

混合人工化学反应优化和狼群算法的特征选择

张亚钊¹ 李浩¹ 宋晨明² 卜荣景¹ 王海宁¹ 康雁¹

¹ 云南大学软件学院 昆明 650500

² 西安交通大学软件学院 西安 710000

(1466463308@qq.com)

摘要 包装器特征选择是一种数据预处理方法,通过筛选出信息量最大的特征来降低原始数据集的维数,同时使分类特征的精度最大化。为提高包装器特征选择能力,提出了一种混合人工化学反应狼群优化算法——ACR-WCA。ACR-WCA算法采用自然策略,模仿狼群的搜索策略,可以快速向解空间靠拢,再采用人工化学反应策略优化狼群的种群行为,快速找到最优解,解决局部最优问题;其次,为有效处理数据特征,在初始化阶段利用转换函数处理成二进制特征问题;之后,结合分类准确率和特征选择数给出算法的适应度函数。同时,采用k最近邻(KNN)分类器对测试数据进行训练,并通过K-折交叉验证来克服过拟合问题。实验基于21个著名的不同维度数据集训练,并与4种传统方法和3种接近方法进行比较。实验结果表明,该算法是高效可靠的,它可以对大量特征进行分类任务,具有较高的准确率。

关键词: 特征选择;人工化学反应优化;狼群算法;分类;优化

中图法分类号 TP391

Hybrid Artificial Chemical Reaction Optimization with Wolf Colony Algorithm for Feature Selection

ZHANG Ya-chuan¹, LI Hao¹, SONG Chen-ming², BU Rong-jing¹, WANG Hai-ning¹ and KANG Yan¹

¹ School of Software, Yunnan University, Kunming 650500, China

² School of Software, Xi'an Jiaotong University, Xi'an 710000, China

Abstract Wrapper feature selection is a data preprocessing method for reducing original dataset dimensionality by screening the most informative features to maximize the classification accuracy synchronously. In order to improve the wrapper feature selection ability, this paper proposes a hybrid artificial chemical reaction wolf colony optimization algorithm for selecting feature—ACR-WCA. First, ACR-WCA algorithm adopts natural strategy, imitates the search strategy of wolves, so can quickly approach the solution space. Secondly, in order to deal with data features effectively, the S-shaped transfer function is used to deal with binary features in the initialization stage. Then the fitness function of the algorithm is given by combining classification accuracy and the number of features. Meanwhile, the method uses K-Nearest Neighbor (KNN) classifier for training and tested data by K-fold cross-validation to overcome the over fitting problem. The experiments are trained based on 21 famous and different dimensionality dataset, and compared with four traditional methods and three nearly methods. Experimental results show that the algorithm is efficient and reliable. It can select the most features for classifications tasks with high accuracy.

Keywords Feature selection, Artificial chemical reaction optimization, Wolf colony algorithm, Classification, Optimization

1 引言

在现代都市生活中,社交网络会收集用户喜欢、分享、发布和访问的有关数据;智能手机上的应用程序会收集大量的用户个人信息;购物平台会收集用户购买、浏览和点击网站的行为;在科学研究中,DNA测序、基因芯片、生物信息都需要

大规模的生物数据作为研究基础;超大天文观测技术也是以海量天文数据的产生和分析为基础展开的^[1-2]。数据维数和收集量的增加,会产生大量的冗余信息,进行可视化和推理预测变得越来越困难。图像处理、信息检索和生物信息学技术的发展产生了具有超大规模特征的高维数据集。如何从高维数据中有效提取或选择有用的特征信息,并识别它们的分类,

基金项目:国家自然科学基金(61762092);云南省软件工程重点实验室开放基金项目(2017SE204);云南省重大科技专项(202002AB080001);《材料基因工程-基于Metcloud的集成计算功能模块计算软件开发》(2019CLJY06)

This work was supported by the National Natural Science Foundation of China (61762092), Yunnan Key Laboratory of Software Engineering Open Fund Project(2017SE204), Yunnan Major Science and Technology Project(202002AB080001) and Material Genetic Engineering-Metcloud Based Integrated Computing Function Module Computing Software Development(2019CLJY06).

通信作者:康雁(Kangyan@ynu.edu.cn)

已成为当今信息科学与技术面临的基本问题。

特征选择是数据预处理中过滤冗余特征的一种重要方法。在机器学习中,一个典型的学习任务是通过样本的特征来预测相应的结果。一些算法在高维数据上的性能较差,但特征选择可以提高算法的可用性,选择关键特征数据有助于减少计算训练时间。在深度学习中,减少特征具有重要的现实意义,不仅能减少过度拟合和特征的数量(降维),提高模型的泛化能力,而且可以使得模型更好地表达并增强功能和特征之间的理解,加快模型的训练速度,得到更好的结果。

狼群算法(Wolf Colony Algorithm, WCA)通过模拟狼群的智能捕食行为来解决优化问题。该算法抽象了狼的随机游走、召唤行为和围攻行为^[3]。仿真实验证明,狼群算法比粒子群算法(Particle Swarm Optimization, PSO)^[4]和遗传算法(Genetic Algorithm, GA)^[5]具有更高的求解精度和更快的收敛速度。

人工化学反应优化算法(Artificial Chemical Reaction Optimization, ACRO)是一种新型的元启发式方法,它模拟化学反应中分子之间的相互作用以达到低能稳定状态。仿真结果表明,与现有的元启发式算法相比,人工化学反应优化器具有很强的竞争力,在某些情况下更优。在现实问题中,人工化学反应算法的性能较好,但是目前没有研究者对人工化学反应进行广泛的使用^[6]。一方面,人工化学反应优化很少用于特征选择;另一方面,狼群算法可以快速得到最优解。本文在已有研究的基础上,将这两个概念结合在一起,提出一种混合人工化学反应优化和狼群算法的特征选择算法(Artificial Chemical Reaction Optimization with Wolf Colony Algorithm, ACR-WCA),同时对特征选择结果进行讨论,从各个方面对比并分析该方法性能。

本文的主要贡献如下:

- (1)利用传统的狼群算法,随机初始化,并将搜索过程划分为几个阶段进行;
- (2)采用改进的人工化学反应对狼群个体进行快速变异。该算法可以快速更新狼群搜索空间的位置,从而克服了传统机制的缺点;
- (3)利用化学反应对狼的迭代策略进行优化和改变,促进了人工狼之间的位置信息交互,使人工狼能够更好地收集搜索空间的信息,提高算法的搜索能力,快速逼近猎物;
- (4)将该方法与7种不同方法进行了比较,并对21个数据集的算法结果进行了分析,结果表明该算法的平均准确率和适应度均优于其他算法,我们的算法更加有效和可靠。

2 相关工作

2.1 特征选择

特征选择(Feature Selection, FS)是指根据一定的准则,从给定的特征集集中选取信息量较高的特征子集集合,使所构建的模型更加完善。特征选择是数据预处理中的关键步骤,如何在大规模数据中发现隐藏的模式或有价值的内容已成为当务之急。特征选择问题是统计领域的一个经典问题,自20世纪60年代以来,受到了机器学习、模式识别等各个领

域的众多学者前所未有的关注,具有重要的学术意义和实用价值^[7-8]。总的来说,特征选择算法经历了5次主要的发展:(1)从基于阈值的单一特征选择到多特征选择的组合算法;(2)特征选择从完全决策到不完全决策;(3)特征选择从特征独立原则到特征关联规则;(4)特征选择从平衡导向数据向非平衡导向数据转变;(5)从监督学习到无监督学习的转变。

特征选择算法可以看作是搜索技术和评价指标的结合。前者提供候选的新特征子集,而后者对不同的特征子集进行评分。通过选择不同的评价指标,特征选择算法可以分为过滤、包装、嵌入和集成4类。过滤方法是利用数理统计中一些数据的指标值进行过滤。其特点是在数据处理过程中对特征进行了选择,增强了特征与类之间的相关性。同时,它具有较高的通用性。例如,非常经典的拉普拉斯分数^[8]和相关特征^[9]都是标准的滤波特征选择方法,这对特征选择的入门非常有指导意义。包装器方法将特征选择算法作为学习算法的一个组成部分,直接将分类性能作为特征重要程度的评价指标。这种方法倾向于为模型找到性能最佳的特性集。然而,缺点也很明显:这需要大量的计算;此外,所得结果可能是局部最优解。嵌入式方法首先使用一些学习算法和模型进行训练,得到每个特征的权重系数,最后根据权重系数从高到低选取特征。它不仅解决了滤波器冗余过大的问题,而且改善了包装器时间复杂度过大的问题。最常用的方法之一是正则化。集成方法借鉴了集成学习的思想,通过训练多种特征选择方法来集成最优结果^[10]。

2.2 元启发式算法

2.2.1 元启发式算法分类

元启发式算法(Meta-Heuristic Algorithm)是启发式算法的改进,它是随机算法与局部搜索算法相结合的产物,其由于具有明显的优越性而被应用到多个领域^[11]。元启发式可以是自然启发的,也可以是非自然启发的。受自然启发的元启发式算法可以大致分为4类:进化算法、基于物理的算法、基于群体的算法和基于人的算法,如图1所示。

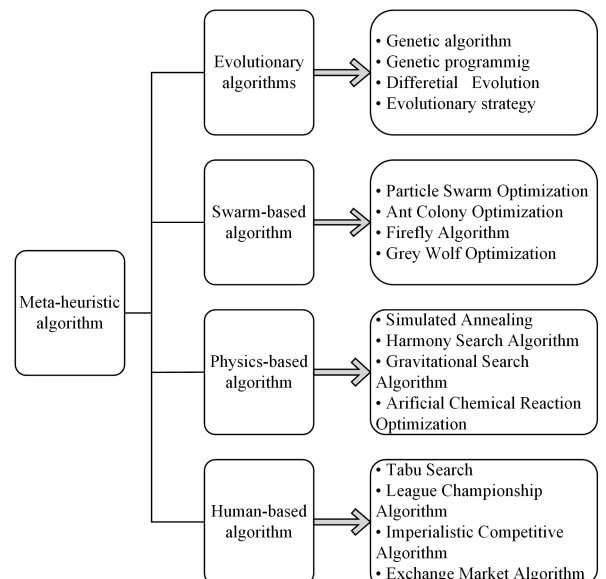


图1 元启发式算法分类

Fig. 1 Meta-heuristics classification

(1)进化算法:进化算法的灵感来自于大自然的生物进化,这是一种成熟的具有高鲁棒性和广泛适用性的全局优化方法,具有自组织、自适应、自学习的特性,能够不受问题性质的限制,有效地处理传统优化算法难以解决的复杂问题。该算法主要包括了基因重组、突变和自然选择,例如遗传算法(GA)、进化策略(ES)^[12]、差分进化和遗传规划(GP)^[13]。

(2)基于群体的算法:该算法是基于整个群体的生存行为而设计的。在群体中,个体和个体之间、个体和外界需要进行信息交互,在获取到知识的基础上不断提高自己。主要包括粒子群算法、蚁群算法^[14]、蝙蝠启发算法和萤火虫算法^[15]。

(3)基于物理的算法:主要是基于某些物理现象而设计出的算法,例如重力、惯性和电磁场等。这类算法主要包括模拟退火算法^[16]、引力搜索算法(GSA)^[17]、和声搜索算法(HS)^[18]和射线优化算法(RO)^[19]。

(4)基于人类的算法:受到人类活动行为的启发而产生的算法,主要包括禁忌搜索算法、帝国主义竞争算法(ICA)^[20]、联赛冠军算法(LCA)^[21]和外汇市场算法(EMA)^[22]。

2.2.2 化学反应算法

人工化学反应优化算法是由 Alatas B^[23] 提出受化学反应启发而提出的优化算法。它通过模拟化学反应中分子之间的相互作用进行优化,并没有其他算法那样复杂的运算符和参数设置,而是将化学反应与优化合并在一起。在很短的时间内,ACRO 已经被用来解决许多问题,并取得了良好的效果。ACRO 不使用全局搜索方法来优化搜索,但它具有全局搜索和局部搜索的能力。在算法的初始化时刻中,不需要指定许多参数,只需确定启动反应所需的初始反应物数量。由于初始反应物分布在一个可行的全局搜索区域内,因此可以用较少的迭代得到最优解,从而大大减少了计算时间。近年来,人工化学反应优化算法已被成功地用于解决许多复杂问题,并被发现优于其他许多基于进化群体的智能算法,但原始的 ACRO 也存在局部搜索能力较差的问题,并且人工化学反应的特点非常适合离散的特征选择,因此我们考虑将局部搜索能力较强的灰狼算法与其混合。目前,人工化学反应优化算法在特征选择、多序列比对、数据挖掘、分类发现等方面应用得还较少。

Bechikh 等^[24] 提出了一种多目标的化学反应优化方法,其被称为非支配排序化学反应优化,试图利用化学反应优化的特征来解决涉及多个冲突准则的问题^[25]。文中提出了一种新的化学反应优化算法和贪婪策略算法(CROG)求解 0-1 背包问题,同时还利用贪婪策略和随机选择相结合的修复函数来修复不可行解。

2.2.3 混合策略优化算法

传统元启发式方法通常面临着陷入局部最优的问题,利用混合策略可以增强和补充原算法的搜索能力,所以基于混合策略启发式算法是弥补原始算法缺陷的一种快速有效的方法。

文献^[26]提出了一种基于粒子群优化的混合特征选择方法,该方法使用一种全新的局部搜索策略来增强搜索在全局

最优值附近的能力,并有效地找到弱相关性的区分性特征。Yu^[27] 提出了一种改进的蝙蝠算法(Bat Algorithm)来缓解原始算法中易陷入局部最优的问题,该算法通过阈值控制混沌映射的步骤,并使用速度惯性权重来同步多个智能体的速度,可以快速提高蝙蝠算法的稳定性和收敛速度。Mafarja 等人^[28] 提出了两种基于鲸鱼优化算法(WOA)的混合模型。在第一个模型中,模拟退火算法被嵌入到鲸鱼算法中,而在第二个模型中,模拟退火算法用来改进鲸鱼算法每次迭代后所找到的最优解。通过鲸鱼算法的搜索定位到最有效的区域,模拟退火算法能够增强算法的探索能力。

通过对相关问题的研究,我们发现混合策略可以很好地解决原始算法的固有缺陷问题,有效提升算法的综合性能。所以本文将延续这一算法混合策略,提升特征选择实验结果。

3 ACR-WCA 算法

本文将化学反应优化算法与灰狼算法相结合,形成了一种新的封装器的特征选择方法——ACR-WCA。现阶段,人工化学反应算法已成功地应用于各种领域,解决了许多复杂问题,但却很少用于特征选择。本文结合化学人工反应稳定的变异能力改进狼群算法,进一步提升混合算法的性能,并将其应用于特征选择中。

3.1 特征初始化

前文所描述的启发式算法最初是为解决连续优化问题而开发的,根据特征选择问题的需要重新设计的算法。在特征选择问题中,算法的重点是选择或不选择数据集中的某个特征,从而最大限度地提高最有效特征的分类精度。原始的狼群算法及人工化学反应算法在解决问题时与特征选择问题所需要的二维性不一致,不能利用这些算法直接解决特征选择问题。所以要将该算法中的每个解向量转换为二进制形式,即仅由 0 和 1 组成,然后继续算法。为了实现这种转换,使用了 S-形转换函数,这个函数给出了在解向量中选择特定特征的概率,如图 2 所示。

$$S(x) = \frac{1}{1 + e^{-x}} \quad (1)$$

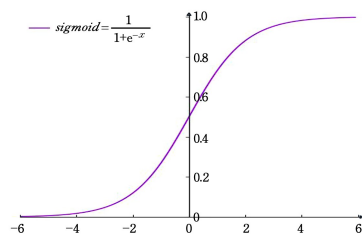


图 2 S-形转换函数

Fig. 2 S-shaped transfer function

每个特征被视为一个二进制向量,其中第 i 个位置对应第 i 个特征的选择情况。在基本参数初始化后,随机生成一组整数 0 或 1 的初始分子,它们参与什么样的化学反应由概率函数决定。在每种情况下,选择的特征的值均为 1,未选择的值为 0。每个分子代表一个可行的解决方案,并且具有等于特征数目的维度 n 。随机分子如图 3 所示, $f_i = 0$ 表示未选择第 i 个特征, $f_i = 1$ 表示选择第 i 个特征。

1	2	3	4	5	6	...	i	n
1	0	0	1	0	1	1	0

图3 二进制特征

Fig. 3 Binary feature

3.2 狼群算法

狼群算法是一种群体智能优化算法。该算法是学者受生物世界中狼的狩猎行为启发而提出的一种优化搜索方法。狼群通常以群居生活为主,拥有紧密的组织系统和协作捕猎方式,它们之间能够互相团结,为狼群的生存和发展承担各自的责任。狼群根据社会分工的专业性,分为以下3种。

(1)头狼:头狼是狼群中最凶猛和聪明的狼,负责对周围的信息做出决策并指挥整个狼群尽快捕捉猎物。通常,将群体中目标函数值最优的人工头狼标记为头狼。每次迭代后的最优狼目标函数值与前一代中的头狼进行对比,若更优则更新头狼位置。头狼不会执行随机游走、传唤行为和围攻行为这3种智能行为中的任何一种,而是直接进入迭代,直到它被更好的人工狼取代。

(2)探狼:寻找猎物的狼作为探狼。它们根据空气中猎物残留的气味进行决策,气味浓度越高,表示距离猎物越近,探狼始终向着气味浓度最高的方向探寻猎物。

(3)猛狼:探狼在发现猎物后会立即向头狼报告,头狼会通过嚎叫来召唤周围的猛狼对猎物进行攻击。周围的猛狼在听到叫声后,会朝着探狼的方向奔进,逼向猎物。

狼群的整个捕猎行为可以抽象概括为以下3种方式,并通过一些数学模型来表示狼群游走、召唤和包围猎物的过程。

3.2.1 随机游走

探狼在每个方向移动相应的步长 h 。同时,它们记录前进一步感知到的猎物气味浓度,然后回到原来的位置 $p(p=1,2,\dots,h)$ 。在向这个方向移动之后,探狼 β 在维度空间中的位置是:

$$X_{\beta d}^k = X_{\beta d} + \sin\left(2\pi \times \frac{p}{h}\right) \times step_p^d \quad (2)$$

探狼收集标记为 $X_{\beta p}$ 的猎物的气味,选择气味大于当前位置 $Y_{\beta o}$ 且气味最大的方向,更新探狼的位置 X_{β} 。重复之前的行走行为,直到其中一只狼感觉到猎物的气味 $Y_{\beta} > Y_{\alpha}$,或者直到行走行为的数量达到最大 T_{\max} 。

3.2.2 传唤行为

头狼通过吼叫来发起传唤行为,听到呼唤后,猛狼就会通过 $step_b^d$ 的步长走到头狼的位置。在 $k+1$ 次迭代中,猛狼在 d 维上的位置是:

$$X_{\beta d}^{k+1} = X_{\beta d}^k + \frac{step_b^d \cdot (g_d^k - X_{\beta d}^k)}{|g_d^k - X_{\beta d}^k|} \quad (3)$$

其中, g_d^k 表示狼的 k 代群体在 d 维空间中的位置。在寻找的途中,如果猛狼嗅到了强烈的气味 $Y_{\beta} > Y_{\alpha}$,那么 $Y_{\alpha} = Y_{\beta}$,猛狼就会变成领头狼并召唤。否则,它继续搜索,直到头狼之间的距离满足 $d_{bl} < d_{near}$ 情况,开始围攻行为。该距离由以下公式计算:

$$d_{near} = \frac{1}{D \cdot \omega} \cdot \sum_{d=1}^D |d_{\max} - d_{\min}| \quad (4)$$

3.2.3 围攻行为

对于 k 代狼来说,猎物在 d 维空间中的位置是 G_d^k ,围攻行为可以表示为:

$$X_{\beta d}^{k+1} = X_{\beta d}^k + \lambda \cdot step_p^d \cdot |d_{\max} - d_{\min}| \quad (5)$$

其中, λ 是在 $[-1,1]$ 随机数之间的均匀分布, $step_p$ 是攻击步长。如果在围攻后猛狼感觉到猎物气味的浓度大于其原来位置,则更新猛狼的位置;否则,狼的位置保持不变。

游走步长 $step_a$ 、奔袭步长 $step_b$ 和攻击步长 $step_c$ 满足以下步长关系:

$$step_a^d = \frac{step_b^d}{2} = 2 \cdot step_c^d = \frac{|d_{\max} - d_{\min}|}{S} \quad (6)$$

其中, S 代表步长因子,表示人工狼在解空间中搜索最优解的精细程度。

3.2.4 更新规则

狼群的更新机制是“强者生存”,猎物也按照“由强到弱”的原则进行分配,所以在算法中去除目标函数值最差的 R 匹人工狼,同时随机产生 R 匹人工狼。 R 越大则新产生的人工狼越多,有利于维护狼群个体的多样性。 R 是来自

$\left[\frac{n}{2 \cdot \mu}, \frac{n}{\mu}\right]$ 的随机整数, μ 表示种群更新比例因子。

3.3 改进的化学反应算法

人工化学反应优化是Alatas于2012年提出的一种新的自适应启发式优化算法,其通过模拟容器内的化学反应过程,解决了全局和局部搜索优化问题;以反应中分子的焓和熵作为目标函数,解决了最小化和最大化问题。该方法的主要步骤为:首先初始化形成反应物的数量;在迭代阶段,当终止条件不满足时,根据条件变量选择化学反应算子执行。反应物的更新等同于可逆反应,有助于目标函数向最优方向发展。

在这一部分中,我们针对每一次人工狼在搜索空间位置的迭代改变提出了5种不同的反应算法,其化学反应的子方法分别为:改进的Redox1反应、改进的分解反应、改进的Redox2反应、改进的置换反应和改进的合成反应。通过在狼群算法中加入化学反应,能够对处在随机行走、召唤行为和包围行为过程中的人工狼进行快速活动,实现了人工狼之间的位置信息交互,使狼群能够更好地收集搜索空间的信息,提高了算法的探索能力。

3.3.1 氧化还原反应 Redox1

Redox1反应,只需要一个分子直接参与反应,即一匹狼。根据初始反应物位置,通过突变反应随机选择某个位置的值进行取反。如果值为0,则更改为1;如果值为1,则更改为0。

该反应的二进制编码示例如图4所示,选择原始狼位置中的0,变为1,其余部分保持不变。

Wolf	1	1	0	0	0	1	0	1	1	1
New wolf	1	1	0	1	0	1	0	1	1	1

图4 二进制氧化还原1反应示例

Fig. 4 Example of binary Redox1 reaction

3.3.2 分解反应

在改进的分解反应中,只需要一个分子参与反应,即一匹狼。在反应过程中,会随机选择一处位点,通过对该位点和其左右3个位置的值进行取反重新产生一个新分子。其余部分

的反应物与原反应物相同。

二进制编码的分解反应示例如图 5 所示。在人工狼上随机选择 3 个索引,新生狼的索引位 4-6 的数值取反,其余部分数值保持不变,与原始狼相同。

Wolf	1	1	0	1	0	1	0	1	1	1
New wolf	1	1	0	0	1	0	0	1	1	1

图 5 二进制分解反应示例

Fig. 5 Example of binary decomposition reaction

3.3.3 氧化还原反应 Redox2

在 Redox2 反应中,需要两个反应物同时参加反应,最后形成两个新的反应物,其中的主要步骤是交换两个分子的部分索引片段。该反应的二进制编码示例如图 6 所示。

Wolf 1	1	0	1	0	0	1	0	1	0	0
Wolf 2	0	1	0	1	1	1	0	1	0	1
New wolf 1	0	1	1	1	1	0	1	0	1	0
New wolf 2	1	0	0	0	0	1	0	1	0	0

图 6 二进制 Redox2 反应示例

Fig. 6 Example of binary Redox2 reaction

假设随机生成的索引序列是 4,5,6 和 7。新生狼 1 的位置 4-7 之间的数值来源于狼 2,而新生狼 2 的位置 4-7 之间的数值来源于狼 1,由此可以产生完全不同于原始狼的两个新生狼。

3.3.4 置换反应

两种反应物通过置换反应可以生成两种新生成物,该反应中最重要的是需要一个和反应物长度相同的随机二元掩码。

在掩码的第 i 索引位为 0 的情况下,新反应物 1 的第 i 比特与反应物 1 的第 i 位保持一直,而新反应物 2 的第 i 比特也一样。否则,掩码的第 i 位值为 1,新反应物 1 的第 i 位是反应物 2 的第 i 位的取值,新反应物 2 的第 i 位是反应物 1 的第 i 位的取值。置换反应的二进制编码示例如图 7 所示。

Wolf 1	1	0	1	0	0	1	0	1	0	0
Wolf 2	0	1	0	1	1	1	0	1	0	1
Mask bit	1	1	0	0	1	0	1	0	0	1
New wolf 1	0	1	1	1	1	0	1	0	1	0
New wolf 2	1	0	0	0	0	1	0	1	0	0

图 7 二进制置换反应示例

Fig. 7 Example of binary displacement reaction

3.3.5 合成反应

在合成反应中,由两个反应物参与反应,但只生成一个生成物。首先,两个反应物先确定出不匹配的位。然后,第一个反应物的非匹配位和第二反应物的非匹配位形成新的反应物。合成反应的二进制编码例子如图 8 所示。

Wolf 1	1	1	1	0	0	1	0	1	1	1
Wolf 2	0	1	0	1	0	1	0	1	0	1
New wolf	1	1	1	0	1	0	1	1	1	1

图 8 二进制合成反应示例

Fig. 8 Example of binary synthesis reaction

狼 1、狼 2 代表两种不同的反应物,每种反应物的大小均为 10 比特位。有 1,3,4 和 6 个位置的非匹配位。根据描述,不匹配的位依次取自狼 1 和狼 2 作为新反应物的对应值,其余的匹配位置保持不变。

在 ACR-WCA 算法中,我们对 5 种化学反应做出了调整改进,在执行完狼群算法之后,将头狼和探狼加入了化学反应迭代中。通过对上述反应进行改进,选择人工狼作为反应分子,并根据概率选择执行哪一种子反应。通过该算法空间位置的快速改变,让狼群个体在解空间中快速逼近最优解,加快整个算法的迭代速度。

3.4 两阶段变异

为了增强 ACR-WCO 算法的探索性,该算法采用两阶段变异来保持人工狼的活性。化学反应后,新个体的位置可能会远离最优解。基于上述伪代码的描述,将对人工狼空间中的二进制位置进行快速改变。第一阶段变异的目的是在保持高精度度的同时减少所选特征的数量,第二阶段变异是为多样性添加更多维护的信息特性。

算法 1 两阶段变异算法

输入: the best artificial wolf X_α from each iteration

输出: the improved X_α

1. fitness = the fitness of X_α ;
2. /* 第一阶段的变异 */
3. Define a vector one_positions to store the binary locations of the selected features in X_α ;
4. Define X_{once}
5. $X_{\text{once}} = X_\alpha$;
6. for i in range(1, len(one_positions)):
7. generate a random number r ;
8. if $r < M_p$;
9. $X_{\text{first}}[\text{one_positions}[i]] = 0$;
10. fitness_{mutated} = the fitness of X_{first} ;
11. if fitness_{mutated} < fitness;
12. fitness = fitness_{mutated};
13. $X_\alpha = X_{\text{first}}$
14. /* 第二阶段变异 */
15. Define a vector zero_positions to store the binary locations of the selected features in X_α ;
16. Define X_{twice}
17. $X_{\text{twice}} = X_\alpha$;
18. for i in range(1, len(zero_positions)):
19. generate a random number r ;
20. if $r < M_p$;
21. $X_{\text{twice}}[\text{zero_positions}[j]] = 1$;
22. fitness_{mutated} = the fitness of X_{twice} ;
23. if fitness_{mutated} < fitness;
24. fitness = fitness_{mutated};
25. $X_\alpha = X_{\text{twice}}$

4 实验结果分析与讨论

4.1 数据集与参数设置

为了评估算法的有效性和准确性,我们将对 21 个著名的

¹⁾ <https://www.openml.org/>

数据集进行特征选择。数据集从 UCI 机器学习存储库下载¹⁾。表 1 列出了所用数据集的所有信息,提供了每个数据集的 4 个属性数据集的名称、特征(#F)、样本(#S)、类别(#C)和领域。

根据数据集所含的特征数,我们将以上 21 个数据集划分为大数据集(#F>30)、中数据集(15<#F≤30)和小数据集(0<#F≤15)3类(如表 2 所列),方便后续实验结果的分析 and 描述。

将 ACR-WCA 的性能与一些经典的和新的算法进行比较,所对比的算法时间跨度从 1995 年到 2020 年,具体算法的描述如下所述。

(1)PSO:粒子群算法,Kennedy 等人受到候鸟迁徙和觅食过程中的聚类行为启发而设计出来的算法(Kennedy and Eberhart^[29])。

(2)FA:花粉算法,Yang 等人受花朵的授粉过程启发的设计出的新算法^[30]。

(3)BA:二进制蝙蝠算法,模仿蝙蝠的回声定位行为进行全局优化的启发式算法等^[31]。

(4)MVO:多节优化器,该算法的主要灵感来自于宇宙中的白洞、黑洞和虫洞概念。通过建立了这 3 个概念的数学模型,分别用于进行勘探、开发和局部搜索而设计出的算法^[32]。

(5)BCSA:二进制乌鸦搜索算法,于 2018 年提出的一种基于乌鸦智能行为的生物启发式算法,称为乌鸦搜索算法^[33]。

(6)bWOA:二进制鲸鱼优化算法,于 2019 年提出,属于鲸鱼优化算法的一种新的二进制版本,用于为降维和分类问题选择最优特征子集^[34]。

(7)TMGWO:二阶段变异的灰狼算法,于 2020 年提出的一种新的灰狼优化算法,它结合了两阶段变异来解决特征选择问题^[35]。

表 1 数据集介绍

Table 1 Introduction to datasets

No.	Dataset	#F	#S	#C	Area
1	Breast cancer tissue	9	106	6	Medical
2	Glass	9	214	7	Physical
3	Stock	9	950	2	Business
4	Page blocks	10	5473	2	Computer
5	Vowel	10	990	2	Life
6	WineEW	13	178	3	Physical
7	Australian	14	690	2	Financial
8	Zoo	16	101	7	Life
9	Lymphography	18	148	4	Medical
10	Vehicle	18	846	4	Life
11	Segment	19	2310	7	Life
12	Climate	20	540	2	Physical
13	SpectEW	22	267	2	Medical
14	WDBC	30	569	2	Life
15	IonosphereEW	34	351	2	Physical
16	SonarEW	60	208	2	Physical
17	Robot-failures-lp1	90	88	4	Physical
18	Robot-failures-lp2	90	47	5	Physical
19	Robot-failures-lp3	90	47	4	Physical
20	Robot-failures-lp4	90	117	3	Physical
21	Robot-failures-lp5	90	164	5	Physical

表 2 大中小数据集分类

Table 2 Classification of large, medium, small datasets

Large	Medium	Small
IonosphereEW	Zoo	Breastcancer tissue
SonarEW	Lymphography	Glass
Robot-failures-lp1	Vehicle	Stock
Robot-failures-lp2	Segment	Page blocks
Robot-failures-lp3	Climate	Vowel
Robot-failures-lp4	SpectEW	WineEW
Robot-failures-lp5	WDBC	Australian

实验所需的参数值都展示在表 3 中。一方面,为了保证这些算法之间的合理比较,我们不对参数进行调整;另一方面,这些数据更有针对性地突出了我们算法的性能和特点。

表 3 提出方法的参数设置

Table 3 Parameters setting of the proposed algorithm

Parameter	Value
Number of independent runs	20
Number of iterations	20
Dimension D	Number of features
K -neighbors	5
K -fold cross-validation	10
α	0.01
β	0.99
M_p	0.5

4.2 评价函数

在设计完算法后,需要通过一些评价指标来衡量算法的优劣程度。就算法的稳定性而言,其可以看作算法在增加或删除新的训练样本时产生一致的特征子集的一致性。同时,算法必须达到以下两点:(1)最小化选择的特征数;(2)最大化给定分类器的分类准确度。在上述分析的基础上,设计了评价解的适应度函数,以达到以下两个目标之间的平衡。在这一部分中,我们使用几个度量函数来比较 ACR-WCA 的性能。适应度函数值描述如下:

$$fitness = \alpha \rho_s(D) + \beta \frac{|L|}{|d|} \quad (7)$$

其中, $\alpha \in [0, 1]$, $\beta = 1 - \alpha$ 。 α 和 β 是对应于分类精度的重要性和所选特征子集长度的权重参数。 D 是数据集 S 的决策。 $\rho_s(D)$ 通过分类器 KKN 计算出的分类错误率得到。 $|L|$ 表示选定子集的长度。 $|d|$ 是原始数据集每个样本中找到的所有特征的数量。最优(最佳)适应值是基于式(8)分别运行 ACR-WCA 算法 20 次所获得的最小适应值。

$$fitness = \min_{i=1}^{20} f_i \quad (8)$$

算法的平均适应值是 20 次训练结果中每一次最优适应值的平均值:

$$\overline{fitness} = \frac{1}{20} \sum_{i=1}^{20} f_i \quad (9)$$

特征选择率描述在算法执行过后,算法选择出来的特征数量和原始数据集特征数量的比值,数值越低表示选择出的高信息量的特征越少,算法的筛选越高效。

$$S = \frac{\#F'}{Num_of_feature} \times 100\% \quad (10)$$

4.3 结果比较与分析

在每个数据集运行 20 次之后,计算每个数据集独立运行后 21 个数据集的相关指数,表 4 中显示了关于执行 ACR-WCA 算法后的大致情况,从表中可以直观看出各个数据集在 20 次迭代里找出的特征数量、适应度和准确率的最好、最

坏和平均值。与原始数据集的特征相比,即使是最差的选择特征数量也明显减少。

随着迭代次数的递增,最佳精确度值的收敛如图 9 所示。很明显,80%的数据集的精确定率收敛很快,ACR-WCA 算法也可以在 10 次迭代内实现最终的精度。这是因为加入了化学

反应算法之后,由于初始值分布在一个可行域中,因此能够通过较少的迭代尽快得到最优解,所以该算法能快速找到最优准确率值,收敛速度更快。在某些情况下,如果一开始随机初始化时就找到了最优解,那么整个期间的迭代就会呈现一条直线,直至迭代结束。

表 4 21 个数据集的特征数量、适应度和准确率

Table 4 Number of features, fitness and accuracy of 21 datasets

Datasets		# F'	fitness	accuracy	Datasets		# F'	fitness	accuracy
Australian	Worst	6	0.2449	0.7751	Robot-faliures-lp5	Worst	31	0.2657	0.7770
	Average	4	0.1444	0.8541		Average	26	0.2449	0.7514
	Best	3	0.1337	0.8678		Best	19	0.2242	0.7275
Breast cancer tissue	Worst	8	0.5468	0.4533	Segment	Worst	9	0.0916	0.9119
	Average	6	0.5008	0.4929		Average	5	0.0374	0.9647
	Best	3	0.2926	0.7089		Best	4	0.0300	0.9727
Climate	Worst	10	0.0827	0.9184	SonarEW	Worst	23	0.2460	0.7544
	Average	5	0.0765	0.9239		Average	15	0.2047	0.7864
	Best	2	0.7185	0.9389		Best	9	0.1027	0.8365
Glass	Worst	7	0.2942	0.6950	SpectEW	Worst	12	0.2544	0.7342
	Average	6	0.2807	0.7185		Average	9	0.2286	0.7671
	Best	4	0.2722	0.7318		Best	3	0.2100	0.7933
IonosphereEW	Worst	17	0.08818	0.9088	Stock	Worst	4	0.0917	0.8972
	Average	12	0.0727	0.9257		Average	4	0.0914	0.9051
	Best	8	0.0581	0.9408		Best	3	0.0905	0.9119
Lymphography	Worst	10	0.1503	0.8409	Vehicle	Worst	12	0.2832	0.7173
	Average	9	0.1270	0.8710		Average	9	0.2704	0.7288
	Best	6	0.1107	0.8934		Best	6	0.2576	0.7432
Page blocks	Worst	5	0.0524	0.0948	Vowel	Worst	9	0.0312	0.9727
	Average	4	0.0443	0.9577		Average	6	0.0150	0.9884
	Best	3	0.0411	0.9624		Best	4	0.0110	0.9929
Robot-faliures-lp1	Worst	21	0.0868	0.8956	WineEw	Worst	7	0.0535	0.9341
	Average	13	0.0661	0.9292		Average	5	0.0491	0.9516
	Best	6	0.0248	0.9532		Best	2	0.0473	0.9570
Robot-faliures-lp2	Worst	18	0.2555	0.7600	WDBC	Worst	6	0.0691	0.9308
	Average	13	0.1993	0.8000		Average	4	0.0526	0.9480
	Best	7	0.1772	0.8633		Best	3	0.0481	0.9525
Robot-faliures-lp3	Worst	7	0.2652	0.7333	Zoo	Worst	7	0.0692	0.9726
	Average	10	0.2157	0.7826		Average	6	0.0183	0.9843
	Best	18	0.1675	0.8317		Best	3	0.0120	0.9917
Robot-faliures-lp4	Worst	29	0.0726	0.9280					
	Average	17	0.0458	0.9412					
	Best	11	0.4529	0.9565					

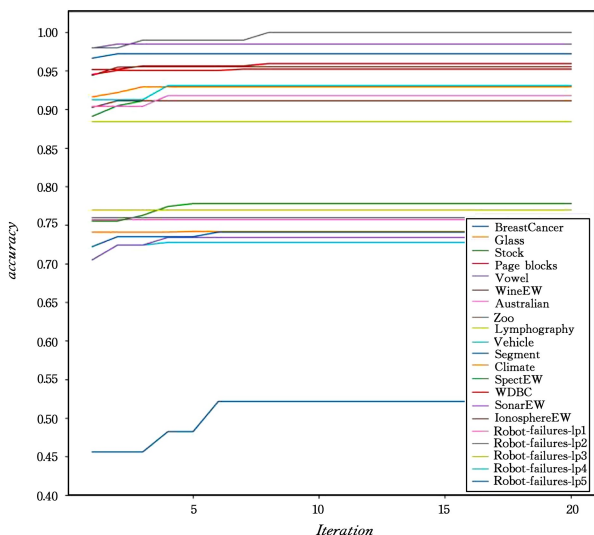


图 9 利用 ACR-WCA 算法的分类精度的变化

Fig. 9 Variation of classification accuracy obtained using the ACR-WCA algorithm

图 10 中比较了基于不同算法的所有数据集的平均选择特征数。很明显, MVO 算法所选择出的特征数最少。本文提出的 ACR-WCA 算法也属于二进制优化算法,取值为 9,与其他几种算法相差不多,但是并不是最优。

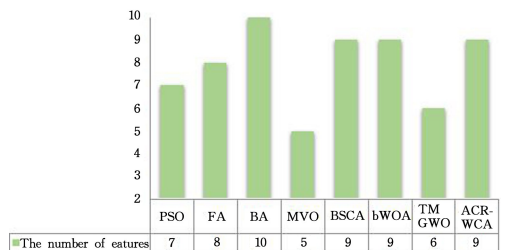


图 10 8 种算法的平均特征数量对比

Fig. 10 Comparison of the average number of features of the 8 algorithms

表 5 列出了大、中、小 3 种数据集经过 ACR-WCA 算法的平均选择特征数量情况。通过实验结果不难看出,该算法对于较大的数据集(即特征数量大于 30),选择出的有效特征数大量减少,选择率均在 36% 以下,选择数量降低较

多。对于中型数据集,特征选择的数量减少在 50% 以下,意味着选择出的数量普遍减少了一半甚至更多。对于小

型数据集来说,因为原本特征数就在 15 个以下,对特征过滤程度不高。

表 5 大、中、小数据集的平均特征筛选率

Table 5 Average feature screening rate of large, medium and small datasets

Large Datasets	# F'	S/%	Medium Datasets	# F'	S/%	Small Datasets	# F'	S/%
IonosphereEW	12	35.3	Zoo	6	37.5	Breast cancer tissue	6	66.7
SonarEW	15	25	Lymphography	9	50	Glass	6	66.7
Robot-faliures-lp1	13	14.4	Vehicle	9	50	Stock	4	44.4
Robot-faliures-lp2	13	14.4	Segment	5	26.3	Page blocks	4	40
Robot-faliures-lp3	10	11.1	Climate	5	25	Vowel	6	60
Robot-faliures-lp4	17	20	SpectEW	9	40.9	WineEw	5	38.5
Robot-faliures-lp5	26	28.9	WDBC	4	13.3	Australian	4	28.6

表 6 的数值描述了 8 种算法经过训练后的平均分类准确率。实验数据表明,本文提出的 ACR-WCA 算法具有较高的准确率,弥补了以往算法的不足,为特征选择提供了一种很好的方法。

对于所有的数据集和迭代,ACR-WCA 都有较好的分类精度质量,特别是在 Breast cancer tissue 和 Zoo 等数据集上,ACR-WCA 算法大幅度优于其他算法。在 Pageblocks 等数

据集上,其他经典算法还是表现出了良好的准确率,但是 ACR-WCA 实验结果与其相差不多。因此,该算法在准确率上总体具有较好的选择精度,在准确率这一方面是可靠的。几种对比算法的总平均准确率如图 11 所示。ACR-WCA 算法整体效果较好,而且略高于其他排名靠前的算法。通过分析可以看出,ACR-WCA 算法优于大多数算法,比最差算法提升较多。

表 6 21 个数据集的平均准确率

Table 6 Average accuracy of 21 datasets

No.	Dataset	PSO	FA	BA	MVO	BCSA	bWOA	TMGWO	ACR-WCA
1	Australian	0.8333	0.8318	0.8312	0.8318	0.8304	0.8391	0.8463	0.8543
2	Breast cancer tissue	0.3200	0.3300	0.3300	0.3300	0.3300	0.3200	0.3300	0.4929
3	Climate	0.9203	0.9277	0.9184	0.9277	0.9416	0.9222	0.9314	0.9243
4	Glass	0.7380	0.7380	0.7342	0.7380	0.7380	0.7380	0.7380	0.7182
5	IonosphereEW	0.9142	0.9057	0.8765	0.9057	0.9085	0.8942	0.9314	0.9244
6	Lymphography	0.8785	0.8571	0.8775	0.8857	0.8785	0.8714	0.9000	0.8701
7	Page blocks	0.9643	0.9612	0.9163	0.9610	0.9559	0.9643	0.9639	0.9577
8	Robot-failures-lp1	0.8875	0.8750	0.8924	0.9125	0.8875	0.8750	0.9125	0.9268
9	Robot-failures-lp2	0.7500	0.7500	0.7500	0.7750	0.7750	0.7250	0.7750	0.8002
10	Robot-failures-lp3	0.7000	0.7000	0.7250	0.7500	0.7500	0.7250	0.7750	0.7785
11	Robot-failures-lp4	0.9000	0.8909	0.8750	0.9272	0.8818	0.8818	0.9363	0.9422
12	Robot-failures-lp5	0.6375	0.6437	0.6374	0.6875	0.6250	0.6312	0.6750	0.7510
13	Segment	0.9658	0.9709	0.9714	0.9744	0.9679	0.9679	0.9744	0.9647
14	SonarEW	0.6800	0.6600	0.6550	0.7300	0.6799	0.6600	0.7400	0.7855
15	SpectEW	0.7346	0.7436	0.7423	0.7576	0.7346	0.7230	0.7615	0.7671
16	Stock	0.9326	0.9326	0.9326	0.9326	0.9326	0.9326	0.9326	0.9051
17	Vehicle	0.7621	0.7273	0.7261	0.7345	0.7321	0.7309	0.7380	0.7295
18	Vowel	0.9888	0.9888	0.9888	0.9888	0.9534	0.9672	0.9888	0.9884
19	WineEW	0.9352	0.9411	0.9411	0.9470	0.9325	0.9411	0.9470	0.9512
20	WDBC	0.9482	0.9482	0.9482	0.9482	0.9464	0.9482	0.9482	0.9475
21	Zoo	0.9600	0.9600	0.9600	0.9600	0.9600	0.9600	0.9600	0.9841

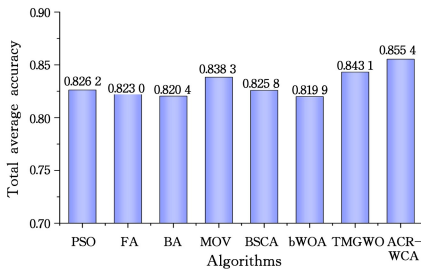


图 11 8 种算法的平均特征准确率对比

Fig. 11 Comparison of total average accuracy of the 8 algorithms

根据上面的描述,较低的适应度表示算法较好。对于所有数据集,实验在测试完成后计算所有的平均适应度,结果如表 7 所列。结果表明,ACR-WCA 的适应度值在大部分数据集表现较低,优于之前的经典算法。对于 Glass, Segment, Vowel 和 Stock 数据集,和最优的值差异不大。就适应度方

面而言,ACR-WCA 算法有所优势。

如图 12 所示,本文算法的适应度比所有算法都低。ACR-WCA 比 TMGWO 更低,性能改善明显。这也表明,随着化学反应的加入,对人工狼的诱导改变再产生的最优解具有更好的性能。

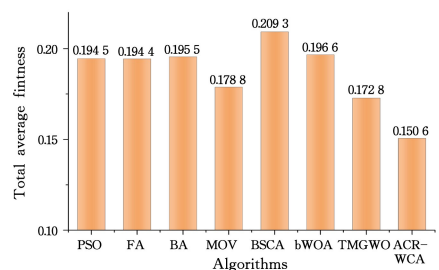


图 12 8 种算法的平均特征适应度对比

Fig. 12 Comparison of total average fitness of the 8 algorithms

表 7 21 个数据集的平均适应度
Table 7 Average fitness of 21 datasets

No.	Dataset	PSO	FA	BA	MVO	BCSA	bWOA	TMGWO	ACR-WCA
1	Australian	0.1900	0.1819	0.1781	0.1759	0.2222	0.1882	0.1686	0.1443
2	Breast cancer tissue	0.7014	0.6844	0.6955	0.6788	0.7313	0.7007	0.6817	0.5001
3	Climate	0.0861	0.0827	0.0849	0.0844	0.0774	0.0845	0.0804	0.0765
4	Glass	0.2687	0.2678	0.2691	0.2652	0.2741	0.2683	0.2647	0.2808
5	IonosphereEW	0.1175	0.1146	0.1257	0.1118	0.1202	0.1190	0.0906	0.0737
6	Lymphography	0.1615	0.1577	0.1587	0.1441	0.1796	0.1531	0.1400	0.1283
7	Page blocks	0.0469	0.0472	0.0458	0.0455	0.0509	0.0485	0.0455	0.0443
8	Robot-failures-lp1	0.1403	0.1412	0.1464	0.1138	0.2014	0.1487	0.0909	0.0680
9	Robot-failures-lp2	0.2657	0.2759	0.2739	0.2489	0.2903	0.2835	0.2415	0.1992
10	Robot-failures-lp3	0.3056	0.3058	0.3026	0.2716	0.3255	0.3061	0.2493	0.2190
11	Robot-failures-lp4	0.1303	0.1405	0.1464	0.0912	0.1563	0.1400	0.0826	0.0563
12	Robot-failures-lp5	0.3874	0.3849	0.3860	0.3460	0.4043	0.3866	0.3430	0.2804
13	Segment	0.0415	0.0413	0.0381	0.0311	0.0464	0.0419	0.0284	0.0343
14	SonarEW	0.3795	0.4055	0.3957	0.3347	0.3914	0.3876	0.3337	0.2424
15	SpectEW	0.2833	0.2876	0.2858	0.2727	0.3053	0.2880	0.2644	0.2286
16	Stock	0.0773	0.0786	0.0776	0.0723	0.0191	0.0775	0.0730	0.1936
17	Vehicle	0.2855	0.2868	0.2885	0.2808	0.3074	0.2833	0.2731	0.2569
18	Vowel	0.0224	0.0109	0.0109	0.0109	0.0437	0.0313	0.0109	0.0150
19	WineEW	0.0798	0.0742	0.0824	0.0706	0.1211	0.0800	0.0638	0.0494
20	WDBC	0.0605	0.0604	0.0602	0.0559	0.0700	0.0598	0.0535	0.0531
21	Zoo	0.0529	0.0518	0.0541	0.0488	0.0570	0.0517	0.0485	0.0186

结束语 本文提出了一种全新的二进制人工化学反应狼群优化的特征选择算法。该算法会先执行狼群的 3 种群体智能行为,之后经 5 种不同的化学应对人工狼位置进行改变,最后成功获得较高的分类准确率和最少的选择特征数。同时也证明了,混合策略算法可以作为一种思路来协同改进一些经典算法。但是,所提算法对特征较少的小数据集的优化不够明显。未来,我们希望继续提出新的基于自然的群体智能优化算法,针对不同类型数据集提出更准确的选择方法。

参考文献

[1] ZHONG X, MA Z P, ZHANG B. Overview of Data Mining[J]. Pattern Recognition and Artificial Intelligence, 2001(1):50-57.

[2] HAN J W, KAMBER M. Data mining concepts and techniques [M]Beijing: China Machine Press, 2012.

[3] WU H S, ZHANG F M, WU L S. New swarm intelligence algorithm-wolf pack algorithm[J]. Systems Engineering and Electronics, 2013(11):204-212.

[4] KENNEDY J, EBERHART R C. Particle swarm optimization [C]//Proc IEEE International Conference on Neural Networks. IV, Perth, Australia, 1995:1942-1948.

[5] HOLLAND J H. Adaptation in natural and artificial system [M]. Ann Arbor: The University of Michigan Press, 1975:68-73.

[6] LAM A Y S, LI V O K. Cheical-Reaction-Inspired Metaheuristic for Optimization[J]. IEEE Transactions on Evolutionary Computation, 2010, 14(3):381-399.

[7] KOZODOI N, LESSMANN S, PAPA KONSTANTINO U K, et al. A multi-objective approach for profit-driven feature selection in credit scoring[J]. Decision Support Systems, 2019, 120: 106-117.

[8] HE X, CAI D, NIYOGI P. Laplacian score for feature selection [C] // Advances in Neural Information Processing Systems. 2006:507-514.

[9] ARAUZO-AZOFRA A, BENITEZ J M, CASTRO J L. A feature set measure based on relief[C]//Proceedings of the fifth inter-

national conference on Recent Advances in Soft Computing. 2004:104-109.

[10] SHI W F, HU X G, YU K. K-part Lasso based on feature selection algorithm for high-dimensional data [J]. Computer Engineering and Application, 2012, 48(1):157-161.

[11] WANG Q T. Application of meta heuristic algorithm in discrete location [D]. Nanjing: Nanjing University of Aeronautics and Astronautics, 2010.

[12] BEYER H G, SCHWEFEL H P. Evolution strategies- A comprehensive introduction[J]. Natural Computing, 2002, 1(1):3-52.

[13] KOZA J R. Genetic programming as a means for programming computers by natural selection[J]. Statistics and Computing, 1994, 4(2):87-112.

[14] DORIGO M, BIRATTARI M, STÜTZLE T. Ant Colony Optimization[J]. IEEE Computational Intelligence Magazine, 2006, 1(4):28-39.

[15] YANG X S. Firefly Algorithm, Stochastic Test Functions and Design Optimisation [J]. International Journal of Bio Inspired Computation, 2010, 2(2):78-84.

[16] KIRKPATRICK S, GELATT C D, VECCHI M P. Optimization by simulated annealing[J]. Science, 1983, 220(4598):671-680.

[17] RASHEDI E, NEZAMABADI-POUR H, SARYAZDIS. GSA: a Gravitational Search Algorithm[J]. Information Sciences, 2009, 179(13):2232-2248.

[18] GEEM Z W, KIM J H, LOGANATHANG V. A New Heuristic Optimization Algorithm: Harmony Search[J]. Simulation, 2001, 2(2):60-68.

[19] KAVEH A, KHAYATAZADM. A new meta-heuristic method: Ray Optimization[J]. Computers & Structures, 2012, 112-113 (DEC.):283-294.

[20] ATASHPAZ-GARGARI E, LUCAS C. Imperialist competitive algorithm: An algorithm for optimization inspired by imperialistic competition[C]//IEEE Congress on Evolutionary Computation. IEEE, 2008.

[16] HUANG H H, CHEN H H. Contingency and comparison relation labeling and structure prediction in Chinese sentences[C]// Proceedings of the 13th Annual Meeting of the Special Interest Group on Discourse and Dialogue. 2012:261-269.

[17] YANG J C, CHEN Z Z, SHEN X J, et al. Automatic identification of the relational category of marked complex sentences in the non-filling state of two sentences[J]. Application Research of Computers, 2017, 34(10): 2950-2953.

[18] YANG J C, WANG Y Y, CAO Y, et al. Feature fusion CNN relation recognition method of relative words non-filled complex sentences[J]. Computer Systems & Applications, 2020, 29(6): 224-229.

[19] SUN K L, DENG Z H, LI Y, et al. Recognition of Chinese Complex Sentence Relations Based on Multi-channel CNN Based on

In-Sentence Attention Mechanism[J]. Journal of Chinese Information Processing, 2020, 34(6): 9-17, 26.

[20] LAI S, XU L, LIU K, et al. Recurrent convolutional neural networks for text classification[C]// Twenty-ninth AAAI Conference on Artificial Intelligence. 2015.



YANG Jin-cai, born in 1967, doctor, professor, doctoral supervisor, is a member of China Computer Federation. His main research interests include modern database and information system, Chinese information processing, artificial intelligence and natural language processing.

(上接第 101 页)

[21] KASHAN A H. League Championship Algorithm: A New Algorithm for Numerical Function Optimization[C]// International Conference of Soft Computing and Pattern Recognition. Malacca, 2009: 43-48.

[22] GHORBANI N, BABAEI E. Exchange market algorithm[J]. Applied Soft Computing Journal, 2014, 19: 177-187.

[23] ALATAS B. ACROA: Artificial Chemical Reaction Optimization Algorithm for global optimization[J]. Expert Systems with Applications, 2011, 38(10): 13170-13180.

[24] BECHIKH S, CHAABANI A, BEN SAID L. An Efficient Chemical Reaction Optimization Algorithm for Multiobjective Optimization[J]. IEEE Trans. Cybern., 2015, 45(10): 2051-2064.

[25] TRUONG T K, LI K, XU Y. Chemical reaction optimization with greedy strategy for the 0-1 knapsack problem[M]. Elsevier Science Publishers B. V., 2013.

[26] MORADI P, GHOLAMPOUR M. A hybrid particle swarm optimization for feature subset selection by integrating a novel local search strategy[J]. Applied Soft Computing, 2016, 43(C): 117-130.

[27] YU H, ZHAO N, WANG P, et al. Chaos-enhanced synchronized bat optimizer[J]. Applied Mathematical Modelling, 2019, 77.

[28] MAFARJA M M, MIRJALILI S. Hybrid Whale Optimization Algorithm with simulated annealing for feature selection[J]. Neurocomputing, 2017, 260: 302-312.

[29] KENNEDY J, EBERHART R C. Particle Swarm Optimization [C]// Proceedings of IEEE Conference on Neural Networks. Perth: IEEE, 1995: 1942-1948.

[30] YANG X S. Flower Pollination Algorithm for Global Optimization [C] // International Conference on Unconventional Compu-

ting and Natural Computation. Berlin: Springer, 2012.

[31] MIRJALILI S, MIRJALILI S M, YANG X S. Binary bat algorithm[J]. Neural Computing & Applications, 2014, 25(3): 663-618.

[32] MIRJALILI S, MIRJALILI S M, HATAMLOU A. Multi-Verse Optimizer: a nature-inspired algorithm for global optimization [J]. Neural Computing and Applications, 2015, 27(2): 495-513.

[33] SOUZA R C T D, COELHO L D S, MACEDO C A D, et al. A V-Shaped Binary Crow Search Algorithm for Feature Selection [C]// 2018 IEEE Congress on Evolutionary Computation (CEC). IEEE, 2018.

[34] HUSSEIN A G, HASSANIEN A E, HOUSSEIN E H, et al. S-shaped Binary Whale Optimization Algorithm for Feature Selection[M]// Recent Trends in Signal and Image Processing. Berlin: Springer, 2019: 79-87.

[35] ABDEL-BASSET M, EL-SHAHAT D, EL-HENAWY I, et al. A new fusion of grey wolf optimizer algorithm with a two-phase mutation for feature selection[J]. Expert Systems with Applications, 2020, 139: 112824.



ZHANG Ya-chuan, born in 1996, post-graduate. Her main research interests include swarm intelligence optimization and data mining.



KANG Yan, born in 1972, Ph.D, associate professor. Her main research interests include transfer learning, deep learning and integrated learning.