



计算机科学

COMPUTER SCIENCE

一种基于超网络的多目标回归方法

孙开伟, 郭豪, 曾雅苑, 方阳, 刘期烈

引用本文

孙开伟, 郭豪, 曾雅苑, 方阳, 刘期烈. 一种基于超网络的多目标回归方法[J]. 计算机科学, 2022, 49(11A): 211000205-9.

SUN Kai-wei, GUO Hao, ZENG Ya-yuan, FANG Yang, LIU Qi-lie. [Multi-target Regression Method Based on Hypernetwork](#) [J]. Computer Science, 2022, 49(11A): 211000205-9.

相似文章推荐 (请使用火狐或 IE 浏览器查看文章)

Similar articles recommended (Please use Firefox or IE to view the article)

[基于机器学习的剩余使用寿命预测实证研究](#)

Empirical Research on Remaining Useful Life Prediction Based on Machine Learning
计算机科学, 2022, 49(11A): 211100285-9. <https://doi.org/10.11896/jsjcx.211100285>

[基于三支聚类的云任务优化调度](#)

Optimal Scheduling of Cloud Task Based on Three-way Clustering
计算机科学, 2022, 49(11A): 211100139-7. <https://doi.org/10.11896/jsjcx.211100139>

[对抗性网络流量的生成与应用综述](#)

Generation and Application of Adversarial Network Traffic:A Survey
计算机科学, 2022, 49(11A): 211000039-11. <https://doi.org/10.11896/jsjcx.211000039>

[基于改进的SLIC和聚类算法结合的高分辨率遥感海冰图像分割](#)

High-resolution Remote Sensing Sea Ice Image Segmentation Based on Combination of ImprovedSLIC Algorithm and Clustering Algorithm
计算机科学, 2022, 49(11A): 211200100-6. <https://doi.org/10.11896/jsjcx.211200100>

[基于t-SVD的结构保持多视图子空间聚类](#)

Structure Preserved Multi-view Subspace Clustering Based on t-SVD
计算机科学, 2022, 49(11A): 210800215-6. <https://doi.org/10.11896/jsjcx.210800215>

一种基于超网络的多目标回归方法

孙开伟¹ 郭豪¹ 曾雅苑¹ 方阳¹ 刘期烈²

¹ 重庆邮电大学数据工程与可视计算重点实验室 重庆 400065

² 重庆邮电大学通信与信息工程学院 重庆 400065

摘要 多目标回归(Multi-target Regression, MTR)是一种同时预测多个相互关联的连续型输出目标的机器学习问题。在多目标回归中,多个输出目标共享同一个特征表示,其主要挑战在于如何有效地发掘和利用输出目标之间的关联,以提高所有输出目标的预测准确性。文中提出了一种基于超网络的多目标回归方法(Multi-target Regression Method based on Hypernetwork, MTR-HN)。首先采用 k -means 算法对每个连续型输出目标进行一维聚类,然后根据聚类结果将多目标回归问题转化成多类别多标签分类问题,最后采用超网络模型对多类别多标签分类问题进行建模,构建最终的多目标回归预测模型。MTR-HN 方法的优点在于:1)对输出空间离散化,能够降低模型过拟合的风险;2)采用超网络模型,能更有效地对输出目标之间的关联进行建模。在 18 个多目标回归数据集上进行的对比实验表明,文中提出的 MTR-HN 方法能够取得比现有方法更高的预测准确性。

关键词: 机器学习; 多目标回归; 聚类; 多标签分类; 超网络

中图分类号 TP181

Multi-target Regression Method Based on Hypernetwork

SUN Kai-wei¹, GUO Hao¹, ZENG Ya-yuan¹, FANG Yang¹ and LIU Qi-lie²

¹ Key Laboratory of Data Engineering and Visual Computing, Chongqing University of Posts and Telecommunications, Chongqing 400065, China

² School of Communication and Information Engineering, Chongqing University of Posts and Telecommunications, Chongqing 400065, China

Abstract Multi-target regression(MTR) is a kind of machine learning problem which predicts multiple relevant continuous output targets simultaneously. In MTR, multiple output targets share the same input feature representation, the main challenge of MTR lies in how to effectively explore and utilize the correlations among multiple output targets to improve the prediction accuracy of all output targets. In this paper, a multi-target regression method based on hypernetwork(MTR-HN) is proposed. First, the k -means method is applied to each output target to divide it into multiple clusters. Then, according to the clustering results, MTR problem is transformed into a multi-class multi-label classification problem. Finally, hypernetwork model is utilized to model the multi-class multi-label classification problem, and the final prediction model for MTR is built based on hypernetwork. The main merits of MTR-HN lie in: 1) discretizing the output space, can reduce the risk of overfitting; 2) hypernetwork can model the inter-target correlations more effectively. Comparative experiments on 18 multi-target regression datasets show that the proposed MTR-HN achieves better prediction performance than existing state-of-the-art multi-target regression methods.

Keywords Machine learning, Multi-target regression, Clustering, Multi-label classification, Hypernetwork

1 引言

在机器学习领域,回归分析是一类被广泛研究的学习框架。传统的回归分析方法通过一组输入特征变量预测单个连续型变量的输出。与传统的回归问题不同,多目标回归^[1-2]通过一组相同的输入特征变量同时预测多个连续型的输出目标。在多目标回归中,各输出目标之间并不是相互独立的,

可能是相互关联的。由于其能够同时对多个相互关联的输出目标进行建模和预测,多目标回归已经在数据挖掘^[3]、经济学^[4]、生态建模^[5]、计算机视觉^[6]以及医学图像分析^[7]等领域得到了广泛的应用。

在多目标回归中,输出目标并不是相互独立的,而是相互关联的,输出目标之间的关联关系为构建预测模型提供了额外的有价值的信息。因此,如何有效地发掘和利用输出目标

基金项目:国家自然科学基金(61806033);重庆市自然科学基金面上项目(cstc2019jcyj-msxmX0021);重庆市科技创新领军人才支持计划(CSTCCXLJRC201908);重庆市自然基金重点项目(cstc2019jcyj-zdxm0008);重庆市教委重点项目(KJZD-K201900605);重庆市教委“成渝地区双城经济圈建设”科技创新项目(KJXCZD2020027)

This work was supported by the National Natural Science Foundation of China (61806033), Natural Science Foundation of Chongqing (cstc2019jcyj-msxmX0021), Chongqing Science and Technology Innovation Leading Talent Support Program (CSTCCXLJRC201908), Basic and Advanced Research Projects of CSTC (cstc2019jcyj-zdxmX0008), Science and Technology Research Program of Chongqing Municipal Education Commission (KJZD-K201900605) and Science Innovation Program of Chengdu - Chongqing Economic Circle in Southwest China (KJXCZD2020027).

通信作者:孙开伟(sunkw@cqupt.edu.cn)

之间的关联以提高多目标回归的整体预测性能,成为多目标回归面临的最大的挑战^[8]。为了发掘输出目标之间的关联,一些多目标回归方法^[9]将目标变量的预测输出作为元输入,扩展原始输入空间,通过将其他目标变量的预测输出作为预测模型的输入变量,达到利用输出目标之间的关联提升预测性能的目的。此外,一些多目标回归方法^[8,10]通过在回归模型的参数矩阵上添加正则项,使得模型的参数矩阵具有稀疏性和低秩性,从而使得相互关联的输出目标在预测模型上具有相似的模型参数。现有的多目标回归方法都是在连续型的输出目标空间中处理输出目标之间的关联问题。然而,输出目标的取值分布在实数域上,容易受到噪声的影响,很可能导致预测模型过拟合。多标签分类(Multi-label Classification, MLC)^[11-12]是一种与多目标回归密切相关的多输出分类问题,两者的区别在于多目标回归中输出目标的取值是连续的,而多标签分类中输出标签的取值是离散的;两者的核心问题都在于如何有效利用输出变量之间的关联提升整体的预测性能。由于两者在学习问题本质上具有相似性,一些多标签分类方法在稍加改进后就可以应用于多目标回归问题。与多目标回归相比,多标签分类对离散输出变量之间的联合分布进行建模更直接、更有效。因此,通过对多目标回归的连续型输出变量进行离散化处理,不仅可以降低模型过拟合的风险,还可以更有效地对多个输出目标之间的关联关系进行建模。Spyromitros-Xioufifis 等^[13]提出了基于输出空间离散化的多目标回归方法(Multi-target Regression via Quantization, MRQ),将多目标回归问题转化为多分类问题,并且指出在离散的输出空间上比在连续的输出空间上能更有效地表示输出目标之间的关联性。

超网络(Hypernetwork)^[14]是 Zhang 等在生物分子网络的启发下,根据认知计算的相关理论,以超图为表现形式提出的一种用于认知学习和记忆的概率图模型。超网络具有结构灵活、分类模型简单、能够有效表示对象之间高阶关联等优点。在多标签分类领域,超网络模型已经取得了成功的应用。Sun 等^[15]将传统的超网络模型进行改进,提出了一种多标签演化超网络学习方法。在文献^[16]中,Sun 等提出了一种基于协同演化的多标签演化超网络学习方法,该方法能够对标签之间的高阶关联进行有效建模,并且对标签数量具有很好的扩展性。此外,针对多标签分类问题中的类别标记分布不平衡的问题,Sun 等^[17]提出了一种基于两阶段学习的多标签演化超网络模型,通过有效利用标签之间的关联,来提升不平衡标签的分类准确性。

文中提出了一种基于超网络的多目标回归方法(MTR-HN)。首先采用 k-means 算法对每个连续型输出目标进行一维聚类;然后根据聚类结果将多目标回归问题转化成多类别多标签分类问题;最后采用超网络模型对多类别多标签分类问题进行建模,构建最终的多目标回归预测模型。与现有的多目标回归方法相比,MTR-HN 具有以下优点:1)通过对连续型输出目标进行离散化,把多目标回归问题转换为多类别多标签分类问题,能够在离散空间对输出目标之间的联合分布进行建模,提高多目标回归模型的鲁棒性,降低过拟合的风险;2)采用超网络模型处理多类别多标签分类问题,能够有效地表示离散输出目标之间的高阶关联,提高预测的准确性;

3)超网络模型的学习结果表现为带权值的超图,模型具有很好的可解释性。

本文第 2 节介绍多目标回归的相关工作;第 3 节说明基于超网络模型的多目标回归方法的具体细节;第 4 节介绍在多目标回归数据集上的对比实验研究,并对实验结果进行分析;最后总结全文。

2 相关工作

多目标回归是多输出学习^[18]框架中的一个研究热点,它与多标签分类问题有着密切的关系,同时也被认为是多任务学习(Multi-task Learning, MTL)^[19]的一个特例,三者之间相互联系、相互促进。针对多标签分类和多任务学习开展的研究,可以为解决多目标回归问题提供宝贵的经验。本节从多标签分类、多任务学习和多目标回归 3 个方面分别介绍与多目标回归相关的研究工作。

针对多标签分类问题,目前已经涌现大量的学习算法。总体上来看,现有的多标签学习算法可以大致分为“问题转化”方法和“算法适应”方法两类^[11-12]。问题转化方法的基本思想是通过对多标签训练样本进行处理,将多标签学习问题转化成其他已知的机器学习问题进行求解。二元关系法(Binary Relevance, BR)^[20]是一种典型的问题转化多标签学习方法,它将每个类别标签的分类看作一个独立的二分类问题,为每个类别标签训练一个独立的分类器。基准标记排序(Calibrated Label Ranking, CLR)^[21]方法将多标签学习问题转化成“标签排序(Label Ranking, LR)”问题进行求解。标签幂集法(Label Powerset, LP)^[20]将每个样本的类别标签集合看成一个单标签,训练数据中所有不同的类别标签集合组成一个新的标签集合,通过这种方式将多标签分类问题转化成新标签集合下的单标记多分类问题。随机 k 标记子集(Random k-labelsets, RAKEL)^[22]是 LP 方法的一种集成学习方法,该方法组合多个 LP 分类器,每个 LP 分类器都是在随机生成且互不相同的小规模标签子集上训练而成。分类器链集成方法(Ensemble of Classifier Chains, ECC)^[23]采用分类器链的方式将多个二分类器连成一条链,训练样本每经过一个二分类器,就将其预测结果添加到样本特征向量中,继续下一个二分类器的训练,并组合多条随机产生的不同标记序列的分类器链,以减轻不同标签序列给分类结果带来的不利影响。算法适应方法的主要思想是对常用的传统机器学习方法进行改进,将其直接应用于多标签数据学习。其中较为典型的有多标记 k 近邻方法(Multi-label kNN, ML-kNN)^[24]、反向传播多标签学习方法(Back Propagation for Multi-label Learning, BP-MLL)^[25]、排序支持向量机(Rank-support Vector Machine, Rank-SVM)^[26]等。

多任务学习同样与多目标回归有着密切的联系。实际上,多目标回归通常被认为是多任务学习的一个特例。多任务学习主要研究如何利用多个学习任务之间的关联提升所有学习任务的泛化性能。在多任务学习中,多个学习任务可以具有不同的输入特征表示,也可以共享一个相同的输入特征表示。当不同学习任务共享同一个特征表示的时候,多任务学习实际上就演变为多标签分类或者多目标回归问题。与多标签分类和多目标回归一样,多任务学习的核心问题也在于

3.1 问题描述

令 $X = \{x_1, x_2, \dots, x_d\}$ 表示 d 个输入特征的集合, $Y = \{y_1, y_2, \dots, y_m\}$ 表示 m 个输出目标的集合。在多目标回归中, 一个训练样本 (x_i, y_i) 由两部分组成, 即一个 d 维的特征向量 $x_i = [x_{i1}, x_{i2}, \dots, x_{id}]$ 和一个 m 维的目标输出向量 $y_i = [y_{i1}, y_{i2}, \dots, y_{im}]$ 。给定一个包含 N 个训练样本的数据集 $D = \{(x_i, y_i) \mid 1 \leq i \leq N\}$, 令 $X \in R^{N \times d}$ 表示由 N 个训练样本组成的输入特征矩阵, $Y \in R^{N \times m}$ 表示由 N 个训练样本组成的输出目标矩阵, 多目标回归方法的任务就是从训练集 D 中学习得到一个预测模型 $h: X \rightarrow Y$ 将输入特征矩阵映射到输出目标矩阵。对于一个由 d 维特征向量表示的数据样本 x , 预测模型 h 输出一个 m 维的向量 y 作为多目标预测输出。

3.2 输出目标离散化

在多目标回归问题中, 多个输出目标共享一个输入特征表示, 每个训练样本同时具有多个连续型的输出目标, 对应一个连续型的目标输出向量。为了更有效地表示多目标回归中各输出目标之间的关联, 采用向量量化 (Vector Quantization, VQ) 方法将连续型的输出目标进行离散化处理。向量量化方法起源于信息论, 最初被用于模拟信号-数字信号的转化以及数据压缩。向量量化的目标就是在保持最小损失的条件下, 将高维实数向量用一组离散的取值来近似表示。具体来讲, 对于任意输出目标 $y_i \in Y$, 首先将训练集中所有样本在输出目标 y_i 上的取值组成一个 N 维的向量 $t = [y_{1i}, y_{2i}, \dots, y_{Ni}]$, 即向量 t 是由训练集的输出目标矩阵 Y 的第 i 列元素组成, 其中 N 表示训练样本的个数。然后采用 k-means 算法对向量 t 进行一维聚类, 得到 s 个聚类中心 $C_i = \{c_{ij}\}_{j=1}^s$ 和 s 个簇标签 $l_i = \{0, 1, \dots, s-1\}$ 。在聚类之后, 输出目标 y_i 就可以通过 s 个离散的值来近似表示, 并被离散化为 s 个类别。为了给每个输出目标确定一个合适的聚类数 s , 让 s 从 1 开始进行迭代, 并按照式 (1) 计算聚类中心与训练样本的平均离差 $disp(Y, C_i)$, 当相邻两次迭代的平均离差之差的绝对值小于给定的阈值时, 停止迭代, 并取当前迭代的 s 值作为最终的聚类个数。输出目标离散化的过程如算法 1 所示。

$$disp(Y, C_i) = \frac{1}{N} \sum_{j=1}^N \sum_{n=1}^s (y_{ji} - c_{in})^2 \quad (1)$$

算法 1 输出目标离散化

输入: 连续型的目标输出矩阵 Y , 其中一行表示一个训练样本的 m 个目标输出值; 最大聚类个数 S

输出: 离散化的目标输出矩阵 L , 其中一行表示一个训练样本离散化之后的目标输出; 输出目标对应的聚类中心组成的集合 $C = \{C_1, C_2, \dots, C_m\}$

1. 为每个输出目标确定聚类个数

1.1. for $i=1$ to m

1.2. 将矩阵 Y 的第 i 列组成一个 N 维的向量 t , 初始化数组 $cluster_num$ 为零, 用于保存输出目标的最优聚类个数

1.3. for $j=1$ to S

1.4. 采用 k-means 算法将向量 t 聚类为 j 个簇, 簇中心分别为 $C_i = \{C_{i1}, C_{i2}, \dots, C_{ij}\}$, 并根据式 (1) 计算本次迭代的平均离差 $disp_i$

1.5. if $|disp_i - disp_{i-1}| \leq 5 \times 10^{-3}$

1.6. $cluster_num[i] = j$

1.7. break

2. 采用 k-means 算法将输出目标进行离散化

2.1. 初始化矩阵 $L^{N \times m}$ 为零; 初始化聚类中心集合 C 为空

2.2. for $i=1$ to m

2.3. 将矩阵 Y 的第 i 列组成向量 t

2.4. 采用 k-means 算法将向量 t 聚类为 $cluster_num[i]$ 个簇, 得到聚类中心 C_i 和聚类标签 l_i

2.5. 将矩阵 L 的第 i 列 $L[:, i]$ 设置为 l_i

2.6. 将聚类中心 C_i 加入到集合 C 中

3. 返回离散化的目标输出矩阵 L 和聚类中心 C

经过离散化处理, 原始连续型输出被转化为离散型的输出, 每个训练样本同样具有 m 个输出目标, 但是每个输出目标被几个离散的取值替代, 每个离散的取值对应输出目标中的一个类别。由此, 多目标回归问题被转化为一个多类别多标签分类问题。与原始的连续型输出相比, 离散化之后的输出由于通过多个离散取值进行近似表示, 不仅提高了对输出噪声的鲁棒性, 而且能够更有效地表示输出目标之间的联合概率分布。

由于在连续型输出空间上发掘输出目标之间关联的难度较大, 且容易过拟合, 一些对输出空间进行离散化处理的多目标回归方法^[13]被提出。这些方法将多个输出目标组成的输出向量进行聚类, 把多目标回归问题转化为多分类问题。与这类方法使用的策略不同, 本文将每个输出目标独立地进行离散化处理, 聚类为多个类别, 把多目标回归问题转化为多类别多标签分类问题, 能够更真实地反映各个输出目标的分布特征。

3.3 超网络模型

在输出目标离散化之后, 多目标回归问题被转化为多类别多标签分类问题。由于超网络模型在多标签分类问题中已经取得了成功的应用, 本文采用超网络模型对多类别多标签分类问题进行建模, 构建多目标回归的预测模型。超网络模型是一种基于超图的概率图模型, 可以通过一个三元组 $H = (V, E, W)$ 来形式化表示, 其中 $V = \{v_1, v_2, \dots, v_m\}$ 是顶点的集合, $E = \{e_1, e_2, \dots, e_n\}$ 是超边的集合, $W \in R^{n \times m}$ 是超边权值向量组成的矩阵, 矩阵中一行表示一条超边的权值向量。在超网络模型中, 每个顶点对应一个特征变量, 每条超边对应多个特征的组合, 超边的权值则反映了特征组合对于模型预测的重要性。由于超边可以连接两个以上的顶点, 超边可以表示特征变量的高阶组合对预测结果的区分能力。图 1 展示了超网络模型的结构。在超网络模型中, 一条超边包含多个特征变量组成的顶点集合、一个匹配阈值、一个输出标签向量以及一个权值向量。例如, 图 1 中的超边表示在以特征 x_1, x_2, x_7 组成的特征子向量 $[12, 5.1, 5.7]$ 为中心, 匹配阈值 $\delta = 0.25$ 为半径的子区域中, 样本在输出目标上的类别取值 $y_1 = 2, y_2 = 1, y_m = 0$ 的可能性分别为 0.90, 0.88, 0.08。权值越大表示可能性越大, 如果多个输出目标对应的权值都比较大, 则表示它们在该区域上具有很强的关联性, 例如 $y_1 = 2, y_2 = 1$ 的权值分别为 0.90 和 0.88, 表明它们在超边表示的子区域内具有很强的关联性。

超网络模型通过一组超边以及超边的权值来近似表示输入特征矩阵 X 与离散化后的输出目标矩阵 $L = [L_1, L_2, \dots, L_m]^T$ 之间的联合概率分布 $P(X, L_i | W)$, 如式 (2) 所示:

$$P(X, L_i | W) = \frac{1}{Z(W)} \exp(-\epsilon(X, L_i | W)) \quad (2)$$

其中, L_i 表示离散化之后的第 i 个输出目标, W 用来代表超网络

模型, $Z(\mathbf{W})$ 是归一化项, $\epsilon(\mathbf{X}, \mathbf{L}_i | \mathbf{W})$ 表示能量函数, 其定义如式(3)所示。

$$\epsilon(\mathbf{X}, \mathbf{L}_i | \mathbf{W}) = -\ln\left(\sum_{j=1}^m \omega_{ji} I(\mathbf{X}, \mathbf{L}_i; e_j)\right) \quad (3)$$

其中, ω_{ji} 是超边 e_j 中对应于第 i 个输出目标的权值, $I(\mathbf{X}, \mathbf{L}_i; e_j)$ 是超边的匹配函数, 其定义如式(4)所示。

$$I(\mathbf{X}, \mathbf{L}_i; e_j) = \begin{cases} 1, & \text{if } \text{dist}(\mathbf{X}, e_j) \leq \delta \\ 0, & \text{otherwise} \end{cases} \quad (4)$$

如果数据样本与超边 e_j 顶点表示的特征子向量之间的欧氏距离 $\text{dist}(\mathbf{X}, e_j)$ 小于等于超边的匹配阈值 δ , 即训练样本位于超边表示的子区域内时, 数据样本与超边匹配成功, 返回 1; 否则匹配不成功, 返回 0。超边匹配阈值 ϵ 的计算如式(5)所示。

$$\delta = \frac{1}{|G_e|} \sum_{x \in G_e} \|\mathbf{v}_e - \mathbf{x}_e\| \quad (5)$$

其中, G_e 表示超边在训练集中的近邻样本组成的集合, \mathbf{v}_e 和 \mathbf{x}_e 分别表示超边 e 中顶点取值组成的特征子向量和近邻样本 \mathbf{x} 中由超边 e 的顶点组成的特征子向量, $\|\mathbf{v}_e - \mathbf{x}_e\|$ 表示 \mathbf{v}_e 和 \mathbf{x}_e 之间的欧氏距离。

由于超网络模型通过超边及其权值向量表示输入特征与输出目标之间的联合概率分布, 对于一个待预测样本 \mathbf{x} , 可以采用贝叶斯决策方法对其进行预测。首先在训练集中找到待预测样本 \mathbf{x} 的 k 个近邻样本, 然后将 \mathbf{x} 与从近邻样本中生成的所有超边进行匹配, 并根据式(6)计算该样本在每个输出目标上被预测为各个类别的概率, 最后选择概率最大的类别作为输出, 并取该类别对应的聚类中心取值为该目标的预测输出。超网络多目标回归预测的过程如算法 2 所示。

$$P(y_i = l | \mathbf{x}) = \frac{P(\mathbf{x}, y_i = l)}{P(\mathbf{x})} = \frac{\sum_{j=1}^m \omega_{ji} I(\mathbf{x}, y_i = l; e_j)}{\sum_{c \in C_i} \sum_{j=1}^m \omega_{ji} I(\mathbf{x}, y_i = c; e_j)} \quad (6)$$

算法 2 超网络多目标回归预测

输入: 待预测样本 \mathbf{x} , 聚类中心的集合 C

输出: m 维目标输出向量 \mathbf{y}

1. 初始化目标输出向量 \mathbf{y} 为零
2. 从训练集 D 中为待预测样本确定 k 个近邻样本
3. 在超网络模型中找出从 k 个近邻样本生成的超边组成超边集合 U
4. for $i=1$ to m
5. 初始化向量 \mathbf{sum} 为零
6. for e_j in U
7. 将 e_j 与待预测样本 \mathbf{x} 进行匹配
8. 如果匹配成功, $\mathbf{sum}[c] += \omega_{ji}$, c 表示超边 e_j 中第 i 个输出目标的类别, ω_{ji} 表示该类别对应的权值
9. 选取 \mathbf{sum} 向量中最大数据元素对应的下标作为对第 i 个输出目标的类别预测
10. 在第 i 个输出目标所对应的聚类中心 $C_i \in C$ 中, 找到聚类标签为 l 的聚类的中心 c_{il} , 作为预测输出, 即 $\mathbf{y}[i] = c_{il}$
11. 返回目标输出向量 \mathbf{y}

以上介绍了采用超网络模型进行多目标回归预测的过程。为了让超网络准确地表示输入特征和输出目标之间的联合概率分布, 超网络的学习过程需要从训练数据中学习超边的结构以及权值。对于超边结构的学习, 首先从每个训练样本中随机选择几个特征作为超边的顶点, 计算超边的匹配

阈值, 初始化超边的输出目标向量和权值向量。超边的匹配阈值的计算如式(5)所示, 取训练样本的输出目标向量为超边的输出目标向量, 超边的权值初始化为零。由于超边的顶点是随机生成的, 为了得到能够更好地表现新数据分布的超网络模型, 需要对初始化的超边进行替代学习。对于超网络模型中的每条超边, 首先按照式(7)计算适应值, 对于适应值低的超边, 根据相同的训练样本再重新生成一条超边, 如果新生成的超边的适应值大于原超边的适应值, 则用新生成的超边替代原超边。

$$\text{fitness}(e_j) = \sum_{i=1}^m f(e_j, y_i) = \frac{\sum_{i=1}^m \sum_{t=1}^N I(\mathbf{x}_t, y_{it}, e_j)}{\sum_{i=1}^m \sum_{t=1}^N I(\mathbf{x}_t, e_j)} \quad (7)$$

其中, $f(e_j, y_i)$ 表示超边 e_j 中第 i 个输出目标的适应值, 右边分式中分母表示与超边匹配的训练样本的个数, 分子表示与超边匹配且第 i 个输出目标与超边中第 i 个输出目标的类别取值相等的训练样本的个数。

在替代学习后, 超网络的结构已经确定, 接下来需要为每个超边学习得到合适的权值向量。本文采用梯度下降算法, 最小化式(8)所示的代价函数。

$$\begin{aligned} \ell(D, \mathbf{W}) &= \frac{1}{2} \sum_{i=1}^m \sum_{j=1}^m \text{loss}(\mathbf{x}_i, y_{ij}) \\ \text{loss}(\mathbf{x}_i, y_{ij}) &= \sum_{s, l \in C_j} ((P^*(y_{ij} = s | \mathbf{x}_i) - \\ &\quad P(y_{ij} = s | \mathbf{x}_i)) \text{Cost}(s, l))^2 \end{aligned} \quad (8)$$

其中, $P^*(y_{ij} = s | \mathbf{x}_i) = 1$ 表示训练样本 \mathbf{x}_i 的第 j 个输出目标的真实标签为 s , $P(y_{ij} = s | \mathbf{x}_i)$ 表示超网络预测训练样本 \mathbf{x}_i 的第 j 个输出目标的标签为 s 的概率, 其计算如式(6)所示, l 表示超网络预测的第 j 个输出目标的标签。

对于每个输出目标, 由于其不同的类别对应的聚类中心的差别比较大, 因此不同类别之间的错分代价应该是不一样的, 聚类中心越远的两个类别之间的错分代价越高, 预测模型应该尽量避免把聚类中心远的两个类别相互混淆。为此, 在代价函数中加入一个代价敏感函数 $\text{Cost}(s, l)$, 如式(9)所示。

$$\text{Cost}(s, l) = e^{\frac{\|c_{js} - c_{jl}\|^2}{2\sigma}} \quad (9)$$

其中, c_{js} 和 c_{jl} 分别表示第 j 个输出目标的类标签 s 和 l 对应的聚类中心, σ 表示第 j 个输出目标各聚类中心的标准差。

从式(9)可以看出, 真实标签与预测标签对应的聚类中心距离越大, 错分的代价越高, 对代价函数的影响越大。在梯度下降过程中, 对于任意一条超边 e_t , 其第 j 个输出目标对应的权值 ω_{tj} 的更新如式(10)所示。

$$\begin{aligned} \omega_{tj} &= \omega_{tj} - \eta \frac{\partial \ell(D, \mathbf{W})}{\partial \omega_{tj}} \\ \frac{\partial \ell(D, \mathbf{W})}{\partial \omega_{tj}} &= \frac{\partial \ell(D, \mathbf{W})}{\partial P(y_{ij} = s | \mathbf{x}_i)} \frac{\partial P(y_{ij} = s | \mathbf{x}_i)}{\partial \omega_{tj}} \\ &= \frac{\partial \ell(D, \mathbf{W})}{\partial P(y_{ij} = s | \mathbf{x}_i)} = \sum_{i=1}^m ((P^*(y_{ij} = s | \mathbf{x}_i) - \\ &\quad P(y_{ij} = s | \mathbf{x}_i)) \text{Cost}(s, l)) \\ &\quad \frac{\partial P(y_{ij} = s | \mathbf{x}_i)}{\omega_{tj}} = I(\mathbf{x}_i, e_t) \end{aligned} \quad (10)$$

其中, η 表示学习速率。

超网络的学习过程如算法 3 所示。

算法 3 超网络学习过程

输入: 训练数据 $D = \{(\mathbf{x}_i, \mathbf{y}_i)\}_{i=1}^N$, 离散化的目标输出矩阵 $\mathbf{L}^{N \times m}$

输出:超网络模型 $H=(V, E, W)$

1. 初始化超网络模型

1.1. 初始化超边集合 E 为空

1.2. for $i=1$ to N

1.3. 从 \mathbf{x}_i 中随机选取特征变量组成超边顶点集合; 将 \mathbf{L} 的第 i 行 $\mathbf{L}[i,:]$ 作为超边的输出标签向量; 按照式(5)计算超边的匹配阈值

1.4. 将生成的超边加入超边集合 E 中

2. 超边结构学习

2.1. for each e in E

2.2. 生成一条新超边 e_{new}

2.3. 按照式(7)分别计算超边 e 和 e_{new} 的适应值

2.4. 如果 e_{new} 的适应值大于 e 的适应值, 则用 e_{new} 替代 e

2.5. 生成超网络 $H^0(V, E, W)$

3. 超边权值学习

3.1. for $i=1$ to T // T 表示最大迭代次数

3.2. for $j=1$ to N

3.3. 用当前超网络模型 $H^{i-1}(V, E, W)$ 对训练样本 \mathbf{x}_j 进行分类, 得到样本被预测为每个标签的概率 \mathbf{P}

3.4. 根据式(9)计算错分代价

3.5. 根据式(10)更新超边的权值

3.6. 得到第 i 次迭代训练后的超网络模型 $H^i(V, E, W)$

3.7. 结束, 返回最终的超网络模型 $H^T(V, E, W)$

3.4 模型复杂度分析

对于一个包含 N 个训练样本 m 个输出目标的多目标回归问题, MTR-HN 方法首先对输出目标进行离散化, 将多目标回归问题转化为多类别多标签分类问题。MTR-HN 方法在每个输出目标上采用 k -means 算法进行一维聚类, 因此该部分的时间复杂度为 $O(N^2) \times O(m) = O(N^2 \times m)$, 其中 $O(N^2)$ 是 k -means 一维聚类的时间复杂度。然后采用演化超网络对多类别多标签分类问题进行建模, 在学习超网络超边结构的过程中, 需要为每个训练样本确定 k 个近邻, 本文采用 KDTree 算法, 其时间复杂度为 $O(\log^N)$; 在超边权值学习过程中, 采用梯度下降方法更新超边权值, 其时间复杂度为 $O(T \times N \times m \times n)$, 其中 n 表示超边的个数。因此, MTR-HN 方法训练的时间复杂度为 $O(N^2 \times m) + O(\log N^2) + O(T \times N \times m \times n) = O(N^2 \times m + \log^N + T \times N \times m \times n)$ 。对于待预测的数据样本, MTR-HN 方法首先在训练集中找到与其最近的 k 个近邻, 并将其与这些近邻样本产生的超边进行匹配, 根据匹配结果计算每个类别的预测概率, 选取概率最大的类别对应的聚类中心作为多目标回归预测的输出, 其预测时间复杂度为 $O(\log^N + m \times n)$ 。

4 实验

本节首先介绍实验所用的数据集以及评价标准, 然后设计实验对 MTR-HN 方法进行验证并对实验结果进行分析讨论。

4.1 数据集及评价标准

文中采用 18 个多目标回归数据集对 MTR-HN 方法的预测性能进行实验验证和对比分析。文中所用的数据集都来自于 Mulan¹⁾ 多标签学习库, 数据集的统计信息如表 1 所列。

在机器学习领域, 对于回归方法的性能, 通常采用均方误差 (Mean Squared Error, MSE) 来评价。然而, 在多目标回归问题中, 对于一个待预测数据, 需要同时预测多个输出目标的取值。为了准确反映多个输出目标的总体预测性能, 本文采用平均相对均方根误差 (average Relative Root Mean Squared Error, aRRMSE) 作为多目标回归方法预测性能的评价标准。具体来讲, 对于每个输出目标, 先计算其相对均方根误差, 然后对所有的输出目标的相对均方根误差取平均值。aRRMSE 的计算如式(11)所示。

$$aRRMSE(D_T, h) = \frac{1}{m} \sum_{i=1}^m \sqrt{\frac{\sum_{(x,y) \in D_T} (y_i - y_i^*)^2}{\sum_{(x,y) \in D_T} (Y_{\text{mean}}^i - y_i^*)^2}} \quad (11)$$

其中, m 表示输出目标个数, h 表示多目标回归预测模型, D_T 表示测试集, y_i 表示测试样本在第 i 个输出目标上的真实输出, y_i^* 表示模型对测试样本在第 i 个输出目标上的预测输出, Y_{mean}^i 表示第 i 个输出目标在训练集上的平均值。aRRMSE 的取值越小, 模型的预测准确性越高。

表 1 18 个多目标回归数据集的统计信息

Table 1 Statistics of 18 MTR datasets

数据集	样本数 N	特征维数 d	输出目标数 m
edm	154	16	2
enb	768	8	2
jura	359	15	2
scpf	1137	23	3
sf1	323	10	3
sf2	1066	10	3
slump	103	7	3
andro	49	30	6
atp1d	337	411	6
atp7d	296	411	6
rf1	9125	64	8
rf2	9125	576	8
osales	639	413	12
wq	1060	16	14
oes10	403	298	16
oes97	334	263	16
scml d	9803	280	16
scm20d	8966	61	16

4.2 实验设置

将 MTR-HN 方法与以下 5 个典型的多目标回归方法进行对比实验。

(1) Stacked Single-Target^[9], 简称 SST。采用多标签分类方法 BR 的思想, 分别为每个输出目标训练一个预测模型, 在训练模型时将其他输出目标的输出值作为额外的输入, 扩展原始输入特征。

(2) Ensemble of Regression Chains^[9], 简称 ERC。采用多标签分类方法 ECC 的思想, 将多个目标随机组成一条链, 将链上其他输出目标的输出值作为额外的输入, 扩展原始输入特征, 构建一条回归器链, 并对多条回归器链进行集成。

(3) Multi-Task Feature Learning^[27], 简称 MTFI。将多任务学习中的特征学习策略应用到多目标回归问题中。

(4) Multi-layer Multi-target Regression^[8], 简称 MMR。采用与多任务学习中低秩方法类似的策略来学习多目标回归中各输出目标的预测模型的参数。

(5) Multi-target Regression via Quantization^[13], 简称 MRQ。与本文提出的 MTR-HN 类似, 通过聚类方法将多目标

¹⁾ <http://mulan.sourceforge.net/>

回归问题转化为多分类问题,而 MTR-HN 则是将多目标回归问题转化为多类别多标签分类问题。

对每个数据集采用交叉验证的方式进行实验,取交叉验证中评价标准 aRRMSE 的平均值作为各算法在该数据集上的最终结果。本文采用与对比算法相同的交叉验证设置,即对样本数多的 rf1 和 rf2 两个数据集采用 2-折交叉验证,对样本数较多的 scm1d 和 scm20d 数据集采用 5-折交叉验证,其他数据集上采用 10-折交叉验证。

4.3 实验结果及分析

在 18 个数据集上分别采用本文提出的 MTR-HN 方法和 5 种对比算法进行实验,表 2 列出了 MTR-HN 以及对比算法在各个数据集上的预测结果,每个数据集上的最好结果都采用粗体表示。从表 2 中可以看出,MTR-HN 在 11 个数据集上都取得了比其他 5 种对比算法更好的预测准确性。为了验证 MTR-HN 方法与其他 5 种对比算法在预测准确性上是否存在显著的差异,采用 Friedman 检验^[40]对表 2 中的实验结果进行统计分析。首先设定原假设 H_0 为:MTR-HN 与 5 种对比算法之间的预测准确性不存在明显的差别,然后计算 Friedman 检验的统计变量,如表 3 所列。从表 3 中可以看出 Friedman 假设检验的临界值远远小于 Friedman 统计量,因此原假设 H_0 被拒绝,即 MTR-HN 与 5 种对比算法之间的预测准确性确实存在明显的差异。

表 2 MTR-HN 及对比算法在 18 个数据集上的实验结果

Table 2 Experiment results of MTR-HN and comparing algorithms over 18 dataset

数据集	SST	ERC	MTFL	MMR	MRQ	MTR-HN
edm	0.740	0.741	0.851	0.716	0.889	0.656
enb	0.121	0.114	0.316	0.111	0.188	0.114
jura	0.591	0.590	0.608	0.582	0.764	0.585
scpf	0.831	0.830	0.899	0.812	0.985	0.788
sf1	1.068	1.089	1.112	0.958	0.825	0.787
sf2	1.055	1.088	1.127	0.984	0.821	0.791
slump	0.695	0.689	0.681	0.587	0.926	0.652
andro	0.579	0.567	0.803	0.527	0.954	0.364
atp1d	0.372	0.372	0.415	0.332	0.474	0.345
atp7d	0.507	0.512	0.553	0.443	0.591	0.492
rf1	0.094	0.091	0.983	0.089	0.251	0.086
rf2	0.097	0.095	1.103	0.095	0.251	0.091
osales	0.726	0.713	1.682	0.709	0.770	0.686
wq	0.909	0.906	0.962	0.889	1.068	0.902
oes10	0.421	0.420	0.532	0.403	0.507	0.373
oes97	0.524	0.524	0.818	0.497	0.642	0.501
scm1d	0.336	0.330	0.437	0.318	0.455	0.315
scm20d	0.413	0.394	0.643	0.389	0.485	0.381

表 3 Friedman 检验的统计结果

Table 3 Statistics of Friedman test

Friedman 假设检验	
Friedman 统计量	65.916
显著性水平($\alpha=0.05$)下的 F 检验临界值	2.322
拒绝或接受原假设 H_0	拒绝

为了进一步分析 MTR-HN 方法分别与 5 个对比算法在预测性能上的优劣,采用 Nemenyi 检验^[40]分析算法两两间的相对性能差异,并根据式(12)计算算法两两间显著差异的最小临界值(Critical Difference,CD)。

$$CD = q_{\alpha} \sqrt{\frac{S(S+1)}{6N}} \quad (12)$$

其中,S 表示算法的个数,N 表示数据集的个数,在 $\alpha=0.05$ 时, $q_{\alpha}=2.850$ 。

如果任意两个算法在所有数据集上的平均排序之差大于等于 CD 值,表明两者之间有显著的差异,且平均排序靠前的算法的性能优于平均排序靠后的算法的性能。MTR-HN 与 5 个对比算法的平均排序对比如图 2 所示,可以看出 MTR-HN 方法的预测性能明显优于 SST,ERC,MTFL 方法;MTR-HN 方法虽然与 MMR 方法之间没有显著性的性能差异,但是 MTR-HN 方法的平均排序仍稍高于 MMR 算法,说明 MTR-HN 的预测性能略优于 MMR 方法。通过对实验结果进行统计分析,可以得出 MTR-HN 方法能够在多目标回归问题上取得比现有方法更高的预测准确性。

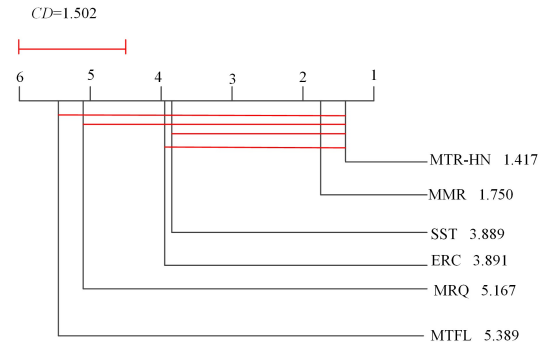


图 2 MTR-HN 方法与 5 个对比算法在平均排序上的比较

Fig. 2 Comparison of average ranks of MTR-HN and 5 comparing algorithms

此外,从表 2 中我们观察到,MRQ 和 MTR-HN 在 sf1 和 sf2 两个数据集上的 aRRMSE 值明显低于其他 5 个对比算法,即 MRQ 和 MTR-HN 在这两个数据集上的预测性能明显高于对比其他方法。在 sf1 和 sf2 两个数据集上,3 个输出目标的取值都是离散型的。其中,sf1 数据集的第一个输出目标的取值有{0,1,2},第二个输出目标的取值有{0,1,2,4},第三个输出目标的取值只有{0,1};sf2 数据集的第一个输出目标的取值有{0,1,2,3,4,5,6,8},第二个输出目标的取值有{0,1,2,3,4,5},第三个输出目标的取值有{0,1,2}。说明 sf1 和 sf2 这两个多目标回归数据集本身就可以被看作多类别多标签分类的数据集。在这两个数据集上采用输出目标离散化方法将多目标回归问题转化为多类别问题和多类别多标签分类问题都能更有效地对输出目标之间的关联关系进行建模,提高预测准确性。

在超网络模型中,为了计算超边的匹配阈值,需要从训练集中为生成超边的样本确定 k 个近邻样本,近邻样本个数 k 的取值对超网络模型有着重要的影响。为了选取合适的 k 值,本文采用交叉验证的方式,分别设置不同的 k 值,从训练集中随机抽取 10% 的样本组成验证集,并在验证集上查看不同 k 值情况下的预测准确性,最后确定 k 的较优取值。图 3 所示的是不同数据集下不同 k 值时的 aRRMSE 的变化情况。从图 3 中可以看出,当 k 值较小时或者 k 值较大时,aRRMSE 都偏高,即预测性能变差。造成这种现象的原因在于当 k 值较小时邻近样本个数少,超边所表示的特征子区域比较狭窄,超网络模型对数据的局部信息过于敏感;当 k 值较大时,邻近样本个数较多,超边所表示的特征子区域较大,很容易受到噪声数据的干扰,降低了超网络的鲁棒性。

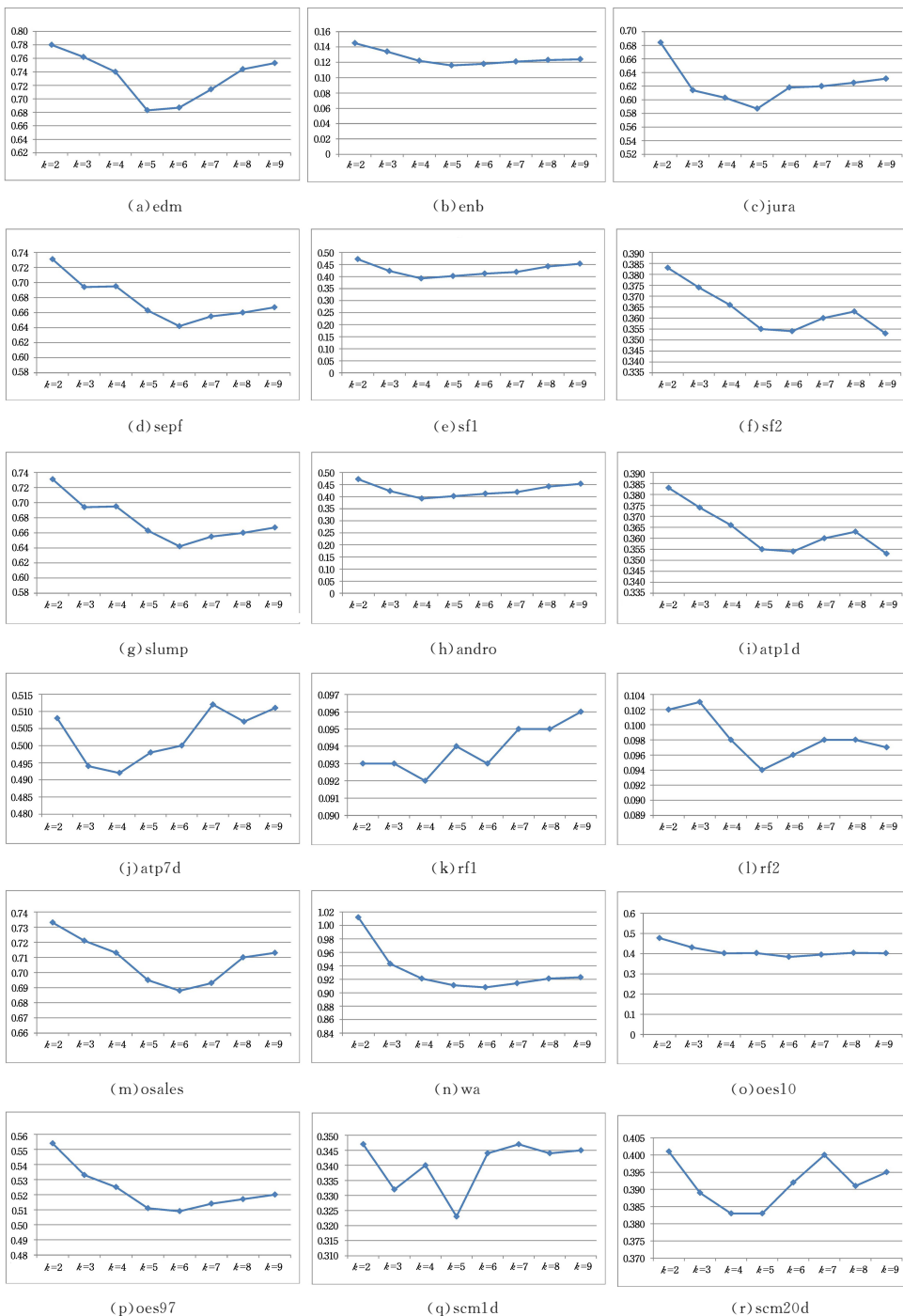


图3 MTR-HN方法在不同k值下18个数据集上的aRRMSE变化
 Fig. 3 aRRMSE of MTR-HN over 18 datasets with different k

结束语 本文提出了一种基于超网络的多目标回归方法MTR-HN, 通过将连续型输出目标进行离散化, 将多目标回归问题转化为多类别多标签分类问题, 并采用超网络模型进行建模, 构建多目标回归预测模型。MTR-HN很好地利用了超网络模型对离散的多输出目标之间高阶关联的建模能力, 提高了多目标回归问题的预测准确性。在18个多目标回归数据集上的对比实验结果也表明了MTR-HN在处理多目标回归问题上的有效性。然而, 本文只在输出空间上探索挖掘输出目标之间的关联, 没有结合输入特征来进一步发掘输出目标之间的关联。在未来的工作中, 将尝试结合目标相关特征构建方法与超网络模型的多目标回归方法, 更好地挖掘和利用输出目标间的关联, 进一步提升多目标回归问题的预测准确性。

参考文献

[1] SUN K W, DENG M X, LI H, et al. Learning local instance correlations for multi-target regression [J]. Applied Intelligence, 2021, 51: 6124-6135.
 [2] XIAO X, XU Y. Multi-target regression via self-parameterized Lasso and refactored target space [J]. Applied Intelligence, 2021, 51: 6743-6751.
 [3] WANG Y, WIPF D P, LING Q, et al. Multitask learning for subspace segmentation [C]// Proceedings of the 32nd International Conference on Machine Learning (ICML). 2015: 1209-1217.
 [4] XIONG T, BAO Y, HU Z. Multiple-output support vector regression with a firefly algorithm for interval-valued stock price

- index forecasting [J]. *Knowledge-Based Systems*, 2014, 55: 87-100.
- [5] HADAVANDI E, SHAHRABI J, SHAMSHIRBAND S. A novel Boosted-neural network ensemble for modeling multi-target regression problems [J]. *Engineering Applications of Artificial Intelligence*, 2015, 45: 204-219.
- [6] YAN Y, RICCI E, SUBRAMANIAN R, et al. A multi-task learning framework for head pose estimation under target motion [J]. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, 2016, 38(6): 1070-1083.
- [7] WANG X, ZHEN X, LI Q, et al. Cognitive assessment prediction in Alzheimer's disease by multi-layer multitarget regression [J]. *Neuroinformatics*, 2018, 16: 285-294.
- [8] ZHEN X, YU M, HE X, et al. Multi-target regression via robust low-rank learning [J]. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, 2018, 40(2): 497-504.
- [9] SPYROMITROS-XIOUFIS E, TSOUMAKAS G, GROVES W, et al. Multi-target regression via input space expansion: treating targets as inputs [J]. *Machine Learning*, 2016, 104(1): 55-98.
- [10] YUAN H, ZHENG J, LAI L, et al. Sparse structural feature selection for multitarget regression [J]. *Knowledge-Based Systems*, 2018, 160: 200-209.
- [11] ZHANG M L, ZHOU Z H. A Review on Multi-Label Learning Algorithms [J]. *IEEE Transactions on Knowledge & Data Engineering*, 2014, 26(8): 1819-1837.
- [12] LV J, WU T, PENG C, et al. Compact Learning for Multi-Label Classification [J]. *Pattern Recognition*, 2021, 113: 107833.
- [13] SPYROMITROS-XIOUFIS E, SECHIDIS K, VLAHAVAS I. Multi-target regression via output space quantization [C]// *International Joint Conference on Neural Networks*. 2020.
- [14] ZHANG B T. Hypernetworks: A Molecular Evolutionary Architecture for Cognitive Learning and Memory [J]. *IEEE Computational Intelligence Magazine*, 2008, 3(3): 49-63.
- [15] SUN K W, LEE C, XIE X X. MLHN: A Hypernetwork Model for Multi-Label Classification [J]. *International Journal of Pattern Recognition and Artificial Intelligence*, 2015, 29(6): 1550020. 1.
- [16] SUN K W, LEE C, WANG J. Multilabel Classification via Co-Evolutionary Multilabel Hypernetwork [J]. *IEEE Transactions on Knowledge & Data Engineering*, 2016, 28(9): 2438-2451.
- [17] SUN K W, LEE C. Addressing class-imbalance in multi-label learning via two-stage multi-label hypernetwork [J]. *Neurocomputing*, 2017, 266(29): 375-389.
- [18] XU D, SHI Y, TSANG I W, et al. A Survey on Multi-output Learning [J]. *IEEE Transactions on Neural Networks and Learning Systems*, 2020, 31(7): 2409-2429
- [19] ZHANG Y, YANG Q. A Survey on Multi-Task Learning [J]. *IEEE Transactions on Knowledge and Data Engineering*, 2021, arXiv: 1707. 08114.
- [20] TSOUMAKAS G, KATAKIS I. Multi-Label Classification: An Overview [J]. *International Journal of Data Warehousing and Mining*, 2009, 3(3): 1-13.
- [21] BRINKER K, MENCIA E L, FUERNKRANZ J, et al. Multilabel classification via calibrated label ranking [J]. *Machine Learning*, 2008, 73(2): 133-153.
- [22] TSOUMAKAS G, KATAKIS I, VLAHAVAS I. Randomk-Labelsets for Multi-Label Classification [J]. *IEEE Transactions on Knowledge and Data Engineering*, 2011, 23(7): 1079-1089.
- [23] READ J, PFAHRINGER B, HOLMES G, et al. Classifier chains for multi-label classification [J]. *Machine Learning*, 2011, 85(3): 333-359.
- [24] ZHANG M L, ZHOU Z H. ML-KNN: A lazy learning approach to multi-label learning [J]. *Pattern Recognition*, 2007, 40(7): 2038-2048.
- [25] ZHANG M L, ZHOU Z H. Multilabel Neural Networks with Applications to Functional Genomics and Text Categorization [J]. *IEEE Transactions on Knowledge & Data Engineering*, 2006, 18(10): 1338-1351.
- [26] WANG H, XU Y. Sparse elastic net multi-label rank support vector machine with pinball loss and its applications [J]. *Applied Soft Computing*, 2021, 104(9): 107232.
- [27] ARGYRIOU A, EVGENIOU T, PONTIL M. Convex multi-task feature learning [J]. *Machine Learning*, 2008, 73(3): 243-272.
- [28] SU F, SHANG H Y, WANG J Y. Low-Rank Deep Convolutional Neural Network for Multi-Task Learning [J]. *Computational Intelligence and Neuroscience*, 2019, 2019: 1-10.
- [29] WANG D, NIE F, HUANG H. Learning Task Relational Structure for Multi-task Feature Learning [C]// *IEEE International Conference on Data Mining*. 2016: 1239-1244.
- [30] BICKEL S, BOGOJESKA J, LENGAUER T, et al. Multi-task learning for HIV therapy screening [C]// *Proceedings of the 25th International Conference on Machine Learning*. 2008: 56-63.
- [31] TSOUMAKAS G, SPYROMITROS X E, VREKOU A, et al. Multi-target regression via random linear target combinations [C]// *Joint European Conference on Machine Learning and Knowledge Discovery in Databases*. 2014: 225-240.
- [32] BRESKVAR M, KOCEV D, DZEROSKI S. Ensembles for multi-target regression with random output selections [J]. *Machine Learning*, 2018, 107(11): 1673-1709.
- [33] ZHANG Y, YEUNG D Y. A convex formulation for learning task relationships in multi-task learning [C]// *Proceedings of UAI*. 2010: 73-742.
- [34] TODOROVSKI L, BLOCKEEL H, DZEROSKI S. Ranking with Predictive Clustering Trees [C]// *European Conference on Machine Learning*. 2002: 444-455.
- [35] OSOJNIK A, PANOV P, DZEROSKI S. Tree-based methods for online multi-target regression [J]. *Journal of Intelligent Information Systems*, 2018, 50: 315-339
- [36] LEVATIC J, CECI M, KOCEV D, et al. Self-training for multi-target regression with tree ensembles [J]. *Knowledge-Based Systems*, 2017, 123: 41-60.
- [37] LIU H T, LI H, WANG J, et al. Multi-Target Regression via Sparse Integration and Label-Specific Features [J]. *Acta Electronica Sinica*, 2020, 48(5): 906-913.
- [38] WANG J, GAO X R, ZHANG R, et al. Multi-Target Regression via Specific Features and Inter-Target Correlations [J]. *Acta Electronica Sinica*, 2020, 48(11): 2092-2100.
- [39] WANG J, CHEN Z L, SUN K W, et al. Multi-target regression via target specific features [J]. *Knowledge-Based Systems*, 2019, 170(15): 70-78.
- [40] DEMSAR J. Statistical comparisons of classifiers over multiple data sets [J]. *Journal of Machine Learning Research*, 2006, 7: 1-30.



SUN Kai-wei, born in 1987, Ph.D, associate professor, master supervisor, is a member of China Computer Federation. His main research interests include machine learning, natural language processing and big data analysis.