

多目标优化在特征选择子集评价中的应用*

万红^{1a, 1b, 2}, 李蒙蒙^{1a, 1b, 2}, 王昊锋^{1a, 1b, 2}, 岳彩通^{1a, 1b}, 王力^{1a, 1b, 2}, 尚志刚^{1a, 1b, 2†}

(1. 郑州大学 a. 电气工程学院; b. 产业技术研究院, 郑州 450001; 2. 河南省脑科学与脑机接口技术重点实验室, 郑州 450001)

摘要: 特征选择是处理高维大数据的常用降维手段, 但其中牵涉到的多个彼此冲突的特征子集评价目标难以平衡。为综合考虑特征选择中多种子集评价方式间的折中, 优化子集性能, 提出一种基于子集评价多目标优化的特征选择框架, 并重点对多目标粒子群优化(MOPSO)在特征子集评价中的应用进行了研究。该框架分别根据子集的稀疏度、分类能力和信息损失度设计多目标优化函数, 继而基于多目标优化算法进行特征权值向量寻优, 并通过权值向量 Pareto 解集膝点选取确定最优向量, 最终实现基于权值向量排序的特征选择。设计实验对比了基于多目标粒子群优化算法的特征选择(FS_MOPSO)与四种经典方法的性能。多个数据集上的结果表明, FS_MOPSO 在低维空间表现出更高的分类精度, 并保证了更少的信息损失。

关键词: 特征选择; 多目标优化; PSO; 稀疏; 分类; 信息损失

中图分类号: TP391 doi: 10.19734/j.issn.1001-3695.2019.03.0043

Application of multi-objective optimization in feature selection subset evaluation

Wan Hong^{1a, 1b, 2}, Li Mengmeng^{1a, 1b, 2}, Wang Haofeng^{1a, 1b, 2}, Yue Caitong^{1a, 1b}, Wang Li^{1a, 1b, 2}, Shang Zhigang^{1a, 1b, 2†}

(1. a. School of Electrical Engineering, b. Industrial Technology Research Institute, Zhengzhou University, Zhengzhou 450001, China; 2. Henan Key Laboratory of Brain Science & Brain-Computer Interface Technology, Zhengzhou 450001, China)

Abstract: Feature selection is a common dimension reduction approach for processing high-dimensional big data, but it often involves multiple conflicting feature subsets evaluation objectives difficult to balance. To reach a compromise among various feature subset evaluations in feature selection and optimize the performance of subset, this paper proposed a subset evaluation multi-objective optimization based feature selection framework and focused on the application of Multi-Objective Particle Swarm Optimization (MOPSO) in feature subset evaluation. The framework used sparsity, classification ability and information loss to design multi-objective optimization functions. Then it optimized the weight vectors of the features based on multi-objective optimization algorithm, and selected the “knee” of Pareto solution set as optimal vector. Finally, the framework realized feature selection based on weight vectors ranking. This paper designed experiments to compare the performance of MOPSO based Feature Selection (FS_MOPSO) with four classical methods. The results on several standard data sets showed that, FS_MOPSO showed higher classification accuracy in low dimensional space while ensuring less information loss.

Key words: feature selection; multi-objective optimization; pso; sparsity; classification; information loss

0 引言

样本数量多、特征维度高, 是海量大数据的典型特征, 这造成了数据处理的困难。相对于解决大样本集问题的难度, 高维样本集的处理更加棘手。数据降维是解决这一问题的关键方法。在有效降低特征空间维度的前提下, 保留原始数据空间中的隐含规律或拓扑结构, 是高效合理的降维方法应遵循的思路^[1]。筛除原始数据集中的冗余或无关特征, 可以达到对可靠有效信息的提取并减少计算机的运算负担, 同时也可以提高学习模型的稳定性与可解释性。特征选择(Feature Selection, FS)是降维的常用手段, 通过确定特征子集评价方式, 它在原有特征中筛选得到最优或次优特征子集, 保留原始数据集中的重要信息, 有效降低了数据处理的复杂度, 有

利于提高学习模型的准确性和泛化能力^[2]。

与实践中的大多数工程和科学问题一样, 特征选择中牵涉到的子集评价问题也可以被描述为不同的目标函数, 并且各个目标之间常彼此冲突, 因此它也可以被看做是多目标优化问题。多目标优化问题的本质在于, 多数情况下, 某目标的改善可能引起其他目标性能的降低, 而同时使多个目标均达到最优又是不可能的。因此只能在各目标之间进行协调权衡和折中处理, 使所有目标函数尽可能达到最优。在这些问题上, 问题的最优解集往往由数量众多, 甚至无穷大的 Pareto 最优解组成^[3]。

多目标优化算法(multi-objective optimization algorithms, MOA)可以实现对互相冲突的目标函数进行协调寻优^[4]。在特征选择问题中引入多目标优化能否更好地实现对特征子集的

收稿日期: 2019-03-21; 修回日期: 2019-05-09 基金项目: 国家自然科学基金资助项目(U1304602, 61673353)

作者简介: 万红(1964-), 女, 辽宁沈阳人, 教授, 博导, 博士, 主要研究方向为信号处理; 李蒙蒙(1990-), 男, 河南商丘人, 博士研究生, 主要研究方向为特征选择、信号处理与模式识别; 王昊锋(1995-), 男, 河南驻马店人, 硕士研究生, 主要研究方向为特征选择; 岳彩通(1990-), 男, 河南濮阳人, 博士研究生, 主要研究方向为多目标优化; 王力(1994-), 男, 山西阳泉人, 硕士研究生, 主要研究方向为优化算法与建模; 尚志刚(1975-), 男, 甘肃兰州人, 教授, 博导, 博士, 主要研究方向为数据挖掘与信号处理(zhigang_shang@zzu.edu.cn)。

多角度评价呢? 针对这一问题, 近年出现了许多基于多目标优化算法的特征选择研究。Hamdani 等人^[5]在特征选择中采用多目标优化方法最小化特征数目和分类错误率两个目标, 但没有考虑其他目标, 如特征间的冗余。Vebkatadri 等人^[6]在考虑不同评价准则的情况下, 应用多目标优化方法寻找近似最优子集, 但目标函数设计中未考虑特征数目。Saroj 等人^[7]构造了三个目标函数, 包括最大信息增益、非冗余度和未使用特征数, 并利用多目标遗传算法优化特征子集。陈翔等人^[8]通过最小化特征子集规模和最大化预测效果进行多目标寻优解决软件缺陷预测问题。然而这些研究均只基于单一的算法进行优化, 所使用的数据集维数一般较低, 且没有关注特征子集的稳定性表现。故如何选择合适的多目标优化算法, 并基于特定算法实现对特征子集评价问题中多种因素的综合考量仍有待研究。

因此, 本文考虑特征子集稀疏性、分类能力及信息量等多个目标函数, 设计了基于子集评价多目标优化的特征选择框架(feature selection framework based on multi-objective algorithms, FS_MOA), 并结合包括多目标粒子群优化(multi-objective particle swarm optimization, MOPSO)算法进行特征权值向量寻优, 通过权值向量 Pareto 解集膝点选取确定最优向量以实现基于排序的特征选择。本文设计数值实验将基于多目标粒子群优化的特征选择(FS_MOPSO)与其他四种经典特征选择方法进行性能评估, 对比结果表明, FS_MOPSO 选择的特征子集表现出其有效性和优越性。

1 特征选择中的多目标子集评价函数构建

1.1 特征选择及子集评价准则

特征选择的基本流程大致可以分为四个部分, 即产生过程、评价函数、停止准则和验证过程^[9]。其中评价函数的作用是搜索得到的特征子集的好坏提供一种可靠的度量方式。根据评价准则的工作机理, 特征选择方法主要可以分为过滤式方法、封装式方法和嵌入式方法三大类^[10]。本文着眼于在监督学习中利用多目标优化算法对特征子集的评价准则进行目标函数式的描述, 因此主要讨论的是过滤式方法^[11]。

在监督学习中, 对一个特征子集进行评价时通常涉及其特征稀疏性^[12]、分类能力^[13]、信息损失度^[14]和稳定性等等。得到具有稀疏性的特征子集是特征选择的首要目的, 这是由特征选择要选出小于全部特征数量的少数高效特征的性质决定的; 分类能力表示的是监督学习中获得的特征子集的预测能力, 直接反映了该子集的应用效果; 信息损失度则反映特征子集对于原始数据集表示能力的强弱; 而稳定性表示该特征子集在不同的应用环境下的适应能力。

多个目标彼此冲突时, 很难找到一个解使所有目标函数同时最优。特征子集评价问题中就存在彼此冲突的优化目标, 而基于子集评价多目标优化的特征选择过程实际上综合考量多种子集评价方式间的折中进而求解 Pareto 最优解集的过程。

1.2 多目标子集评价函数的构建

根据上述提到的特征子集的评价准则, 本文在基于多目标优化的特征子集评价中设计了对应的 3 个目标函数, 即特征空间的稀疏度函数、监督学习中的分类精度评价函数和特征子集的信息量评价函数^[15]。设数据集中的样本空间及其对应的类别标号可标记为 $\{X, y\}$, 而 w 表示所有特征的权值向量, 则上述评价准则可以分别构建函数并表述为

1) 特征空间的稀疏度函数

$$f_1(w) = \|w\|_0 \quad (1)$$

其中 $\|w\|_0$ 表示特征权值向量 w 的 L_0 范数。事实上, 实现稀疏特征选择的最理想途径是 L_0 正则化, 但由于这是个 NP 难问题, 很难求解。而 L_1 正则化, 也就是通常所说的 Lasso 正则化是 L_0 正则化的最优逼近, 表示向量中各元素的绝对值之和, 比 L_0 正则化更容易求解, 因此得到了广泛的应用^[16]。故本文选择 L_1 正则化作为特征选择稀疏度的度量方式, L_1 范数越小则说明对应特征子集的稀疏度越高。

2) 监督学习中的分类精度评价函数

$$f_2(w) = |\text{corr}(X * w, y)| \quad (2)$$

其中 $\text{corr}(\cdot)$ 表示变量间的相关系数, 可以直接反映其相关程度^[17]。如果两个变量是完全线性相关的, 则相关系数的绝对值为 1; 否则, 如果它们完全无关, 则值为 0。根据相关性原则, 如果特征子集中的特征与类别标签的相关程度较大, 那么该特征的预测效果就会较好。根据这一朴素思想, 本文设计了上述的评价函数以衡量特征子集的分类精度。

3) 特征子集的信息量评价函数

$$f_3(w) = \text{std}(X * w) \quad (3)$$

其中 $\text{std}(\cdot)$ 表示变量的方差。本质上, 信息损失即为将原始样本投影压缩到一维变量空间时的样本分布方差, 而这反映了这种最大程度降维所造成的重建误差^[18]。当某维度上的样本分布方差最大时, 降维引起的误差最小, 信息损失最少, 信息量最多; 反之则信息量最小。一个二维空间中的例子如图 1 所示。

从图中易知, 当原始数据投影到方向 w_1 时, 样本分布方差较大, 信息损失较小。因此, w_1 方向对应的特征空间的信息量要大于 w_2 方向对应的特征空间。在特征选择的过程中, 需要考虑尽可能保留更多的信息量, 故而本文设计上述评价函数衡量特征选择造成的信息损失。

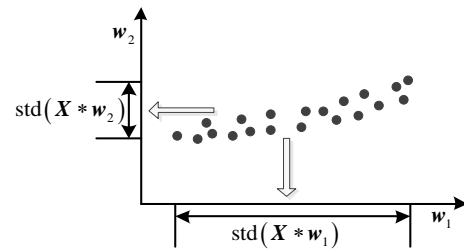


图 1 二维空间中降维后的信息量比较示例

Fig. 1 Information comparison example after dimensionality reduction in two dimensional space

2 基于子集评价多目标优化的特征选择框架

目前许多较有代表性的多目标优化算法, 如改进强度帕累托进化算法(Strength Pareto Evolutionary Algorithm 2, SPEA2), 改进非支配排序遗传算法(Non-dominated Sorting Genetic Algorithms II, NSGAI)和多目标粒子群算法^[19]等, 在涉及多目标优化问题的不同领域都得到了广泛的应用。近年来, 基于这些算法的特征选择研究^[20-23]也层出不穷, 它们通过构建不同的特征子集评价函数进行多目标寻优, 进而实现特征选择。

本文结合多目标优化算法, 设计了基于子集评价多目标优化的特征选择框架 FS_MOA, 它们分别基于子集评价函数进行权值向量多目标寻优得到对应的 Pareto 最优解集, 从中选取膝点(knee)^[24]对应的最佳权值向量, 进而确定特征排序。

FS_MOA 框架流程如图 2 所示。具体步骤描述如下:

- a) 对于给定数据集 X , 构建特征子集多目标评价函数;
- b) 以特征权值向量 w 为决策变量并初始化;

- c) 结合由式(1)~(3)评价函数组合而成的目标函数执行基于多目标优化算法的迭代寻优;
- d) 根据寻优结果确定 Pareto 最优解集;
- e) 利用基于角度的膝点选取方法^[25]确定最佳 Pareto 解, 及其对应的决策空间特征权值向量;
- f) 根据步骤 e) 确定的权值向量进行特征排序, 较大权值对应的特征排序靠前;
- g) 输出特征排序结果。

作为典型的群智能方法, MOPSO 利用计算机模拟飞行和捕食中鸟群个体间相互配合和协作的过程, 实现对整个种群的优化。MOPSO 可以按照并行方式高效搜索到多个 Pareto 最优解, 其良好的通用性又使其适合处理多种类型的目标函数和约束, 从而在解决多目标优化问题时表现出很大优势。加之较小的个体数目、出色的迭代收敛速度、简单方便的运算与计算机实现等优点, 使 MOPSO 在工程优化领域得到广泛的应用。本文综合考虑其出色的多目标优化能力和较高的运算效率, 将其嵌入上述特征选择框架 FS_MOA, 进行特征选择子集评价和对比分析, 基于 MOPSO 的特征选择记作 FS_MOPSO。

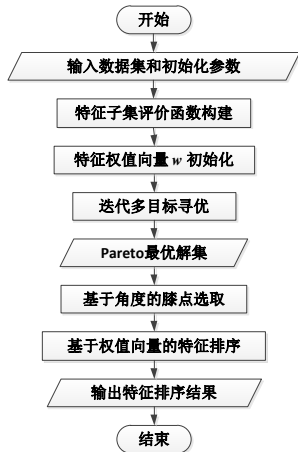


图 2 FS_MOA 框架流程图

Fig. 2 Flow chart of FS_MOA framework

3 FS_MOPSO 与经典特征选择方法的性能对比

为考察 FS_MOPSO 在不同数据集上的特征选择效果, 本文综合考量前述稀疏度、分类性能以及信息量三个评价指

标构建目标函数进行数值实验, 目标函数表示为

$$\begin{cases} \min \|w\|_1 \\ \max |\text{corr}(X * w, y)| \\ \max \text{std}(X * w) \end{cases} \quad (4)$$

本文选择基于 Fisher 分数的方法(FFS), 基于基尼系数(Gini coefficient)的方法(GFS), 基于信息增益(Information Gain, IG)的方法(IGFS)及经典的 ReliefF 特征选择方法作为对比方法开展了分析。文中所有计算均采用编程实现, 实验条件为 Intel(R) Core(TM) i7-4790 CPU 3.60 GHz/8.00 GB/Windows 7/MATLAB R2014a。MOPSO 算法的种群规模设置为 100, 最大迭代次数为 30。考虑在类别数、样本数和特征数上多样化地对比结果而选择的 UCI 标准数据集详细情况如表 1 所示。

表 1 特征选择实验中用到的数据集

Table 1 Data sets used in the feature selection experiments

数据集	类数	实例数	维数
Waveform3	3	5000	21
USPS	10	7291	256
Isolet	26	7797	617
COIL20	20	1440	1024

3.1 监督学习中的分类性能分析

采用 Hold-out 交叉验证对上述数据集进行特征选择, 随机选取每个数据集 70% 的样本作训练集寻优得到对应的特征子集序列, 30% 作测试集并以 20 次随机抽样实验的平均结果比较分类精度以评估方法性能。为测试方法选择出的特征子集在不同分类器上表现的稳定性, 本文选取了两种被广泛使用的分类器, 即朴素贝叶斯(naive Bayes, NB)和支持向量机(support vector machine, SVM)开展实验。分类性能结果对比如图 3 所示。

结果表明, 在不同数据集上, 随入选特征数量的增加, FS_MOPSO 和其他经典方法对应的特征子集在不同分类器上的分类精度均呈现逐渐增加直至基本稳定的状态。与其他经典方法相比, 在 USPS、Isolet 和 COIL20 三个数据集上, FS_MOPSO 的表现明显占优; 而在 Waveform3 数据集上, FS_MOPSO 的表现与 FFS 和 IGFS 基本相当, 但明显优于 GFS 和 ReliefF 方法。上述结果表明通过基于子集评价多目标优化的特征选择, FS_MOPSO 得到了分类性能优异的特征子集。

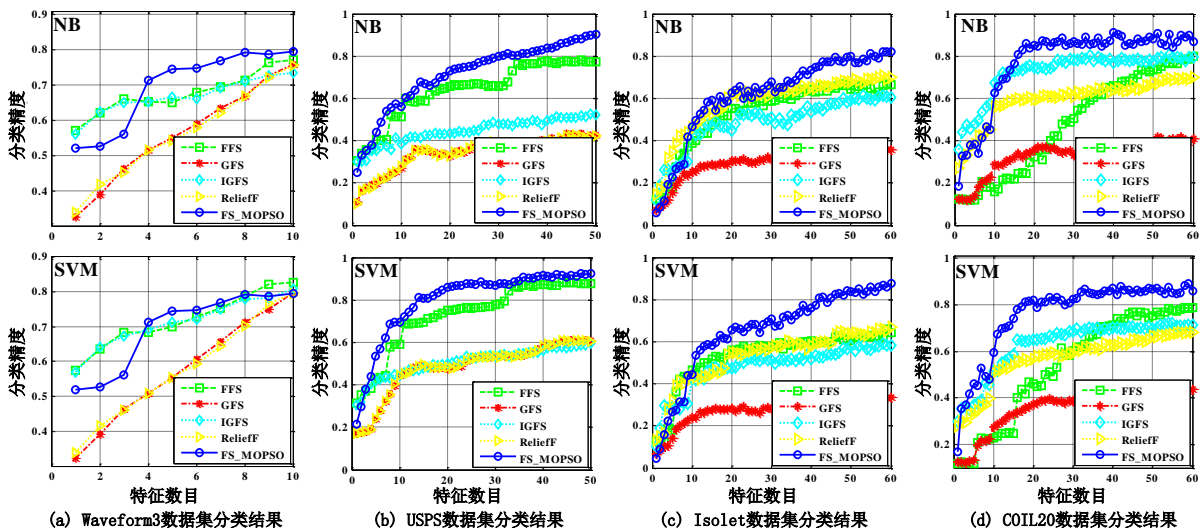


图 3 不同方法在不同数据集上对应特征子集的分类结果

Fig. 3 Classification results of the feature subsets by different methods on different datasets

3.2 特征子集的稳定性分析

通过比较不同分类器的分类精度, 可以分析不同方法对对应特征子集的分类稳定性, 结果如图 4 所示。图中二维平面上任一点对应的坐标值代表一个特征子集在两个分类器上的分类精度, 将不同维度下的结果相连组成各方法对应的连接

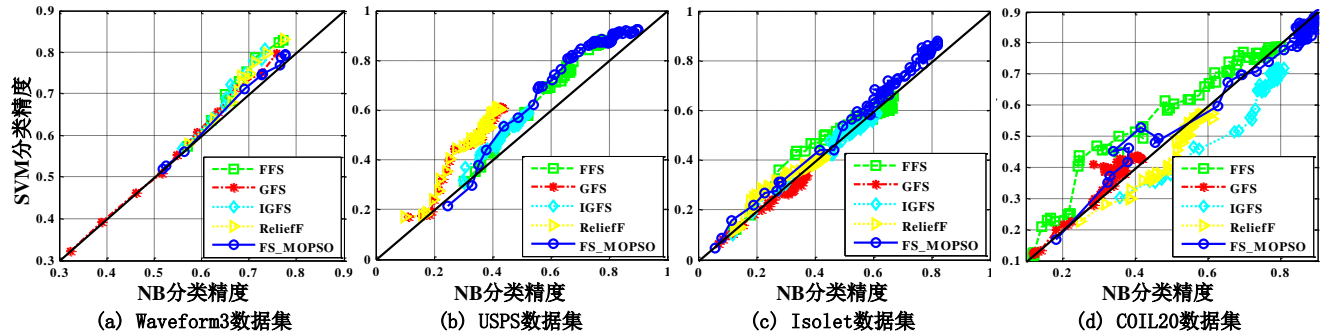


图 4 不同方法在不同数据集上对应特征子集的精度分布

Fig. 4 Distribution of two accuracies of different methods on different sets

由图 4 可以看出, 与其他经典特征选择方法相比, 在 Isolet 和 COIL20 数据集上, FS_MOPSO 对应的特征子集的曲线最接近 $Y=x$ 且其对应的最高分类精度也都高于其他方法, 这既说明 FS_MOPSO 对应的特征子集稳定性最高, 也验证了其优越的分类性能; 在 USPS 数据集上, FS_MOPSO 的稳定性高于 GFS 和 ReliefF, 与 FFS 和 IGFS 相当, 但最高分类精度均高于其他方法; 在 Waveform3 数据集上, FS_MOPSO 的最高分类精度略低于其他方法, 但表现出最佳的稳定性。结果表明, 与其他经典特征选择方法比较, FS_MOPSO 选出的特征子集稳定性相对较高。

3.3 特征选择的信息损失评价分析

根据文献[18]对信息损失程度的定义, 本文引入表示熵 (Representation Entropy, RE) 这一概念衡量特征子集相对于原始特征全集的信息损失程度。假设对于一个 d 维的特征集合, 其协方差矩阵的特征值可以表示为 $\lambda_j (j=1, \dots, d)$ 。令 $\tilde{\lambda}_j = \lambda_j / \sum_{j=1}^d \lambda_j$, 则可以定义一个熵函数, 即表示熵为

$$RE = -\sum_{j=1}^d \tilde{\lambda}_j \log \tilde{\lambda}_j \quad (5)$$

由式(5)可以看出当所有特征值取值全部相同时, RE 取最大值, 这时特征之间的信息均匀分布在特征空间中, 特征之间的冗余度达到最小, 即特征子集对应的信息损失最小; 当所有的特征值中除了一个非零, 其余全为零时, RE 取最小值零, 这时所有特征之间的信息是沿着一个单一的坐标方向的, 特征间的冗余度最大, 信息损失最大。在三目标特征子集评价优化实验中, 信息量评价函数的设计正是希望将特征选择造成的信息损失降到最小。因此, 利用如下的公式定义最终选出的最优特征子集的信息损失度, 即

$$I_{loss} = 10/RE \quad (6)$$

分别计算并记录 FS_MOPSO 与其他经典特征选择方法对应特征子集的信息损失度 I_{loss} , 结果对比如表 2 所示。从表中可以看出, 在所有数据集上, FS_MOPSO 对应特征子集的 I_{loss} 值均最低, 表明 FS_MOPSO 能够较好地保留原始数据的有效信息, 保证较低的信息损失程度。

4 结束语

本文针对特征选择问题中涉及的多种子集评价方式, 分别设计了目标函数, 并引入多目标优化算法进行特征子集的求解, 提出了一种基于特征子集评价多目标优化的特征选择

曲线。由此易知, 对于不同的特征选择方法, 其对应的曲线越靠近二维空间的对角线 $Y=x$, 说明其选出的特征子集在不同分类器上的分类精度越接近, 即说明对应的优选特征子集的稳定性越高。

思路。为评价 FS_MOPSO 方法的特征选择效果并比较与其他经典特征选择方法的性能, 设计了多目标优化的特征选择数值实验, 在不同数据集上使用不同分类器的应用结果表明, 与其他四种经典的特征选择方法相比, FS_MOPSO 选取特征子集在稀疏性、分类精度以及信息损失度三方面具有优越性, 且 FS_MOPSO 对应的特征子集表现出突出的监督学习分类稳定性。

表 2 不同特征选择方法选出特征子集的信息损失度

Table 2 Information loss degree of feature subsets selected by different feature selection methods

数据集	不同特征选择方法的信息损失度				
	FFS	GFS	IGFS	ReliefF	FS_MOPSO
Waveform3	4.68	3.80	4.90	4.39	3.56
USPS	3.35	3.47	4.20	3.01	2.67
Isolet	4.14	3.50	4.99	3.60	2.26
COIL20	4.22	3.67	5.02	4.89	2.46

虽然 FS_MOPSO 基于 MOPSO 算法实现了有效的特征选择, 但在应用中还面临一些问题, 如本文在确定特征权重向量时只选择了一个均衡了所有预设目标函数的 Pareto 解, 事实上, 根据入选特征数目、特征子集分类性能、特征冗余度等不同的实际要求, Pareto 解集中存在更多样化的选择; 其次, 多目标优化算法中初始种群是随机产生的, 如果在产生初始种群时更多地考虑一些先验知识, 变随机搜索方式为启发式搜索, 可能会加快搜索速度, 提高寻优效率; 另外, 如何设置更合理的目标函数去描述特征选择问题中涉及的评价准则也是非常值得研究的问题。故上述几点将成为笔者下阶段的主攻方向。

参考文献:

[1] 林春喜, 徐宏喆, 王谊青, 等. 基于混合频繁模式树的粗糙集属性约减算法的研究与应用 [J]. 计算机应用研究, 2018, 35 (4): 988-991. (Lin Chunxi, Xu Hongzhe, Wang Yiqing, et al. Research and application on rough set attribute reduction model based on mixed frequent pattern tree [J]. Application Research of Computers, 2018, 35 (4): 988-991.)

[2] Wang Lipo, Wang Yaoli, Chang Qing. Feature selection methods for big data bioinformatics: A survey from the search perspective [J]. Methods, 2016, 111: 21-31.

[3] 刘建昌, 李飞, 王洪海, 等. 进化高维多目标优化算法研究综述 [J].

- 控制与决策, 2018, 33 (5): 879-887. (Liu Jianchang, Li Fei, Wang Honghai, *et al.* Survey on evolutionary many-objective optimization algorithms [J]. Control and Decision, 2018, 33 (5): 879-887.)
- [4] Liu Ruochen, Li Jianxia, Fan Jing, *et al.* A coevolutionary technique based on multi-swarm particle swarm optimization for dynamic multi-objective optimization [J]. European Journal of Operational Research, 2017, 261 (3): 1028-1051.
- [5] Hamdani T M, Won J M, Alimi A M, *et al.* Multi-objective feature selection with NSGA II [C]// Proc of International Conference on Adaptive and Natural Computing Algorithms. Berlin: Springer-Verlag, 2007: 240-247.
- [6] Venkatadri M, Rao K S. A multiobjective genetic algorithm for feature selection in data mining [J]. International Journal of Computer Science and Information Technologies, 2010(1): 443-448.
- [7] Saroj J. Multi-objective genetic algorithm approach to feature subset optimization [C]// Proc of IEEE International Advance Computing Conference. New York: IEEE Press, 2014: 544-548.
- [8] 陈翔, 沈宇翔, 孟少卿, 等. 基于多目标优化的软件缺陷预测特征选择方法 [J]. 计算机科学与探索, 2018, 12 (9): 1420-1433. (Chen Xiang, Shen Yuxiang, Meng Shaoqing, *et al.* Multi-objective optimization based feature selection method for software defect prediction [J]. Journal of Frontiers of Computer Science & Technology, 2018, 12 (9): 1420-1433.)
- [9] Dash M, Liu Huan. Feature selection for classification [J]. Intelligent Data Analysis, 1997, 1 (1-4): 131~156.
- [10] Chandrashekar G, Sahin F. A survey on feature selection methods [J]. Computers & Electrical Engineering, 2014, 40 (1): 16-28.
- [11] Darshan S L S, Jaidhar C D. Performance evaluation of filter-based feature selection techniques in classifying portable executable files [J]. Procedia Computer Science, 2018, 125: 346-356.
- [12] Li Zechao, Liu Jing, Yang Yi, *et al.* Clustering-guided sparse structural learning for unsupervised feature selection [J]. IEEE Trans on Knowledge & Data Engineering, 2014, 26 (9): 2138-2150.
- [13] Dhyaram L P, Vishnuvardhan B. Classification performance improvement using random subset feature selection algorithm for data mining [J]. Big Data Research, 2018, 12: 1-12.
- [14] Ververidis D, Kotropoulos C. Information loss of the Mahalanobis distance in high dimensions: Application to feature selection [J]. IEEE Trans on Software Engineering, 2009, 31 (12): 2275-2281.
- [15] Li Mengmeng, Shang Zhigang, Yue Caitong. A feature subset evaluation method based on multi-objective optimization [C]// Proc of Asia-Pacific Conference on Simulated Evolution and Learning. Cham: Springer, 2017: 581-590.
- [16] 王江辉, 吴小俊. 基于 L_1 范数的形状快速匹配算法 [J]. 计算机应用研究, 36(1): 264-267, 272. (Wang Jianghui, Wu Xiaojun. Fast shape retrieval based on L_1 norm [J]. Application Research of Computers, 2019, 36(1): 264-267, 272.))
- [17] 吴佳婧, 贺嘉楠, 王越群, 等. 基于项目属性分类的协同过滤算法研究 [J]. 吉林大学学报:信息科学版, 2018, 36 (4): 470-474. (Wu Jiajing, He Jianan, Wang Yuequn, *et al.* Research on collaborative filtering algorithm based on items'attribute categories [J]. Journal of Jilin University:Information Science Edition, 2018, 36 (4): 470-474.)
- [18] Mitra P, Murthy C A, Pal S K. Unsupervised feature selection using feature similarity [J]. IEEE Trans on Pattern Analysis & Machine Intelligence, 2002, 24 (3): 301-312.
- [19] Coello C A C, Pulido G T, Lechuga M S. Handling multiple objectives with particle swarm optimization [J]. IEEE Trans on Evolutionary Computation, 2004, 8 (3): 256-279.
- [20] Das P, Das A K. An application of strength pareto evolutionary algorithm for feature selection from crime data [C]// Proc of the 8th International Conference on Computing, Communication and Networking Technologies. New York: IEEE Press, 2017: 1-6.
- [21] Singh U, Singh S N. Optimal feature selection via NSGA-II for power quality disturbances classification [J]. IEEE Trans on Industrial Informatics, 2018, 14 (7): 2994-3002.
- [22] Aboud A, Fdhila R, Alimi A M. MOPSO for dynamic feature selection problem based big data fusion [C]// Proc of IEEE International Conference on Systems, Man, and Cybernetics. New York: IEEE Press, 2016: 003918-003923.
- [23] Sohrabi M K, Tajik A. Multi-objective feature selection for warfarin dose prediction [J]. Computational Biology & Chemistry, 2017, 69: 126-133.
- [24] Das I. On characterizing the "knee" of the pareto curve based on normal-boundary intersection [J]. Structural Optimization, 1999, 18 (2-3): 107-115.
- [25] Branke J, Deb K, Dierolf H, *et al.* Finding knees in multi-objective optimization [C]// Proc of the 8th International Conference on Parallel Problem Solving from Nature. Berlin: Springer, 2004: 722-731.