J]. *Applied Intelligence*, 2023, 53(5): 5240-5258.
- [25] Lines J, Davis L M, Hills J, et al. A Shapelet transform for time series classification [C]//*Proceedings of the 18th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. New York: Association for Computing Machinery, 2012: 289-297.
- [26] Grabocka J, Schilling N, Wistuba M, et al. Learning time-series shapelets [C]//*Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. New York: Association for Computing Machinery, 2014: 392-401.
- [27] Guillaume A, Vrain C, Elloumi W. Random dilated shapelet transform: a new approach for time series shapelets [C]//*International Conference on Pattern Recognition and Artificial Intelligence*. Cham: Springer International Publishing, 2022: 653-664.
- [28] Chen J, Wan Y. Localized Shapelets selection for interpretable time series classification [J]. *Applied Intelligence*, 2023, 53(14): 17985-18001.
- [29] Liu H Y, Gao Z Z, Wang Z H, et al. Time series classification with Shapelet and canonical features [J]. *Applied Sciences*, 2022, 12(17): 8685.
- [30] 王威娜, 胡佳利, 任艳. 基于优化 Shapelet 的时间序列分类方法 [J]. *科学技术与工程*, 2023, 23(8): 3345-3353.  
Wang Weina, Hu Jiali, Ren Yan. Time series classification method based on optimized Shapelet [J]. *Science Technology and Engineering*, 2023, 23(8): 3345-3353.
- [31] 李晨, 万源. 基于优化和两阶段筛选的时间序列 Shapelets 提取研究 [J]. *计算机科学*, 2023, 50(2): 146-157.  
Li Chen, Wan Yuan. Research on extraction of time series Shapelets based on optimization and two-stage screening [J]. *Computer Science*, 2023, 50(2): 146-157.
- [32] 杨骏, 敬思远, 钟勇. 面向时间序列有序分类的 Shapelet 抽取算法 [J]. *电子科技大学学报*, 2023, 52(6): 887-896.  
Yang Jun, Jing Siyuan, Zhong Yong. Shapelet extraction algorithm for ordered classification of time series [J]. *Journal of University of Electronic Science and Technology*, 2023, 52(6): 887-896.
- [33] 闫汶和, 李桂玲. 基于 Shapelet 的时间序列分类研究 [J]. *计算机科学*, 2019, 46(1): 29-35.  
Yan Wenhe, Li Guiling. Research on time series classification based on Shapelet [J]. *Computer Science*, 2019, 46(1): 29-35.
- [34] Zhang N, Sun S. Multiview unsupervised Shapelet learning for multivariate time series clustering [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2022, 45(4): 4981-4996.
- [35] Hills J, Lines J, Baranauskas E, et al. Classification of time series by Shapelet transformation [J]. *Data Mining and Knowledge Discovery*, 2014, 28: 851-881.
- [36] Shokoochi-Yekta M, Wang J, Keogh E J. On the non-trivial generalization of dynamic time warping to the multi-dimensional case [C]//*Proceedings of the 2015 SIAM International Conference on Data Mining*. Vancouver, Canada: SIAM, 2015: 289-297.
- [37] Li G, Choi B, Xu J, et al. Shapenet: a Shapelet-neural network approach for multivariate time series classification [C]//*Proceedings of the AAAI Conference on Artificial Intelligence*. Menlo Park: AAAI, 2021: 8375-8383.
- [38] Karlsson I, Papapetrou P, Boström H. Generalized random Shapelet forests [J]. *Data Mining and Knowledge Discovery*, 2016, 30(5): 1053-1085.
- [39] 宋奎勇, 王念滨, 王红滨. 基于 Shapelets 的多变量 DS 证据加权集成分类 [J]. *吉林大学学报(信息科学版)*, 2021, 39(2): 205-214.  
Song Kuiyong, Wang Nianbin, Wang Hongbin. Multivariate DS evidence weighted ensemble classification based on Shapelets [J]. *Journal of Jilin University (Information Science Edition)*, 2021, 39(2): 205-214.
- [40] 詹熙, 黎维, 潘志松. Multi-Shapelet: 一种基于 Shapelet 的多变量时间序列分类方法 [J]. *数据采集与处理*, 2023, 38(2): 386-400.  
Zhan Xi, Li Wei, Pan Zhisong. Multi-Shapelet: a multivariable time series classification method based on Shapelet [J]. *Data Acquisition and Processing*, 2023, 38(2): 386-400.
- [41] Bagnall A, Dau H A, Lines J, et al. The UEA multivariate time series classification archive, 2018 [J]. *ArXiv Preprint ArXiv*, 2018: 1811.00075.
- [42] Zhang X, Gao Y, Lin J, et al. Tapnet: multivariate time series classification with attentional prototypical network [C]//*Proceedings of the AAAI Conference on Artificial Intelligence*. Menlo Park: AAAI, 2020: 6845-6852.
- [43] Zuo J, Zeitouni K, Taher Y. Smate: semi-supervised spatio-temporal representation learning on multivariate time series [C]//*2021 IEEE International Conference on Data Mining (ICDM)*. Piscataway: IEEE, 2021: 1565-1570.
- [44] Chen R, Yan X, Wang S, et al. DA-Net: dual-attention network for multivariate time series classification [J]. *Information Sciences*, 2022, 610: 472-487.
- [45] Du M, Wei Y, Zheng X, et al. Multi-feature based network for multivariate time series classification [J]. *Information Sciences*, 2023, 639: 119009.