

DOI: 10.11992/tis.201709033

网络出版地址: <http://kns.cnki.net/kcms/detail/23.1538.TP.20180418.0946.006.html>

图正则化字典对学习的轻度认知功能障碍预测

魏彩锋^{1,2}, 孙永聪^{1,2}, 曾宪华^{1,2}

(1. 重庆邮电大学 计算机科学与技术学院, 重庆 400065; 2. 重庆邮电大学 计算智能重庆市重点实验室, 重庆 400065)

摘要: 针对字典对学习 (DPL) 方法只考虑了同类子字典的重构误差和不同类表示系数的稀疏性, 没有考虑图像间的几何近邻拓扑关系的问题。通过近邻保持使得在同类近邻投影系数之间的距离较小, 而不同类投影系数之间的距离大, 能够有效提高字典对学习算法的分类性能, 基于此提出了基于几何近邻拓扑关系的图正则化的字典对学习 (GDPL) 算法。在 ADNI1 数据集上对轻度认知功能障碍预测的实验表明, 使用 GDPL 算法学习的编码系数作为特征预测的准确率 (ACC) 和 ROC 曲线下的面积 (AUC) 比使用结合生物标志作为特征预测的准确率提高了 2%~6%, 使用 GDPL 算法比 DPL 算法的实验结果也有提高。

关键词: 图正则化; 字典对学习; 几何近邻关系; 图像分类; 轻度认知功能障碍预测

中图分类号: TP391; R749 文献标志码: A 文章编号: 1673-4785(2019)02-0369-09

中文引用格式: 魏彩锋, 孙永聪, 曾宪华. 图正则化字典对学习的轻度认知功能障碍预测[J]. 智能系统学报, 2019, 14(2): 369-377.

英文引用格式: WEI Caifeng, SUN Yongcong, ZENG Xianhua. Dictionary pair learning with graph regularization for mild cognitive impairment prediction[J]. CAAI transactions on intelligent systems, 2019, 14(2): 369-377.

Dictionary pair learning with graph regularization for mild cognitive impairment prediction

WEI Caifeng^{1,2}, SUN Yongcong^{1,2}, ZENG Xianhua^{1,2}

(1. College of Computer Science and Technology, Chongqing University of Posts and Telecommunication, Chongqing 400065, China; 2. Chongqing Key Laboratory of Computation Intelligence, Chongqing University of Posts and Telecommunications, Chongqing 400065, China)

Abstract: Aiming at dictionary pair learning (DPL) methods only consider the reconstruction error of a sub-dictionary from the same class and the sparseness of coefficients from different classes, and do not consider the geometric neighborhood topological relationships between images. To improve the classification ability of DPL algorithms, we propose a DPL with graph regularization (GDPL) algorithm based on geometric neighborhood topological relationships. This algorithm is based on the idea that keeping the neighborhood relationship makes the distance between the neighborhood projection coefficients of the same kind small, while the distance between projection coefficients of different kinds is large. Experiments on mild cognitive impairment prediction using the ADNI1 dataset show that the coding coefficient learned from the GDPL algorithm is 2%~6% higher than that which uses the combined biomarker as feature prediction, according to accuracy (ACC) and area under curve (AUC) metrics. Moreover, the experimental result obtained using GDPL is also better than that obtained using DPL algorithm.

Keywords: graph regularization; dictionary pair learning; geometric neighborhood relationship; image classification; mild cognitive impairment prediction

目前, 多媒体技术飞速发展, 图像数量呈指数

级增长, 图像分类技术也得到了飞速发展。图像分类方法主要包括分类器的设计和特征提取两个部分, 目前对图像分类的研究主要集中在分类器性能的改进和改善特征提取方法两方面。稀疏编码 (字典学习)^[1-10] 是图像分类的有效技术之一, 用

收稿日期: 2017-09-16. 网络出版日期: 2018-04-18.

基金项目: 国家自然科学基金项目 (61672120); 重庆市科委基础学科和前沿技术研究一般项目 (cstc2015jcyjA40036, cstc2014jcyjA40049).

通信作者: 曾宪华. E-mail: zengxh@cqupt.edu.cn.

于图像分类的字典学习算法也可以分为 2 类: 1) 学习字典把原始图像映射到有利于图像分类的空间上, 用传统的分类器去分类; 2) 用稀疏系数所表示出来的样本属性进行分类。

在用于图像分类的字典学习算法中, 字典和编码系数的设计从中起着关键作用^[3]。基于稀疏表示的分类算法^[4], 将整个训练样本作为字典, 利用测试样本在字典上的重构误差分类。在字典中的每个原子都是一个样本, 由于样本数量的限制, 字典原子中有相同或相似的, 还有一些则不包含在内造成分类错误。而 Aharon M 等提出的 KSVD 算法^[5]学习的字典具有自适应性, 能够更好对信号进行稀疏表示。虽然, 使用重构误差作为损失函数, 可以用于图像分类, 但是还有许多因素没有考虑到。Yang 等^[6]提出 Fisher 字典学习在自适应字典基础上, 考虑到相同的样本之间的类内散度较小, 类间散度大, 在目标函数上加入了识别系数项, 提高分类能力。Jiang 等^[7]提出的 LC-KVD 算法, 在 KSVD 算法^[5]基础上加入了标签约束项。在文献^[8]中, 由于分别学习特征影射矩阵和字典, 会影响特征影射矩阵的判别信息, 提出了同时学习投影矩阵和字典, 为了能够提取更多的判别信息, 加入图正则项。构建近邻图, 近邻且同类相似度为 1, 近邻但不同类相似度为 -1, 其他为 0。在文献^[9]中图正则化中的相似度矩阵定义为两近邻样本的相似度为两样本之间的欧氏距离, 不近邻样本之间的相似度为 0。使用不同的相似矩阵, 分类效果也不同。另一方面, 由于字典学习^[3-9]大多是使用编码系数的 0 或 1 范数作正则项, 在训练和测试阶段效率较低, 运行时间长。2014 年在 NIPS 上 Gu 等^[10]提出了字典对学习, 联合学习一个综合字典和分析字典, 用分析字典去分析编码, 由于综合字典的重构效果比较好, 使用综合字典去重构图像。它不仅减少了在训练和测试阶段的时间复杂度, 也提高了模型分类能力。但是, 字典对学习模型没有考虑图像的几何近邻拓扑关系。为了提高模型分类能力, 基于同类样本的系数之间的距离小, 不同类系数之间的距离大的思想, 提出了基于图正则化字典对学习 (GDPL) 算法。

在最近的综合研究^[11]中, 用十倍交叉验证的方法评估轻度认知功能障碍 (MCI) 到阿尔兹海默症 (AD) 的预测, 只有 4 种方法的实验结果比随机分类的结果好, 因此, 提出新的方法提高 MCI-AD 的预测是十分必要的。在图像分类中, 并不是所有的特征都与识别分类有关, 其中有一些是无关特征, 将这些特征用于分类会出现过拟合现象。由于稀疏表示中表示系数具有稀疏性, 将本

文提出的图正则化字典对学习算法用于轻度认知功能障碍预测。

1 相关工作

1.1 字典学习

设 $X = [X_1 \cdots X_k \cdots X_K] \in \mathbf{R}^{p \times n}$ 是一个 K 类 p 维的训练样本, 第 k 类样本 $X_k = [x_1 \cdots x_i \cdots x_{n_k}]$, $X_k \in \mathbf{R}^{p \times n_k}$, $k = 1, 2, \dots, K$ 是第 k 类训练 n_k 样本, $n = n_1 + n_2 + \dots + n_k + \dots + n_K$ 。字典学习模型分类任务是通过利用训练数据的类标签信息学习一个有效的数据表示模型去进行分类。在使用字典学习算法的图像分类研究中, 如何设计对特征提取有效的字典和编码系数是决定算法性能的关键因素^[1]。对于字典学习算法设计考虑的有三方面: 1) 稀疏表示残差 (重构误差) 小, 使样本在稀疏表示时尽可能的接近原始样本; 2) 对表示系数约束, 使表示系数稀疏; 3) 考虑能够更好提取更多的判别信息的判别项。用于图像分类的判别字典学习模型^[6-7, 12-13]可通常以表示为:

$$\min_{D,A} \|X - DA\|_F^2 + \lambda g(A) + \eta f(D, A, Y) \quad (1)$$

式中: λ 、 η 是常量 (平衡因子); Y 是样本 X 的标签向量; D 是学习的综合字典; A 是 X 在综合字典 D 上的编码系数矩阵。在训练模型时, 数据保真项 $\|X - DA\|_F^2$ 是为了保证字典 D 的表示能力, 使训练数据的重构误差最小, 重构出来的图像尽可能的接近原始样本。正则项用于约束编码系数的稀疏性, 其通常应用范数表示为:

$$g(A) = \|A\|_p \quad (2)$$

式中: $\|\bullet\|_p$ 是编码系数 A 正则项的常用形式 ($p < 2$), 使编码系数 A 在重建误差满足要求的情况下尽可能地稀疏。 $f(D, A, Y)$ 是字典学习用于分类的判别函数项, 保证 D 和 A 的判别能力。

1.2 字典对学习

在 1.1 中描述的判别式字典学习模型是学习一个共享字典, 对系数的约束都是使用 l_0 或者是 l_1 范数对编码系数进行系数正则化约束, 使训练模型在训练阶段和测试阶段效率降低。Gu 等^[10]为了提高模型在训练阶段和测试阶段的效率和整体的识别能力, 提出了联合学习一个综合字典和分析字典的字典对学习 (DPL), 不使用 l_0 或者是 l_1 范数对编码系数进行稀疏正则化, 而是找到一个分析字典 P , 使训练样本 X 在分析字典 P 上线性投影得出编码系数 A , 即 $A = PX$ 保持稀疏性, 重建误差与判别约束保证下使训练样本 X 的稀疏表示变的更加有效且解决了 out-of-sample 问题。字典对模型描述表示如下:

$$\{\mathbf{P}, \mathbf{D}\} = \min_{\mathbf{P}, \mathbf{D}} \|\mathbf{X} - \mathbf{D}\mathbf{P}\mathbf{X}\|_F^2 + f(\mathbf{D}, \mathbf{P}, \mathbf{X}, \mathbf{Y}) \quad (3)$$

式中: $f(\mathbf{D}, \mathbf{P}, \mathbf{X}, \mathbf{Y})$ 表示判别函数。 \mathbf{D} 和 \mathbf{P} 是一个字典对: 分析字典 \mathbf{P} 用于编码 \mathbf{X} , 综合字典 \mathbf{D} 用于重构 \mathbf{X} 。学习一个结构综合字典 $\mathbf{D} = [\mathbf{D}_1 \cdots \mathbf{D}_k \cdots \mathbf{D}_K]$ 和结构分析字典 $\mathbf{P} = [\mathbf{P}_1 \cdots \mathbf{P}_k \cdots \mathbf{P}_K]$, 其中 $\{\mathbf{D}_k, \mathbf{P}_k\}$ 是第 k 类的子字典对。对于分析字典 \mathbf{P} 使其满足第 i ($i \neq k$) 类样本在第 k 类的子字典 \mathbf{P}_k 上的投影系数为空, 即 $\mathbf{P}_k \mathbf{X}_i \approx 0, \forall k \neq i$, 也就是编码系数矩阵 $\mathbf{P}\mathbf{X}$ 接近块正交。对于合成字典 \mathbf{D} , 使其子字典 \mathbf{D}_k 与投影编码矩阵 $\mathbf{P}_k \mathbf{X}_k$ 能够很好的重构样本 \mathbf{X}_k , 要求字典对有最小的重构误差:

$$\min_{\mathbf{P}, \mathbf{D}} \sum_{k=1}^K \|\mathbf{X}_k - \mathbf{D}_k \mathbf{P}_k \mathbf{X}_k\|_F^2 \quad (4)$$

因此, DPL 模型的目标函数可以写为:

$$\{\mathbf{P}, \mathbf{D}\} = \min_{\mathbf{P}, \mathbf{D}} \sum_{k=1}^K \|\mathbf{X}_k - \mathbf{D}_k \mathbf{P}_k \mathbf{X}_k\|_F^2 + \lambda \|\mathbf{P}_k \bar{\mathbf{X}}_k\|_F^2 \quad (5)$$

s.t. $\|d_i\|_2 \leq 1$

式中: $\bar{\mathbf{X}}_k$ 是 \mathbf{X}_k 在整个训练集中的补集; λ 是一个标准常量; d_i 是合成字典的第 i 个原子。

2 图正则化字典对学习

字典对学习只考虑图像的重构误差和不同类的分析字典和样本的投影系数尽可能的小, 没有考虑原始样本的近邻几何结构, 引入图正则化, 使低维表示能够保留原始样本的近邻结构。

2.1 图正则化字典对学习

在流形学习中, 假设两个高维空间的相邻的两点, 在其对应的低维空间仍保持其在高维空间的几何近邻拓扑关系。为了使样本之间可区分, 并同时保持原始图像的几何近邻结构, 且使同类样本在分析字典 \mathbf{P} 上的投影系数之间距离小, 不同类之间的距离大。对编码系数的图正则项约束表示为:

$$f(\mathbf{X}) = \sum_{u=1}^n \sum_{v=1}^n \|\mathbf{P}x_u - \mathbf{P}x_v\|_2^2 S_{uv} \quad (6)$$

式中: n 是所有样本总数; u 和 v 是第 $u(v)$ 训练样

$$\begin{aligned} \{\mathbf{P}^*, \mathbf{D}^*\} &= \arg \min_{\mathbf{P}, \mathbf{D}} \alpha \sum_{k=1}^K (\|\mathbf{X}_k - \mathbf{D}_k \mathbf{A}_k\|_F^2 + \tau \|\mathbf{P}_k \mathbf{X}_k - \mathbf{A}_k\|_F^2 + \lambda \|\mathbf{P}_k \bar{\mathbf{X}}_k\|_F^2) + (1 - \alpha) \beta \sum_{u=1}^n \sum_{v=1}^n \|\mathbf{P}x_u - \mathbf{P}x_v\|_F^2 S_{uv} = \\ \arg \min_{\mathbf{P}, \mathbf{D}} \alpha \sum_{k=1}^K (\|\mathbf{X}_k - \mathbf{D}_k \mathbf{A}_k\|_F^2 + \tau \|\mathbf{P}_k \mathbf{X}_k - \mathbf{A}_k\|_F^2 + \lambda \|\mathbf{P}_k \bar{\mathbf{X}}_k\|_F^2) + (1 - \alpha) \beta \sum_{u=1}^n \left(\sum_{v=1}^n ((\mathbf{P}x_u)^2 - 2\mathbf{P}x_u \mathbf{P}x_v + (\mathbf{P}x_v)^2) S_{uv} \right) &= \\ \arg \min_{\mathbf{P}, \mathbf{D}} \alpha \sum_{k=1}^K (\|\mathbf{X}_k - \mathbf{D}_k \mathbf{A}_k\|_F^2 + \tau \|\mathbf{P}_k \mathbf{X}_k - \mathbf{A}_k\|_F^2 + \lambda \|\mathbf{P}_k \bar{\mathbf{X}}_k\|_F^2) + (1 - \alpha) \beta \sum_{u=1}^n c_u (\mathbf{P}x_u)^2 - \sum_{u=1}^n \sum_{v=1}^n \mathbf{P}x_u \mathbf{P}x_v S_{uv} &= \\ \arg \min_{\mathbf{P}, \mathbf{D}} \alpha \sum_{k=1}^K (\|\mathbf{X}_k - \mathbf{D}_k \mathbf{A}_k\|_F^2 + \tau \|\mathbf{P}_k \mathbf{X}_k - \mathbf{A}_k\|_F^2 + \lambda \|\mathbf{P}_k \bar{\mathbf{X}}_k\|_F^2) + (1 - \alpha) \beta (\mathbf{P}\mathbf{X})\mathbf{C}(\mathbf{P}\mathbf{X})^T - (\mathbf{P}\mathbf{X})\mathbf{S}(\mathbf{P}\mathbf{X})^T &= \\ \arg \min_{\mathbf{P}, \mathbf{D}} \alpha \sum_{k=1}^K (\|\mathbf{X}_k - \mathbf{D}_k \mathbf{A}_k\|_F^2 + \tau \|\mathbf{P}_k \mathbf{X}_k - \mathbf{A}_k\|_F^2 + \lambda \|\mathbf{P}_k \bar{\mathbf{X}}_k\|_F^2) + (1 - \alpha) \beta \text{tr}(\mathbf{P}\mathbf{X}\mathbf{L}\mathbf{X}^T \mathbf{P}^T) & \end{aligned} \quad (9)$$

本; $\mathbf{P}x_u$ 和 $\mathbf{P}x_v$ 是训练样本 x_u 和 x_v 在分析字典 \mathbf{P} 的投影, 即分析编码系数。

在大多数图正则项, 都是对数据集构建近邻图^[9], 用图的边的权值表示数据之间的关联度。但是本文利用样本总数去平衡类内和类间的距离, 已经证明能够提高在字典上的编码系数的判别能力^[12], 因此, 本文中相似度矩阵 \mathbf{S} 是依据 x_u 和 x_v 的标签保证相应系数之间的相似性^[12]。用训练集构建近邻图, 用标签信息判读样本是否近邻。如果 x_u 和 x_v 标签相同, 则为近邻, 相似度为 $S_{uv} = 1/n_{l(x_u)}$; 如果 x_u 和 x_v 标签不同, 则不近邻, 相似度为 $S_{uv} = -1/(n - n_{l(x_u)})$ 。相似度矩阵 \mathbf{S} 的第 u 行 v 列的元素为:

$$S_{uv} = \begin{cases} 1/n_{l(x_u)}, & l(x_u) = l(x_v) \\ -1/(n - n_{l(x_u)}), & \text{otherwise} \end{cases} \quad (7)$$

式中: $l(x_u)$ 和 $l(x_v)$ 表示的是样本 x_u 和 x_v 的类标签; $n_{l(x_u)}$ 是训练样本中标签为 $l(x_u)$ 的样本总数。矩阵 $\mathbf{L} = \mathbf{C} - \mathbf{S}$, 其中 $\mathbf{C} = \text{diag}\{c_1, c_2, \dots, c_n\}$ 是对角矩阵, 且 $c_v = \sum_{v=1}^n S_{uv}$ 。

在 DPL 算法的基础上, 构建近邻图保持原始图像间的近邻关系, 由此提出了图正则化字典对学习 (GDPL) 算法。GDPL 算法的目标函数可写为:

$$\{\mathbf{P}^*, \mathbf{D}^*\} = \arg \min_{\mathbf{P}, \mathbf{D}} \sum_{k=1}^K (\|\mathbf{X}_k - \mathbf{D}_k \mathbf{P}_k \mathbf{X}_k\|_F^2 + \lambda \|\mathbf{P}_k \bar{\mathbf{X}}_k\|_F^2) + \beta \sum_{u=1}^n \sum_{v=1}^n \|\mathbf{P}x_u - \mathbf{P}x_v\|_2^2 S_{uv} \quad (8)$$

式中: λ 和 β 是平衡因子, 在公式 (9) 中的第 1 项是重构误差项, 重构误差项尽可能的小, 使重构出来的图像与原图像尽可能的接近。 $\bar{\mathbf{X}}_k$ 是整个训练样本上的补集, 第 2 项是第 K 类以外的训练样本, 在第 K 类的分析字典上的编码系数尽可能小, 使编码系数尽可能的稀疏。第 3 项是图正则项, 使图像在字典上的编码系数仍旧保持原始图像的近邻关系。

引入变量编码系数 \mathbf{A} , 使 $\mathbf{A} = \mathbf{P}\mathbf{X}$, 目标函数可以写为下面式子:

式中: $\text{tr}(\bullet)$ 表示矩阵的迹。矩阵 $L = C - S$, $C = \text{diag}\{c_1, c_2, \dots, c_n\}$ 该对角矩阵的对角元素是 S 的行(列)的元素之和。

2.2 图正则化字典对学习

在流形学习中,假设两个高维空间的相邻的两随机初始化分析字典 P 和综合字典 D , 之后,迭代更新编码系数 A 和 $\{P, D\}$ 。

1) 固定综合字典 D_k 和分析字典 P_k , 更新编码系数 A_k , 由于含有 A_k 的式子前面有相同的系数 α , 在计算的过程中可以消去, 所以可以利用式 (10) 解出 A_k :

$$A_k = (D_k^T D_k + \tau I)^{-1} (D_k^T X_k + \tau P_k X_k) \quad (10)$$

2) 固定编码系数 A_k , 更新综合字典 D_k 和分析字典 P_k 。

$$\begin{cases} P_k = \arg \min_{P_k} \alpha \sum_{k=1}^K \tau \|P_k X_k - A_k\|_F^2 + \lambda \|P_k X_k\|_F^2 + \beta \text{tr}(P_k X L X^T P_k^T) \\ D_k = \arg \min_{D_k} \alpha \sum_{k=1}^K \|X_k - D_k A_k\|_F^2 \\ \text{s.t. } \|d_i\|_2 \leq 1 \end{cases} \quad (11)$$

对 P_k 封闭求解, 即

$$P_k = \arg \min_{P_k} \alpha \sum_{k=1}^K \tau \|P_k X_k - A_k\|_F^2 + \lambda \|P_k X_k\|_F^2 + (1 - \alpha) \beta \text{tr}(P_k X L X^T P_k^T)$$

对 P_k 求导, 得

$$\frac{\partial}{\partial P_k} = 2P_k(\alpha \tau X_k X_k^T + \alpha \lambda \bar{X}_k \bar{X}_k^T + (1 - \alpha) \beta X L X^T + \gamma I) - 2\alpha \tau A_k X_k^T$$

令 $\frac{\partial}{\partial P_k} = 0$, 得

$$P_k^* = \alpha \tau A_k X_k^T (\alpha \tau X_k X_k^T + \alpha \lambda_3 \bar{X}_k \bar{X}_k^T + \gamma I + (1 - \alpha) \beta X L X^T)^{-1} \quad (12)$$

式中: $\gamma = 10e^{-4}$, I 是单位矩阵, 在公式加入 (一个极小的常数) 是为了保证逆运算有解。

在公式 (13) 中用交替方向乘子算法 (AMMD)^[8] 对 D 优化求解。具体的求解过程如下所示:

$$\begin{cases} D^{(r+1)} = \arg \min_D \sum_{k=1}^K \|X_k - D_k A_k\|_F^2 + \rho \|D_k - S_k^{(r)} + K_k^{(r)}\|_F^2 \\ S^{(r+1)} = \arg \min_S \sum_{k=1}^K \rho \|D_k^{(r+1)} - S_k + T_k^{(r)}\|_F^2 \text{ s.t. } \|s_i\|_2 \leq 1 \\ T^{(r+1)} = T^{(r)} + D_k^{(r+1)} - S_k^{(r+1)} \end{cases} \quad (13)$$

对 D_k 求导:

$$\begin{aligned} \frac{\partial}{\partial D_k} &= 2D_k(A_k A_k^T + \rho I) - 2(X_k A_k^T + \rho(S_k^r - T_k^r)) \\ \text{令 } \frac{\partial}{\partial D_k} &= 0 \text{ 得:} \\ D_k &= (X_k A_k^T + \rho(S_k^r - T_k^r))(A_k A_k^T + \rho I)^{-1} \end{aligned} \quad (14)$$

对 S_k 求导:

$$\frac{\partial}{\partial S_k} = -2\rho(D_k^{(r+1)} + T_k^{(r)}) + 2\rho S_k$$

令 $\frac{\partial}{\partial S_k} = 0$ 得:

$$S_k = D_k^{(r+1)} + T_k^{(r)} \quad (15)$$

2.3 图正则化字典对学习训练算法

由于算法中的图正则项, 需要先用训练样本构建近邻图, 计算出相似矩阵 S 和拉普拉斯矩阵 L 。利用目标函数优化过程进行迭代更新, 学习字典对 $\{P, D\}$ 。图正则化字典对学习算法的训练过程在算法 1 中详细描述。

算法 1 GDPL 训练算法

输入 训练样本 $X = [X_1 X_2 \dots X_K]$, 训练样本标签信息 $Y = [Y_1 Y_2 \dots Y_K]$, 参数 $\alpha, \tau, \lambda, \beta$, 子字典原子数 m 。最大迭代次数 \max 。

1) 利用样本标签信息构建近邻图。根据样本标签信息判断样本是否为近邻点。若 x_u 和 x_v 标签相同, 则为近邻点, 相似度 $S_{uv} = 1/n_{l(x_u)}$; 若 x_u 和 x_v 标签不相同, 则不是近邻点, 相似度 $S_{uv} = -1/(n - n_{l(x_u)})$ 。对角矩阵 $C = \text{diag}\{c_1, c_2, \dots, c_n\}$, 其中 $c_v = \sum_{v=1} S_{uv}$, 拉普拉斯矩阵 $L = C - S$ 。

2) 初始化: 随机生成综合字典 D_0 和分析字典 P_0 , 迭代次数 $t = 0$ 。

3) For $i = 1 : K$

While 未收敛 (具体的在收敛性分析中) 执行下面步骤:

1) 固定字典 D_k 和 P_k , 由目标函数中含有 A_k 的式子求解得出式 (10), 根据式 (10), 更新编码系数 A_k ;

2) 固定编码系数 A_k , 由式 (11) 中 D_k 使用 ADMM 算法得出式 (14) 更新字典 D_k ; 在 ADMM 算法中 D_k, S_k, T_k 根据式 (14)、(15)、(13) 中最后一个式子更新直到达到 ADMM 算法的收敛条件停止。

3) 固定编码系数 A_k , 由式 (11) 中 P_k 解出式 (12), 根据式 (12) 更新字典 P_k ;

End While

End For

输出 综合字典 D , 分析字典 P 。

2.4 分类方法

在图像分类中, 并不是所有的特征都与识别分类有关, 其中有一些是无关特征, 将这些特征用于分类会出现过拟合现象。另一方面, 稀疏表示是在超完备字典中用尽可能少的原子去表示原来的图像信号, 其中的表示系数具有稀疏性。用表示系数作为特征去分类, 可以减少无关特征的影

响, 提高模型的判别能力。用 ADNI1 数据集进行验证, 具体的实验结果在第 4 部分中, 使用编码系数用分类器分类效果比直接使用分类器^[14]的效果好。

在分类模型中, 输入训练样本 X 使用算法 1 训练出来的综合字典 D 和分析字典 P , 由 A , 计算出训练集在分析字典上的投影系数即稀疏表示中的编码系数 $A_{train} = PX$, 由测试样本 X_{test} 和分析字典 P , 可以计算出测试集的编码系数 $A_{test} = PX_{test}$ 。将编码系数作为特征, 用分类器对测试样本进行分类。

3 算法分析

本节主要对提出的图正则化字典对算法 (GDPL) 进行时间复杂度分析和收敛性分析, 该算法通过联合学习一个综合字典和分析字典, 将样本在分析字典上的投影作为表示系数, 减小了算法的时间复杂度。

3.1 时间复杂度分析

在 GDPL 算法中主要是计算训练阶段的时间复杂度。在训练阶段 A_k 和 $\{P_k, D_k\}$ 交替更新, 每次更新过程相同, 在此以一次更新为例。固定 $\{P_k, D_k\}$ 由式 (10) 更新 A_k , A_k 更新的时间复杂度是一个逆矩阵和一个矩阵相乘的时间复杂度即 $m \times m$ 的矩阵, $\{D_k^T D_k + \tau I\}$ 的时间复杂度为 $O(m^3)$, 在 $m \times n_k$ 的矩阵 $\{\tau P_k X_k + D_k^T X_k\}$ 中, $P_k X_k$ 的时间复杂度为 $O(m p n_k)$, $D_k^T X_k$ 的时间复杂度也是 $O(m p n_k)$, 所以矩阵 $\{\tau P_k X_k + D_k^T X_k\}$ 的时间复杂度为 $O(m p n_k)$ 。矩阵 $(D_k^T D_k + \tau I)^{-1}$ 和 $(\tau P_k X_k + D_k^T X_k)$ 相乘的时间复杂度为 $O(m^2 n_k)$, 整个更新 A_k 的时间复杂度为 $O(m^3 + m p n_k + m^2 n_k)$ 。固定 A_k , 更新 $\{P_k, D_k\}$ 。 P_k 由式 (12) 进行更新, 在式 (14) 中 $p \times p$ 的矩阵 $(\tau X_k X_k^T + \lambda_3 \bar{X}_k \bar{X}_k^T + \gamma I + \beta X L X^T)$ 的时间复杂度为 $O(p^2 n_k + p^2(n - n_k) + p n^2 + p^2 n)$, 但是在迭代过程中该矩阵值不变, 在初始化阶段提前计算出来, 在迭代过程中不需要重复计算, 因此不影响更新 P_k 的时间复杂度。只需计算 $(\tau X_k X_k^T + \lambda_3 \bar{X}_k \bar{X}_k^T + \gamma I + \beta X L X^T)$ 逆矩阵的时间复杂度为 $O(p^3)$ 。 $A_k X_k^T$ 的时间复杂度为 $O(m n p)$, 两矩阵相乘的时间复杂度为 $O(m p^2)$, 更新 P_k 的时间复杂度为 $O(m n p + p^3 + m p^2)$ 。使用 ADMM 更新 D_k , 将 ADMM 算法中迭代更新次数记为 H , ADDM 算法中, 主要计算在式 (13) 中 D 更新过程的时间复杂度。 D_k 根据式 (14) 更新的, 式 (14) 中 $X_k A_k^T$ 、 $A_k A_k^T$ 、 $(A_k A_k^T + \rho I)$ 的逆矩阵、 $X_k A_k^T (A_k A_k^T + \rho I)^{-1}$ 的时间复杂度分别为 $O(p n_k m)$ 、 $O(m^2 n_k)$ 、 $O(m^3)$ 、 $O(p m^2)$, 更新 D_k 整个过程的时间复杂度为 $O(H(p m n_k + m^2 n_k + m^3 + p m^2))$ 。

3.2 收敛性分析

在目标函数式 (1) 中对于 $\{(D, P), (A)\}$ 的迭代更新是一个双凸问题。在优化求解的过程中, 固定 A 对 D 和 P 求解是一个凸问题; 固定 D 和 P 对 A 求解也是一个凸问题, 因此目标函数是一个双凸问题。根据文献[10]的补充材料, 收敛时目标函数的能量值达到一个局部最小值。为了避免在更新过程中达到收敛条件, 不同类之间单个样本的能量不在同一个数量级上, 所以, 本文的收敛条件是单个样本的能量到达某一局部最小。目标函数中各项的曲线如图 1 所示。图 1 只是以 GDPL 算法过程中使用其中一类迭代过程中的目标函数值画出曲线为例。算法 1 中 while 后面的收敛条件为前后两次迭代的整个目标函数值之差小于 0, 单个样本的重构误差的绝对值小于 1, 保证目标函数值在每次迭代过程中逐渐减小, 最大迭代次数为 50。

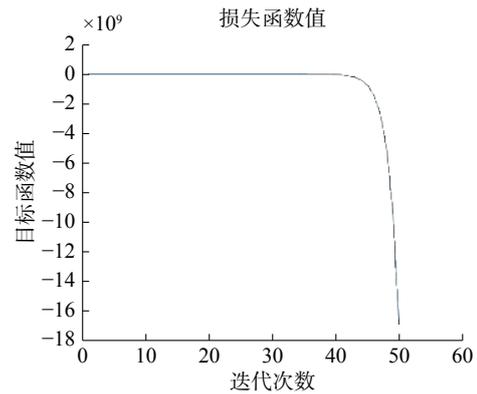


图 1 目标函数值随迭代次数变化

Fig. 1 The value of the objective function varies with the number of iterations

图 1 中单个样本的整个目标函数值在开始到 40 次的变化趋势是平行趋于 0, 之后变成一个比较大的负值, 由于相似度矩阵不是半正定矩阵, 由此使图正则项的值变成负值, 并引起其他项的变化。并且由于图像的损失函数一般都不会小于 0, 所以, 从图 1 可以看出算法在 2~40 次之间是局部收敛的。所以, 有的实验在第 2 次迭代就达到最优值。

4 实验结果与分析

为了验证图正则化字典对学习算法在图像分类中的性能, 本文在 ADNI1 数据集上进行验证。ADNI 由国际老年研究所、生物医学成像和生物工程研究所、美国食品和药物管理局、民营医药企业和非营利组织在 2003 年启动, 资金为 60 万美元。ADNI 的主要目标是测试是否能通过组合

MRI、PET、其他生物标志物，以及临床和神经心理学评估来测定 MCI 和早期 AD 的进展。

在 ADNI1^[14]中，根据 MCI 患者在 36 个月之后的状态，将 36 个月之后仍表现出 MCI 的状态，则记为 SMCI；36 个月之后表现为 AD 的病况，记为 PMCI；没有 36 个月之后的信息的患者，记为 uMCI。实验中只利用表 1 中显示的有标签的 164 个 PMCI 患者信息和 100 个 SMCI 患者信息，对 MCI 进展进行识别，提前预测 AD。

表 1 PMCI 和 SMCI 相关信息
Table 1 PMCI and SMCI related information

Group (组别)	PMCI	SMCI
Number (例数)	164	100
Age (年龄/岁)	74.5±7.0	74.6±7.5
MMSE (简易精神表/分)	27.6±1.7	27.6±1.7
CDR-SB (临床痴呆综合汇总评定量表/分)	1.88±0.97	1.34±0.65
RAVLT (Rey 听觉言语学习试验)	27.1±6.2	35.1±10.2
FAQ (功能评估问卷/分)	5.61±5.11	2.03±3.08
ADAS-cog (阿尔茨海默病评定量表-认知量表/分)	13.20±4.03	9.65±4.11

所有实验是使用中全局生物标志^[14]和表 1 中显示的 Age、MMSE、CDR-SB、RAVLT、FAQ、ADAS-cog 信息的结合生物标志作为特征表示样本，使用算法 1 中学习的字典对 $\{D, P\}$ ，计算出样本的编码系数作为特征。在实验中选择较常用的分类器 SVM^[11, 14-17]、LDA^[17]、KNN 作为比较方法。为了避免实验结果的随机性，运行 100 次 10 倍交叉验证取最终的平均准确率 (ACC) 和平均 ROC 曲线下方的面积 (AUC)，对于分类器的性能进行评价。

为了验证提出的基于图正则化的字典对学习方法的实验效果比其他的实验效果好，直接使用结合生物标志用 SVM、LDA、KNN 分类器分类，使用 DPL 方法训练出来的字典对上的稀疏系数作为特征用分类器 (SVM、LDA、KNN) 做出的实验效果，和使用 GDPL 方法训练出来的字典对上的稀疏系数作为特征用 SVM、LDA、KNN 分类的实验结果。最终的实验结果是取 100 次 10 倍交叉验证结果的平均值，对于具体实验数据的精确度取小数点后 3 位。

在使用字典对算法训练字典的过程中，算法中参数设置是选择一个参数变化，固定其他参数，进行遍历。选取 3 次十倍交叉验证平均准确率最好。经过遍历参数，算法中的参数设置如表 2、3 所示。

表 2 使用 DPL 稀疏系数作为特征的参数设置
Table 2 Parameter setting using DPL algorithm

特征使用算法	DPL 算法		
分类方法	SVM	LDA	KNN
子字典原子数 (m)	60	60	60
τ	0.01	0.05	0.473
λ	0.007	0.5	0.003
η	0.000 1	0.001	0.000 1

表 3 使用 GDPL 稀疏系数作为特征的参数设置
Table 3 Parameter setting of the GDPL algorithm

特征使用算法	GDPL		
分类方法	SVM	LDA	KNN
子字典原子数/m	60	60	60
α	0.8	0.7	0.51
τ	0.01	0.05	0.8
λ	0.007	0.5	0.001
η	0.000 1	0.001 5	0.001
β	0.01	2	0.1

1) 为了验证在基于图正则化的字典对 (GDPL) 算法学习的字典对上的系数作为特征效果比直接用结合生物标志作为特征的实验效果好。本文使用准确率 (ACC) 和 ROC 曲线下的面积 (AUC) 作为评价标准。

从图 2 可以看出直接用结合生物标志作为特征的准确率与使用字典对学习方法的准确率相差较大，使用 SVM 和 GDPL+SVM 的实验结果相差了 2 个百分点，K 近邻 (KNN) 和 GDPL+KNN 相差了 6 个百分点。用 GDPL 方法的准确率比 DPL 方法的准确率高。从图 2 整体上看使用 GDPL 方法的实验效果比 DPL 方法和直接使用综合特征的效果好。图 3 从算法 AUC 性能评价，从中可以看出使用 GDPL 的稀疏系数作为特征的实验中，除了 SVM 外，其他的方法都比使用 DPL 的效果好，其中 KNN 的 AUC 最高到达 92.1%。从图 2、3 中整体的结果可以看出 GDPL 方法的实验结果比 DPL 的实验结果好。

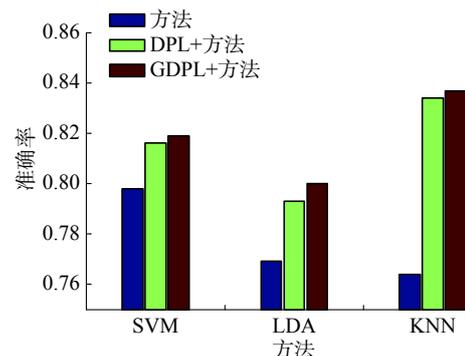


图 2 不同方法上的准确率 (ACC)
Fig. 2 The accuracy of different methods

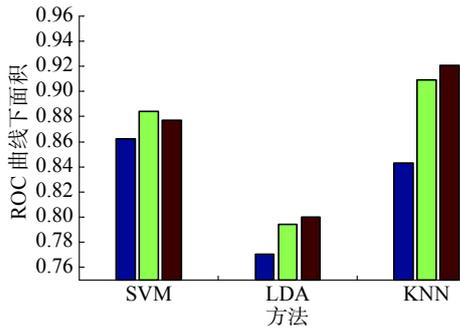


图 3 不同方法上的 ROC 曲线下的面积 (AUC)
Fig. 3 Area under ROC curve in different methods

2) 将结合生物标志直接作为特征用 SVM、LDA、KNN 分类, 使用 DPL 方法的稀疏系数作为特征的分类方法记为 DPL+SVM、DPL+LDA、DPL+KNN, 和使用 GDPL 方法的稀疏系数作为特征的分类方法记为 GDPL+SVM、GDPL+LDA、GDPL+KNN, 所用实验结果都是记录使用 100 次

十倍交叉验证的结果 (平均值) 画出的盒形图。使用准确率 (ACC)、ROC 曲线下面的面积 (AUC) 对于分类器的性能进行评估, 具体实验结果在图 4、5 中显示。

从图 4 中整体上可以看出使用 GDPL 算法的稀疏系数作为特征的平均准确率比 DPL 算法的稍微好一点, 比直接使用结合生物标志作为特征的平均准确率高, GDPL+KNN 算法的平均准确率最高达到了 83.7%。图 5 显示 ROC 曲线下的面积 (AUC) 最好的是 GDPL+KNN 算法, AUC 达到 92.1%。在图 4、5, 100 次实验结果中 GDPL+KNN 算法结果相对集中稳定。整体上看使用 GDPL 方法训练出来的字典对上的稀疏系数作为特征的实验结果比直接使用结合生物标志特征的结果好, GDPL+KNN 实验结果最好, 并且结果比较集中稳定, 每次实验的结果的变化不大。

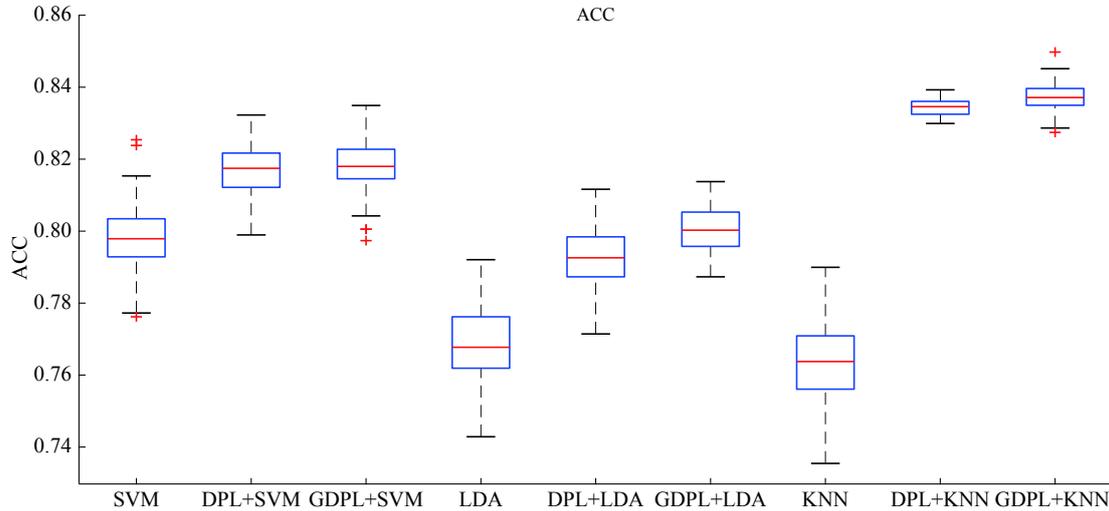


图 4 使用 100 次十倍交叉验证的 ACC 画出的箱型图
Fig. 4 Box plot for the ACC of 100 times 10-fold cross-validation

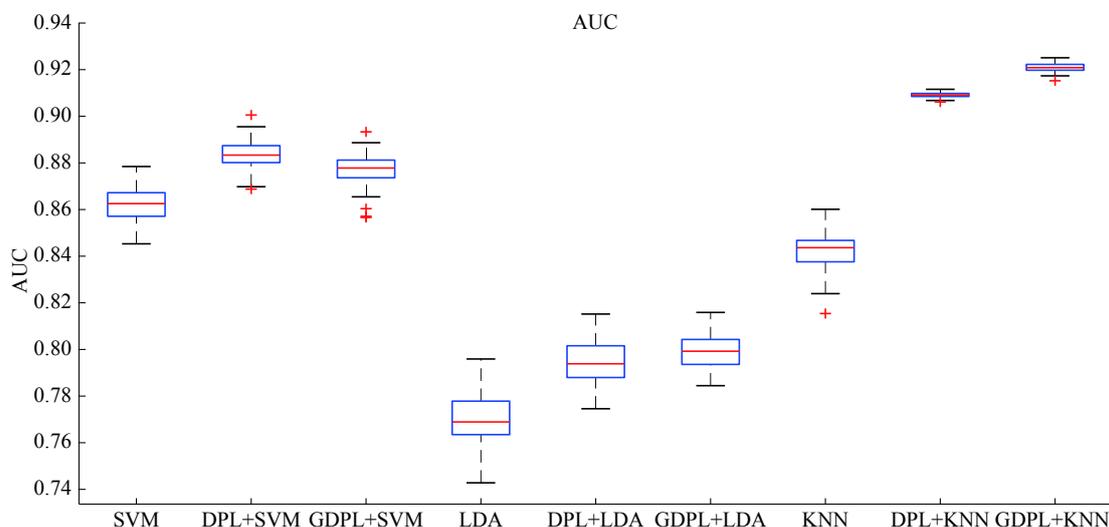


图 5 使用 100 次 10 倍交叉验证的 AUC 画出的箱型图
Fig. 5 Box plot for the AUC of 100 times 10-fold cross-validation

5 结束语

本文实验结果表明在字典对学习算法中加入图正则项,可以提高字典对模型的识别能力,能够更好地对图像进行分类。另一方面,使用在字典对上的稀疏系数作为特征对轻度认知功能障碍预测可以取得更好的效果,并且在 GDPL 算法上的稀疏系数作为特征的实验结果比 DPL 上的实验结果好。随着科技的发展,图像资源越来越丰富,用户想在大量的图像中找到需要的图像信息比较困难。在现实生活中带标签的数据相对于不带标签的数据少很多,获得带标签的图像比较困难。并且训练样本较少不能获得足够的特征去学习字典,下一步可以将无标签的样本用于实验中,学习一个更完备的字典,提高模型的识别能力。可以进行半监督字典对学习,提高模型的判别能力。

参考文献:

- [1] 邵冬华, 施志刚, 史军杰. 二次近邻稀疏重构法及人脸识别[J]. 重庆邮电大学学报(自然科学版), 2017, 29(6): 844-850.
SHAO Donghua, SHI Zhigang, SHI Junjie. Sparse reconstruction algorithm based on secondary nearest neighbor and face recognition[J]. Journal of Chongqing university of posts and telecommunications (natural science edition), 2017, 29(6): 844-850.
- [2] 王佐成, 杨娟, 薛丽霞. 基于二进神经网络的稀疏编码架构的图像无损压缩新方法[J]. 重庆邮电大学学报(自然科学版), 2017, 29(3): 371-376.
WANG Zuocheng, YANG Juan, XUE Lixia. A novel lossless image compression scheme based on binary neural networks with sparse coding[J]. Journal of Chongqing University of Posts and Telecommunications (natural science edition), 2017, 29(3): 371-376.
- [3] WRIGHT J, MA Yi, MAIRAL J, et al. Sparse representation for computer vision and pattern recognition[J]. *Proceedings of the IEEE*, 2010, 98(6): 1031-1044.
- [4] WRIGHT J, YANG A Y, GANESH A, et al. Robust face recognition via sparse representation[J]. *IEEE transactions on pattern analysis and machine intelligence*, 2009, 31(2): 210-227.
- [5] AHARON M, ELAD M, BRUCKSTEIN A. *rmK*-SVD: an algorithm for designing overcomplete dictionaries for sparse representation[J]. *IEEE transactions on signal processing*, 2006, 54(11): 4311-4322.
- [6] YANG Meng, ZHANG Lei, FENG Xiangchu, et al. Fisher discrimination dictionary learning for sparse representation[C]//Proceedings of 2011 International Conference on Computer Vision. Barcelona, Spain, 2011: 543-550.
- [7] JIANG Zhuolin, LIN Zhe, DAVIS L S. Label consistent K-SVD: learning a discriminative dictionary for recognition [J]. *IEEE transactions on pattern analysis and machine intelligence*, 2013, 35(11): 2651-2664.
- [8] LU Jiwen, WANG Gang, DENG Weihong, et al. Simultaneous feature and dictionary learning for image set based face recognition[C]//Proceedings of the 13th European Conference on Computer Vision. Zurich, Switzerland, 2014: 265-280.
- [9] YANG Yang, HUANG Zi, YANG Yi, et al. Local image tagging via graph regularized joint group sparsity[J]. *Pattern recognition*, 2013, 46(5): 1358-1368.
- [10] GU Shuhang, ZHANG Lei, ZUO Wangmeng, et al. Projective dictionary pair learning for pattern classification [C]//Proceedings of the 27th International Conference on Neural Information Processing Systems. Montreal, Canada, 2014: 793-801.
- [11] CUINGNET R, GERARDIN E, TESSIERAS J, et al. Automatic classification of patients with Alzheimer's disease from structural MRI: a comparison of ten methods using the ADNI database[J]. *NeuroImage*, 2011, 56(2): 766-781.
- [12] CHEN Yefei, SU Jianbo. Sparse embedded dictionary learning on face recognition[J]. *Pattern recognition*, 2017, 64: 51-59.
- [13] MAIRAL J, BACH F, PONCE J, et al. Supervised dictionary learning[J]. Paris: INRIA, 2008: 1-8.
- [14] TONG Tong, GAO Qinquan, GUERRERO R, et al. A novel grading biomarker for the prediction of conversion from mild cognitive impairment to Alzheimer's disease[J]. *IEEE transactions on biomedical engineering*, 2017, 64(1): 155-165.
- [15] WYMAN B T, HARVEY D J, CRAWFORD K, et al. Standardization of analysis sets for reporting results from ADNI MRI data[J]. *Alzheimer's & dementia*, 2013, 9(3): 332-337.
- [16] MORADI E, PEPE A, GASER C, et al. Machine learning framework for early MRI-based Alzheimer's conversion prediction in MCI subjects[J]. *NeuroImage*, 2015, 104: 398-412.
- [17] JANOUŠOVÁ E, VOUNOU M, WOLZ R, et al. Biomarker discovery for sparse classification of brain images in Alzheimer's disease[J]. *Annals of the BMVA*, 2012, 2012(2): 1-11.

作者简介:



魏彩锋,女,1989年生,硕士研究生,主要研究方向为字典学习、图像分类。



孙永聪, 男, 1991 年生, 硕士研究生, 主要研究方向为稀疏编码、图像检索。



曾宪华, 男, 1973 年生, 教授, 博士, 中国计算机学会会员, 主要研究方向为流形学习、计算机视觉。主持国家自然科学基金、重庆自然科学基金等省级以上项目 5 项。发表学术论文 30 余篇。

第十七届中国机器学习会议 (CCML 2019)

第十七届中国机器学习会议 (CCML 2019) 由中国人工智能学会和中国计算机学会联合主办, 中国人工智能学会机器学习专业委员会和中国计算机学会人工智能与模式识别专业委员会协办, 贵州大学承办。该系列会议每两年举行一次, 现已成为国内机器学习界最主要的学术活动。此次会议将为机器学习及相关研究领域的学者交流最新研究成果、进行广泛的学术讨论提供便利, 并且将邀请国内机器学习领域的著名学者做精彩报告。

重要时间

会议举办时间: 2019 年 7 月 23 日—25 日

会议投稿截止时间: 2019 年 3 月 25 日

会议录用通知时间: 2019 年 5 月 10 日

征稿范围

机器学习的新理论、新技术与新应用; 特征选择; 聚类; 人类学习的计算模型; 流形学习与降维; 异常检测; 计算学习理论; 信息检索; 演化学习; 监督学习; 生物特征识别; 符号学习; 非监督学习; 生物信息学; 多 Agent 学习; 半监督学习; 距离度量学习; 机器学习应用; 强化学习; 基于案例的推理; 主动学习; 多示例学习; 增量学习与在线学习; 多标记学习; 神经网络; 对复杂结构数据的学习; 模式识别; 集成学习; 增强学习系统可理解性; 大数据学习; 多任务学习; 数据挖掘与知识发现

投稿要求

论文必须未公开发表过, 一般不超过 6 000 字; 只接受中文稿;

论文应包括题目、作者姓名、作者单位、摘要、关键字、正文和参考文献; 另附作者通讯;

地址、邮编、电话及 E-mail 地址;

学生 (不包括博士后和在职博士生) 第一作者的论文稿件请在首页脚注中注明, 否则将不具有参选“优秀学位论文”的资格。

论文出版

会议录用论文将被推荐到《中国科学》、《软件学报》、《计算机科学与探索》、《模式识别与人工智能》、《智能系统学报》、《数据采集与处理》、《计算机科学》、《计算机应用》、《小型微型计算机系统》、《计算机工程与应用》、《南京大学学报(自然科学版)》、《山东大学学报(工学版)》、《南京师范大学学报》等期刊正刊发表。

投稿方式

本会议采用在线投稿方式, 在线投稿网址为: <https://easychair.org/conferences/?conf=ccml2019>

咨询电话

16608512304(龙老师)、18586818471(徐老师)

会议邮箱: CCML2019@126.com

会议网址: <http://bdiri.gzu.edu.cn/ccml2019>