

Evolutionary Machine Learning: A Survey

AKBAR TELIKANI, 澳大利亚卧龙岗大学

AMIRHESSAM TAHMASSEBI, 佛罗里达州立大学

WOLFGANG BANZHAF, 密歇根州立大学

AMIR H. GANDOMI, 澳大利亚悉尼科技大学

进化计算 (EC) 方法受自然启发, 以随机方式解决优化问题。它们可以提供可靠且有效的方法来解决实际应用中的复杂问题, 最近已被用于提高机器学习 (ML) 模型的性能及其结果的质量。进化方法可用于 ML 的三个部分: 预处理 (例如, 特征选择和重采样)、学习 (例如, 参数设置、隶属函数和神经网络拓扑) 和后处理 (例如, 规则优化、决策树/支持) 向量修剪和集成学习)。本文研究了 EC 算法在解决不同 ML 挑战。

我们在这里不提供对进化 ML 方法的全面回顾; 相反, 我们讨论了 EC 算法如何通过解决人工智能和 ML 社区的传统挑战来为 ML 做出贡献。我们在九个子领域研究 EC 对 ML 的贡献: 特征选择、重采样、分类器、神经网络、强化学习、聚类、关联规则挖掘和集成方法

对于每个类别, 我们从三个方面讨论进化机器学习: 问题制定、搜索机制和适应度值计算。我们还考虑应在未来工作中解决的未解决问题和挑战。

CCS 概念: • 一般和参考 → 调查和概述; • 计算方法 → 机器学习; 人工智能;

其他关键词和短语: 进化计算、学习优化、群体智能

ACM 参考格式:

Akbar Telikani, Amirhessam Tahmassebi, Wolfgang Banzhaf, and Amir H. Gandomi. 2021. Evolutionary Machine Learning: A Survey. ACM Comput. Surv. 54, 8, Article 161 (October 2021), 35 pages

<https://doi.org/10.1145/3467477>

1. 介绍

在数据中寻找模式是**机器学习 (Machine Learning, 简称ML)**的核心和最重要的一步。

毫无疑问, 图灵用它来帮助破解纳粹军方令人烦恼的 Enigma 机器是它的原理最成功的早期应用之一, 其方法是建造一台可以快速分类数百万种猜测代码的可能性的机器。然后在 1950 年, 艾伦·图灵提出了一种称为“学习机”的方法来实现进化。如今, 最新、最强大的 ML 技术受到自然的启发, 被称为自然计算领域。自然计算的概念和术语有两个重要来源: (1) 从自然中汲取灵感, (2) 使用计算机。该术语可用于模拟自然现象、使用天然材料或开发新技术来解决问题。作为人工智能的一部分, **进化计算 (Evolutionary Computation, 简称EC)**方法被认为是该领域的一个类别, 它们的力量源于用于生产智能生物的自然过程。EC 中应用的过程受到自然进化的启发, 而最佳解决方案自然已经进化了数百万年。因此, 可以预期 EC 技术是高效和有效的。EC 算法通常适用于与要解决的特定问题相关的个体群体。

进化和学习是自然和人工系统适应的两个方面, **可以根据个体的寿命**来辨别它们。如果一个个体在其寿命中**适应了某个问题领域**, 我们称之为**学习适应**。如果一个个体是那些特征在几代个体的过程中不断变化的个体们的遗传序列的一部分, 我们称之为**进化适应**。类似于进化和学习相辅相成的自然系统, 进化计算方法和学习算法之间的计算存在双向关系, 因此它们可以结合起来共同解决不同领域的复杂优化问题, 例如能源、机械、医疗、工程和制药行业。一方面, 学习算法被集成到进化技术中, 以解决进化计算方法的问题, 例如陷入局部最优和过早收敛。在这方面提出的一些工作是, 例如, 使用 Q-learning [96, 97]、自适应学习 [94]、Taguchi 方法 [95] 和平衡学习策略 [98] 改进的 Cuckoo Search 算法。学习算法也被用于**粒子群优化 (Particle Swarm Optimization, 简称PSO)** [148] 和大象放牧优化 [99, 100]。

另一方面，进化算法可用于改进机器学习算法，这是本文的主要主题。现实世界应用程序中的大多数问题都包含不准确、嘈杂、离散和复杂的数据，进化计算算法凭借其通用性和随机搜索方法，提供了很好的优化机会[183]。近年来，许多研究人员已将 EC 方法集成到 ML 过程的不同阶段（即预处理、学习和后处理），以解决传统方法的局限性。这些新的混合方法被称为**进化机器学习（Evolutionary Machine Learning, EML）**，ML 学习阶段的 EC 还指进化的 AutoML 概念，其中 ML 模型的不同专家设计的组件，如架构和超参数，是使用 EC 方法自动确定的。此外，如基于梯度的训练算法等优化算法被 EC 算法取代，甚至由 EC 方法发明 [103, 144]。

许多调查和评论论文已经发表，涵盖了 EML 的特定方面，例如，Al-Sahaf 等人[3]发表了一篇评论论文，讨论了主要的 EML 任务，如分类、回归和聚类。Badhon 等人[15]发表了一篇综述文章，论述了关联规则挖掘（ARM）的多目标进化算法（MOEA）。此外，Telikani 等人[169]发表了一篇关于 EC 技术在 ARM 中的应用的综述论文。除了最近发表的关于进化特征选择的论文[183]之外，Barros 等人[17]还发表了一份针对决策树（DT）归纳的进化算法的调查报告。发表了四篇调查论文，以解决进化聚类问题，其中引用了文献[61、64、129、133]。Mukhopadhyay 等人[129]发表了一份多目标进化聚类技术的调查报告，从不同的方面来看，包括表示技术、目标函数、进化操作、维持非支配个体的策略以及最终的个体选择。Darwish 等人[38]回顾了群体智能和 EC 方法在深度学习中的应用。Mukhopadhyay 等人[130,131]出版了一份两部分的综述讨论了用于数据挖掘问题的多目标进化算法的最新发展，如特征选择、分类、聚类和 ARM。

本文的重点和内容与这些调查略有不同。然而，其他调查侧重于为特定任务/方面设计的 EC 算法，如特征选择[183]、DT[17]、ARM[15、169]、聚类[61、64、129、133]和深度学习[38]，本文研究了如何将 EC 算法应用于 ML 的不同方面，以及如何利用进化搜索机制对 ML 问题进行优化的更一般问题。我们从以下三个方面来讨论 EML：问题表述\个体表征，搜索机制和适应度函数。

本文的其余部分组织如下：第2节提供了 ML 和 EC 的背景信息。第3节考虑了 EC 在 ML 不同部分的应用，并概述了 EML 方法。第4节回顾了 EML 方法的应用。第5节讨论了当前的问题和挑战。最后，第6节根据当前的问题和挑战，对当前的最新技术进行了重要总结。

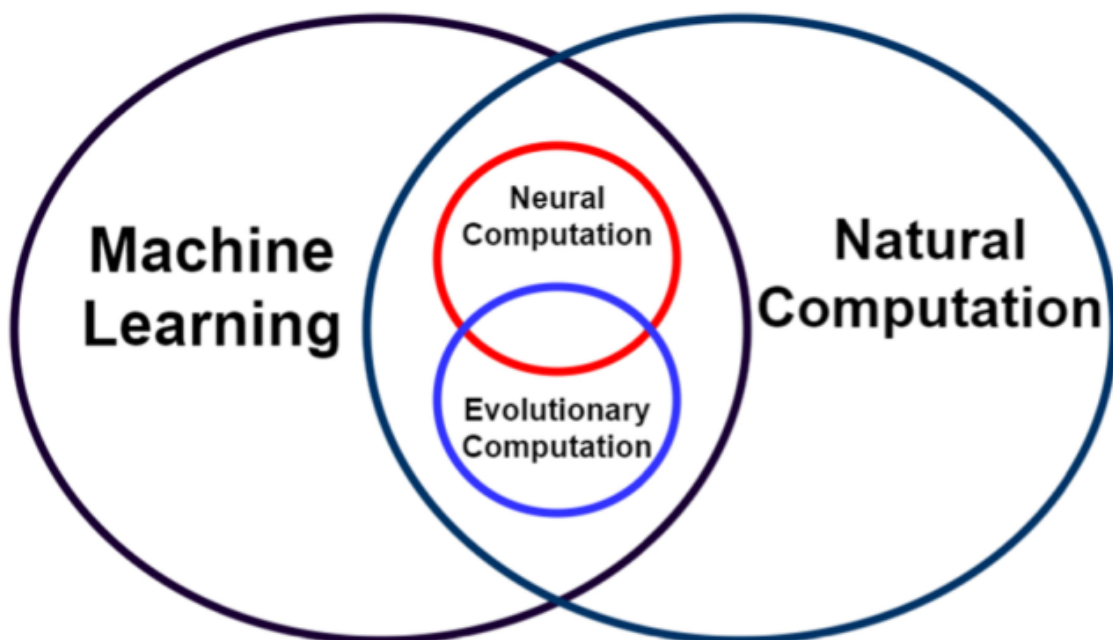


图1：机器学习与自然计算背景下自然计算与进化计算的交叉

2.基本概念

接下来的小节提供了机器学习和进化计算的分类和特征。图1显示了神经和进化计算概念如何与 EC 相交，因为它们与 ML 和自然计算相关。

1 机器学习

机器学习是人工智能技术的一个子集，通过使用数学、统计学、优化和知识发现方法应用算法来提取模式。

图2说明了ML的基本分类，包括三个主要类别：（1）监督学习，（2）无监督学习，和（3）强化学习（Reinforcement Learning，简称RL）[21]。

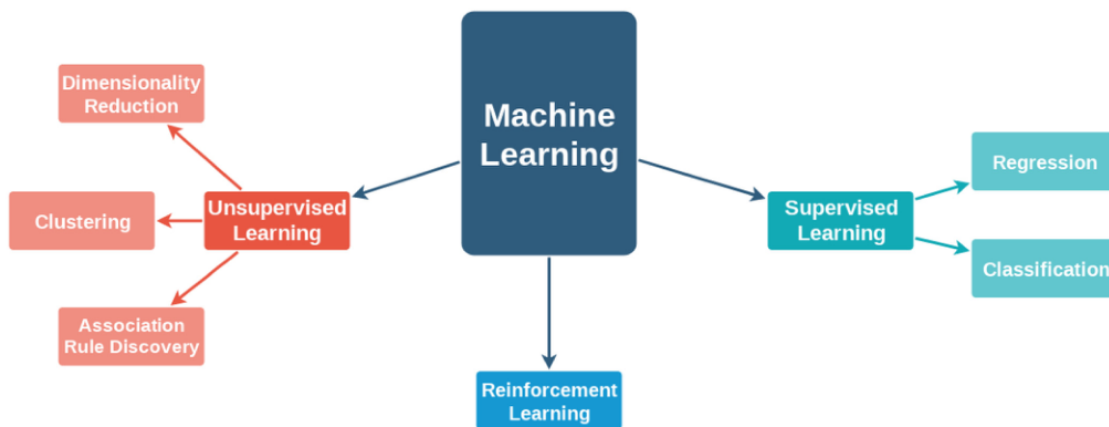


图2：机器学习分类

监督学习是最著名的ML数据处理任务，它试图找到为训练系统提供的一组输入和输出之间的关系[177]。在训练过程最后，应用一个从输入 x 到输出 y 的最佳估计的映射函数（ $f:x \rightarrow y$ ）。监督学习算法建立一个模型来表示输入特征之间的关系，用于预测目标输出。[84]这些算法包括两大类：（1）分类（离散建模）和（2）回归（连续建模）。这两类都是预测建模技术，唯一的区别是它们的目标（响应）变量。在分类中，目标变量采用类别（类别标签）的形式，如二元类或多类问题。然而，在回归中，目标变量是连续的，图3显示了使用ML算法构建模型的示意图。

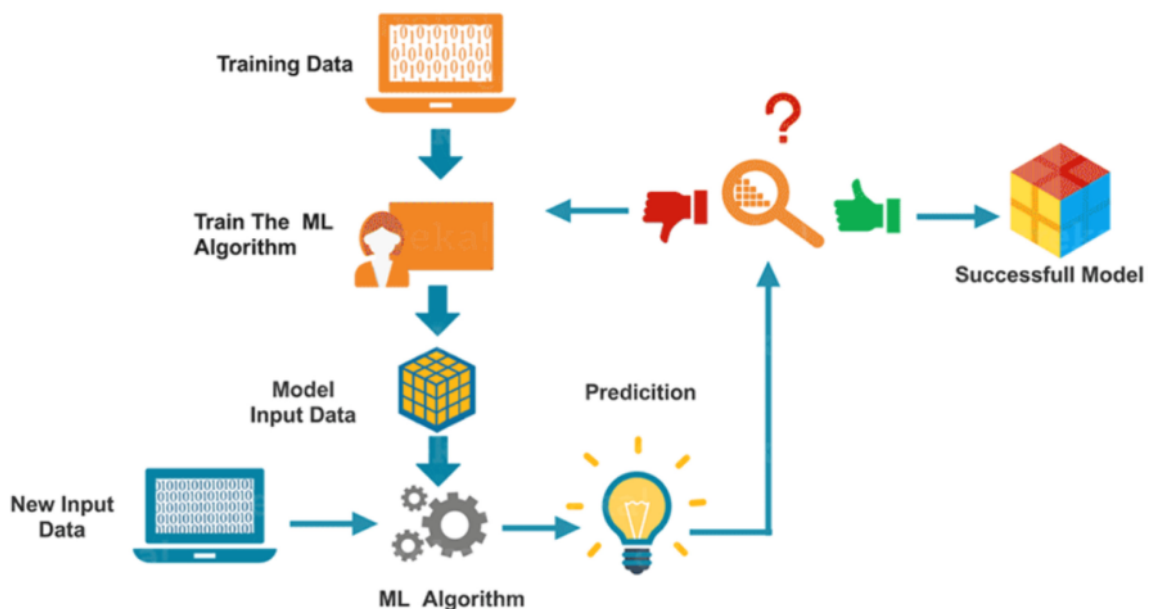


图3. 使用 ML 构建模型的示意图

原始数据是无监督学习的唯一输入，与监督学习不同，无监督学习没有可用于监督学习过程的目标变量。无监督学习可以分为三个主要类别：（1）聚类，（2）关联规则发现，和（3）降维。这将在第3节中更详细地讨论。

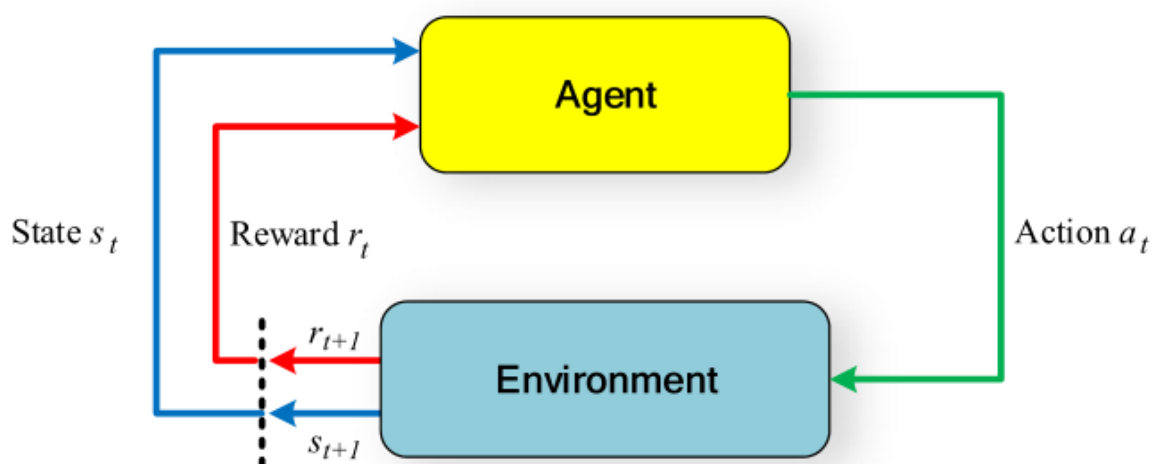


图4.强化学习过程示意图

第三类是强化学习，广泛用于解决马尔可夫决策过程。在 RL 中，一个**智能体(agent)**通过通过环境的互动,学习在它的环境中按照自己的最佳策略行动.RL专注于通过在环境中的行动最大化对智能体的奖励[177].RL的本质涉及一个自治智能体，如图4所示，例如人、动物、机器人或软件智能体，它可以以最大化数字奖励为目标,在不确定的环境中导航.然而，这种奖励并不是在一次行动之后立即产生的，而是在一系列已经导致智能体所处的环境改变的行动之后产生的。体育是RL的一个很好的例子，我们的自主智能体必须处理体育赛事（如网球比赛）中发生的策略和持续行动。在网球比赛中，智能体必须考虑发球、回击和截击等动作。这些即时动作改变了当前集合所描述的游戏状态，当前领先的玩家，以及网球规则手册中的类似状态变量。执行每个动作都是为了获得未来的奖励，例如赢得一分，从而赢得一场游戏、一盘或一场比赛。智能体需要遵循一项原则或一组标准、规则和策略，以最大限度地提高游戏结束时的最终得分。需要解决的一个重要问题是当智能体的行为改变环境状态时的智能体如何建模游戏。从一开始，模型的初始输入是产生最大预期回报的状态和相应的行动【128】。在多次尝试中，RL根据环境的反应和收到的奖励来完善这些反应。

2 进化计算

EC方法受到自然进化原则的启发。EC方法根据个体对问题进行编码，以提高问题解决方案的质量。遗传算子，包括交叉、变异和选择，被用来产生新的个体。基于差异适应度生存机制下，只有最好的个体会进行进一步变异。EC算法使用迭代启发式程序探索搜索空间，以获得逐渐更好的解决方案【143】。在详细讨论之前，我们需要澄清一点：这些元启发式算法有两种类型：基于群体的和基于个体。前一种方法使用一组初始随机（或以其他方式创建）解启动进化过程。例如，**遗传算法（Genetic Algorithm，简称GA）** [63]、**蚁群优化（Ant Colony Optimization，简称ACO）** [119]和**粒子群优化（Particle Swarm Optimization，简称PSO）** [74]。我们将在这里更详细讨论这些内容。然后是基于单一解决方案的方法，称为轨迹优化，它从一个初始随机个体开始。禁忌搜索[51]是基于单一解的算法的一个例子，模拟退火也是如此。图5显示了基于群体的方法的分类，分为四类：生物启发、物理启发、地理启发和文化启发。

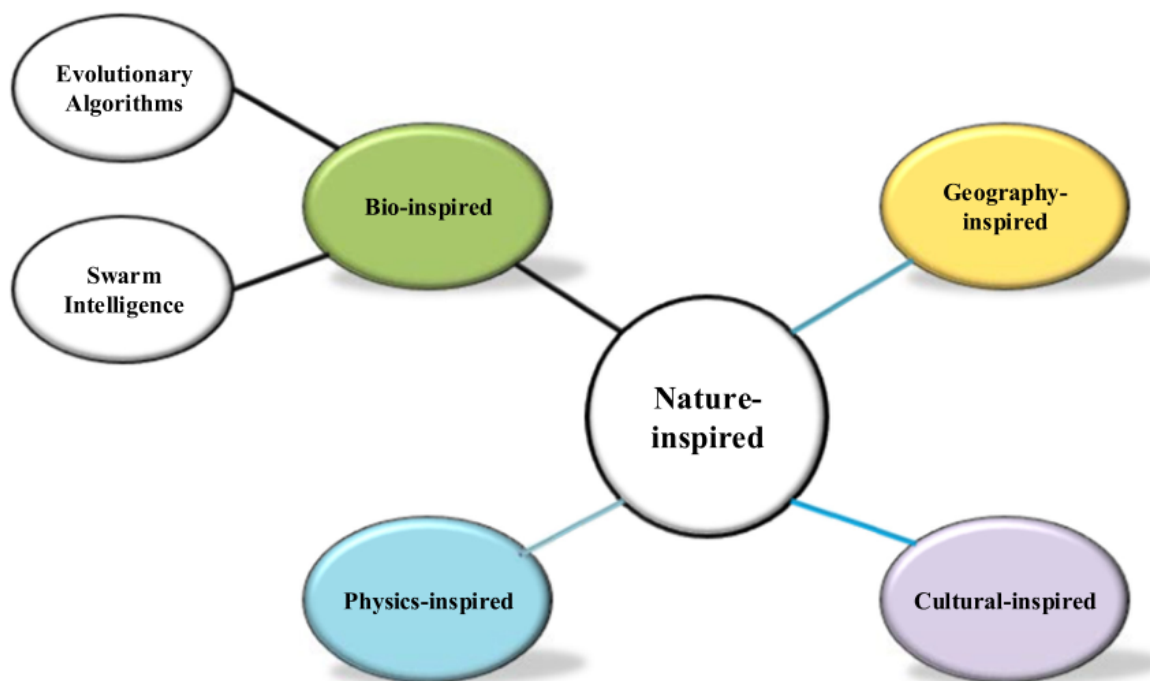


图5：自然启发算法的分类，进化算法是其中的一个分支。

(1) **生物启发 (Bio-inspired)**：这一类别包括基于**群体智能 (swarm intelligence, SI)**的方法和进化启发算法，它们起源于生物体的自然行为。SI模拟蜂群（例如鸟类、鱼类和昆虫）在群体生活中的行为。群体实体可以协作执行生存所需的许多复杂任务。自组织 (Self-organization) 和分散控制 (decentralized control) 是基于群体的系统的两个主要特征，由于群体智能体之间的局部交互，导致突现行为【138】。蚁群算法和粒子群算法是SI的两大主流算法。进化算法的起源在于达尔文的自然进化原理，这使得生物能够很好地适应环境。自组织和强适应性是这些方法的两个主要特点。在这些算法中，整个种群可以被选择、交叉和变异等算子从一代替换到下一代。**遗传算法 (Genetic algorithms, GAs)**、**进化策略 (evolution strategies, ES)**、**进化规划 (evolutionary programming, EP)**和**遗传规划 (genetic programming, GP)**是四种主要的进化启发机制。

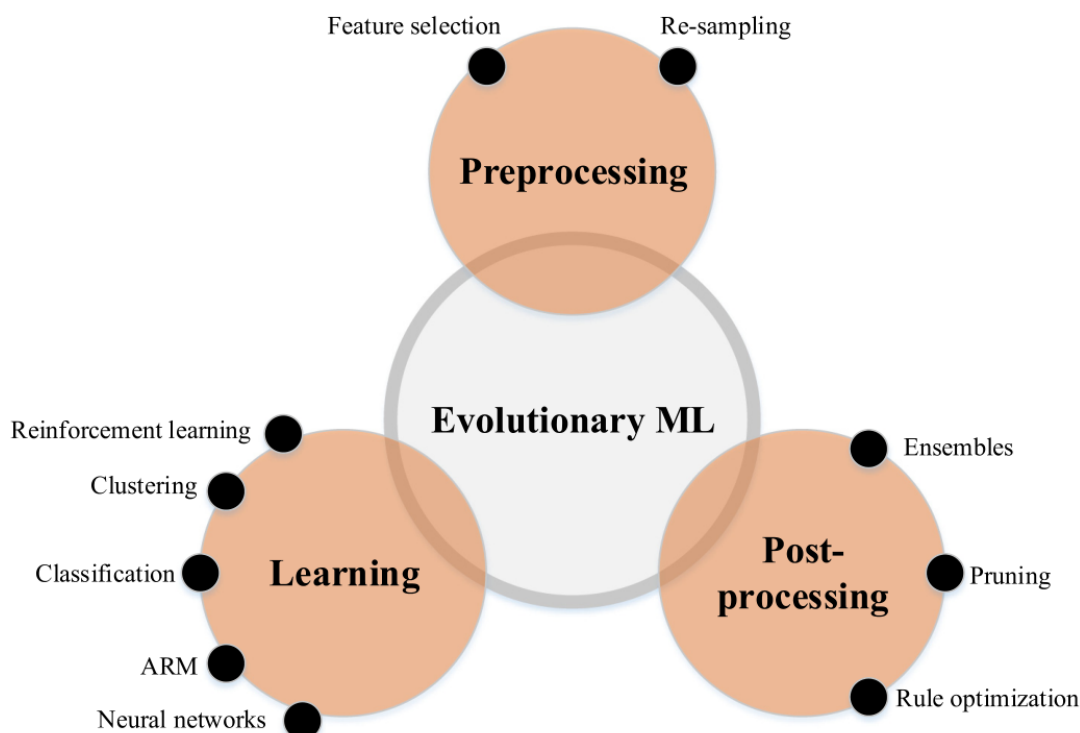
(2) **物理启发(Physics-inspired)**：物理启发算法的起源在于物理/化学规则。例如，引力搜索算法就是这类算法。

(3) **地理启发(Geography-inspired)**：这些算法在地理搜索空间中生成随机解；禁忌搜索属于这一类。

(4) **文化启发(Cultural-inspired)**：这些算法的灵感来自于人类在与他人进行文化互动时的行为。观察他人的自然和固有行为有助于个人学习新知识并改善自己的行为。文化基因算法可以被认为是通过局部启发式模拟变异过程的方法之一。

3.进化机器学习

进化计算在ML中有着广泛的应用。下面总结了EC在不同ML领域的最重要的研究贡献。三个主要方面是：(i) 如何将ML问题转为以个体为表示形式的优化问题(ii)使用哪种搜索机制来解决特定的ML问题，和(iii)如何计算生成新一代的解的质量。我们将EML工作分为九个子领域，重点放在EC做出贡献的特定ML任务上。图6给出了本节所述考虑事项所涉及主题的图解概述。



1 进化特征选择/构造

现实生活应用中的一些数据集，例如基因选择，包含数千个甚至数十万个维度。这不仅是对一般机器学习的挑战，也是对统计学和生物学的挑战。这个问题可以通过特征选择和特征构造方法来提高特征空间质量解决。前者仅从原始特征集中选择信息特征，而后者创建新的高级特征[172]。如果原始特征信息量不足，特征构造可以比特特征选择获得更好的性能[183]。EC方法在特征选择中的应用被称为包装方法。通过利用 ML 作为适应度函数 [23]，实现了贪婪搜索策略以找到一组合适的特征。图 7 展示了进化特征选择的一般框架。

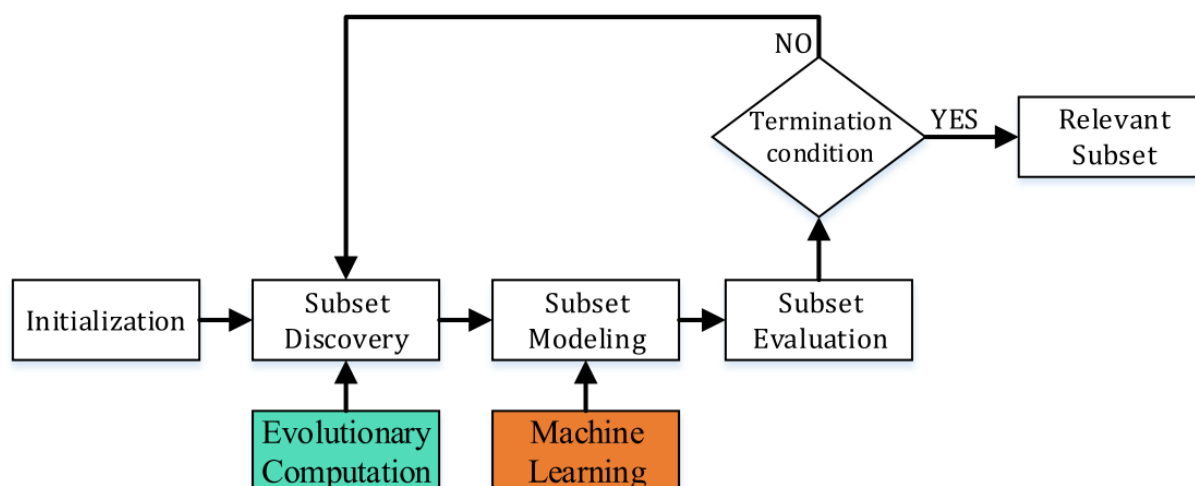
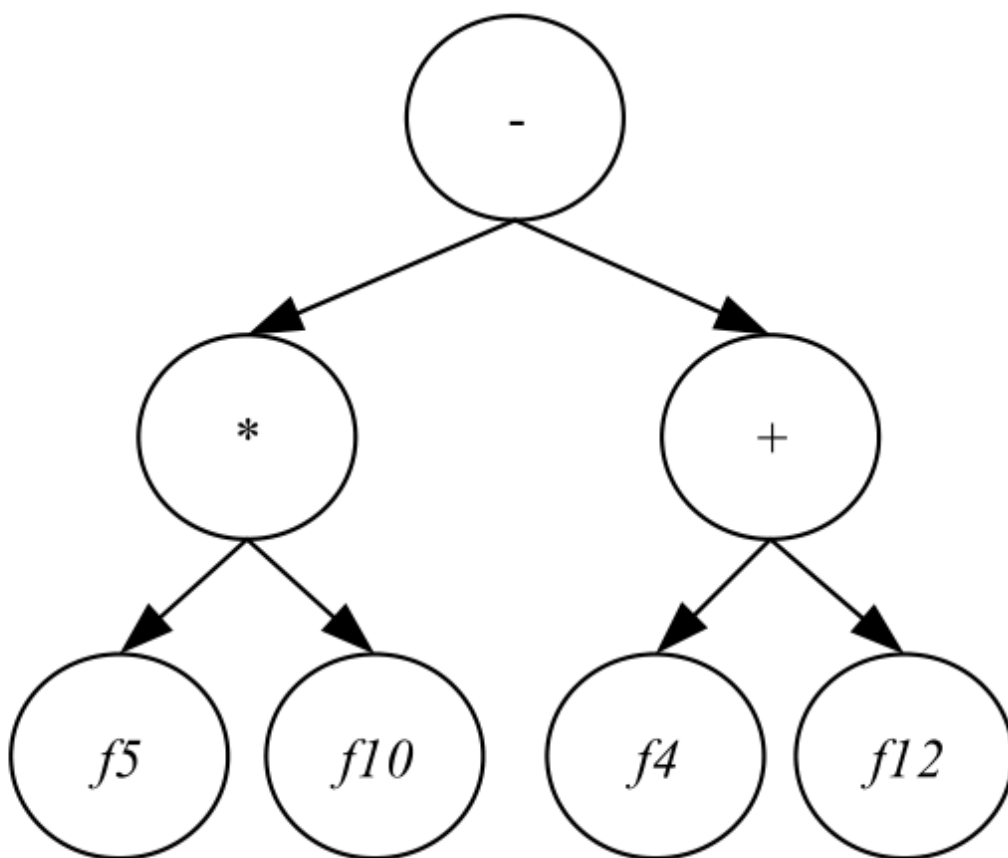


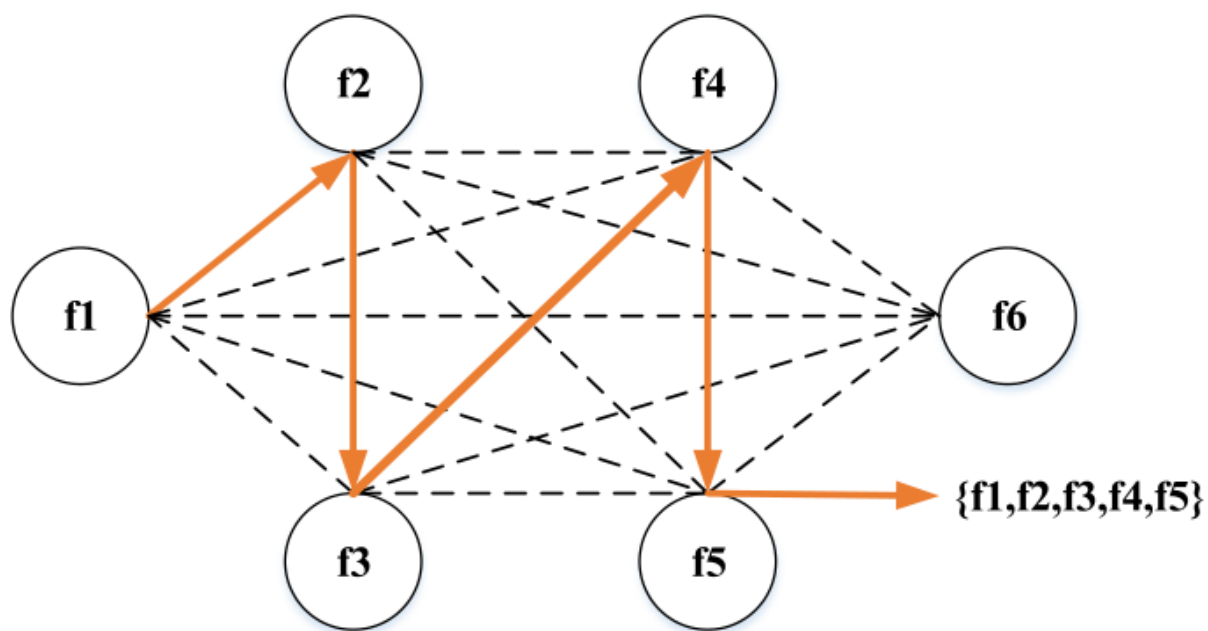
图 7. 一般进化特征选择过程

第 1 步 - **编码(Encoding)**: 二进制编码通常用于特征选择问题。每个解决方案都是一个包含 N 位的位串表示，代表数据集中的特征数量。“1”表示选择了相应的特征，而“0”表示没有选择相应的特征。相比之下，大多数 GP 工作使用如下一种表现形式，即被使用的特征会出现（例如，在树表示中，作为叶节点，在线性表示中，作为寄存器），并且被认为是最终的特征集。GP 能够处理大规模特征选择，因为单个表示不需要有关所有特征选择的信息（例如，它们的索引）。此外，GP 中不需要预定义的解的结构来产生最佳解决方案 [160]。在 ACO 中，特征选择问题由一个图表

示，其中每个特征都被视为图模型的一个节点。如果蚂蚁访问该节点，则该节点被选为所选特征之一。图 8 显示了基于树（图 8(a)）和基于图的编码（图 8(b)）的示例。



(a)基于树的特征选择f4, f5, f10, f12



(b)基于图的特征选择f1, f2, f4, f5

图 8. 特征选择/构建的个体表示示例

第 2 步 - **初始化 (Initialization)**：确定起始特征子集很重要，因为初始化直接影响搜索策略的性能。前向选择和后向选择是两种典型的初始化策略。进化过程从前者的空集开始；然而，在大搜索空间中缺少一些特征是前向选择的主要缺点。后一种策略从一整个解空间开始，并不断迭代地删除一些。冗长的计算时间是后向选择的主要缺点 [182]。伯努利过程是另一种众所周知的生成初始种群的技术。该技术使用从 $[0, D]$ 生成随机数的函数来选择相应的特征，其中 D 是个体的大小。

第 3 步 - **搜索策略 (Search Strategy)**：由于大多数问题的搜索空间很大，因此使用许多启发式方法是不切实际的。然而，诸如 GA 之类的 EC 算法可用于通过进化连续种群来执行搜索过程 [58]。GP 被认为是过滤器方法 (Filter Approach) 中一种有用的搜索机制，它主要用作搜索算法，而在包装器方法中，它可以用作搜索策略和分类技术

第 4 步 - **子集建模 (Subset Modeling)**：使用 ML 算法使用 EC 算法选择的特征子集来构建分类/预测模型。

第 5 步 - **模型评估 (Model Evaluation)**：模型的评估方法分为三组：包装器 (Wrapper)、过滤器 (filter) 和嵌入 (embedded) 方法。包装器方法使用 ML 算法的性能作为其评估标准，而过滤器方法使用数据的内在特征。包装器方法的目标特征通常比过滤器方法选择/构建的特征执行得更好，但包装器方法的计算时间也更长。嵌入式方法同时选择/构建特征并学习分类器。只有 GP 和学习分类器系统 (LCS) 可以执行嵌入式特征选择/构建 [183]

2 进化重采样 (Evolutionary Resampling)

如果一个类的实例数与另一个类的实例数相比要高得多，则数据集被称为“不平衡”或“偏斜”。偏态分布会影响 ML 模型的有效性，这些模型偏向于多数类。重采样是平衡数据分布的最常用方法，其在预处理阶段执行。重采样方法有两种：下采样 (undersampling) 和过采样 (oversampling)。前者删除属于多数类的实例，而后者为少数类生成额外的样本。一方面，下采样可能会删除有关多数类的有用信息。另一方面，过采样增加了训练集的大小，这使得训练更加复杂和繁重。此外，少数实例的随机重复使得过采样策略容易过拟合 [168]。图 9 描述了进化重采样过程。

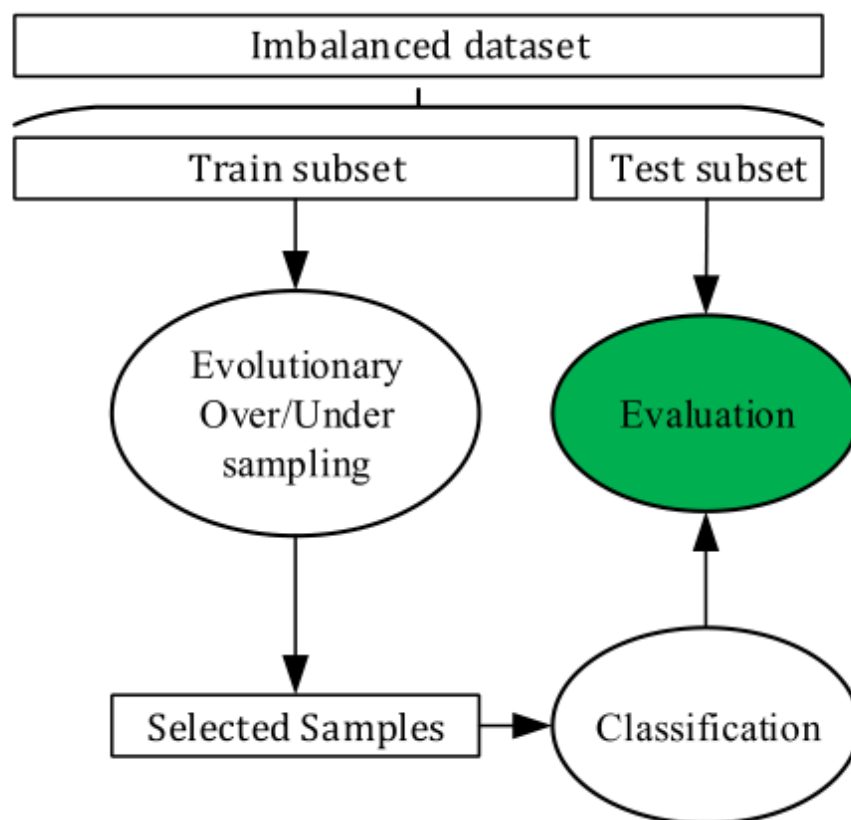


图 9. 进化重采样过程 [49]

传统的进化重采样方法 [50] 使用二进制表示，其中“1”的值表示选择了一条记录，“0”的值代表缺失训练集中的一个实例。然而，这些方法在面对大型数据集时表现不佳，因为个体的长度和搜索空间随着数据集的大小成比例地增加 [105]。相比之下，现代方法试图通过引入仅包含所选多数类样本的索引的稀疏表示来规避大型搜索空间 [47]。图10显示了二进制表示和稀疏表示之间的差异。

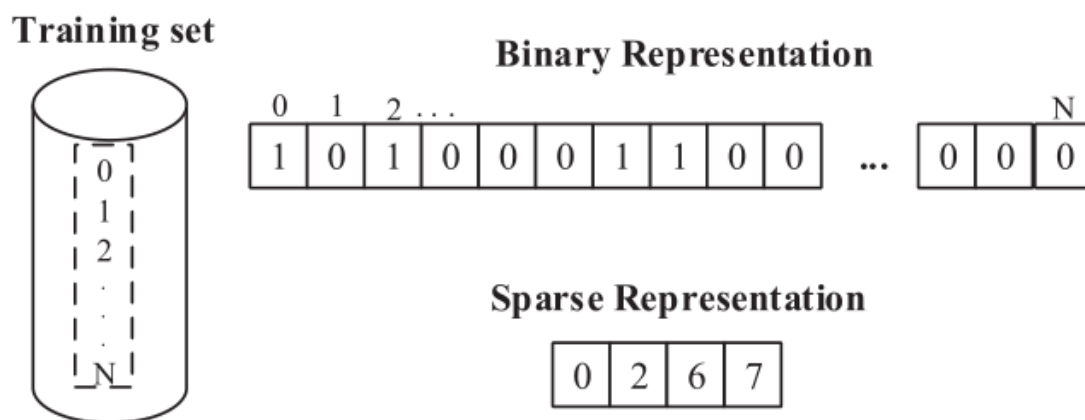


图 10. 重采样方法的二进制和稀疏表示之间的区别 [173]

关于适应度函数，准确率等性能指标不适用于评估所获得模型的质量，因为这两个类的性能权重不同。F1 分数更适用于类别不平衡的问题，因为它同时考虑了精度和召回率，从而生成了一个可用于衡量性能的单一指标 [17]。

3 进化分类器 (Evolutionary Classifiers)

数据平衡、特征选择/构造和数据清洗等数据预处理方法可以提供适当的数据作为数据分类器的输入。然而，这些方法是独立于分类器的，并且有其自身的挑战，例如由重采样方法引起的过度拟合和泛化性差 [19]。因此，接下来考虑通过 EC 方法修改分类器。本节讨论 EC 算法在众所周知的分类器中的应用，例如决策树、支持向量机 (Support Vector Machine, SVM)、k-最近邻 (k-Nearest Neighbor, k-NN) 算法。

3.1 进化决策树

决策树 (DT) 是最广泛使用的 ML 表示之一，因为它们的解释简单且无需领域知识即可快速构建。经典的启发式方法使用贪心方法来选择用于子树构造的节点。因此，这些方法应用局部最优的“测试和失败”来收敛到全局最优解。EC 方法可以两种方式用于 DT 归纳 (图 11)：DT 的进化归纳和 DT 组件的进化设计。在前者中每个个体是一个 DT，而在后者中个体是 DT 分类器的组成部分。

使用每个节点的单个属性或进化分类树中属性的线性或非线性组合来拆分训练数据。与基于多属性的 DT 相比，由于基于单个属性的 DT 易于解释，他们也更常见。然而，基于多属性的 DT 更准确、更小，尽管它们需要更多的计算时间和松散的可理解性。回归树可以被认为是一种特定类型的 DT，其中树的每个叶节点处的目标值是连续值，而不是离散值或标称值 [137]。

关于问题编码，基于树的编码和线性个体（即固定长度的字符串表示）是进化 DT 归纳中用于对个体进行编码的两种常用方法。为非二进制 DT 实现线性个体是很棘手的，因此在大多数研究中，这种类型的 DT 在应用 EC 算法之前被转换为二叉树。编码为线性个体的 DT 比使用树编码技术编码的 DT 更容易处理。然而固定长度字符串编码的一些缺点是需要对基因型和表型之间的恒定映射进行适应性评估，以及处理非二进制 DT 和定义最大位数。以前的一些研究使用动态长度字符串；这产生了不必要的复杂性，因为类似于交叉的进化操作可能需要修改。对于适应度函数，使用单目标优化和多目标优化来评估 DT 的质量。分类准确度是单目标优化最常用的衡量标准。对于多目标适应度函数可以制定一些其他标准，例如准确性、树大小、节点数、灵敏度和特异性。

摆脱局部最优并执行稳健的全局搜索是进化 DT 算法的主要优势，与贪婪 DT 方法相比，它能够更好地处理属性交互。进化 DT 的另一个好处是它们能够在多目标优化中应用不同的度量。然而，进化的 DTs 也引入了一些负面特征。其中一方面是，用于 DT 的 EC 算法对于大规模数据的计算成本很高，因为它们通常会评估每一代群体中的所有候选解[69]。幸运的是，EC 方法可以很容易地并行化，搜索机制和适应度评估都可以在不同的并行和分布式平台上执行，例如 GPU 和 MapReduce。

3.2 进化支持向量机

支持向量机的思想是基于一个最优分离的超平面。SVM 中的原始模式空间首先通过非线性函数映射到高维特征空间；然后，生成特征空间的最佳分离超平面[65]。首先，支持向量机成功地应用于二进制分类问题。对于多类分类问题，通过分解的方式将问题分解为多个二元子问题。然后通过 SVM 解决每个子问题，并将所有预测变量的输出组合起来 [108]。支持向量机随后被用于回归预测和时间序列预测。

对于边际较小的分类器，可以预期更高的风险。如果数据不能线性分离，就会产生一些松弛变量。因此，应该解决凸二次规划问题以构造最大边距[14]。当在线性空间中无法解决获得最优分离平面的问题时，SVM 中的输入空间被映射到高维点积空间。在这种情况下，使用核函数（“核技巧”）在高维空间中找到超平面，而不会显着增加计算成本。径向基函数 (RBF)（等式 (1)）是 SVM 中常用的核函数技术。

$$K(x_i, x_j) = \exp\left(-\frac{\|x_i - x_j\|^2}{2\sigma^2}\right) \quad (1)$$

核参数 σ 影响数据映射过程并改变高维特征空间的数据分布 [65]。总体而言，SVM 的高性能源于三个因素：核函数的选择、核参数的选择和参数 C 。在 SVM 中制定了一个优化问题，以构造一个最大边距分类器，如等式 (2)：

$$\begin{cases} \text{minimize} & \frac{1}{2}\|w\|^2 + C \sum_{i=1}^k \xi_i \\ \text{subject to} & y_i(w_i x_i + b) > 1 - \xi_i \end{cases} \quad (2)$$

C 是一个惩罚参数，在训练误差和泛化之间进行权衡。如果选择的 C 值太大或太小，可能会降低 SVM 的泛化能力。如果在多类问题中使用分解方法，则该参数的选择变得更加困难，因为参数的数量随着每个二元分类器的增加而增加。项 k 、 ξ 、 w 、 b 分别是数据点的数量、松弛因子、法线向量和标量。

可以使用几种方法来调整超参数：网格搜索算法、试错法、交叉验证、泛化误差估计和梯度下降以及进化算法。使用网格搜索算法既复杂又耗时。试错程序非常耗时，而且结果也不可靠。交叉验证方法需要漫长而复杂的计算[56]。SVM 中的梯度下降算法对初始参数很敏感。使用 EC 算法来解决上述挑战的参数优化受到了更多关注 [14]。

在 SVM 问题中对超参数进行编码时采用实际的编码表示，从而避免了后交叉过载问题。在这种情况下，单个 X 表示为 $X = \{C, \sigma\}$ ， C 和 σ 表示上述惩罚和核函数参数。图 12 显示了具有三个类别的分类问题的染色体表示示例 [108]。

SVM 参数选择任务通常通过保留参数的最佳组合来执行。使用详尽的过程来探索参数空间可能会产生良好的结果，尽管出于明显的实际原因应该避免这种策略。因此，优化技术是防止对参数进行详尽或随机探索的好选择，因为它们使用所选目标函数的良好值来探索搜索空间。然而，这些技术的一个缺点是，它们必须从搜索空间均匀采样的随机设置开始。这会使收敛速度变慢，并且算法可能会陷入局部最小值。元学习是解决 SVM 参数选择的有用策略，并将此过程视为监督学习任务 [52]。此策略根据在先前类似问题中成功确定的参数设置推荐 SVM 参数值。图 13 展示了基于元学习的进化 SVM 的一般框架。

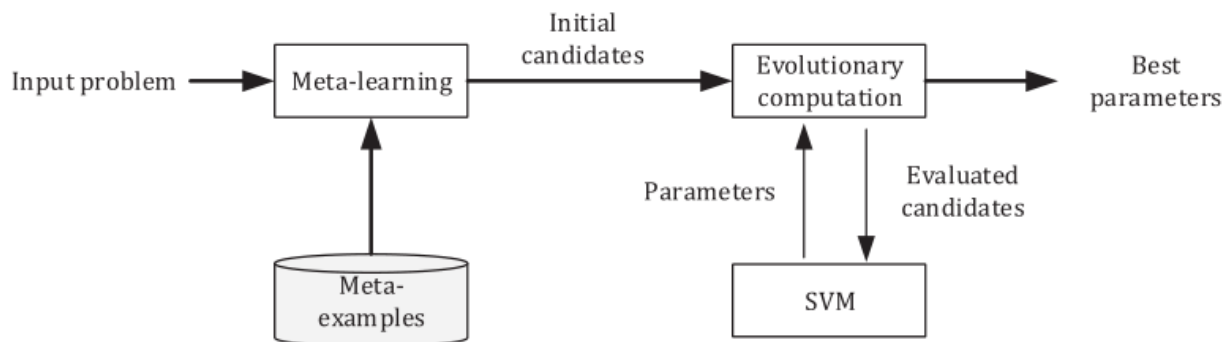


图 13. 进化 SVM 元学习的一般框架

SVM算法往往会生成很多支持向量，这增加了计算决策函数的计算时间。后剪枝是一种可用于消除由标准算法生成的不适当支持向量的策略。每个个体的长度等于二进制表示（广泛使用）中支持向量的数量。如果位等于“1”，则第 i 个支持向量包含在决策函数中，如果位等于“0”，则排除第 i 个支持向量。

3.3 进化k最近邻

最近邻技术[33]及其衍生方法是惰性学习方法的一个子集。k-NN 算法是最近邻算法 [174] 的扩展版本。k-NN 算法是一种非参数分类器，这意味着它不依赖于任何关于数据分布的先前假设。k-NN 算法通过其 k 个邻居的多数票对对象进行分类，其中 k 是用户定义的参数。输出类是通过一个投票度量获得的，该度量应用于测试模式和训练模式之间的所有距离向量。定义邻居的数量 k 是具有挑战性的，因为 k 的某个值可能会导致一个分类问题的性能良好而另一个分类问题的性能不佳，这取决于特征空间中类的分布。结果表明，当 $k=1$ 且训练样本数为 n 时 $\rightarrow \infty$ ，k-NN 分类不准确的概率最多是贝叶斯分类器风险的两倍【33】。但是，如果可用的训练实例数量有限，这就不适用了。除了 k 和距离函数外，邻居、类和特征的重要性也会影响 k-NN 算法的性能。与神经网络和支持向量机类似，k-NN 算法的精度得益于训练阶段的权重优化。这些权重可以分配给邻居、类或特征，每种类型的权重对算法的性能都有特殊影响。特定于类的权重为 k-NN 算法提供了关于类属性的额外知识；属性特定权重可用于消除噪声和冗余特征的影响【22】。加权方案的目的是使用一个良好的度量，该度量将导致对给定的原始原型集的高分类精度。先前发表了一项关于加权系统中数据不同方面差异演化的研究【10】。

在神经网络中执行数据约简的两种常用技术是原型选择和原型生成。原型选择通过移除冗余和噪声样本来从原始训练集中选择实例的子集[27]。原型生成方法不仅能够选择数据，而且能够生成原始数据并将其替换为新的人工数据[174]。原型选择和原型生成都是组合优化问题，因此 EC 方法可以用来解决这类问题并产生很好的结果。原型选择和原型生成可以分别编码为二进制或连续空间搜索问题。原型生成的 EC 技术基于原型的位置调整，优化了原型的位置。然而，EC 技术的一个缺点是它们通常依赖于从训练集中提取的原型的初始子集。此外，扩展到大型数据集是原型选择中的一个挑战，因为它会导致过多的存储要求、更高的时间复杂性和更低的泛化精度。原型选择算法需要搜索所有可用的实例来分类新的输入向量，因此在分类过程中速度很慢[27]。

4 进化神经网络与深度学习

标准的神经网络由许多相连的称为神经元的处理器组成。输入神经元接收来自环境的值，而其他神经元通过来自先前活跃的神经元的加权连接来接收激活值。学习的主要焦点是找到一个最优或足够接近最优的连接权重集。神经网络的成功在很大程度上取决于结构、训练算法和训练中使用的特征的选择。反向传播学习需要调整参数，如学习速率、动量和预定结构。由于其梯度性质，误差反向传播遇到了诸如收敛速度慢和陷入局部极小等挑战[70]。具有手动定义的参数的梯度下降在更深的网络中表现不佳，导致训练数据的不足或过度拟合[91]。因此，为了应用而调整接近最优的神经网络的参数和结构是具有挑战性的[186]。

进化计算可以通过学习神经网络的构建块(例如, 激活函数)、超参数(例如, 学习率)、体系结构(例如, 层的数量和每层中的神经元), 甚至学习规则来应用于神经网络。在20世纪80年代早期, 研究人员专注于只进化具有受限体系结构的网络的权重, 包括固定数量的层和神经元[155]。但在1994年, 参考文献[54]开发了一种人工开发系统, 用于自动生成复杂的神经网络。

在进化神经网络中, 权值矩阵被编码为个体, 并通过交叉、变异等进化操作进行优化。将神经网络产生的误差作为适应度度量。进化神经网络有两种编码方案: 直接编码和间接编码。前者表示节点之间已有的连接。该方法需要背景知识来定义拓扑(例如, 层的数目和隐藏单元的数目)。一旦建立了包括 N 个节点的拓扑, 就为每个神经元分配一个数字, 并生成二进制2D结构 $N \times N$ 。值“1”表示两个神经元之间存在连接。前馈连接可以通过仅启用层 i 和层 $i+1$ 中的单元之间的连接来保证。关于网络拓扑的假设的必要性是直接编码方案的主要缺点, 带来 $O(N^2)$ 复杂性[16]。然而, 间接编码只考虑神经网络拓扑的某些重要特征, 而不是完全连接模式, 导致与直接编码相比, 编码更紧凑。间接编码可分为三种主要方法: (1)指定参数并描述神经网络的拓扑和结构的连通性参数; (2)用于构建拓扑的发展规则(例如, 递归方程或产生式规则); (3)受某些生物发育过程启发的连通性的分形表示[185]。**神经进化的扩充拓扑**

(**NeuroEvolution of Augmenting Topologies, NEAT**)[157]是一个众所周知的算法, 它使用遗传算法来进化神经网络的结构和连接参数。Neat使用了两个向量的直接编码, 一个用于节点, 一个用于连接。每个基因定义了两个节点之间的连接权重; 因此, NEAT只适用于小型网络。HyperNEAT[156]使用间接编码为更复杂的网络优化了NEAT。

在深度学习中, 进化计算的使用由来已久, 在深度学习开始受到重视后迅速开始。2011年, Cheung和Sable提出了一种用于深度神经网络的早期神经进化方法[32], 其中EC被用来寻找**卷积神经网络(Convolutional Neural Network, CNN)**的结构参数的最佳值。CoDeepNet[127]是对Nat[157]的增强, 用于优化**长短期记忆网络(Long-Short-Term Memory, LSTM)**的拓扑、组件和超参数。硬件方面的重大进步使得更深层次的体系结构的使用越来越流行, 这导致了具有多层和超参数的更复杂的神经网络模型。

尽管进化学习被成功地应用于神经网络的自动设计, 但进化神经网络的一个主要缺点是在优化过程中消耗了大量的资源。通常, 数以千计的不同个体是进化了的, 每个个体代表着具有复杂体系结构和评估的深度学习模型的完整训练阶段。早期的研究表明, 进化训练通常是计算密集型的, 而且比反向传播慢[80]。由于缺乏GPU等计算资源, 这些算法直到2012年才开始实用[53]。制造用于深度学习的特定芯片组和产品线是当前用于应对这一挑战的技术趋势。这类技术的例子包括Google Cloud张量处理单元[41]、Amazon EC2 P3实例[11], 以及大型人工智能超级计算机, 如NVIDIA的DGX SATURNV, 它由125台服务器和1,000个针对深度学习而优化的强大GPU组成[134]。

今天, 神经体系结构搜索(NAS)领域正在蓬勃发展, 自2016年左右以来, 该领域的研究呈爆炸式增长[43]。在NAS中, 经常使用混合方法, 其中只使用进化来优化结构超参数, 而将学习留给梯度方法。多目标进化算法最近显示出巨大的成功, 如文献[112]所示。在许多这样的应用程序中, 搜索复杂性最低的体系结构, 从而尽可能准确地执行。这允许以更少的训练计算时间获得更好的泛化性能。可以通过使用代理适应度函数[111]来获得进一步的加速。

5 进化强化学习

解决强化学习问题的三种方法是价值函数 (value functions)、策略搜索 (policy search) 和行动者-批评者 (actor-critic)。价值函数法旨在估计处于每种状态的期望值。相反, 策略搜索方法不需要价值估计模型, 而是直接搜索最优策略。行动者-批评者方法采用了上述两种方法[12]。尽管这些方法在RL中取得了成功, 但仍然存在三个主要问题: 回报稀疏的时态信用分配, 缺乏有效的探索, 以及对超参数选择极其敏感的脆性收敛性质。EC算法非常适合处理每一个问题[75]。使用适应度函数在整个事件中整合回报, 使EC算法对具有长期视野的稀疏奖励不变。以人口为基础的方法可以带来多样化的探索。最后, 人口的固有冗余性也增强了复原力和可持续的趋同特性, 特别是在与精英主义相结合时[35]。

在进化RL算法中, 个体的适应值是个体在其环境中操作后获得的累积奖励。图14显示了基于行动者-批评者的进化RL方法, 其中强化学习者使用群体生成的数据经验。策略梯度法通常用于以损失函数的最小值的形式最大化收益。该方法使用行动者-批评者体系结构来维护确定性策略和动作-值功能批评者。在传统的RL中, 一旦个体执行了一个动作, 就获得了单一的奖励; 然而, 在进化的RL中, 个体在种群群的生命周期结束时或在一系列动作(插曲)之后被考虑

适应值(回报)。EC方法的这一特性使其能够直接适用于间歇性RL任务，例如玩游戏，其中EC算法搜索最优函数值或最优策略[126]。

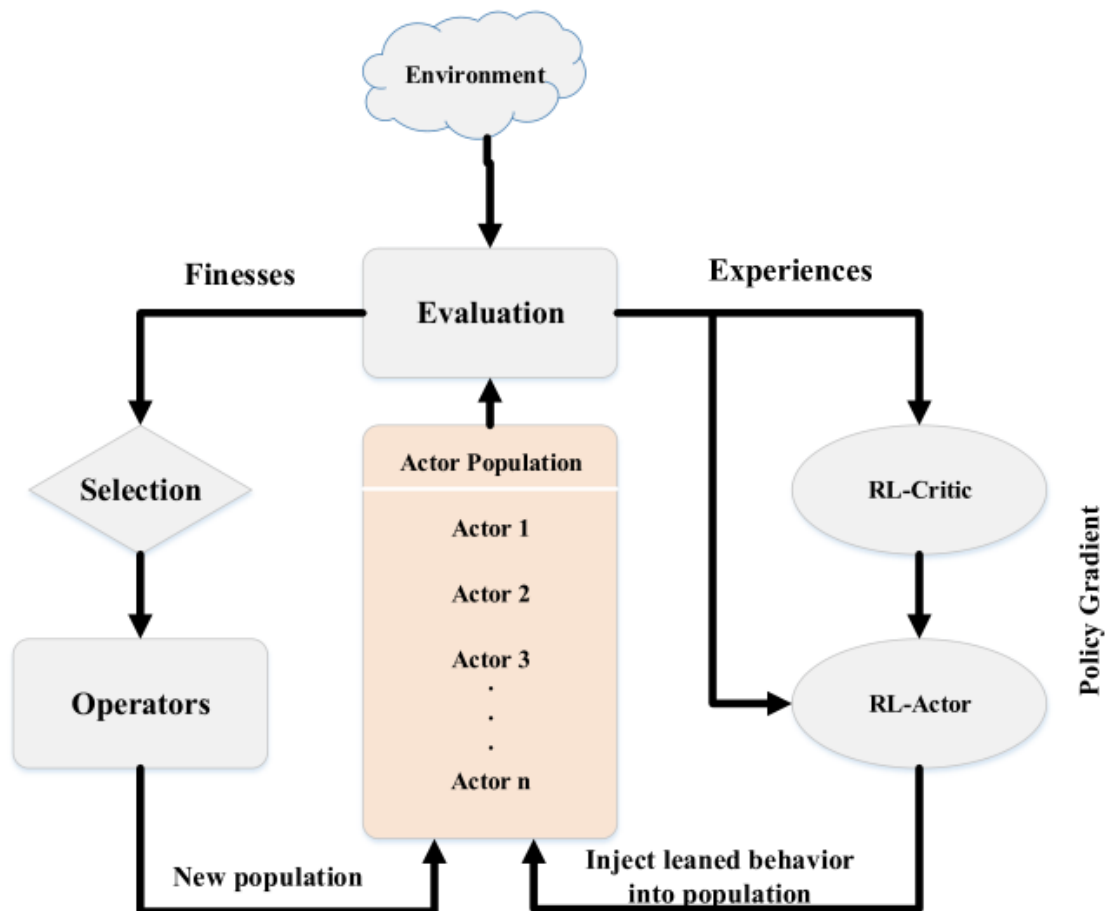


图14.强调将EC基于种群的学习并入基于梯度的优化的进化强化学习示意图(灵感来自参考文献[75])。

RL的主要问题之一是高维输入空间，例如视觉输入。这种维度问题可以通过两种方式来缓解：(1)可以应用预处理器(压缩器)将高维输入空间转换为低维特征空间，以及(2)可以压缩神经网络控制器的表示[82]。主要的方法是间接编码，通过复数映射将小型神经网络转换为任意大小的网络。另一种选择是将行动学习与无监督学习压缩器结合，以提供低维特征向量作为智能体的输入[34]。文献[34, 82]提出了无监督学习和进化RL的结合。相反，当进化方法使用神经网络权重的压缩表示来训练大型网络时，不需要执行压缩阶段。这项技术的使用引入了第一个深度神经网络，直接从高维视觉输入学习RL任务[12]。

另一种将EC方法应用于RL任务的全新方法也于2017年引入，即**纠缠程序图(Tangled Program Graph, TPG)**方法[72]，其中使用一组线性遗传规划作为一个团队来解决RL任务。TPGS可以直接在高维视频输入上工作，并已在各种游戏环境中进行了测试。通过深度网络强化学习获得的效率已被用于允许该方法接近多任务学习[73]。诸如此类的。[158]针对RL难以解决的大量Atari博弈(如Q-学习或策略梯度)，研究了GA在深度强化学习上的性能。作者发现，DNN和GA的结合可以解决稀疏奖励函数和高维问题。

6 进化聚类

聚类是一种无监督学习方法，它根据未标记数据对象之间的相似性将它们划分为几个组[64]。集群的主要特征是不需要关于数据分布的先验知识[67]。划分聚类和层次聚类是两类主要的聚类算法。分区聚类方法基于预定义迭代次数的适应度度量将数据集划分为特定的组[7,133]。简单和计算成本低是分区聚类算法的两个主要优点，例如k-Means[107]。然而，这些算法存在两个主要问题。首先，它们对初始化和陷入局部最优的概率非常敏感。其次，在运

行聚类算法之前，必须确定聚类的数量。少量的集群可能会导致关键隐藏信息的丢失。相反，大量的集群可能会导致集群的高度同质性。

在层次聚类方法中，使用树拓扑来表示聚类集之间的关系。分层方法可以使用分裂方法或聚合法。前一种方法将较小的集群合并为较大的集群，而后一种方法将较大的集群拆分成较小的集群[7]。与分区聚类相比，层次聚类具有优势，因为不需要预先指定聚类的数量。然而，层次聚类的缺点是每个元素只能分配给一个簇[4]，如果对重叠的簇进行分离，则性能会受到影响。

从优化的角度来看，聚类是一个NP-Hard问题。进化数据聚类方法或者使用数据聚类的优化技术，或者在现有的聚类算法中添加优化技术。EC方法试图最小化或最大化目标函数。在聚类环境中，簇内距离应该最小化，而簇间距离应该最大化[4]。在聚类中，EC算法有两个主要目标：确定聚类数目和指定聚类中心。在进化聚类中有两种类型的个体表示：基于原型的和基于点的。与应用基于点的表示相比，应用基于原型的表示时，个体的大小通常更小且冗余更少。然而，基于原型的编码往往倾向于圆形集群，其中每个集群由单个原型表示。相反，基于点的表示允许捕获具有任意形状的簇。

EC可以使用多种聚类有效性度量作为适应度函数来评估个体。一些研究侧重于最小化数据集中N个对象与编码到个体中的拟阵之间的距离之和(式(3))

$$F = \sum_{i=0}^N d(x_i + m) \quad (3)$$

其中m表示最接近对象 x_i 的中心点。这种度量适用于基于中心点的表示。

使对象与其各自的聚类平均值的欧几里得平方距离之和最小化是可以使用的另一种适合度测量(式(4))：

$$f(C_1, \dots, C_k) = \sum_{j=1}^k \sum_{x_i \in C_j} \|x_i - z_j\|^2 \quad (4)$$

其中 x_i 是数据集中的对象，而 z_j 是簇 C_j 的平均向量。该标准适用于基于质心的编码

6.1 固定簇

一些进化的聚类算法使用预定义数量的簇(K)。该技术可能特别适用于其中有关于簇数的信息的应用。进化聚类算法专注于解决与基于原型的聚类相关的挑战，这意味着代表集群的质心、中心点或其他向量进行了优化。进化算法包括使用概率规则来探索搜索空间并选择具有更高概率的更适合的分区的算子。EC的并行性质还允许直接处理具有不同距离标准和适应度函数的多个个体。

在进化聚类中有三种编码方案：二进制、整数和实数。在二进制编码方案中，每个解的长度等于数据集中的实例数。每个比特对应一个实例，即第i比特表示第i个实例。如果第i位是“1”，则第i个实例是原型。图15(A)示出了具有四个集群(K)的二进制表示的示例，其中对象1、5、7和10是集群原型。选择这四个对象，并计算其他对象与这些实例的相似度。矩阵编码是另一种类型的个体表示，其中矩阵的大小为 $k \times N$ (图15(B))。这种类型表示对于二进制字符串编码需要 $O(KN)$ 存储空间而不是 $O(N)$ 空间。然而，矩阵编码的计算距离相似度的计算成本低于字符串编码，因为在计算适应值时只考虑所选择的对象。

整数编码有两种类型：基于标签的表示(图15(c))和基于中心点的表示(图15(d))。前者是N个位置的向量，其中N是对象的数量。每个位都有一个介于1和k之间的值。后者使用由k个元素组成的数组提供基于中心点的数据集表示。个体的长度等于k，每个元素代表1到N之间的对象的索引。基于标签的表示的复杂度是 $O(N)$ ，而基于中心点的表示的复杂度是 $O(k)$ 。然而，与基于中心点的编码不同，基于标签的表示不需要额外的处理来恢复在个体中编码的集群。

在实数编码中，分区的每个特征的质心以基于质心的表示形式表示。这种编码涉及一个长度为nk的向量，其中n是属性的数量，k是簇的数量。图15(e)显示了具有两个变量和四个集群的数据集的真实表示示例。

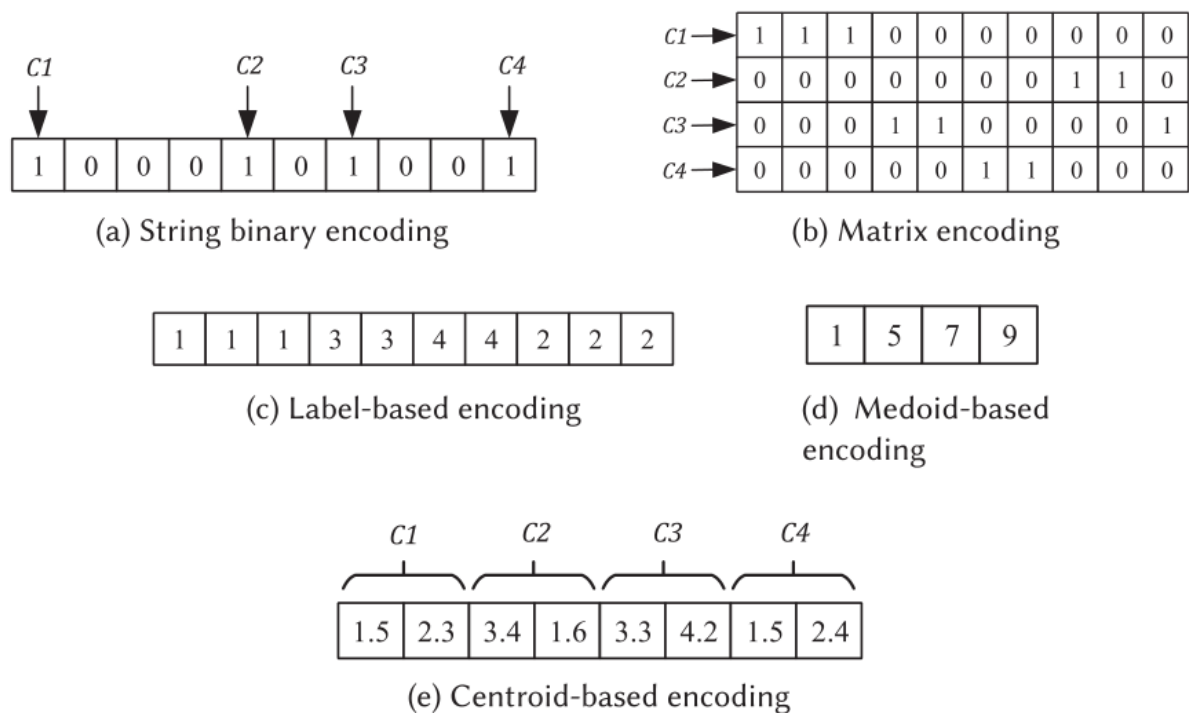


图 15. 进化聚类中不同类型的编码

6.2 可变簇

进化聚类算法的一个主要好处是它们可以自动对数据进行分区，而无需预先指定数量的聚类 and 聚类中心 [89] 自动聚类很有帮助，因为不需要有关于聚类数量的先验 (priori) 信息。进化算法旨在优化集群的数量 (k)。某些编码方案，例如固定聚类算法中采用的二进制编码 (图 15(a)) 和基于标签的编码 (图 15(c)) 可用于对变量聚类问题进行编码此外，还提出了一种不同类型的编码，其中有一组轴对齐的超矩形规则 (图 16)。每个规则由 n 个位置组成，其中 n 是属性的数量，对应变量的边界被编码在每个位置: l_i 和 u_i 是下限和上限。根据图 16，示例规则可以为: if ($1 \leq A_1 \leq 6$) AND ($2 \leq A_2 \leq 5$) THEN (实例属于集群 1)。

Rule 1 (Cluster 1)		Rule 2 (Cluster 2)	
1	6	2	5
l_1	u_1	l_2	u_2

图 16. 基于规则的编码

Bezdek 等人提出了基于 GA 的进化聚类。[18] 并且是 EC 算法在聚类中最早成功的应用之一。作者利用 GA 的探索性和剥削性特征来发现最佳质心。Sarkar 和 Yegnarayana [149] 提出了一种聚类算法，它使用进化规划来确定聚类和聚类中心的数量。用于分区聚类的 EC 方法在参考文献 [133] 中进行了回顾。

6.3 进化模糊簇

模糊聚类方法中的每个数据对象都属于具有模糊隶属度等级的多个聚类。为了将模糊聚类转换为清晰聚类，这种方法将每个数据点分配给具有最高成员值的聚类 [133]。大多数模糊聚类方法都有几个固有的缺点，例如 (1) 用户需要先验知识才能使用聚类方法; (2) 使用随机初始选择可以生成不同的聚类解决方案; (3) 一种基于函数的目标算法使用梯度方法来搜索最优值，这可能导致陷入局部最小值 [42]。EC 方法的另一个应用是优化模糊聚类算法的目标函数

数。

7 进化关联规则挖掘

关联规则挖掘 (Association rule mining, ARM) 旨在推导交易数据中项目之间的关系 [1]。ARM 已成功应用于不同领域，例如市场分析、推荐系统或医学。例如，ARM 提取的模式可以洞察客户经常一起购买哪些商品，从而帮助零售商制定营销策略。经典的 ARM 方法可以分为两大类：Level-wise 和 pattern-growth。使用**广度优先搜索 (Breadth-First Search, BFS)** 和**深度优先搜索 (Depth-First Search, DFS)** 来计算项目集的支持值的 Level-wise 算法的两个示例分别是 Eclat [187] 和 d Apriori [2]。Apriori 可以生成高精度的关联规则；但是，对于大型数据集 [9]，它需要大量的计算时间。FP-growth 算法 [57] 是一种基于模式增长的算法，它使用“分而治之”的策略来提取关联规则，而无需候选生成步骤 [170]。

事实证明，从事务数据集中提取频繁模式是一个 NP-Hard 问题。无论是通过用户还是自动过程，传统的 ARM 方法在应用算法之前都依赖于数据的预处理进行离散化。ARM 可能是一个有损信息发现过程，因为由于预定义参数和分区 [123]，区间之间的边界很明显。Fuzzy ARM 通过使用模糊集在集合的成员和非成员之间创建平滑过渡来处理这个问题。然而，在模糊 ARM 中找到一组合适的隶属函数 (MF) 是主要挑战之一。总体而言，定量值区间之间的清晰边界和模糊集区间的区分隶属度是启发式 ARM 算法的主要缺点之一。

针对传统规则发现方法的不足，引入了基于进化方法的频繁模式挖掘。规则挖掘是在不离散化连续属性的情况下执行的，因此在进化阶段获得间隔，以减轻进化 ARM 中尖锐边界的影响 [121]。EC 方法在 ARM 中的一个应用是规则优化，EC 应用于传统 ARM 算法的后处理阶段，因此 ARM 算法可以提取有意义的规则。ARM 的表示取决于执行的 ARM 类型，匹兹堡和密歇根方法通常用于编码二进制数据集 [63]。匹兹堡方法在个体中编码不同的模式，而密歇根方法只代表个体中的一种模式。匹兹堡技术对 ARM 类更有用，那里的目标是识别一组好的模式。相比之下，密歇根策略很好地挖掘了一组好的模式，因此更擅长发现频繁模式或罕见事件的高质量预测。与匹兹堡方法相比，密歇根方法由于编码固定长度的关联规则，因此简单、直观，并且句法简短。另一种类型的编码是二进制向量表示，其中每一位表示属性值的存在或不存在。尽管这样的二进制字符串需要转换为“If-Then”规则，但它降低了处理速度 [139]。这种类型的编码适用于 MFS 表示。每个解决方案基于项目的范围对隶属函数的中心和跨度进行编码。

在对 ARM 进行遗传网络编程时，判断节点的函数对应于判断节点的项及其值。判断节点的连接表示关联规则。如果满足某个判断节点，则将属性移动到该节点；否则，将该属性移动到另一个处理节点。**语法指导的遗传编程 (Grammar-Guided Genetic Programming, G3P)** 是对遗传网络编程的改进，在遗传网络编程中，使用语法来对 GP 树施加约束 [6]。在基于 G3P 的 ARM 的情况下，通过应用一组产生式规则来创建语法约束。

评估进化臂中个体的性能指标可能会相互冲突，没有一个个体能同时优化所有功能。但解决方案的质量可以通过他们的支持和他们在冲突时的信心来评估。一组非支配解决方案用于在相互冲突的目标之间提供折衷 [130]。与根据用户优先级从集合中选择单个个体的分类和聚类任务不同，所有非支配解都被考虑在多目标臂中作为最终集合。

8 进化集成学习

集成学习方法是一种强大的技术，通过组合多个模型的输出来生成最终预测 [146]。集成学习已被成功地用于解决许多不同应用中的不平衡数据集。在这种情况下，重采样集成技术被广泛使用。当应用一组准确和多样化的集成成员时，集成性能可以得到改善 [55]。

集成方法包括几个分类任务；每个分类任务由数据集、诱导器和分类器组成 [145]。构建预测模型集成需要三个步骤：成员生成、成员选择和成员组合。第一个目标是建立多样化的基础模型。第二个步骤是成员选择，这是一个可选步骤，它使用启发式方法来修剪模型池。第三步，成员组合，负责通过组合预测来生成乐团的最终输出。这三个步骤可以被描述为一个优化问题，并通过 EC 机制来求解。进化成员一代的目标是通过创建一批候选模型来形成一个整体。预测分数和候选者的复杂性值被认为是最重要的标准。在进化成员选择中，EC 对候选模型进行剪枝，以构建最优的集成模型。通过进化成员组合，可以得到加权平均集成的每个候选模型的最优权重。

进化集成员生成已被用于时间序列预测[25]、不平衡数据分类、图像分类[5]和故障诊断[115]。因为EC算法使用个体的总体，所以它们是将潜在的个体模型构建成集合模型的自然选择。然后，集成学习策略应该总是在准确和多样化的模型之间提供折衷，这可以通过错误-歧义分解来总结。这意味着系统的泛化误差是由所有个别错误和歧义的加权平均值所生成。然后，可以尝试通过减小概括误差并增加每个个体的模糊性来减小总体概括误差，这增加了个体的预测误差[151]。在成员生成步骤中创建不同集合的方法可以分为三组：使用不同的训练数据、使用不同的学习算法、以及使用不同的权值或参数来学习模型。装袋[24]和提升[125]是两种用于准备不同训练集的技术；前者使用随机抽样，而后者操纵从原始训练集中选择训练数据的概率。EC方法可以在两种技术中使用二进制表示来选择训练数据的子集。图17示出了进化集成员生成的过程，其中在验证数据集上评估由训练数据集生成的预测模型的性能。然后，关于学习算法的超参数，通过EC算法来优化预测模型。或者，它从原始数据集中选择要素和实例。

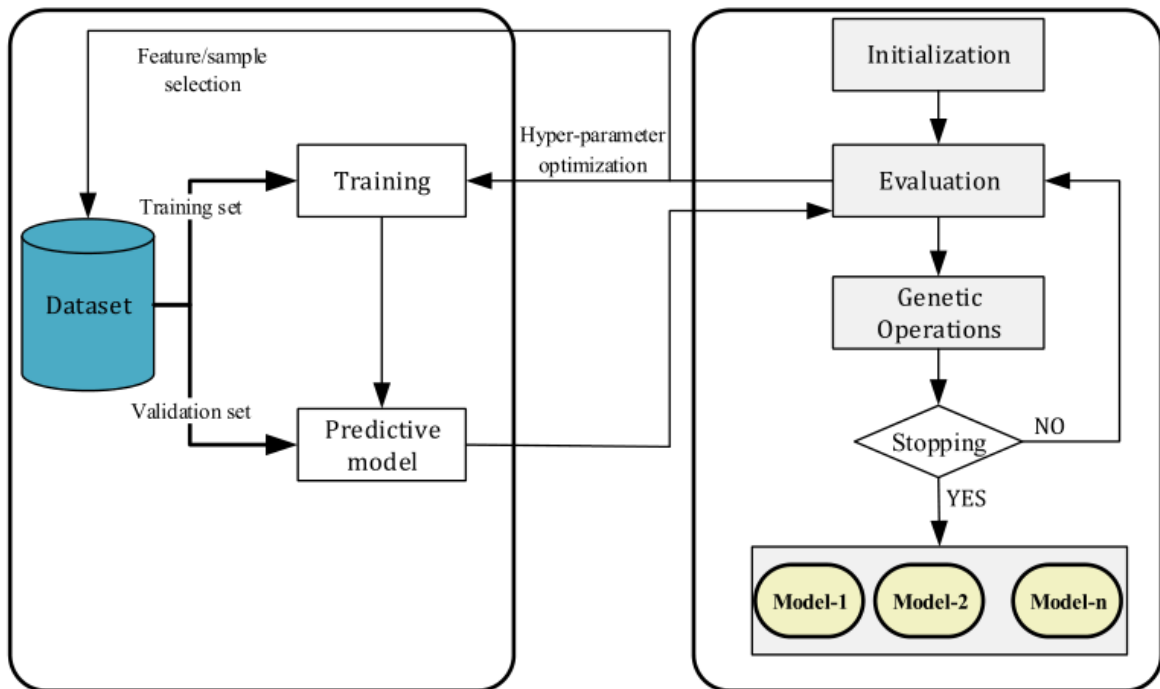


图17.进化集成员生成。

当对集成员选择进行编码时，决策变量通常是二进制向量，其中每个比特表示基础模型的选择或不选择。该技术已应用于情感分析[136]和电力变压器溶解气体含量预测[140]。图18显示了进化合奏成员选择过程的一般视图。基本学习器池预测来自验证数据集的输出，并且池的修剪优化EC算法生成的预测分数。这之后是对选择进行优先排序以选择较佳模型集的另一步骤。

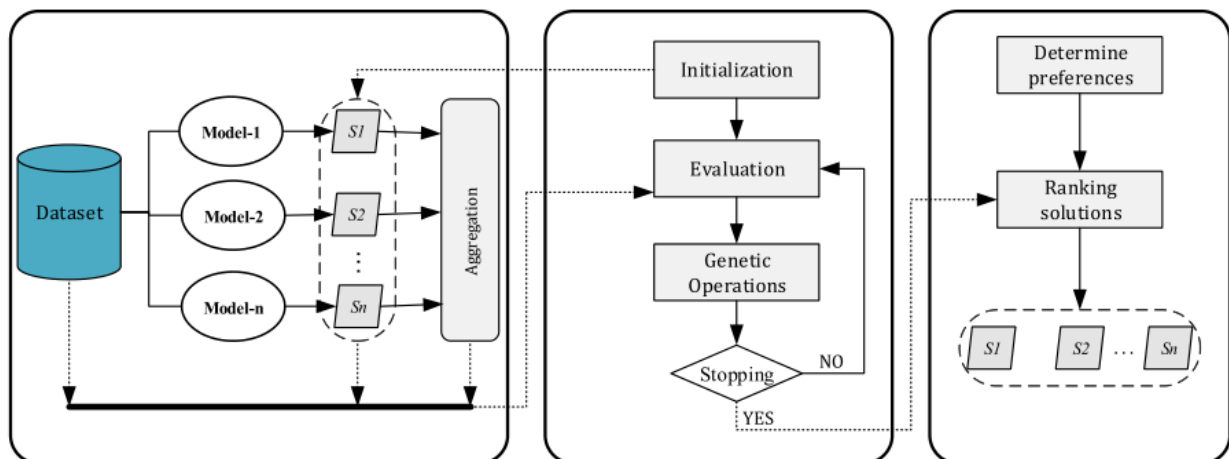


图18.进化集成成员选择。

在进化集成成员组合的过程中，使用EC算法对基本模型进行加权，从而执行加权多数投票方案。该方法通过调整基本模型的权重来提高整个集成的预测性能。图19显示了进化合奏成员组合的总体概况。EC算法用于对建立在验证数据集上的一组更好的模型进行优先排序和选择。

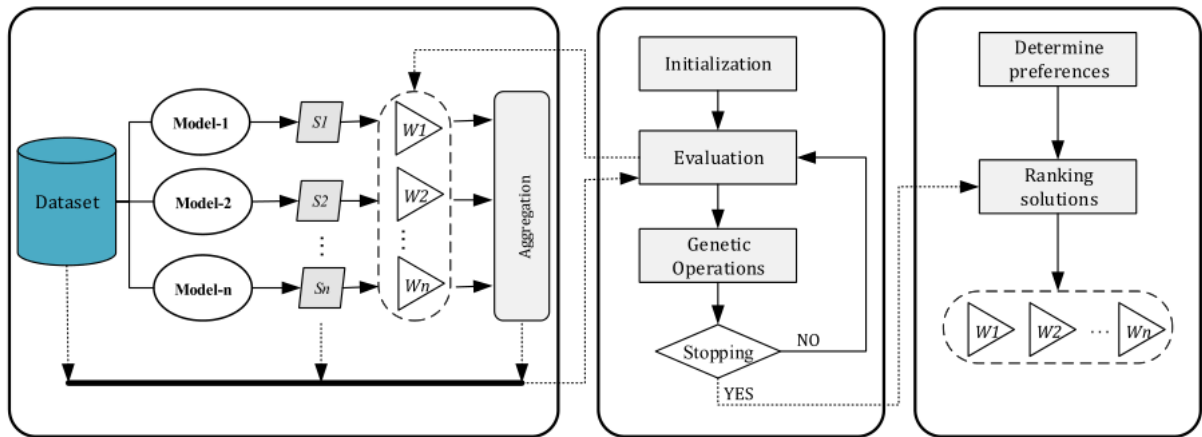


图19.进化集成成员组合

使用集成技术和重采样方法(例如，欠采样和过采样)相结合来解决类不平衡问题已被证明能够提高对少数类的正确分类。基于随机抽样的集成将不会有足够的表现。事实上，多数类的潜在有用样本可以被拒绝，这可能对学习过程很重要。当不平衡比率增加时，这一点更加明显。进化抽样是一种策略，其中强调有利于最多多样化的个体的分类器之间的多样性[46]。进化抽样和集成方法允许适应度函数促进过采样或欠采样数据集的多样性，从而在处理高度不平衡的数据集时获得更准确的结果。

9 进化模型优化

EC 机制可用于 ML 的后处理阶段，当优化传统 ML 算法构建的模型时。在 DT 分类中，EC 算法可以作为进化组件用于修剪生成的树，以删除所有可能受噪声或不精确数据影响的部分，这将防止 DT 模型过度拟合并降低最终 DT 的复杂性。然而，要在修剪水平和预测精度之间找到正确的权衡并不容易。过度修剪会显着扭曲 DT，因此只表示一小部分训练数据。相比之下，欠剪枝可能会导致 DT 对训练数据进行过拟合。树优化的两个主要策略是预剪枝和后剪枝。为每个样本实施阈值是预剪枝策略中的常见解决方案，如果模型性能低于预定阈值，则限制每次扩展。与预剪枝不同，后剪枝需要长出一棵完整的树。首先通过过度拟合训练集来构建完整的 DT。然后修剪树以提高其性能并最小化其大小。在实践中，后剪枝比预剪枝表现更好[117]。同样，二进制编码是一种众所周知的表示技术：解的长度等于 DT 中分支节点的数量。值为“1”表示为结果树选择了分支节点；否则，将不会被选中。

4.进化机器学习的应用

人工智能和机器学习有可能迎来另一场能够自动构建智能系统的“工业革命”。这不仅支持许多工业和专业流程，而且还具有改善日常生活的潜力。不同的情况会降低传统机器学习在实际应用中的机器学习性能。缺乏有效运行传统 ML 的专家知识是行业和企业面临的主要挑战，因为 ML 结果的质量严重依赖于专家经验来确定超参数和其他有关模型设计的调整。如果领域知识不容易获得，EML 方法可能是一种有用的替代方法。例如，传统神经网络中的两个主要问题是网络拓扑的定义和超参数的调整；这些都需要大量的用例背景知识。例如，在将神经网络应用于癌症检测等领域时，患者诊断需要高精度。不恰当的选择会影响性能检测系统，并可能危及患者的预后。

在实践中，由于时间或预算等资源的限制，优化对于最小化或最大化目标至关重要，这在所有行业和其他商业活动中都是相关的。几乎所有的最大似然问题都可以转化为显式优化问题。用进化优化方法训练 ML 模型应该改进与应用程序相关的目标。进化算法可以与损失函数协同更新 ML 模型的参数。表1显示了 EML 用于解决不同实际问题的一些应用程序。这些应用涵盖不同的领域，如计算机网络、商业、计算机视觉和机器人技术。

表1.进化机器学习在实际应用中的总结

种类	应用
计算机网络	网络安全[40、106、110、116、150、179]、电子邮件垃圾邮件检测[153]、无线传感器网络[142]、Web挖掘[59]、网络钓鱼检测[167]
商业	营销[44]、购物篮分析[62]、推荐系统[190]、电子商务[20]、工作流分析[104]、协作过滤[176]
机器人学	自动车辆导航[154]、机器人学[192]
医学	疾病诊断[37]、医学[90,171]、癌症分类[29]、基因表达数据分析[114]、心血管疾病检测[181]、医疗保健[88]、胸外科[118]、胃肠道感染预测[152]
计算机视觉	人脸识别[178]、手写识别[135]、说话人识别[188]、人员识别[30]、字符识别[60]、行人检测[159]、手写数字分类[76]、图像分割[184]、图像聚类[39]、文档聚类[87]
工业	金融[86,132]、软件工程[78]、建筑业[31]、服装业[92]、产品设计[45、66]、产品服务体系[191]
环境	分析臭氧含量[121]、交通拥堵预测[180]、道路交通预测[101]、大气污染[122]、预测臭氧[120]
其他	天文学[28]、教育系统[113、147]、停车场占用预测[26]、能源价格[141]、能源消耗预测[8]、智慧城市[85]

计算机网络：EML可以在计算机网络中使用，以提高ML模型在不同领域的性能，如网络安全、传感器网络或Web挖掘。根据文献，由于可能威胁隐私、完整性和网络资源可用性的恶意活动，保护网络安全是将EML方法应用于计算机网络的主要焦点。进化方法旨在提高ML算法的性能，例如ARM、聚类和分类(例如，深度学习分类器)，主要用于保护计算机网络免受入侵、电子邮件垃圾邮件和网络钓鱼等攻击。

商业：商业是ARM和预测等EML方法的初始目标。购物篮分析发现购买的商品之间的关系，以支持关于商店布局的决策和营销策略。电子商务网站(如亚马逊和eBay)分析客户活动，以提取个性化的偏好和兴趣，并识别用户趋势。企业中的推荐系统可以基于EML技术生成的知识构建，使用信息发现技术向潜在客户提供商品。进化臂是协作过滤中使用的技术之一，在协作过滤中，用户对感兴趣项目的偏好表示为评级。

计算机视觉：计算机视觉是ML技术最具挑战性的应用之一。多媒体任务中巨大的搜索空间使传统的最大似然方法难以实现(如陷入局部最优和/或计算代价较高)，因此进化特征选择/构造/提取方法已成为该领域的主流方法。深度学习中的进化应用已成功地应用于计算机视觉和语音识别。EML在计算机视觉中的应用可以分为两类：应用领域，如医学或机器人；目标任务，如人脸识别或图像分割。例如，阈值的定义和测量是图像分割中的两个具有挑战性的任务，可以通过进化算法来解决。

机器人学：用于机器人的RL可以通过与环境的试错交互来帮助自主发现最佳行为。然而，当在高维和稀疏奖励的环境中将RL应用于机器人时，传统的RL技术不能改善行为学习。进化RL在机器人学中的应用使得自主机器人(例如，车辆或生产线)能够以最少的人类交互来学习行为技能。事实上，ML和进化优化的结合极大地提高了决策系统的决策质量和学习能力。进化优化的全部潜力在Robotics中还没有达到，因为传统的ML方法已经显示出在可靠的数据集可用的情况下具有竞争能力。但复杂的环境和低效的启发式优化函数为机器人中的EML技术提供了机会。

硬件方面的最新进展，如云计算和GPU设备，使得以前不可能的EML任务变得可以实现。像谷歌、微软、优步或IBM这样的大公司已经投资了EML方法，并积极寻求现实生活中的解决方案。

5.讨论与挑战

1 讨论

EC算法已被应用于ML技术，以缓解与传统和启发式ML技术相关的问题。图20显示了EML任务的分类以及与EC方法试图解决的任务相关的挑战。与特征选择/构造和重采样方法相关的任务是EC算法在预处理阶段的主要贡献。前者的贡献集中在通过选择子集特征或通过从原始特征构造新的特征集来生成新的特征空间。后者贡献是通过进化欠采样和过采样技术实现的。

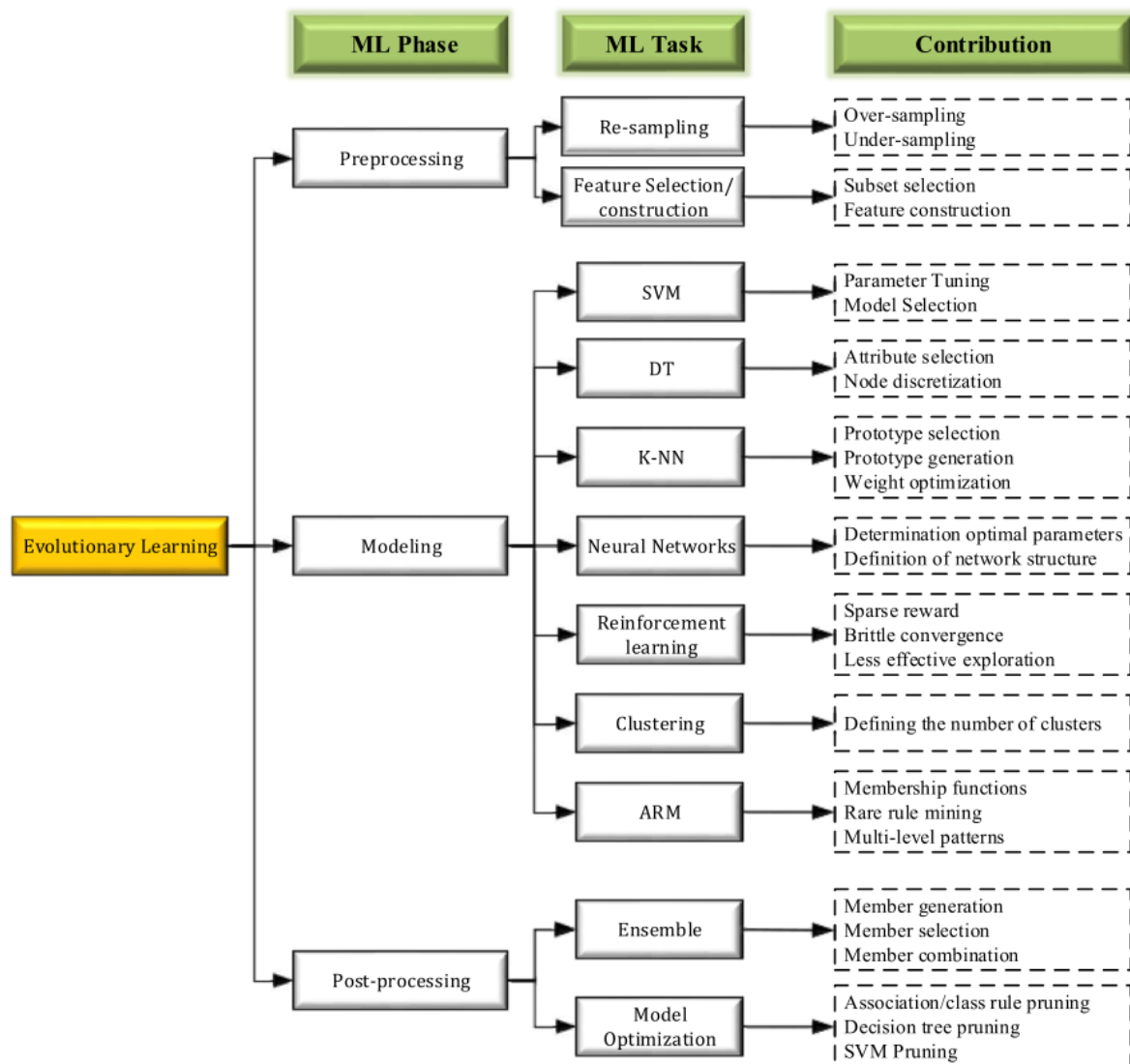


图20.进化机器学习的分类

但EC算法的主要关注点是提高学习算法的性能。EML算法可以分为三类：有监督进化学习、无监督进化学习和强化进化学习。进化分类/预测和进化集成是监督学习的主要贡献。在无监督学习领域，EC可用于特征选择、聚类、降维和异常检测等任务。通过将EC方法集成到强化学习中，可以解决RL的不同挑战，如长时间视野、稀疏奖励、需要补充纠正机制、以及高维动作和状态空间。特别地，EC技术经常被用作RL中的策略搜索机制。

ML加权是一种常用的技术，用于强调数据的某些特征，从而改进结果模型。例如，可以使用加权系统来概述某些特定实例或特征的重要性，或者在合奏的上下文中对一组技术进行排名[124]。神经网络、支持向量机和k-NN是受益于权重的最常见技术。加权系统的主要目标是在训练阶段优化一组模型权重。权重也可以应用于k-NN的投票系统，EC方法可以用作ML中的权重优化器。

在训练数据集上模型的性能较高，但在不可见数据上模型性能较低的分类器中，过拟合是不可忽视的，这会导致泛化能力较差。深度学习模型及其网络拓扑和神经结构导致的巨大复杂性，以及不平衡或高维数据集，是过度拟合问题背后的一些原因。虽然进化算法可以帮助解决这些因素中的每一个，但它们的两个主要贡献是为分类器提供适当平衡和特征减少的输入集。此外，神经网络的学习机制通常收敛到局部极小，因为损失函数几乎总是非凸的[38]。进化编码的代价敏感度是一种改进损失函数以增加分类器对不平衡数据集的稳健性的策略。在神经网络和深度学习的优化中使用多目标方法可以进一步平衡精度和泛化。支持向量机中的惩罚参数C提供了泛化和训练误差之间的折衷[108]。支持向量机进化方法的主要目标之一是优化参数C以提高泛化能力。决策树和支持向量机模型中使用的优化和修剪技术导致了更通用的模型[117]。

在模糊ML中，隶属度函数用于将数值转换为语言术语。隶属度函数的选择会影响模糊最大似然模型的模式发现，因此，学习或调整隶属度函数是有益的。在传统的模糊最大似然方法中，假设隶属度函数是事先已知的。然而，要获得覆盖所有数值变量域的最有效模糊集的先验信息是不可能的。使用EC算法提取隶属度函数是进化ARM任务的主要趋势。

用于模型优化的剪枝策略已成功地应用于DT、支持向量机和ARM。在这里，使用传统的最大似然算法建立模型，然后使用EC进行模型优化。大多数传统的ARM算法可以提取大量的规则，这些规则往往包含冗余和不相关的信息。例如，树修剪是一种ML技术，用于最小化DT的大小，以降低分类器的复杂性并提高其预测精度。在树修剪过程中，DT的一些子树被替换为树叶。支持向量机算法往往会产生巨大的支持向量，导致决策函数收敛速度降低。此外，由于过度拟合的影响，所得到的支持向量机模型可能会适应训练集中的噪声而不是真实的底层数据分布，从而无法正确地对未见样本进行分类。对训练好的支持向量机进行剪枝可以得到更快、更准确的支持向量机。EC算法是对ML算法提取的模型和模式进行剪枝的最重要的技术。

EC方法使我们能够开发ARM算法来提取关联规则，而不需要频繁项集挖掘步骤，从而降低了计算复杂度。然而，进化ARM算法的主要重点是处理定量数据，其中既考虑了将值离散到适当的区间中，也考虑了利用EC方法推导隶属函数来模糊化定量交易。

EC方法与层次聚类算法的结合在文献中还没有涉及。这可能是因为定义一个能够引导进化的适应度函数并不是一件简单的事情。据我们所知，只有几项研究讨论了这个问题[68,109]

2 挑战与未来洞察

在过去的几十年里，各种EC算法已经被应用到ML任务中，但一些严重的问题仍然没有得到足够的研究。下面将描述这些主要研究差距中的一些。

缺乏实验结果：各种调查试图使用研究方法对以EML方法发表的论文进行分类。由于运行时的数学分析、收敛保证和参数配置是一项基本需求，这使得为现实世界的应用选择合适的EML算法成为组织和实践者的一项具有挑战性的任务。EC算法可以成功地用于神经网络和支持向量机的参数整定，但对于初学者来说，很难判断使用哪种算法来完成特定的任务。大多数提案都集中在将EML算法与非进化或传统技术进行比较。然而，这些研究都没有系统地比较不同EC算法在多ML任务中的性能。我们无法找到一项比较研究，确定哪种EC技术在参数设置、结构设计、计算复杂性和其他方面取得了更好的性能。无法获得此类调查可能是由于各种原因，包括缺乏用于进化ML方法的公开可用的源代码、编码技术的变化、不同的目标函数和进化算子。但是，提供这样一个系统的比较应该有助于新的ML用户为特定的应用程序选择合适的方法。

缺乏关于进化机器学习的调查：关于进化ML的不同方面，如集群、DT、神经网络和深度学习，已经发表了各种调查论文。然而，对于EC算法在RL、重采样、分类、支持向量机和集成等领域的应用还缺乏全面的研究。例如，已经发表了不同的关于RL的综述论文，集中在不同的方面，如机器人中的RL[81]、深度RL[13]和安全RL[48]。然而，到目前为止，进化性RL尚未被综述。应该考虑和进一步研究这一领域的研究。

模块化EML：大多数EML模型是专门为解决特定问题而开发的，不能应用于不同的领域。模块化学习是一种将ML模型应用于不同问题的可能性，其中建立了各种通用的模型，并且可以通过较小的自主模块进行学习。每个独立训练的模型都旨在解决大量ML问题中常见的特定子任务。当不同模式之间有合作时，一个大问题就可以解决。训练模型自主地意味着它们可以在其他领域中被重用。在模块学习中应考虑不同的问题，例如确定子任务和确定其具体模块，从以前学习的一组模块中确定候选者，以及通过连接模块创建连贯和有效的模式。多任务学习是解决这个问题的一种方法[71]。

迁移学习：迁移学习的主要思想是重用以前学习的模型来解决新问题。这在ML社区是一个相对较新的研究领域。随着问题的不断增加，这个想法最近变得更加重要。例如，手写字符识别模型可用于识别数字化图书中的字符。总而言之，不同的ML问题有某些共同的方面，需要能够将为一个问题获得的一些专业知识转化为另一个问题。

进化CNN：CNN为大型应用做出了贡献，并已成功地应用于许多领域。然而，进化的CNN一直是一个未被探索的领域，直到最近才受到关注。这是一条很有前途的研究路线，为研究人员提供了各种机会。CNN拓扑的自动进化设计是一个非常前途的领域，需要进一步研究。

多目标EML：标准的EML算法在模型开发过程中通常只优化一个目标，而大多数ML问题有不同的目标需要优化。例如，ARM问题具有支持、信心和可理解性等目标，所有这些目标都必须同时优化。目标函数的选择是多目标EML中的一个重要问题。大多数算法同时优化两个目标，只有少数算法可以同时优化三个以上的目标函数。NSGA-II、PESA-II和SPEA2等多目标算法在解决四个以上目标的问题时面临困难。目前，这种方法在ML中的使用在文献中很少引起注意。此外，进化算法使用选择、交叉和变异等操作符。选择算子主要受多目标进化算法的影响，多目标进化算法是最大似然算法的优化器。相反，交叉和变异操作符通常由编码策略决定。选择最终解的方法是多目标最大似然问题中最重要的任务之一。多目标进化算法在最后一代中提供了一组非支配解，从这一组选择一个解是很重要的。基于目标、基于拐点和基于集合的方法是三种主要的选择技术。帕累托优势用于在多目标问题中寻找关系和比较解决方案[162–164]。然而，随着目标的数量增加到三个以上，单独的帕累托优势不再令人满意[93]。这种需要增加算法复杂性的问题被称为“多目标优化问题”[36]。它们出现在不同的现实世界领域，如空中交通管制或护士排班。将多目标进化方法整合到ML模型中可以解决各种各样的应用问题。

大数据上的EML：大数据为ML提供了新的机遇，但也带来了计算成本、巨大的高维样本量、存储僵局和错误程度等挑战[161]。进化ML的大多数研究只关注质量ML模型，而计算效率是严重大规模ML问题中的一个关键问题，却很少受到关注。搜索机制和适应度值计算的成本是大规模EML过程中的主要挑战，因为在EML方法中每一代都会评估个体群体。当数据集大小增加时，EML算法应该显示出良好的可扩展性。这些类型的数据集需要大内存和长计算时间。可扩展性问题可能会限制EML算法在大规模问题上的适用性。使用大数据处理技术（如主/从、MapReduce和CPU/GPU架构）的并行/分布式进化ML是处理大规模EML的主要解决方案之一[165]。

进化成本敏感学习：在某些分类问题中，错误分类错误之间的成本差异可能相当高。例如，在每个类别代表一个人是否患有癌症的癌症诊断系统中，与将健康人分类为患者相比，错误地将患者分类为健康将导致更大的成本。因此，错误的诊断可能会导致治疗延误或患者死亡[166]。成本敏感学习是一种最小化学习总成本的策略，它以这样一种方式创建学习模型，使训练过程对成本较高的类更敏感。

除了错误分类成本之外，测试成本是实际应用中的另一种重要成本类型，包括金钱、时间或其他资源。最近在成本敏感的学习领域提出了一些方法，试图将特定类别的成本集成到ML算法中，例如深度学习[49, 77]和DTs[83, 102]。然而，迄今为止，tEC算法和成本敏感学习在ML分类器中的集成很少受到关注。由于缺乏先验知识，错误分类成本通常是未知的，并且在实践中难以选择。最近，已经开发了一种进化成本敏感的DBN，其中采用了自适应DE来优化成本函数中使用的错误分类成本[189]。

总结

进化计算算法一直专注于解决传统ML任务的特殊挑战。在本文中，我们从问题编码、搜索机制、适应度函数以及EC算法试图解决的不同挑战等方面综述了EC算法在ML任务设计中的重要性。我们研究了EC算法做出重大贡献的九个不同任务。EML中的ML问题可以用三种主要表示形式表示：图(适用于蚁群算法)、树(适用于规范遗传编程)和向量(大多数EC算法，如GA、PSO和ABC)。搜索机制可用于基于单个解或基于多个解的群体来寻找最优解。每项任务都有具体的评估措施，这些措施是以适应度函数的形式制定的。例如，准确度、召回率、敏感度、特异度和精确度是分类应用的主要目标。评价措施可视为单目标或多目标。

我们描述了应用现有进化机器学习算法的各个领域，包括医学（胸外科和疾病诊断）、计算机网络（入侵检测、流量分类和垃圾邮件检测）、图像和视频处理（人脸识别和手写识别）和环境（例如大气污染、分析臭氧含量和预测臭氧）。EML技术似乎可以在未来的AI和ML中发挥重要作用，并有望进一步扩大其应用范围。

EML仍然存在一些尚未解决的问题。看来，对于进化成本敏感的ML、模块化EML、迁移学习、大数据EML和多目标EML，主要的研究工作是必要的。预计在未来几年内，EC算法与深度学习的集成将在平衡准确性的同时加快训练过程。

同样仍然缺乏有助于评估 EC 方法在不同应用和 ML 任务中的有效性的比较研究。人们经常担心特定 EC 算法在解决各种 ML 问题中的效用。应进行不同的统计检验。此外，一些调查似乎在 EML 领域很有用，例如进化 RL、进化重采样、进化分类、进化 SVM 和进化集成。

鉴于 ML 算法在实际应用中的广泛适用性，学术研究界、行业和制造商必须始终如一地积极应对传统 ML 的挑战。到目前为止，使用进化算法优化 ML 的研究大多在学术出版物中进行。未来，EML 可能会出现在许多行业的多个软件包中，并将进一步融入我们的日常生活。ML 在各种应用中的重要性不断增长；因此，我们很可能会看到基于云的尖端技术，例如**机器学习即服务 (Machine Learning-as-a-Service, MLaaS)**，其中进化优化也可以发挥重要作用。

参考文献

- [1] Rakesh Agrawal, Tomasz Imieliński, and Arun Swami. 1993. Mining association rules between sets of items in large databases. In *ACM SIGMOD Record*, Vol. 22. ACM, 207 – 216.
- [2] Rakesh Agrawal, Ramakrishnan Srikant, et al. 1994. Fast algorithms for mining association rules. In *Proceedings of the 20th International Conference on Very Large Data Bases, VLDB*, Vol. 1215. 487 – 499.
- [3] Harith Al-Sahaf, Ying Bi, Qi Chen, Andrew Lensen, Yi Mei, Yanan Sun, Binh Tran, Bing Xue, and Mengjie Zhang. A survey on evolutionary machine learning. *J. Roy. Soc. New Zeal.* 49, 2 (2019), 205 – 228.
- [4] Shafiq Alam, Gillian Dobbie, Yun Sing Koh, Patricia Riddle, and Saeed Ur Rehman. 2014. Research on particle swarm optimization based clustering: A systematic review of literature and techniques. *Swarm Evolut. Comput.* 17 (2014), 1 – 13.
- [5] Wissam A. Albukhanajer, Yaochu Jin, and Johann A. Briffa. 2017. Classifier ensembles for image identification using multi-objective Pareto features. *Neurocomputing* 238 (2017), 316 – 327.
- [6] Hamid Ali, Waseem Shahzad, and Farrukh Aslam Khan. 2012. Energy-efficient clustering in mobile ad-hoc networks using multi-objective particle swarm optimization. *Appl. Soft Comput.* 12, 7 (2012), 1913 – 1928.
- [7] Ibrahim Aljarah, Majdi Mafarja, Ali Asghar Heidari, Hossam Faris, and Seyedali Mirjalili. 2019. Clustering analysis using a novel locality-informed grey wolf-inspired clustering approach. *Knowl. Inf. Syst.* 62, 2 (2019), 1 – 33.
- [8] Abdulaziz Almalaq and Jun Jason Zhang. 2018. Evolutionary deep learning-based energy consumption prediction for buildings. *IEEE Access* 7 (2018), 1520 – 1531.
- [9] Mehrdad Almasi and Mohammad Saniee Abadeh. 2015. Rare-PEARs: A new multi objective evolutionary algorithm to mine rare and non-redundant quantitative association rules. *Knowl.-based Syst.* 89 (2015), 366 – 384.
- [10] Akram AlSukker, Rami Khushaba, and Ahmed Al-Ani. 2010. Optimizing the k-nn metric weights using differential evolution. In *Proceedings of the International Conference on Multimedia Computing and Information Technology (MCIT)*. IEEE, 89 – 92.
- [11] Amazon. 2017. Amazon EC2 P3 Instances. Retrieved from <https://aws.amazon.com/es/ec2/instance-types/p3>.
- [12] Kai Arulkumaran, Marc Peter Deisenroth, Miles Brundage, and Anil Anthony Bharath. 2017. A brief survey of deep reinforcement learning. *arXiv preprint arXiv:1708.05866* (2017).
- [13] Kai Arulkumaran, Marc Peter Deisenroth, Miles Brundage, and Anil Anthony Bharath. 2017. Deep reinforcement learning: A brief survey. *IEEE Sig. Process. Mag.* 34, 6 (2017), 26 – 38.
- [14] İlhan Aydin, Mehmet Karaköse, and Erhan Akin. 2011. A multi-objective artificial immune algorithm for parameter optimization in support vector machine. *Appl. Soft Comput.* 11, 1 (2011), 120 – 129.

- [15] Bodrunnessa Badhon, Mir Md Jahangir Kabir, Shuxiang Xu, and Monika Kabir. 2019. A survey on association rule mining based on evolutionary algorithms. *Int. J. Comput. Applic.* 41, 1 (2019), 1 – 11.
- [16] Alejandro Baldominos, Yago Saez, and Pedro Isasi. 2020. On the automated, evolutionary design of neural networks: past, present, and future. *Neural Comput. Applic.* 32, 2 (2020), 1 – 27.
- [17] Rodrigo Coelho Barros, Márcio Porto Basgalupp, Andre C. P. L. F. De Carvalho, and Alex A. Freitas. 2011. A survey of evolutionary algorithms for decision-tree induction. *IEEE Trans. Syst., Man, Cyber., Part C (Appl. Rev.)* 42, 3 (2011), 291 – 312.
- [18] James C. Bezdek, Srinivas Boggavarapu, Lawrence O. Hall, and Amine Bensaid. 1994. Genetic algorithm guided clustering. In *Proceedings of the 1st IEEE Conference on Evolutionary Computation. IEEE World Congress on Computational Intelligence*. IEEE, 34 – 39.
- [19] Urvesh Bhowan, Mark Johnston, and Mengjie Zhang. 2011. Developing new fitness functions in genetic programming for classification with unbalanced data. *IEEE Trans. Syst., Man, Cyber., Part B (Cyber.)* 42, 2 (2011), 406 – 421.
- [20] Cosimo Birtolo, Diego De Chiara, Simona Losito, Pierluigi Ritrovato, and Mario Veniero. 2013. Searching optimal product bundles by means of GA-based Engine and Market Basket Analysis. In *Proceedings of the Joint IFSA World Congress and NAFIPS Meeting (IFSA/NAFIPS)*. IEEE, 448 – 453.
- [21] Christopher M. Bishop. 2006. *Pattern Recognition and Machine Learning*. Springer.
- [22] Nimagna Biswas, Saurajit Chakraborty, Sankha Subhra Mullick, and Swagatam Das. 2018. A parameter independent fuzzy weighted k-nearest neighbor classifier. *Pattern Recog. Lett.* 101 (2018), 80 – 87.
- [23] Veronica Bolon-Canedo, Noelia Sanchez-Marono, and Amparo Alonso-Betanzos. 2011. Feature selection and classification in multiple class datasets: An application to KDD Cup 99 dataset. *Exp. Syst. Applic.* 38, 5 (2011), 5947 – 5957.
- [24] Leo Breiman. 1996. Bagging predictors. *Mach. Learn.* 24, 2 (1996), 123 – 140.
- [25] Lam Thu Bui, Thi Thu Huong Dinh, et al. 2018. A novel evolutionary multi-objective ensemble learning approach for forecasting currency exchange rates. *Data Knowl. Eng.* 114 (2018), 40 – 66.
- [26] Andrés Camero, Jamal Toutouh, Daniel H. Stolfi, and Enrique Alba. 2018. Evolutionary deep learning for car park occupancy prediction in smart cities. In *Proceedings of the International Conference on Learning and Intelligent Optimization*. Springer, 386 – 401.
- [27] José Ramón Cano, Francisco Herrera, and Manuel Lozano. 2005. Stratification for scaling up evolutionary prototype selection. *Pattern Recog. Lett.* 26, 7 (2005), 953 – 963.
- [28] Erick Cantú-Paz and Chandrika Kamath. 2000. Combining evolutionary algorithms with oblique decision trees to detect bent-double galaxies. In *Applications and Science of Neural Networks, Fuzzy Systems, and Evolutionary Computation III*, Vol. 4120. International Society for Optics and Photonics, 63 – 71.
- [29] P. A. Castillo, Maribel García Arenas, Juan J. Merelo, V. M. Rivas, and Gustavo Romero. 2006. Multiobjective optimization of ensembles of multilayer perceptrons for pattern classification. In *Parallel Problem Solving from Nature-PPSNIX*. Springer, 453 – 462.
- [30] Kingshuk Chakravarty, Diptesh Das, Aniruddha Sinha, and Amit Konar. 2013. Feature selection by differential evolution algorithm-a case study in personnel identification. In *Proceedings of the IEEE Congress on Evolutionary Computation*. IEEE, 892 – 899.

- [31] Yusi Cheng, Qiming Li, et al. 2015. GA-based multi-level association rule mining approach for defect analysis in the construction industry. *Autom. Construct.* 51 (2015), 78 – 91.
- [32] Brian Cheung and Carl Sable. 2011. Hybrid evolution of convolutional networks. In *Proceedings of the 10th International Conference on Machine Learning and Applications and Workshops*, Vol. 1. IEEE, 293 – 297.
- [33] Thomas Cover and Peter Hart. 1967. Nearest neighbor pattern classification. *IEEE Trans. Inf. Theor.* 13, 1 (1967), 21 – 27.
- [34] Giuseppe Cuccu, Matthew Luci, Jürgen Schmidhuber, and Faustino Gomez. 2011. Intrinsically motivated neuroevolution for vision-based reinforcement learning. In *Proceedings of the IEEE International Conference on Development and Learning (ICDL)*, Vol. 2. IEEE, 1 – 7.
- [35] Antoine Cully, Jeff Clune, Danesh Tarapore, and Jean-Baptiste Mouret. 2015. Robots that can adapt like animals. *Nature* 521, 7553 (2015), 503 – 507.
- [36] David M. Curry and Cihan H. Dagli. 2014. Computational complexity measures for many-objective optimization problems. *Procedia Comput. Sci.* 36 (2014), 185 – 191.
- [37] Sérgio Francisco Da Silva, Marcela Xavier Ribeiro, João do E. S. Batista Neto, Caetano Traina-Jr, and Agma J. M. Traina. 2011. Improving the ranking quality of medical image retrieval using a genetic feature selection method. *Decis. Supp. Syst.* 51, 4 (2011), 810 – 820.
- [38] Ashraf Darwish, Aboul Ella Hassanien, and Swagatam Das. 2020. A survey of swarm and evolutionary computing approaches for deep learning. *Artif. Intell. Rev.* 53, 3 (2020), 1767-1812.
- [39] Swagatam Das and Amit Konar. 2009. Automatic image pixel clustering with an improved differential evolution. *Appl. Soft Comput.* 9, 1 (2009), 226-236.
- [40] Emiro De la Hoz, Eduardo De La Hoz, Andrés Ortiz, Julio Ortega, and Antonio Martinez-Alvarez. 2014. Feature selection by multi-objective optimisation: Application to network anomaly detection by hierarchical self-organising maps. *Knowl.-based Syst.* 71 (2014), 322-338.
- [41] Jeff Dean and U. Holzle. 2017. Build and train machine learning models on our new Google Cloud TPUs. Retrieved from <https://www.blog.google/topics/google-cloud/google-cloud-offer-tpus-machine-learning>.
- [42] Hongbin Dong, Yuxin Dong, Cheng Zhou, Guisheng Yin, and Wei Hou. 2009. A fuzzy clustering algorithm based on evolutionary programming. *Exp. Syst. Applic.* 36, 9 (2009), 11792-11800.
- [43] Thomas Elsken, Jan Hendrik Metzen, and Frank Hutter. 2019. Neural architecture search: A survey. *J. Mach. Learn. Res.* 20, 55 (2019), 1-21.
- [44] Zhiwei Fu, Bruce L. Golden, Shreevardhan Lele, S. Raghavan, and Edward A. Wasil. 2003. A genetic algorithm-based approach for building accurate decision trees. *INFORMS J. Comput.* 15, 1 (2003), 3-22.
- [45] K. Y. Fung, C. K. Kwong, Kin Wai Michael Siu, and K. M. Yu. 2012. A multi-objective genetic algorithm approach to rule mining for affective product design. *Exp. Syst. Applic.* 39, 8 (2012), 7411-7419.
- [46] Mikel Galar, Alberto Fernández, Edurne Barrenechea, and Francisco Herrera. 2013. EUSBoost: Enhancing ensembles for highly imbalanced data-sets by evolutionary undersampling. *Pattern Recog.* 46, 12 (2013), 3460-3471.
- [47] Salvador Garcí, Isaac Triguero, Cristóbal J. Carmona, Francisco Herrera, et al. 2012. Evolutionary-based selection of generalized instances for imbalanced classification. *Knowl.-based Syst.* 25, 1 (2012), 3-12.

- [48] Javier Garcia and Fernando Fernández. 2015. A comprehensive survey on safe reinforcement learning. *J. Mach. Learn. Res.* 16, 1 (2015), 1437-1480.
- [49] Salvador Garcia, Alberto Fernández, and Francisco Herrera. 2009. Enhancing the effectiveness and interpretability of decision tree and rule induction classifiers with evolutionary training set selection over imbalanced problems. *Appl. Soft Comput.* 9, 4 (2009), 1304-1314.
- [50] Salvador Garcia and Francisco Herrera. 2009. Evolutionary undersampling for classification with imbalanced datasets: Proposals and taxonomy. *Evolut. Comput.* 17, 3 (2009), 275-306.
- [51] Fred Glover. 1986. Future paths for integer programming and links to artificial intelligence. *Comput. Oper. Res.* 13, 5(1986), 533-549.
- [52] Taciana A. F. Gomes, Ricardo B. C. Prudêncio, Carlos Soares, André L. D. Rossi, and André Carvalho. 2012. Combining meta-learning and search techniques to select parameters for support vector machines. *Neurocomputing* 75, 1 (2012), 3-13.
- [53] Ian Goodfellow, Yoshua Bengio, and Aaron Courville. 2016. *Deep Learning*. The MIT Press.
- [54] Frédéric Gruau. 1994. Automatic definition of modular neural networks. *Adapt. Behav.* 3, 2 (1994), 151-183.
- [55] Shenkai Gu and Yaochu Jin. 2014. Generating diverse and accurate classifier ensembles using multi-objective optimization. In *Proceedings of the IEEE Symposium on Computational Intelligence in Multi-Criteria Decision-Making (MCDM)*. IEEE, 9-15.
- [56] X. C. Guo, J. H. Yang, C. G. Wu, C. Y. Wang, and Y. C. Liang. 2008. A novel LS-SVMs hyper-parameter selection based on particle swarm optimization. *Neurocomputing* 71, 16-18 (2008), 3211-3215.
- [57] Jiawei Han, Jian Pei, and Yiwen Yin. 2000. Mining frequent patterns without candidate generation. In *ACM Sigmod Record*, Vol. 29. ACM, 1-12.
- [58] Emrah Hancer, Bing Xue, Dervis Karaboga, and Mengjie Zhang. 2015. A binary ABC algorithm based on advanced similarity scheme for feature selection. *Appl. Soft Comput.* 36 (2015), 334-348.
- [59] Julia Handl and Bernd Meyer. 2002. Improved ant-based clustering and sorting in a document retrieval interface. In *Proceedings of the International Conference on Parallel Problem Solving from Nature*. Springer, 913-923.
- [60] Shigeru Haruyama and Qiangfu Zhao. 2002. Designing smaller decision trees using multiple objective optimization based GPS. In *Proceedings of the IEEE International Conference on Systems, Man and Cybernetics*, Vol. 6. IEEE.
- [61] Mohamed Jafar Abul Hasan and Sivakumar Ramakrishnan. 2011. A survey: Hybrid evolutionary algorithms for cluster analysis. *Artif. Intell. Rev.* 36, 3 (2011), 179-204.
- [62] Majeed Heydari and Amir Yousefli. 2017. A new optimization model for market basket analysis with allocation considerations: A genetic algorithm solution approach. *Manag. Market. Chall. Knowl. Soc.* 12, 1 (2017), 1-11.
- [63] John H. Holland. 1992. *Adaptation in Natural and Artificial Systems: An Introductory Analysis with Applications to Biology, Control, and Artificial Intelligence*. The MIT Press.
- [64] Eduardo Raul Hruschka, Ricardo J. G. B. Campello, Alex A. Freitas, et al. 2009. A survey of evolutionary algorithms for clustering. *IEEE Trans. Syst., Man, Cyber., Part C (Appl. Rev.)* 39, 2 (2009), 133-155.
- [65] Jian Huang, Xiaoguang Hu, and Fan Yang. 2011. Support vector machine with genetic algorithm for machinery fault diagnosis of high voltage circuit breaker. *Measurement* 44, 6 (2011), 1018-1027.

- [66] Huimin Jiang, C. K. Kwong, W. Y. Park, and K. M. Yu. 2018. A multi-objective PSO approach of mining association rules for affective design based on online customer reviews. *J. Eng. Des.* 29, 7 (2018), 381-403.
- [67] Hua Jiang, Shenghe Yi, Jing Li, Fengqin Yang, and Xin Hu. 2010. Ant clustering algorithm with K-harmonic meansclustering. *Exp. Syst. Applic.* 37, 12 (2010), 8679-8684.
- [68] Raja Jothi, Elena Zotenko, Asba Tasneem, and Teresa M. Przytycka. 2006. COCO-CL: Hierarchical clustering of homology relations based on evolutionary correlations. *Bioinformatics* 22, 7 (2006), 779-788.
- [69] Krzysztof Jurczuk, Marcin Czajkowski, and Marek Kretowski. 2017. Evolutionary induction of a decision tree for large-scale data: A GPU-based approach. *Soft Comput.* 21, 24 (2017), 7363-7379.
- [70] Dervis Karaboga, Bahriye Akay, and Celal Ozturk. 2007. Artificial bee colony (ABC) optimization algorithm for training feed-forward neural networks. In *Proceedings of the International Conference on Modeling Decisions for Artificial Intelligence*. Springer, 318-329.
- [71] Stephen Kelly, Wolfgang Banzhaf, and Cedric Gondro. 2021. Evolving hierarchical memory-prediction machines in multi-task reinforcement learning. *arXiv preprint arXiv:2106.12659* (2021).
- [72] Stephen Kelly and Malcolm I. Heywood. 2017. Emergent tangled graph representations for Atari game playing agents. In *Proceedings of the European Conference on Genetic Programming*. Springer, 64-79.
- [73] Stephen Kelly and Malcolm I. Heywood. 2018. Emergent solutions to high-dimensional multitask reinforcement learning. *Evolut. Comput.* 26, 3 (2018), 347-380.
- [74] James Kennedy and Russell Eberhart. 1995. Particle swarm optimization. In *Proceedings of the International Conference on Neural Networks*, Vol. 4. IEEE, 1942-1948.
- [75] Shauharda Khadka and Kagan Tumer. 2018. Evolution-guided policy gradient in reinforcement learning. In *Proceedings of the International Conference on Advances in Neural Information Processing Systems*. 1188-1200.
- [76] Mujahid H. Khalifa, Marwa Ammar, Wael Ouarda, and Adel M. Alimi. 2017. Particle swarm optimization for deep learning of convolution neural network. In *Proceedings of the Sudan Conference on Computer Science and Information Technology (SCCSIT)*. IEEE, 1-5.
- [77] Salman H. Khan, Munawar Hayat, Mohammed Bennamoun, Ferdous A. Sohel, and Roberto Togneri. 2017. Costsensitive learning of deep feature representations from imbalanced data. *IEEE Trans. Neural Netw. Learn. Syst.* 29, 8(2017), 3573-3587.
- [78] Taghi M. Khoshgoftaar and Yi Liu. 2007. A multi-objective software quality classification model using genetic programming. *IEEE Trans. Reliab.* 56, 2 (2007), 237-245.
- [79] Scott Kirkpatrick, C. Daniel Gelatt, and Mario P. Vecchi. 1983. Optimization by simulated annealing. *Science* 220, 4598(1983), 671-680.
- [80] Hiroaki Kitano. 1990. Empirical studies on the speed of convergence of neural network training using genetic algorithms. In *Proceedings of the AAAI Conference on Artificial Intelligence*. 789-795.
- [81] Jens Kober, J. Andrew Bagnell, and Jan Peters. 2013. Reinforcement learning in robotics: A survey. *Int. J. Robot. Res.* 32, 11 (2013), 1238-1274.
- [82] Jan Koutnik, Jürgen Schmidhuber, and Faustino Gomez. 2014. Evolving deep unsupervised convolutional networks for vision-based reinforcement learning. In *Proceedings of the Conference on Genetic and Evolutionary Computation*. 541-548.

- [83] Bartosz Krawczyk, Michal Wozniak, and Gerald Schaefer. 2014. Cost-sensitive decision tree ensembles for effective imbalanced classification. *Appl. Soft Comput.* 14 (2014), 554-562.
- [84] D. Praveen Kumar, Tarachand Amgoth, and Chandra Sekhara Rao Annavarapu. 2019. Machine learning algorithms for wireless sensor networks: A survey. *Inf. Fus.* 49 (2019), 1-25.
- [85] Pardeep Kumar and Amit Kumar Singh. 2019. Efficient generation of association rules from numeric data using genetic algorithm for smart cities. In *Security in Smart Cities: Models, Applications, and Challenges*. Springer, 323-343.
- [86] Chan-Sheng Kuo, Tzung-Pei Hong, and Chuen-Lung Chen. 2007. Applying genetic programming technique in classification trees. *Soft Comput.* 11, 12 (2007), 1165-1172.
- [87] R.J. Kuo and L. M. Lin. 2010. Application of a hybrid of genetic algorithm and particle swarm optimization algorithm for order clustering. *Decis. Supp. Syst.* 49, 4 (2010), 451-462.
- [88] R. J. Kuo, S. Y. Lin, and C. W. Shih. 2007. Mining association rules through integration of clustering analysis and ant colony system for health insurance database in Taiwan. *Exp. Syst. Applic.* 33, 3 (2007), 794-808.
- [89] R.J. Kuo, Y.J. Syu, Zhen-Yao Chen, and Fang-Chih Tien. 2012. Integration of particle swarm optimization and genetic algorithm for dynamic clustering. *Inf. Sci.* 195 (2012), 124-140.
- [90] Halina Kwasnicka and Kajetan Switalski. 2006. Discovery of association rules from medical data-classical and evolutionary approaches. *Annales Universitatis Mariae Curie-Sklodowska, Sectio AI-Informatica* 4, 1 (2006), 204-217.
- [91] Yann A. LeCun, Léon Bottou, Genevieve B. Orr, and Klaus-Robert Müller. 2012. Efficient backprop. In *Neural Networks: Tricks of the Trade*. Springer, 9-48.
- [92] C. K. H. Lee, King Lun Choy, George T. S. Ho, and Cathy H. Y. Lam. 2016. A slippery genetic algorithm-based process mining system for achieving better quality assurance in the garment industry. *Exp. Syst. Applic.* 46 (2016), 236-248.
- [93] Bingdong Li, Jinlong Li, Ke Tang, and Xin Yao. 2015. Many-objective evolutionary algorithms: A survey. *ACM Comput. Surv.* 48, 1 (2015), 1-35.
- [94] Juan Li, Yuan-xiang Li, Sha-sha Tian, and Jie-lin Xia. 2019. An improved cuckoo search algorithm with self-adaptive knowledge learning. *Neural Comput. Applic.* 32, 16 (2019), 1-31.
- [95] Juan Li, Yuan-xiang Li, Sha-sha Tian, and Jie Zou. 2019. Dynamic cuckoo search algorithm based on Taguchi opposition-based search. *Int. J. Bio-insp. Comput.* 13, 1 (2019), 59-69.
- [96] Juan Li, Dan-dan Xiao, Hong Lei, Ting Zhang, and Tian Tian. 2020. Using cuckoo search algorithm with q-learning and genetic operation to solve the problem of logistics distribution center location. *Mathematics* 8, 2 (2020), 149.
- [97] Juan Li, Dan-dan Xiao, Ting Zhang, Chun Liu, Yuan-xiang Li, and Gai-ge Wang. 2021. Multi-swarm cuckoo search algorithm with q-learning model. *Comput. J.* 64, 1 (2021), 108-131.
- [98] Juan Li, Yuan-Hua Yang, Hong Lei, and Gai-Ge Wang. 2020. Solving logistics distribution center location with improved cuckoo search algorithm. *Int. J. Comput. Intell. Syst.* 14, 1 (2020), 676-692.
- [99] Wei Li and Gai-Ge Wang. 2021. Elephant herding optimization using dynamic topology and biogeography-based optimization based on learning for numerical optimization. *Eng. Comput.* (2021), 1-29.
- [100] Wei Li, Gai-Ge Wang, and Amir H. Alavi. 2020. Learning-based elephant herding optimization algorithm for solving numerical optimization problems. *Knowl.-based Syst.* 195 (2020), 105675.

- [101] Xianneng Li, Shingo Mabu, Huiyu Zhou, Kaoru Shimada, and Kotaro Hirasawa. 2010. Genetic network programming with estimation of distribution algorithms for class association rule mining in traffic prediction. In Proceedings of the IEEE Congress on Evolutionary Computation. IEEE, 1-8.
- [102] Xiangju Li, Hong Zhao, and William Zhu. 2015. A cost sensitive decision tree algorithm with two adaptive mechanisms. *Knowl.-based Syst.* 88 (2015), 24-33.
- [103] Jason Liang, Elliot Meyerson, Babak Hodjat, Dan Fink, Karl Mutch, and Risto Miikkulainen. 2019. Evolutionary neural automl for deep learning. In Proceedings of the Genetic and Evolutionary Computation Conference. 401-409.
- [104] Amy H. L. Lim, Chien-Sing Lee, and Murali Raman. 2012. Hybrid genetic algorithm and association rules for mining workflow best practices. *Exp. Syst. Applic.* 39, 12 (2012), 10544-10551.
- [105] Pin Lim, Chi Keong Goh, and Kay Chen Tan. 2016. Evolutionary cluster-based synthetic oversampling ensemble (eco-ensemble) for imbalance learning. *IEEE Trans. Cyber.* 47, 9 (2016), 2850-2861.
- [106] Yongguo Liu, Kefei Chen, Xiaofeng Liao, and Wei Zhang. 2004. A genetic clustering method for intrusion detection. *Pattern Recog.* 37, 5 (2004), 927-942.
- [107] Stuart Lloyd. 1982. Least squares quantization in PCM. *IEEE Trans. Inf. Theor.* 28, 2 (1982), 129-137.
- [108] Ana Carolina Lorena and Andre C. P. L. F. De Carvalho. 2008. Evolutionary tuning of SVM parameter values in multiclass problems. *Neurocomputing* 71, 16-18 (2008), 3326-3334.
- [109] José Antonio Lozano and Pedro Larranaga. 1999. Applying genetic algorithms to search for the best hierarchical clustering of a dataset. *Pattern Recog. Lett.* 20, 9 (1999), 911-918.
- [110] Nannan Lu, Shingo Mabu, Tuo Wang, and Kotaro Hirasawa. 2013. An efficient class association rule-pruning method for unified intrusion detection system using genetic algorithm. *IEEE Trans. Electric. Electron. Eng.* 8, 2 (2013), 164-172.
- [111] Zhichao Lu, Kalyanmoy Deb, Erik Goodman, Wolfgang Banzhaf, and Vishnu Naresh Boddeti. 2020. Nsganetv2: Evolutionary multi-objective surrogate-assisted neural architecture search. In Proceedings of the European Conference on Computer Vision. Springer, 35-51.
- [112] Zhichao Lu, Ian Whalen, Vishnu Boddeti, Yashesh Dhebar, Kalyanmoy Deb, Erik Goodman, and Wolfgang Banzhaf. 2019. NSGA-Net: Neural architecture search using multi-objective genetic algorithm. In Proceedings of the Genetic and Evolutionary Computation Conference. 419-427.
- [113] José María Luna, Cristóbal Romero, José Raúl Romero, and Sebastián Ventura. 2015. An evolutionary algorithm for the discovery of rare class association rules in learning management systems. *Appl. Intell.* 42, 3 (2015), 501-513.
- [114] Patrick C. H. Ma, Keith C. C. Chan, Xin Yao, and David K. Y. Chiu. 2006. An evolutionary clustering algorithm for gene expression microarray data analysis. *IEEE Trans. Evolut. Comput.* 10, 3 (2006), 296-314.
- [115] Sai Ma and Fulei Chu. 2019. Ensemble deep learning-based fault diagnosis of rotor bearing systems. *Comput. Industr.* 105 (2019), 143-152.
- [116] Shingo Mabu, Ci Chen, Nannan Lu, Kaoru Shimada, and Kotaro Hirasawa. 2010. An intrusion-detection model based on fuzzy class-association-rule mining using genetic network programming. *IEEE Trans. Syst., Man, Cyber., Part C (Appl. Rev.)* 41, 1 (2010), 130-139.
- [117] Arif Jamal Malik and Farrukh Aslam Khan. 2018. A hybrid technique using binary particle swarm optimization and decision tree pruning for network intrusion detection. *Cluster Comput.* 21, 1 (2018), 667-680.

- [118] Veenu Mangat and Renu Vig. 2014. Novel associative classifier based on dynamic adaptive PSO: Application to determining candidates for thoracic surgery. *Exp. Syst. Applic.* 41, 18 (2014), 8234-8244.
- [119] Vittorio Maniezzo, Luca Maria Gambardella, and Fabio De Luigi. 2004. Ant colony optimization. In *New Optimization Techniques in Engineering*. Springer, 101-121.
- [120] María Martínez-Ballesteros, Francisco Martínez-Alvarez, A. Troncoso, and José C. Riquelme. 2009. Quantitative association rules applied to climatological time series forecasting. In *Proceedings of the International Conference on Intelligent Data Engineering and Automated Learning*. Springer, 284-291.
- [121] María Martínez-Ballesteros, Francisco Martínez-Alvarez, Alicia Troncoso, and José C Riquelme. 2011. An evolutionary algorithm to discover quantitative association rules in multidimensional time series. *Soft Comput.* 15, 10 (2011), 2065.
- [122] María Martínez-Ballesteros, A. Troncoso, Francisco Martínez-Alvarez, and José C. Riquelme. 2010. Mining quantitative association rules based on evolutionary computation and its application to atmospheric pollution. *Integr. Comput. -aided Eng.* 17, 3 (2010), 227-242.
- [123] Jacinto Mata, José-Luis Alvarez, and José-Cristobal Riquelme. 2002. Discovering numeric association rules via evolutionary algorithm. In *Pacific-Asia Conference on Knowledge Discovery and Data Mining*. Springer, 40-51.
- [124] Daniel Mateos-Garcia, Jorge Garcia-Gutiérrez, and José C. Riquelme-Santos. 2016. An evolutionary voting for knearest neighbours. *Exp. Syst. Applic.* 43 (2016), 9-14.
- [125] Ron Meir and Gunnar Rätsch. 2003. An introduction to boosting and leveraging. In *Advanced Lectures on Machine Learning*. Springer, 118-183.
- [126] Jan Hendrik Metzen, Mark Edgington, Yohannes Kassahun, and Frank Kirchner. 2008. Analysis of an evolutionary reinforcement learning method in a multiagent domain. In *Proceedings of the 7th International Joint Conference on Autonomous Agents and Multiagent Systems-Volume 1*. Citeseer, 291-298.
- [127] Risto Miikkulainen, Jason Liang, Elliot Meyerson, Aditya Rawal, Daniel Fink, Olivier Francon, Bala Raju, HormozShahrzad, Arshak Navruzian, Nigel Duffy, et al. 2019. Evolving deep neural networks. In *Artificial Intelligence in the Age of Neural Networks and Brain Computing*. Elsevier, 293-312.
- [128] Volodymyr Mnih, Koray Kavukcuoglu, David Silver, Andrei A. Rusu, Joel Veness, Marc G. Bellemare, Alex Graves, Martin Riedmiller, Andreas K. Fiedjeland, Georg Ostrovski, et al. 2015. Human-level control through deep reinforcement learning. *Nature* 518, 7540 (2015), 529.
- [129] Anirban Mukhopadhyay, Ujjwal Maulik, and Sanghamitra Bandyopadhyay. 2015. A survey of multiobjective evolutionary clustering. *ACM Comput. Surv.* 47, 4 (2015), 1-46.
- [130] Anirban Mukhopadhyay, Ujjwal Maulik, Sanghamitra Bandyopadhyay, and Carlos Artemio Coello Coello. 2013. A survey of multiobjective evolutionary algorithms for data mining: Part I. *IEEE Trans. Evolut. Comput.* 18, 1 (2013), 4-19.
- [131] Anirban Mukhopadhyay, Ujjwal Maulik, Sanghamitra Bandyopadhyay, and Carlos A. Coello Coello. 2013. Survey of multiobjective evolutionary algorithms for data mining: Part II. *IEEE Trans. Evolut. Comput.* 18, 1 (2013), 20-35.
- [132] S. R. Nanda, Biswajit Mahanty, and M. K. Tiwari. 2010. Clustering Indian stock market data for portfolio management. *Exp. Syst. Applic.* 37, 12 (2010), 8793-8798.
- [133] Satyasai Jagannath Nanda and Ganapati Panda. 2014. A survey on nature inspired metaheuristic algorithms for partitional clustering. *Swarm Evolut. Comput.* 16 (2014), 1-18.

- [134] NVIDIA. 2017. The world's most efficient supercomputer for AI and deep learning. Retrieved from <http://images.nvidia.com/content/pdf/infographic/dgx-saturnv-infographic.pdf>.
- [135] Luiz S. Oliveira, Robert Sabourin, Flávio Bortolozzi, and Ching Y. Suen. 2002. Feature selection using multi-objective genetic algorithms for handwritten digit recognition. In *Object Recognition Supported by User Interaction for Service Robots*, Vol. 1. IEEE, 568-571.
- [136] Aytuğ Onan, Serdar Korukoğlu, and Hasan Bulut. 2017. A hybrid ensemble pruning approach based on consensusclustering and multi-objective evolutionary algorithm for sentiment classification. *Inf. Process. Manag.* 53, 4 (2017), 814-833.
- [137] Fernando E. B. Otero, Alex A. Freitas, and Colin G. Johnson. 2012. Inducing decision trees with an ant colony optimization algorithm. *Appl. Soft Comput.* 12, 11 (2012), 3615-3626.
- [138] Rafael S. Parpinelli and Heitor S. Lopes. 2011. New inspirations in swarm intelligence: A survey. *Int. J. Bio-insp.Comput.* 3, 1 (2011), 1-16.
- [139] Russel Pears and Yun Sing Koh. 2011. Weighted association rule mining using particle swarm optimization. In *PacificAsia Conference on Knowledge Discovery and Data Mining*. Springer, 327-338.
- [140] Abdolrahman Peimankar, Stephen John Weddell, Thahirah Jalal, and Andrew Craig Laphorn. 2018. Multi-objective ensemble forecasting with an application to power transformers. *Appl. Soft Comput.* 68 (2018), 233-248.
- [141] Lu Peng, Shan Liu, Rui Liu, and Lin Wang. 2018. Effective long short-term memory with differential evolution algorithm for electricity price prediction. *Energy* 162 (2018), 1301-1314.
- [142] Pyari Mohan Pradhan and Ganapati Panda. 2012. Connectivity constrained wireless sensor deployment using multiobjective evolutionary algorithms and fuzzy decision making. *Ad Hoc Netw.* 10, 6 (2012), 1134-1145.
- [143] Esmat Rashedi, Elaheh Rashedi, and Hossein Nezamabadi-pour. 2018. A comprehensive survey on gravitationalsearch algorithm. *Swarm Evolut. Comput.* 41 (2018), 141-158.
- [144] Esteban Real, Chen Liang, David So, and Quoc Le. 2020. AutoML-zero: Evolving machine learning algorithms from scratch. In *Proceedings of the International Conference on Machine Learning*. PMLR, 8007-8019.
- [145] Victor Henrique Alves Ribeiro and Gilberto Reynoso-Meza. 2019. A holistic multi-objective optimization design procedure for ensemble member generation and selection. *Appl. Soft Comput.* 83 (2019), 105664.
- [146] Victor Henrique Alves Ribeiro and Gilberto Reynoso-Meza. 2020. Ensemble learning by means of a multi-objective optimization design approach for dealing with imbalanced data sets. *Exp. Syst. Applic.* 147 (2020), 113232.
- [147] Cristóbal Romero, Amelia Zafra, Jose Maria Luna, and Sebastián Ventura. 2013. Association rule mining using genetic programming to provide feedback to instructors from multiple-choice quiz data. *Exp. Syst.* 30, 2 (2013), 162-172.
- [148] Hussein Samma, Chee Peng Lim, and Junita Mohamad Saleh. 2016. A new reinforcement learning-based memetic particle swarm optimizer. *Appl. Soft Comput.* 43 (2016), 276-297.
- [149] Manish Sarkar, B. Yegnanarayana, and Deepak Khemani. 1997. A clustering algorithm using an evolutionary programming-based approach. *Pattern Recog. Lett.* 18, 10 (1997), 975-986.
- [150] Mansour Sheikhan and Maryam Sharifi Rad. 2013. Using particle swarm optimization in fuzzy association rules-based feature selection and fuzzy ARTMAP-based attack recognition. *Secur. Commun. Netw.* 6, 7 (2013), 797-811.

- [151] Christopher Smith and Yaochu Jin. 2014. Evolutionary multi-objective generation of recurrent neural network ensembles for time series prediction. *Neurocomputing* 143 (2014), 302-311.
- [152] Qin Song, Yu-Jun Zheng, Yu Xue, Wei-Guo Sheng, and Mei-Rong Zhao. 2017. An evolutionary deep neural network for predicting morbidity of gastrointestinal infections by food contamination. *Neurocomputing* 226 (2017), 16-22.
- [153] Pedro Sousa, Paulo Cortez, Rui Vaz, Miguel Rocha, and Miguel Rio. 2013. Email spam detection: A symbiotic feature selection approach fostered by evolutionary computation. *Int. J. Inf. Technol. Decis. Mak.* 12, 04 (2013), 863-884.
- [154] Andreas Stafylopatis and Konstantinos Blekas. 1998. Autonomous vehicle navigation using evolutionary reinforcement learning. *Eur. J. Oper. Res.* 108, 2 (1998), 306-318.
- [155] Kenneth O. Stanley, Jeff Clune, Joel Lehman, and Risto Miikkulainen. 2019. Designing neural networks through neuroevolution. *Nat. Mach. Intell.* 1, 1 (2019), 24-35.
- [156] Kenneth O. Stanley, David B. D'Ambrosio, and Jason Gauci. 2009. A hypercube-based encoding for evolving largescale neural networks. *Artif. Life* 15, 2 (2009), 185-212.
- [157] Kenneth O. Stanley and Risto Miikkulainen. 2002. Evolving neural networks through augmenting topologies. *Evolut.Comput.* 10, 2 (2002), 99-127.
- [158] Felipe Petroski Such, Vashisht Madhavan, Edoardo Conti, Joel Lehman, Kenneth O. Stanley, and Jeff Clune. 2017. Deep neuroevolution: Genetic algorithms are a competitive alternative for training deep neural networks for reinforcement learning. *arXiv preprint arXiv:1712.06567* (2017).
- [159] Thorsten Suttrop and Christian Igel. 2006. Multi-objective optimization of support vector machines. In *Multi-objective Machine Learning*. Springer, 199-220.
- [160] Amirhessam Tahmassebi and Amir H. Gandomi. 2018. Building energy consumption forecast using multi-objective genetic programming. *Measurement* 118 (2018), 164-171.
- [161] Amirhessam Tahmassebi and Amir H. Gandomi. 2018. Genetic programming based on error decomposition: A big data approach. In *Genetic Programming Theory and Practice XV*. Springer, 135-147.
- [162] Amirhessam Tahmassebi, Amir H. Gandomi, and Anke Meyer-Baese. 2018. An evolutionary online framework for MOOC performance using EEG data. In *Proceedings of the IEEE Congress on Evolutionary Computation (CEC)*. IEEE, 1-8.
- [163] Amirhessam Tahmassebi, Amir H. Gandomi, and Anke Meyer-Baese. 2018. A Pareto front based evolutionary model for airfoil self-noise prediction. In *Proceedings of the IEEE Congress on Evolutionary Computation (CEC)*. IEEE, 1-8.
- [164] Amirhessam Tahmassebi, Behshad Mohebbi, Anke Meyer-Baese, and Amir H. Gandomi. 2020. Multiobjective genetic programming for reinforced concrete beam modeling. *Appl. AI Lett.* 1, 1 (2020), e9.
- [165] Amirhessam Tahmassebi and Trace Smith. 2021. SlickML: Slick Machine Learning in Python. Retrieved from <https://github.com/slickml/slick-ml>.
- [166] Pınar Tapkan, Lale Özbakır, Sinem Kulluk, and Adil Baykasoğlu. 2016. A cost-sensitive classification algorithm: BEEMiner. *Knowl.-based Syst.* 95 (2016), 99-113.
- [167] Kshitij Tayal and Vaddamani Ravi. 2016. Particle swarm optimization trained class association rule mining: Application to phishing detection. In *Proceedings of the International Conference on Informatics and Analytics*. 1-8.

- [168] Akbar Telikani and Amir H. Gandomi. 2019. Cost-sensitive stacked auto-encoders for intrusion detection in the Internet of Things. *Internet of Things* 14 (2019), 100122.
- [169] Akbar Telikani, Amir H. Gandomi, and Asadollah Shahbahrami. 2020. A survey of evolutionary computation for association rule mining. *Inf. Sci.* 524 (2020).
- [170] Akbar Telikani and Asadollah Shahbahrami. 2018. Data sanitization in association rule mining: An analytical review. *Exp. Syst. Applic.* 96 (2018), 406-426.
- [171] Cuong To and Tuan D. Pham. 2009. Analysis of cardiac imaging data using decision tree based parallel genetic programming. In *Proceedings of the 6th International Symposium on Image and Signal Processing and Analysis*. IEEE, 317-320.
- [172] Binh Tran, Bing Xue, and Mengjie Zhang. 2016. Genetic programming for feature construction and selection in classification on high-dimensional data. *Mem. Comput.* 8, 1 (2016), 3-15.
- [173] Isaac Triguero, Mikel Galar, Humberto Bustince, and Francisco Herrera. 2017. A first attempt on global evolutionary undersampling for imbalanced big data. In *Proceedings of the IEEE Congress on Evolutionary Computation (CEC)*. IEEE, 2054-2061.
- [174] Isaac Triguero, Salvador Garcia, and Francisco Herrera. 2011. Differential evolution for optimizing the positioning of prototypes in nearest neighbor classification. *Pattern Recog.* 44, 4 (2011), 901-916.
- [175] Alan M. Turing. 1950. Computing machinery and intelligence. *Mind* 59, 236 (1950), 433-460.
- [176] Shweta Tyagi and Kamal K. Bharadwaj. 2013. Enhancing collaborative filtering recommendations by utilizing multiobjective particle swarm optimization embedded association rule mining. *Swarm Evolut. Comput.* 13 (2013), 1-12.
- [177] Joost Verbraeken, Matthijs Wolting, Jonathan Katzy, Jeroen Kloppenburg, Tim Verbelen, and Jan S. Rellermeyer. 2020. A survey on distributed machine learning. *ACM Comput. Surv.* 53, 2 (2020), 1-33.
- [178] Leandro D. Vignolo, Diego H. Milone, and Jacob Scharcanski. 2013. Feature selection for face recognition based on multi-objective evolutionary wrappers. *Exp. Syst. Applic.* 40, 13 (2013), 5077-5084.
- [179] Wengdong Wang and Susan M. Bridges. 2000. Genetic algorithm optimization of membership functions for mining fuzzy association rules. *Depart. Comput. Sci. Mississ. State Univ.* 2 (2000).
- [180] Feng Wen, Guo Zhang, Lingfeng Sun, Xingqiao Wang, and Xiaowei Xu. 2019. A hybrid temporal association rules mining method for traffic congestion prediction. *Comput. Industr. Eng.* 130 (2019), 779-787.
- [181] Chun-Hui Wu, Ta-Cheng Chen, Yi-Chih Hsieh, and Huei-Ling Tsao. 2019. A hybrid rule mining approach for cardiovascular disease detection in traditional Chinese medicine. *J. Intell. Fuzz. Syst.* 36, 2 (2019), 861-870.
- [182] Bing Xue, Mengjie Zhang, and Will N. Browne. 2014. Particle swarm optimisation for feature selection in classification: Novel initialisation and updating mechanisms. *Appl. Soft Comput.* 18 (2014), 261-276.
- [183] Bing Xue, Mengjie Zhang, Will N. Browne, and Xin Yao. 2015. A survey on evolutionary computation approaches to feature selection. *IEEE Trans. Evolut. Comput.* 20, 4 (2015), 606-626.
- [184] Dongdong Yang, Licheng Jiao, Maoguo Gong, and Fang Liu. 2011. Artificial immune multi-objective SAR image segmentation with fused complementary features. *Inf. Sci.* 181, 13 (2011), 2797-2812.
- [185] Xin Yao. 1993. A review of evolutionary artificial neural networks. *Int. J. Intell. Syst.* 8, 4 (1993), 539-567.

- [186] Jianbo Yu, Lifeng Xi, and Shijin Wang. 2007. An improved particle swarm optimization for evolving feedforward artificial neural networks. *Neural Process. Lett.* 26, 3 (2007), 217-231.
- [187] Mohammed Javeed Zaki. 2000. Scalable algorithms for association mining. *IEEE Trans. Knowl. Data Eng.* 12, 3 (2000), 372-390.
- [188] Maider Zamalloa, Germán Bordel, Luis Javier Rodriguez, and Mikel Penagarikano. 2006. Feature selection based on genetic algorithms for speaker recognition. In *Proceedings of the IEEE Odyssey-The Speaker and Language Recognition Workshop*. IEEE, 1-8.
- [189] Chong Zhang, Kay Chen Tan, Haizhou Li, and Geok Soon Hong. 2018. A cost-sensitive deep belief network for imbalanced classification. *IEEE Trans. Neural Netw. Learn. Syst.* 30, 1 (2018), 109-122.
- [190] Lei Zhang, Guanglong Fu, Fan Cheng, Jianfeng Qiu, and Yansen Su. 2018. A multi-objective evolutionary approach for mining frequent and high utility itemsets. *Appl. Soft Comput.* 62 (2018), 974-986.
- [191] Zaifang Zhang, Nana Chai, Egon Ostrosi, and Yuliang Shang. 2019. Extraction of association rules in the schematic design of product service system based on Pareto-MODGDFA. *Comput. Industr. Eng.* 129 (2019), 392-403.
- [192] Changjiu Zhou. 2002. Robot learning with GA-based fuzzy reinforcement learning agents. *Inf. Sci.* 145, 1-2 (2002), 45-68.

Received December 2020; revised April 2021; accepted May 2021