基于信息融合的半监督有序分类框架

冉炯宇，汤梦姿，解庆，刘永坚
（武汉理工大学计算机与人工智能学院，湖北 武汉 430070）

摘　要：有序分类属于分类的一种，其要求类标签存在自然顺序，在很多领域例如电影分级、年龄估计都得到了广泛的研究。目前，大部分有序分类方法假设所有样本都被标记。但由于数据的特殊性，在实践中往往难以收集大量的标记数据，影响有序分类的性能。针对以上问题，提出一种结合额外信息的半监督有序分类框架。首先，利用未标记样本的顺序关系生成额外的偏序信息，并将偏序信息构建为有向图网络；然后使用图神经网络（GNN）聚合邻居信息，丰富节点表示，同时捕捉节点间的顺序关系，利用学习到的表示恢复偏序信息间的全局排名；接着使用高斯混合加权的方法对数据特征根据全局排名进行加权，并使用聚类方法为全局排名赋予伪标签，从而将这些信息合并到有序信息中；最后，使用有监督学习的有序分类模型进行年龄估计。在 FGNET、Adience、UTKFace 3 个数据集上的实验结果表明，该框架使用较少的标记数据便能够取得可靠的性能，在平均绝对误差（MAE）、准确率（Accuracy）2 个指标上相较于半监督学习基线方法均有提升：MAE 在 3 个数据集上分别降低了 0.05、0.04、0.04，Accuracy 在 3 个数据集上分别提高了 4.8、4.5、3.5 百分点。

关键词：有序分类；图神经网络；特征加权；信息融合；半监督学习
中图分类号：TP18　　**文献标志码：**A　　**DOI：**10.19678/j.issn.1000-3428.0069814

Semi-Supervised Ordinal Classification Framework Based on Information Fusion

RAN Tongyu, TANG Mengzi, XIE Qing, LIU Yongjian
(School of Computer Science and Artificial Intelligence, Wuhan University of Technology, Wuhan 430070, Hubei, China)

【Abstract】 In ordinal classification, class labels have a natural order. It has been widely studied in various fields, such as movie ratings and age estimation. Most existing methods assume that all samples are labeled. However, the unique nature of data often makes the collection of extensive labeled data challenging, thereby affecting the performance of ordinal classification. This study proposes a semi-supervised ordinal classification framework that incorporates additional information. The framework starts by generating partial order information from the relationships among unlabeled samples and constructing a directed graph network. Then, it uses Graph Neural Network (GNN) to aggregate neighbor information, enrich node representations, and capture the order between nodes, thereby recovering global rankings from partial order information. Subsequently, the method applies a Gaussian mixture model for feature weighting according to global rankings and employs clustering to assign pseudo labels by integrating this information into ordered information. Finally, the framework uses supervised learning models for ordinal classification tasks such as age estimation. Experiments on the FGNET, Adience, and UTKFace datasets show that the framework achieves reliable performance with fewer labeled data. It performs better than semi-supervised learning baselines in terms of Mean Absolute Error (MAE) and Accuracy. Specifically, MAE decreases by 0.05, 0.04, and 0.04, and Accuracy increases by 4.8, 4.5, and 3.5 percentage points on the three datasets, respectively.

【Key words】 ordinal classification; Graph Neural Network (GNN); feature weighting; information fusion; semi-supervised learning

0　引言

在众多机器学习应用中，分类问题占据了核心地位。传统的分类类别标签之间没有内在的顺序关系，例如在图像识别任务中，目标变量分为猫、狗、鸟等类别，这些类别之间并不存在自然排序。而有序分类的类别标签存在明确的自然顺序，如满意度等级可以分为 1～5 颗星，影视评分可以分为 1～10 分。与一般分类相比，有序分类的误分类代价是不同的。如在满意度调查中，一般将满意度分为“非常满意”“满意”“一般”“不满意”“非常不满意”，而将“满意”误分类为“非常满意”与将其误分类为“不满

基金项目：武汉理工大学自主创新研究基金（2024IVA036）。
作者简介：冉炯宇，男，硕士研究生，主研方向为机器学习；汤梦姿（通信作者），讲师、博士；解庆，副教授、博士；刘永坚，教授。
收稿日期：2024-05-06　　**修回日期：**2024-08-28
E-mail：tangmz@whut.edu.cn

意”,这样不同的分类是有不同后果的。因此,在医疗、信息评级、年龄估计等领域中,如何合理地利用标签之间的顺序是非常重要的。

在早期的研究中,有序分类模型大多依赖于传统的有监督分类方法进行预测,例如支持向量机(SVM)[1-3]、二元分类[4-6]、神经网络[7-9]等算法。SVM 由于其良好的泛化能力,广泛应用于有序分类问题,但其也存在着计算复杂度高等问题。也有人将有序分类问题转化为一系列二元分类问题,即将目标变量分解为多个二元变量,然后通过单个或多个模型进行预测,而由于忽略了不同二元分类器之间的关系,此类方法无法确保保留全局序数关系。为解决这个问题,文献[5]将有序分类任务分解为一系列递归二元分类问题,进而细分相邻类别。文献[6]则扩展了可微决策树模型并结合卷积神经网络(CNN)从而获得了精确的全局序数关系。在神经网络方面,近年来,受益于 CNN 等神经网络模型的提出,越来越多的神经网络模型也被应用于有序分类任务中。

虽然上述方法充分利用了数据间的顺序关系,并且在很多任务中取得了显著的成功,但其面临的一个重要挑战是需要大量的有序信息(即带样本标签的样本数据,也被称为绝对信息)以训练模型。在现实世界中,获得大规模的有序信息可能非常昂贵或者耗时,如果没有足够的有序信息,监督学习模型的性能就会受到影响。

为了解决以上问题,基于半监督学习的有序分类方法应运而生。文献[10-12]引入未标记的数据从而增强有序分类性能,不过大多数半监督有序分类算法由于采用批量学习算法,计算成本都很高。另一种方案是利用额外的辅助信息进行有序分类,根据文献[13]所提出的理论,对于人工标注来说,获取偏序信息(即带有偏序关系的样本对,也被称为相对信息)相对于获得有序信息容易得多。例如在年龄估计领域中,与直接判断一个人的年龄相比,判断两个人年龄之间的相对大小要容易得多,因此可以很轻易地获得大量的偏序信息。

现有的融合偏序信息的半监督有序方法其中之一为基于偏序关系建模的有序分类方法。具体来说,文献[14]提出一种改进的有序回归模型来考虑偏序信息,即在原凸优化目标函数中引入线性限制进行改进,然而该模型引入了许多参数使求解变得更为复杂。针对此问题,文献[15]提出在使用经典的有序信息训练方法的基础上,同时使用偏序信息以统一的思路引入对应的线性限制,将这些限制加入原目标函数中来限制函数的优化过程,使学习得到的模型能够同时处理两类信息。这两种方法虽然有效融入了偏序信息,缓解了有序信息稀缺的问题,但忽略了有序信息与偏序信息之间存在的信息差异性。

从反映局部信息的成对比较中恢复全局排名是排名学习中信息检索的一个基本问题[16],其中代表性的方法是使用特定相似度矩阵的费德勒(Fiedler)向量恢复全局排名[17]。而在融合偏序信息的有序分类问题中拥有大量的偏序信息。受此启发,可以将偏序信息看作为成对比较从而将其进行全局排名,从而与有序信息融合进行有序分类。

本文在年龄估计的背景下[18]提出了一种基于数据融合的有序分类框架,将偏序信息和有序信息融合进行数据增强以解决有序分类问题。首先针对偏序信息,通过构建有向图,使用图卷积网络(GCN)学习图中复杂节点特征,从而计算相似度矩阵得到每个节点的排名,然后将由偏序信息得到的全局排名与有序信息进行聚类得到新的训练样本,最后通过现有先进的有序分类方法进行预测。

1 相关工作

1.1 有序分类

融入偏序信息处理半监督有序分类任务的工作大致可以分为如下几类:基于偏序关系建模的有序分类方法[14-15]、基于偏序信息多样性的有序分类方法[19-20]和基于偏序信息距离度量学习的有序分类方法[21-22]。

基于偏序关系建模的有序分类方法是在有序信息基础上融入偏序信息对偏序关系进行建模的一类半监督有序分类方法。SADER 等[14]提出一种改进的有序回归模型来考虑偏序信息,即在原凸优化目标函数中引入线性限制进行改进,然而该模型引入了许多参数使求解变得复杂且不适合处理大数据。当面临有序信息时,经典的有序分类方法包括基于概率的比例优势模型、基于支持向量的有序回归方法、基于显式限制的支持向量回归方法、基于隐式限制的支持向量回归方法、线性判别有序回归模型等。当同时面临有序与偏序信息时,文献[15]指出可在使用有序信息训练这些方法的基础上,同时使用偏序信息以统一的思路引入对应的线性限制,将这些限制加入原目标函数中来限制函数的优化过程,使学习到的模型能够同时处理两类信息。

基于偏序信息多样性的有序分类方法是在有序

信息基础上考虑偏序信息多样性进行半监督有序分类任务处理的一类方法。在现实情况中，通常不会仅收集类型单一的偏序信息，而往往会收集带频率分布的偏序信息。这种带分布的偏序信息更加符合实际中数据收集的习惯，例如可通过咨询一系列新手针对一个样本对收集对应的偏序关系分布，进而得出带频率分布的偏序信息。文献[19-20]引入这种信息自然地拓展现有处理偏序信息的有序分类方法，以达到同时结合有序信息、不带频率分布的偏序信息或者带频率分布的偏序信息等多种信息的目的。

基于偏序信息距离度量学习的有序分类方法是在有序信息基础上融入偏序信息进行距离度量学习的一类半监督有序分类方法。当面临有序信息时，k 近邻方法潜在地假设相邻的样本有相似的类别标签；而当同时面临有序与偏序信息时，自然地，可将针对有序信息中关于样本的假设推广到偏序信息中的样本对，即假设偏序信息中相邻的样本对也有相似的偏序关系。使用该假设可改进传统的 k 近邻方法来同时融合有序与偏序信息处理半监督有序分类任务[21]。然而，k 近邻方法中使用的欧氏距离只适合特征简单且维数同等重要的数据，无法满足一些给定数据的要求。文献[22]引入了距离度量学习，同时结合有序与偏序信息设置三元距离限制，通过求解带限制的凸目标函数，学习一个更加合适的距离度量，提高了有序分类的性能。

1.2 排名学习

排名学习在很多领域中都有应用，如足球比赛结果建模、电影推荐系统、搜索引擎等。在分析大规模数据集时，人们经常寻求各种形式的数据排名，以识别最重要的条目、计算搜索排序操作或提取主要特征。早期的成对比较排名恢复算法主要有 BT(Bradley-Terry)[23]、BTL(Bradley-Terry-Luce)[24]、PL(Plackett-Luce)[25] 及利用 EM(Expectation-Maximization)算法的演化版本[26]。由于 BTL 模型在成对比较数据模型中拥有良好的统计性质，因此该模型已应用到了生物医学、经济学、社会学等很多领域。

为了更好地处理复杂的数据关系，越来越多的研究开始使用图等网络结构来进行排名任务，旨在量化图节点最重要的程度[27]。文献[28]提出了 SyncRank 模型，将不完全噪声成对信息的排序问题转化为平面旋转群 SO(2)上的群同步问题的一个实例。文献[29]受到物理模型的启发，提出通过求解线性方程组来推断有向图网络中节点的层次排名 SpringRank 算法。文献[30]引入了基于奇异值分解的简单排序和同步算法，并给出了理论保证。文献[17]提出了 SerialRank 算法，首先根据某个相似度矩阵计算拉普拉斯算子，然后计算相似度矩阵相应的 Fiedler 向量作为最终的排名估计，但 SerialRank 算法性能在很大程度上取决于相似性矩阵的质量。

图神经网络(GNN)在图结构数据建模方面展现出巨大潜力。文献[31]应用神经网络方法进行偏好学习以确定在给定数据对中排在第一位的数据。文献[32]放宽了 RankNet[33] 中使用的一些约束条件，并使用了更优化的算法。文献[34]提出了第一个基于 GNN 的模型，用于近似介数中心性和接近中心性，有助于在图中定位具有信息传播和连接性方面影响力的节点。但这些工作中很少考虑成对比较的方向。

基于上述相关工作的分析，现有的融合偏序信息的半监督有序学习模型忽略了有序信息与偏序信息之间的信息差异，并且没有充分挖掘复杂的偏序信息。本文在 GNNRank[35] 的基础上，将偏序信息视为成对比较，并引入 GCN，提高模型对复杂数据结构的理解能力，充分挖掘偏序信息，在得到全局排名后，根据改进的聚类结果生成伪标签，将两种信息融合后得到新的样本集，并使用先进的有监督有序分类模型，形成一个综合的半监督有序分类框架，提高模型的整体性能与泛化能力。

2 模型框架

2.1 数据描述

在介绍模型框架之前，先介绍本文中涉及的一些基本概念和符号。

本文研究中数据包括有序信息和偏序信息，其中偏序信息的数量远大于有序信息。令 $D=\{(x_1,y_1),(x_2,y_2),\cdots,(x_n,y_n)\}$ 为数据集，输入集 $X=\{x_1,x_2,\cdots,x_n\}$，y_i 的标签集合记为 $L=\{l_1,l_2,\cdots,l_m\}$，其中，$l_1<l_2<\cdots<l_m$，$<$ 反映了数据标签之间的顺序关系。

有序信息通过集合 $D_0=\{(x_i,y_i)\mid(x_i,y_i)\in D\}$ 表示，其中数据 $x_i\in X$ 带有标签 $y_i\in L$，且标签有自然顺序；偏序信息由集合 $P=\{(x_i,x_j)\mid x_i,x_j\in X\}$ 表示，其中每个偏序信息对 (x_i,x_j) 由数据集 X 中的数据点组成，这些数据的标签信息未知，而偏序信息对 (x_i,x_j) 中的数据顺序反映了偏序关系 $\succ$，表明 x_i 的标签在顺序上先于 x_j 的标签。

将偏序信息 P 构建为有向图 $G=(V,E)$，其中，节点集 V 为偏序信息中样本 x 的集合，边集 E

由偏序信息中的偏序关系定义。对于 G 的邻接矩阵 $\boldsymbol{A}$，若 $v_i \succ v_j$，则 $A_{i,j}=1$，否则 $A_{i,j}=0$。

2.2 总体框架

本文提出一个用于处理相对信息和绝对信息的半监督有序学习框架，如图 1 所示（彩色效果见《计算机工程》官网 HTML 版，下同）。该框架主要由 6 个部分组成：1）数据集模块在原始数据集的基础上，设计出有序信息与偏序信息两个数据集，并为偏序信息构建有向图和邻接矩阵以进行训练；2）特征提取模块使用预训练好的 CNN 模型对图像数据提取特征，同时为偏序信息提取到的特征组成特征矩阵以进行训练；3）信息传播模块利用 GNN 模型学习深层次的节点表示；4）排名预测模块利用节点嵌入计算相对信息中节点排名；5）数据增强模块利用改进的聚类结果为相对信息赋予伪标签，并与有序信息进行数据融合得到新的有序信息；6）有序分类模块利用最终得到的新有序信息训练有序分类模型，以此进行年龄估计任务。

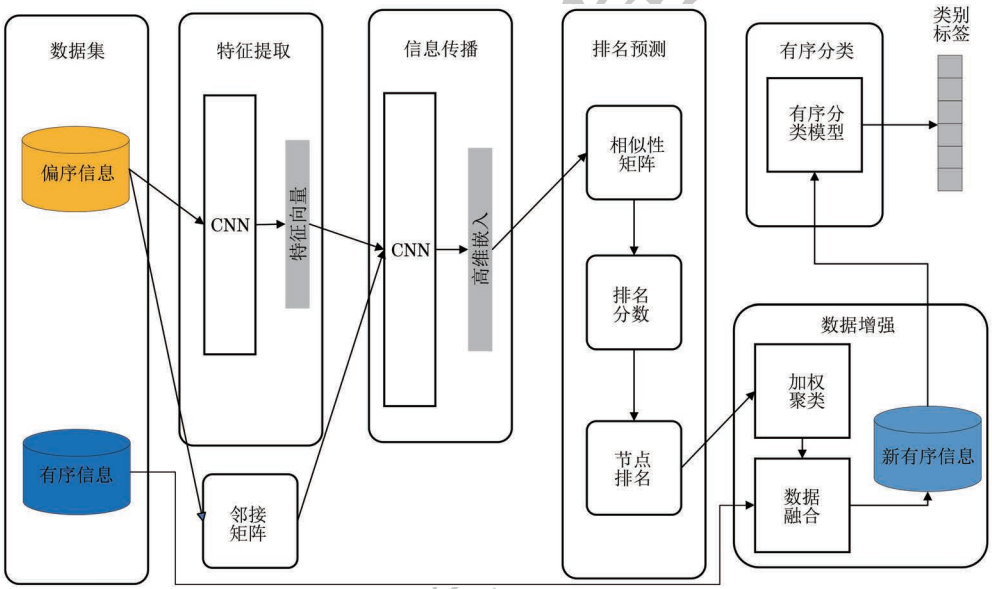

图 1 框架流程图

Fig. 1 Framework flowchart

2.3 特征提取

图像数据往往具有很高的维度，处理如此高维度的数据不仅计算成本高，还容易引起维数灾难。特征提取旨在从图像数据中提取深层次特征表示，可以显著降低数据的维度，使得处理变得更加高效，提高模型性能，同时避免维数灾难。

基于 CNN 的模型 ResNet 广泛应用于图像处理研究领域，它能够自动地从数据中学习复杂的特征，免去了繁复的人工手动设计，并通过残差网络的设计，避免了梯度消失和梯度爆炸问题。尽管 CNN 在特征提取方面表现出色，但本身并不直接获取数据中的偏序信息。采用 CNN 的目的主要是为了提取特征，以便后续的分析和处理。

在本文的特征提取模块中，去除了 ResNet 末端的全连接层，并在池化层中同时使用全局平均池化（GAP）与全局最大池化（GMP），同时保留数据的整体特征和局部关键信息，增强模型泛化能力。

图像的特征向量表示如下：

$$\boldsymbol{z}'_i = f(x_i) \tag{1}$$

式中：$\boldsymbol{z}'_i \in \mathbb{R}^d$，$x_i \in X$，$\boldsymbol{z}'_i$ 为图像 x_i 通过 ResNet 模型输出的特征向量，d 是特征向量的维度。

在将偏序信息构造为有向图并获取邻接矩阵后，构造的节点嵌入如下：

$$\boldsymbol{Z}' = (\boldsymbol{z}'_1, \boldsymbol{z}'_2, \cdots, \boldsymbol{z}'_n) \tag{2}$$

式中：$\boldsymbol{Z}' \in \mathbb{R}^{n \times d}$，$n$ 为节点数量。

2.4 信息聚合

之前的排名方法如 SerialRank，通过计算相似性矩阵的拉普拉斯算子，进而计算对应的 Fiedler 向量用作最终排名估计。该方法虽然有效，但其表现极其依赖相似性矩阵的质量。受此启发，引入 GNN 模型，通过递归聚合邻接节点信息来获取具有丰富语义和结构信息的新节点表示，该模型可以捕捉到节点间的顺序关系，并利用节点嵌入进行后续的预测。

GCN 作为 GNN 广泛使用的技术，利用图的拓扑结构，通过聚合结构信息和节点特征实现深层的特征学习。当 GCN 推广到有向图时，基于谱的 GCN 直接将有向图转化为无向图来获得拉普拉斯矩阵，不仅导致信息传递和聚合出现误差，也剥夺了有向图的结构特征。DiGCN（Digraph inception

Convolutional Networks)[36]是一种针对有向图设计的图神经网络框架,它通过 Inception-like 模块并行处理多个尺度的图结构,以捕获复杂的节点间关系。该方法利用个性化 PageRank 算法优化信息传播,确保节点表示能够综合反映各种邻域信息。

更新后的节点嵌入如下:

$$\boldsymbol{Z}=\mathrm{DiGCN}(\boldsymbol{A},\boldsymbol{Z}') \tag{3}$$

式中:$\boldsymbol{Z}\in\mathbb{R}^{n\times d}$。

2.5 排名预测

有很多方法可以实现由嵌入 $\boldsymbol{Z}$ 获得最终的排名,例如机器学习、图算法等。相较于其他方法,图算法更能充分利用嵌入空间中的结构特征,因此本文使用图谱中的 Fiedler 向量法。

首先,由节点嵌入 $\boldsymbol{Z}$ 通过高斯核函数计算相似性矩阵 $\boldsymbol{S}$ 和初始排名向量 $\boldsymbol{r}'$:

$$S_{i,j}=\exp\left(-\frac{\left|\boldsymbol{z}_j-\boldsymbol{z}_i\right|_2^2}{d\sigma^2}\right) \tag{4}$$

$$\boldsymbol{r}'_i=\mathrm{Sigmoid}(\boldsymbol{z}_i\cdot\boldsymbol{a}+b) \tag{5}$$

式中:$\sigma\in\mathbb{R}^d$ 是可训练的参数;$\boldsymbol{a}\in\mathbb{R}^d$ 为可学习向量;$b\in\mathbb{R}$ 为可训练的偏差。

接着,计算对角矩阵 $\boldsymbol{D}$ 以及拉普拉斯矩阵 $\boldsymbol{L}$:

$$D_{i,i}=\sum_j S_{i,j} \tag{6}$$

$$\boldsymbol{L}=\boldsymbol{D}-\boldsymbol{S} \tag{7}$$

Fiedler 向量是拉普拉斯矩阵的第二小的特征向量。寻找 Fiedler 向量可以形式化为求解一个最优化问题:

$$\begin{aligned}&\min_{\boldsymbol{r}}\ \boldsymbol{r}^{\mathrm{T}}\boldsymbol{L}\boldsymbol{r}\\&\text{s. t. }\ \|\boldsymbol{r}\|_2^2=1,\ \boldsymbol{r}^{\mathrm{T}}1=0\end{aligned} \tag{8}$$

式中:$\boldsymbol{r}$ 为 Fiedler 向量。为了简化问题的求解过程,通过旋转问题和约束使超平面成为基数,将一维固定为 0 并在其他维度上求解,最终简化为:

$$\begin{aligned}&\min_{\boldsymbol{y}}\ \boldsymbol{y}^{\mathrm{T}}\boldsymbol{Q}\boldsymbol{L}\boldsymbol{Q}^{\mathrm{T}}\boldsymbol{y}\\&\text{s. t. }\ \|\boldsymbol{r}\|_2^2=1,\ y_1=0\end{aligned} \tag{9}$$

式中:$\boldsymbol{Q}$ 为正交矩阵,$\boldsymbol{y}=\boldsymbol{Q}\boldsymbol{r}$。一种可能的正交矩阵 $\boldsymbol{Q}$ 为上 Hessenberg 矩阵。

使用比例失误损失作为损失函数。这个损失函数考虑了预测分数与输入邻接矩阵之间的比例差异,定义为:

$$\boldsymbol{L}_{\text{upset ratio}}=\frac{\|\boldsymbol{B}'-\boldsymbol{M}\|_{\mathrm{F}}^2}{\mathrm{tr}(\boldsymbol{M})} \tag{10}$$

式中:$\boldsymbol{M}=\dfrac{\boldsymbol{A}-\boldsymbol{A}^{\mathrm{T}}}{\boldsymbol{A}+\boldsymbol{A}^{\mathrm{T}}}$;$B_{i,j}=\dfrac{r_i-r_j}{r_i+r_j}$;$\boldsymbol{B}'$是经过调整的$\boldsymbol{B}$矩阵;$M_{i,j}=0$;$B'_{i,j}=0$;$M_{i,j}\neq 0$;$B'_{i,j}=B_{i,j}$;$\|\cdot\|_{\mathrm{F}}$表示 Frobenius 范数;$\mathrm{tr}(\boldsymbol{M})$为矩阵 $\boldsymbol{M}$ 的迹。

最后,应用近端梯度法来逼近相似性矩阵 $\boldsymbol{S}$ 的 Fiedler 向量,以此作为排名得分向量。

2.6 数据增强

在得到排名过的偏序信息后,如何有效利用这些信息就成为了需要解决的问题。在半监督学习领域,一个有效且常见的方法是为这些数据赋予伪标签从而与标记数据一起进行后续任务。使用模型进行伪标签预测生成是当前常见的方法,但这个方法对模型质量的要求较高,对计算资源的需求较高。

另一种常见的为未标记数据赋予伪标签的方法是聚类。根据文献[37]的理论,k-means[38]方法对有序数据进行聚类效果良好,该方法以质心的形式确定类簇中心,并基于最小化样本到其最近质心的距离作为聚类准则。

但传统的 k-means 方法并不能利用数据间的排序信息,为了充分利用数据间的排序信息,为每个数据设计基于高斯混合的权重,对数据特征进行加权,提高不同类别数据间的相似度,以提高聚类的性能。

高斯混合权重计算公式如下:

$$w_i=\sum_{j=1}^{m}\mathrm{e}^{-\frac{\left(k(i)-\mu_j\right)^2}{2\sigma_j^2}} \tag{11}$$

式中:$k(i)$为样本 x_i 的排序;μ 为高斯分布的中心,代表着数据中重要的中心;σ 为高斯分布的标准差,决定了每个中心的影响范围。μ 和σ 两个参数均由有序信息在不同类别中数量的比例确定。

特征加权公式如下:

$$x_i=w_i\times x_i \tag{12}$$

在本文中,使用欧几里得距离作为距离计算公式:

$$D_{\text{dis}}(x,o)=\sqrt{\sum_{i=1}^{d}(x_i-o_i)^2} \tag{13}$$

式中:o 为簇的质心,初始质心在有序信息的每个类别中随机抽取。

在聚类完成后,为每一个样本 x_i 赋予伪标签 $y_i\in L$,并根据伪标签与有序信息融合,形成新的样本集 $D_1=\{(x_i,y_i)\mid x_i\in X,\ y_i\in L\}$,实现数据增强的效果。

尽管伪标签无法完全代表样本的真实类别,但它们能够在无标注数据上提供一个基础的、基于数据分布的类别近似,为进一步的模型训练与验证提供了便利。

2.7 有序分类

最终得到的新数据集可以用于目前各种先进的

有监督有序分类方法。要估计一个物体的具体值，相比于直接预测，如果有一些参照物会更加容易。例如，估计一个物体的长度，如果有一个 1 m 长的木棒作为参考，这个物体的长度将更加容易估计。

基于此原理，移动窗口回归(MWR)[39]引入相对排名的概念，用于表示输入样本相对于参考样本的排名，这种表示不直接依赖于绝对排名值，而是关注实例间的相对位置关系，并将输入样本的估计设置为下一个窗口的中心，如此迭代直到收敛。

MWR 在不同数据集上的综合表现良好，因此本文使用 MWR 进行最终预测。

3 实验结果和分析

3.1 数据集

为了验证模型框架的有效性，在年龄估计场景下基于公共数据集 FGNET、Adience 和 UTKFace 进行了大量实验。

FGNET 是一个用于年龄估计的经典小型数据集，含有 82 位参与者的 1 002 张人脸图像，这些图像在姿势、表情和光照条件上都有明显的差异。每个人拥有十几张图像，涵盖了 0～69 岁的年龄范围。

Adience 是一个用于年龄估计的大型数据集，包含 2 284 人的 26 580 张人脸图像，每张图像都有 1 个二元的性别标签和 8 个不同年龄(0～2 岁，4～6 岁，8～13 岁，15～20 岁，25～32 岁，38～43 岁，48～53 岁，60 岁以上)中的一个标签，数据集被分成 5 个部分以便进行交叉验证。该数据集设计的主要原则是尽可能贴近真实世界的条件，包括外观、姿势、光照条件、图像质量等各种变化。

UTKFace 是一个年龄跨度非常大的大型数据集，其年龄覆盖范围为 0～116 岁，由 23 708 张人脸图像组成。这些图像涵盖了姿势、面部表情、照明、遮挡、分辨率等方面的巨大变化。

由于 FGNET 数据集中年龄分布较为不均衡，难以准确估计年龄，因此将 FGNET 年龄标签划分为 6 个不同年龄组(0～2 岁，3～10 岁，11～15 岁，16～23 岁，24～39 岁，40 岁以上)，分别赋予 0～5 岁的年龄组 ID 标签以具有自然顺序。而初始 Adience 数据集就拥有 8 个不同的年龄组，分别赋予 0～7 岁的年龄组 ID 标签。接着对 Adience 原始数据集进行数据清洗，过滤年龄为空的数据，并对一些不符合任何年龄组范围的单个年龄数据归入最接近的年龄组。类似的，对 UTKFace 数据集根据年龄标签划分为 14 个不同的年龄组(0～2 岁，3～6 岁，7～12 岁，13～19 岁，20～23 岁，24～27 岁，28～30 岁，31～38 岁，39～45 岁，46～55 岁，56～65 岁，66～73 岁，74～80 岁，81～117 岁)，并赋予 0～13 岁的年龄组 ID 标签。对于 FGNET 数据集，将预处理后的数据集 80%作为训练集，20%作为测试集；对于 Adience 数据集，将预处理后的数据集 80%作为训练集，20%作为测试集；对于 UTKFace 数据集，同样将预处理后的数据集 80%作为训练集，20%作为测试集。所有的特征提取方法均采用在 ImageNet 数据集上预训练的模型 ResNet50 从人脸图片中提取图像特征，并采用五倍交叉验证来评估实验的性能。

然而，这些数据只包括有序信息，处理这些数据的过程如图 2 所示。

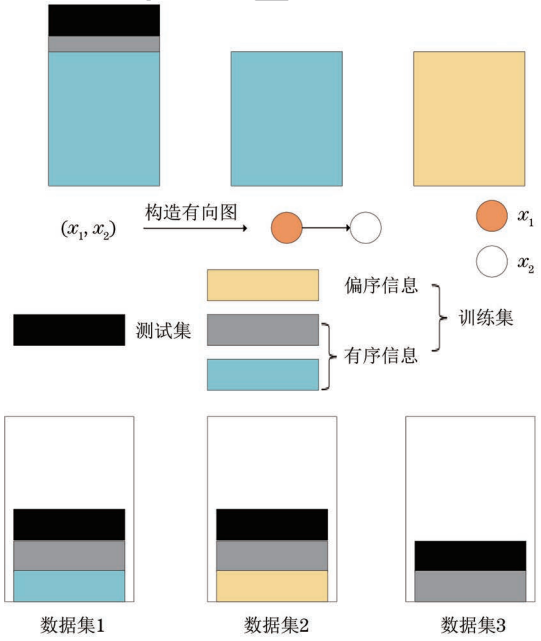

图 2 数据集的划分

Fig. 2 Dataset partitioning

图 2 顶部展示了原始数据集拆分的过程。在原始数据集中划分测试集(黑色部分)以及训练集(其余部分)，此时都为有序信息；接着在整个训练集中的每个类别中抽取固定比例的有序信息(蓝色部分)，用于生成偏序信息(黄色部分)，以满足相同分布假设。详细的生成过程为：随机抽取其中的两个样本并比较它们的类标签，以获取相应的顺序关系。例如抽取到的两个样本 x_1，x_2 的类标签 y_1，y_2 分别为 1、5，则它们类标签之间的顺序关系为 $y_1 \succ y_2$，那么会生成 (x_1, x_2) 这样的偏序信息，反之，如果 $y_1 \succ y_2$，则会生成 (x_2, x_1) 这样的数据对，也就是说，偏序信息位置的前后能够反映出它们之间的偏序关系。

图 2 中间部分描述了由偏序信息如何构建有向图的过程：每一个偏序信息对(x_1,x_2)中都蕴含着偏序关系，偏序信息中的每一个独特的元素构成了有向图的节点集，偏序关系则构成了有向图的边集。

为了验证偏序信息的重要性，针对每个初始数据集，参考文献[20]的方法，生成 3 个不同版本的数据集。图 2 最下方展示了 3 个数据集的构成：3 个数据集中的测试集与处理后数据的测试集保持一致，其余数据被用来生成各自不同的训练集，第 1 个数据集包含全部的有序信息（即原始训练集），第 2 个数据集包含小部分有序信息，其余部分是大量额外的偏序信息，而第 3 个数据集只包含有限的有序信息。数据集统计数据如表 1 所示。

表 1 数据集统计数据

Table 1 Dataset statistics

数据集	版本	有序信息数量/个	偏序信息数量/个	节点数量/个	边数量/条	密度/%
FGNET	F1	1 002	0	0	0	0
	F2	281	50 000	703	50 000	10.13
	F3	281	0	0	0	0
Adience	A1	15 966	0	0	0	0
	A2	4 997	1 500 000	10 498	1 500 000	2.72
	A3	4 997	0	0	0	0
UTKFace	U1	23 708	0	0	0	0
	U2	6 638	1 500 000	16 914	1 500 000	1.48
	U3	6 638	0	0	0	0

3.2 实验设置

本文的评估指标选用在年龄估计中常用的平均绝对误差（MAE，e_{MAE}）和准确率（Accuracy，A_{Accuracy}），以及在聚类方法中常用的轮廓系数（SI，s_{SI}）和方差比准则（CHI，r_{CHI}）。

MAE 表示模型预测值与实际值差异的平均绝对值，衡量的是预测误差的平均大小。MAE 是一个非负数，它的值越小，表示模型的预测准确性越高。MAE 计算公式如下：

$$e_{\mathrm{MAE}}=\frac{1}{n}\sum_{i=1}^{n}|y_i-\hat{y}_i| \tag{14}$$

式中：n 是样本数量；y_i 表示第 i 个样本的实际标签值；$\hat{y}_i$ 是第 i 个样本的预测标签值。

Accuracy 表示模型正确分类的样本占所有样本的比例，通常适用于类别平衡的情况。Accuracy 计算公式如下：

$$A_{\mathrm{Accuracy}}=\frac{m}{n} \tag{15}$$

式中：m 为正确预测的数量。

SI 结合内聚度和分离度两种因素，其值范围为[−1，1]，越接近 1 表示聚类效果越好。SI 的定义如下：

$$s(i)=\frac{b(i)-a(i)}{\max\{a(i),b(i)\}} \tag{16}$$

整个聚类结果的 SI 结果为：

$$s_{\mathrm{SI}}=\frac{1}{n}\sum_{i=1}^{n}s(i) \tag{17}$$

式中：$s(i)$是第 i 个对象的 SI；$a(i)$是第 i 个对象与其同一簇中所有其他对象的平均距离；$b(i)$是第 i 个对象与最近的另一个簇中的所有对象的平均距离，表示分离度。

CHI 是另一种聚类评价指标，表示簇间距离与簇内距离的比值，值越高表示聚类效果越好。CHI 计算公式如下：

$$r_{\mathrm{CHI}}=\frac{S_{\mathrm{BCSS}}/(k-1)}{S_{\mathrm{WCSS}}/(n-k)} \tag{18}$$

式中：k 为簇的数量；S_{BCSS} 是每个聚类质心与整体数据质心之间距离的加权平方和；S_{WCSS} 是数据与其各自的聚类质心之间距离的平方和。

本文提出的半监督有序分类框架使用 DiGCN[36]作为框架中的 GNN 模型，采用 k-means 方法对数据进行聚类，将 MWR[39]作为框架中的有序分类模型。由于相关文献中较少讨论基于半监督学习的年龄估计的情况，因此将本文框架与下列有代表性的有监督基线模型以及少量半监督基线模型进行比较。

MWR：使用移动窗口迭代地更新窗口内样本的顺序，针对有序的样本，提出了一种新的回归方法估计样本的标签值。

CNNPOR[40]：使用 CNN 框架自动提取高级特征，并通过传统的反向传播方法学习最优解。

Ord2Seq[6]：将一个有序分类任务分解为一系列递归二元分类任务，从而巧妙地区分相邻类别。

DRFs[41]：将节点连接到 CNN 的全连接层，并通过在该节点联合学习与输入相关的数据和在叶节点联合学习数据抽象来处理异构数据。

SimPLE[42]：通过增强标记数据和非标记数据之间的一致性约束提高分类准确率。

FixMatch[43]：利用模型对高置信度的未标记图像的预测生成伪标签，然后对模型进行训练，使其能够在同一图像的增强版本中预测伪标签。

本文框架使用了两个模型以及 k-means 聚类方法。DiGCN 的 GCN 层数设置为 3，初始学习率为 0.01。MWR 学习率为 0.001，k-means 簇数为 6、8 和 14。使用随机水平翻转和随机裁剪到 224×224 像素的大小来进行数据增强。其他基线模型都使用原文中超参数。所有实验代码都使用 PyTorch 框架实现。

3.3 结果比较

本文框架与基线方法的比较结果如表 2、表 3 所示，每个数据集中指标的最优结果用黑体加粗表示，有监督的基线模型（MWR、CNNPOR、Ord2Seq 和 DRFs）使用 F1、A1、U1 数据集（即原始数据集）以及 F3、A3、U3 数据集（即只有少量标记数据的数据集），本文框架与半监督基线模型（SimPLE、FixMatch）使用 F2、A2、U3 数据集。在半监督基线模型中，并没有采用同样使用偏序信息的其他方法[15, 20-21]，原因之一是应用领域不同，其方法是为低维数据所设计，并没有适配图像等高维数据，影响算法性能。

表 2 不同数据集上不同模型的实验结果（MAE）

Table 2 Experimental results of different models on different datasets (MAE)

数据集	有监督基线方法				半监督基线方法		Ours
	MWR	CNNPOR	Ord2Seq	DRFs	SimPLE	FixMatch	
F1	**0.55**	0.81	0.66	0.78	—	—	—
F2	—	—	—	—	1.10	0.98	**0.93**
F3	**1.19**	1.37	1.33	1.54	—	—	—
A1	0.45	0.55	**0.43**	0.59	—	—	—
A2	—	—	—	—	0.78	0.64	0.60
A3	0.89	0.85	**0.82**	1.06	—	—	—
U1	**0.77**	0.85	1.08	1.12	—	—	—
U2	—	—	—	—	1.70	1.75	**1.66**
U3	**1.87**	2.06	2.15	2.30	—	—	—

表 3 不同数据集上不同模型的实验结果（Accuracy）

Table 3 Experimental results of different models on different datasets (Accuracy) %

数据集	有监督基线方法				半监督基线方法		Ours
	MWR	CNNPOR	Ord2Seq	DRFs	SimPLE	FixMatch	
F1	**60.4**	49.7	58.9	50.8	—	—	—
F2	—	—	—	—	33.5	42.2	**47.0**
F3	**30.9**	25.6	27.1	24.3	—	—	—
A1	62.6	57.4	**63.5**	55.3	—	—	—
A2	—	—	—	—	48.1	49.6	**54.1**
A3	38.5	38.2	**40.8**	36.2	—	—	—
U1	**85.7**	85.1	80.8	81.3	—	—	—
U2	—	—	—	—	67.0	66.3	**70.5**
U3	46.4	**47.9**	45.1	40.8	—	—	—

从表 2 和表 3 中可以看出，在有序信息充足时，有监督模型的表现是最好的，但当有序信息稀缺时，有监督模型的性能则是最差的。其中，在 F3 数据集中，MWR 在数据稀缺情况下表现仍好于其他有监督模型，而有监督基线模型在 A3、U3 数据集与 A1、U1 数据集上有相似的表现。原因之一可能是 F3 数据集数据过小，只要数据能较好地代表总体分布，MWR 使用到的 k 近邻算法仍然可以进行相对合理的预测，而其他方法使用到的深度学习模型相较于 k 近邻算法在数据稀缺情况下更加敏感。

在 F2、A2、U2 数据集中，在相同的有序信息情况下，由于本文框架使用额外的偏序信息，使模型获得更多信息，表现好于半监督基线模型，并且取得与 F1、A1、U1 数据集中基线模型接近的效果，减少了

对有序信息的依赖。这展现出本文框架在有序信息有限情况下的潜力。

3.4 不同排序模型对聚类的影响

为了评估偏序数据恢复全局排名后对聚类效果的影响，如图 2 所示，将两个数据集划分出的数据集 1 的蓝色部分作为原始数据(Baseline)，两个数据集划分出的数据集 2 的黄色部分作为偏序数据，以保证实验中的数据完全相同，实验均对数据特征进行加权，接着再使用 k-means 作为聚类方法。实验结果如表 4 所示。

表 4 不同模型产生排名的聚类结果

Table 4 Clustering results of different models for ranking

数据集	评估指标	Baseline	DIGRAC	SyncRank	SerialRank	DiGCN
FGNET	SI	0.123	0.107	0.091	0.095	**0.145**
	CHI	204.9	202.5	198.7	199.5	**209.9**
Adience	SI	0.242	0.239	0.227	0.233	**0.286**
	CHI	717.4	715.3	701.9	703.6	**791.8**
UTKFace	SI	0.165	0.170	0.181	0.183	**0.200**
	CHI	835.4	936.7	911.3	912.3	**1 008.8**

从表 4 中可以看出，相对于原始数据，经过 DiGCN 模型排名后的数据聚类表现最好，而其他方法特别是 SyncRank、SerialRank 甚至不如原始数据。这显示出相对于传统的方法，使用 GNN 框架模型的排名效果较好。同时也表明数据间的有序性有助于提升聚类效果，但错误的排名反而有可能导致聚类效果的下降。

在评估指标方面，FGNET、Adience 与 UTKFace 的 SI 值都比较小(接近 0)，说明数据集都存在聚类边界不明显的情况，有很多点处于聚类边界附近。但 CHI 值都比较大，特别是 Adience 和 UTKFace，表示其聚类类别区分度很高，内部较为紧密，如图 3 所示。这可能是造成 Adience 和 UTKFace 数据集在框架中整体表现比 FGNET 数据集效果好的原因之一。

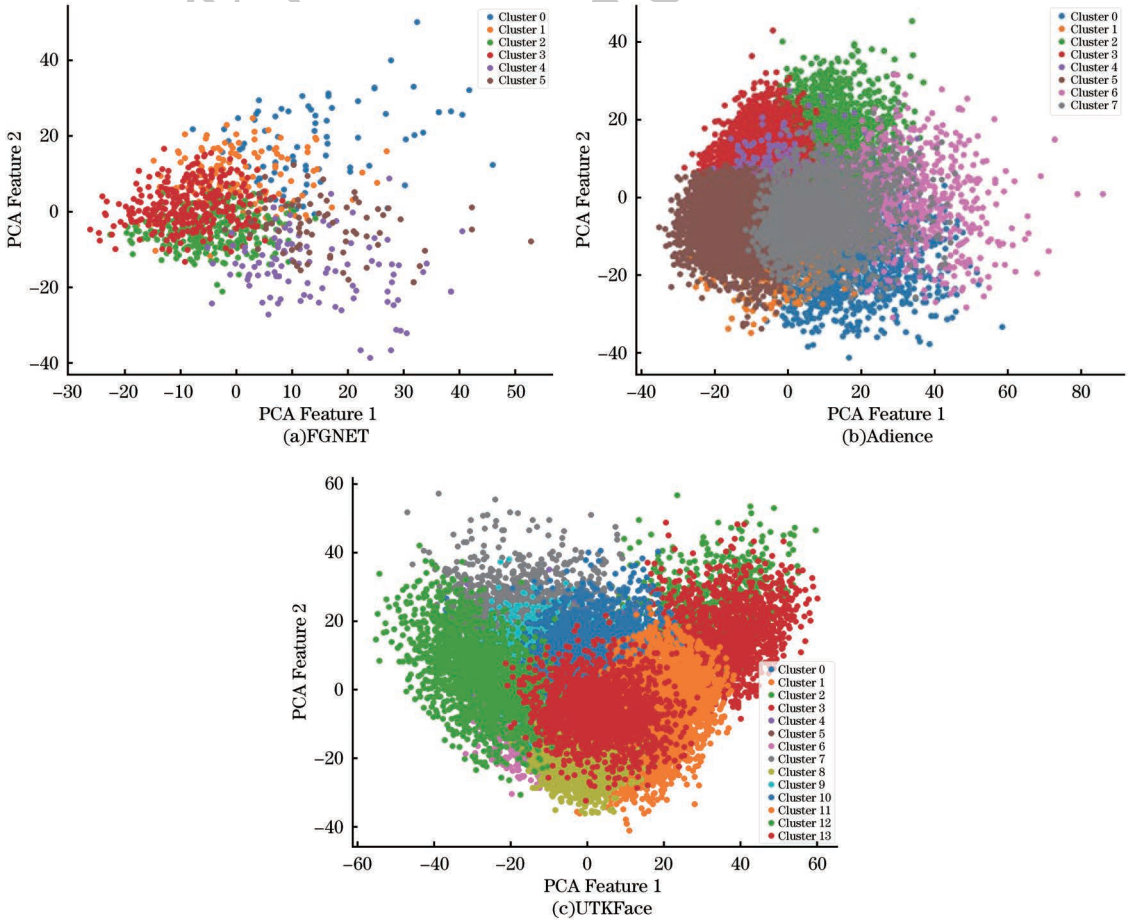

图 3 FGNET、Adience 与 UTKFace 聚类结果

Fig.3 Clustering results for FGNET, Adience and UTKFace

3.5 加入偏序信息的影响

为了评估加入偏序信息对有序分类框架带来的影响，针对每个数据集分别设计了 3 个不同的数据集：数据集 1 仅包含有序信息、无偏序信息；数据集 2 包含少量有序信息和大量偏序信息；数据集 3 仅包含少量有序信息。针对上述 3 个数据集对应进行 3 个实验，实验结果如图 4 所示。

从实验结果可以看出，在不同的数据集上，数据集 1 性能最好，数据集 3 则性能最差，数据集 2 性能处于两者之间，这符合实验的预期。在缺少足量的有序信息情况下，偏序信息的加入可以获得更好的预测效果；同时，也可以看出加入偏序信息对 FGNET 数据集影响更大，这可能是因为相对于 Adience 和 UTKFace 数据集，FGNET 数据更加不平衡，同时规模也更小，导致在有序信息稀缺的情况下，难以获得关于所有类别的足够信息。另外，从相同数据集在不同指标上的表现来看，并不一定 MAE 表现越好，Accuracy 就表现越好，这可能是因为数据本身比较不平衡或者边界模糊，即使 MAE 低，Accuracy 也不一定高。

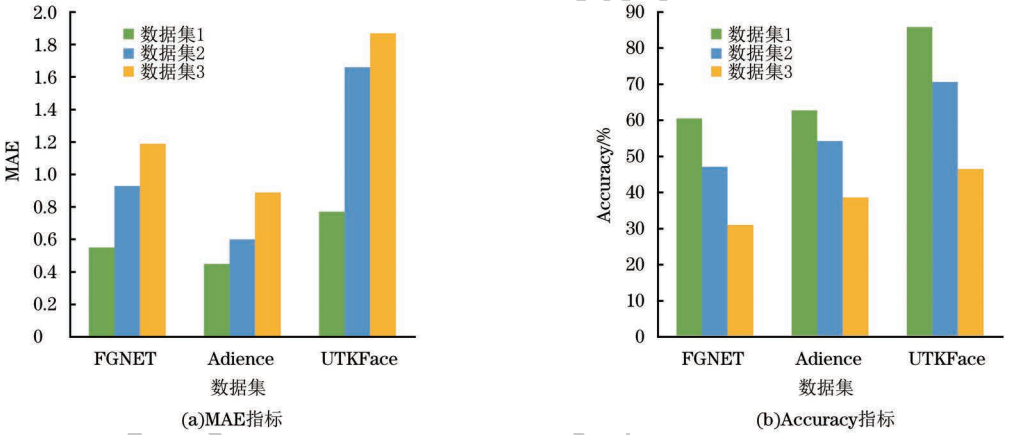

图 4 不同数据集的实验结果

Fig. 4 Experimental results of different datasets

3.6 偏序信息数量的影响

为了评估偏序信息的数量对性能的影响，保持有序信息占训练集比例不变，将偏序信息数量设为参数 n_{num}。根据数据集的不同，在数据集 2 的基础上，分别对 FGNET、Adience 与 UTKFace 设置不同的 n_{num} 值，其中 50 000、1 500 000、1 500 000 为 3 个数据集中偏序信息的初始数量。实验结果如图 5 和图 6 所示。

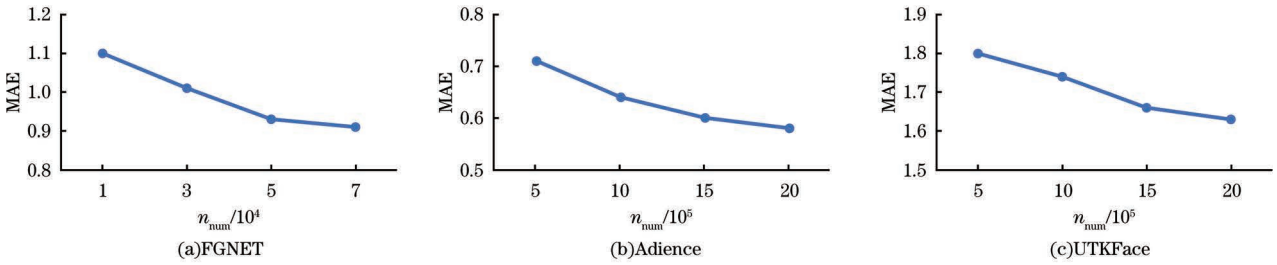

图 5 不同数量偏序信息的实验结果(MAE)

Fig. 5 Experimental results with different amounts of partial order information (MAE)

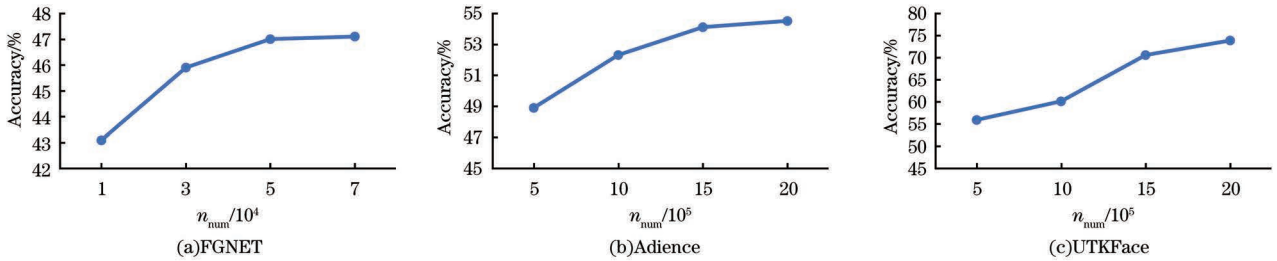

图 6 不同数量偏序信息的实验结果(Accuracy)

Fig. 6 Experimental results with different amounts of partial order information (Accuracy)

从图 5 和图 6 中可以看出，在所有的数据集中，随着偏序信息数量的增加，不同方法的性能都有提升，这符合一般规律。具体来说，随着偏序信息的增多，有向图中节点数量增多，不同节点间的信息表达

更加丰富，充足的信息使得深度学习模型效果更好。从增幅上看，在偏序信息较少时，随着偏序信息的增多，性能增幅较大，而当偏序信息增加到一定程度之后，性能增幅逐渐变小。这说明不是所有的偏序信息对性能提升都是有用的，在偏序信息过多时，可以适当减少偏序信息的数量以加快运行速度。

3.7 有序信息数量的影响

为了评估有序信息的数量对性能的影响，将有序信息样本量占训练集的比例设置为 α，并保持偏序信息的数量不变（如果有的话）。一般来说，在半监督学习中，有序信息越多，能够为不同类别提供的信息越多，有序分类任务效果越好。在数据集 2 的基础上，设置了 4 个不同的 α 值，分别为 10%、30%、50%、100%（其中，10%为初始比例，100%即为数据集 1，即原始数据集），使用这 4 个不同比例的数据集进行实验，结果如图 7 和图 8 所示。

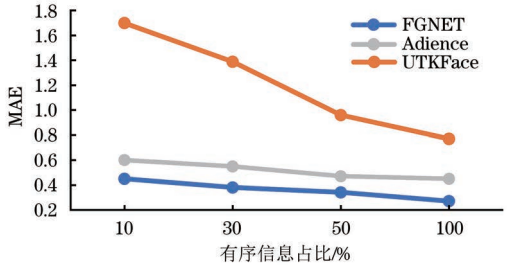

图 7 不同有序信息占比下的实验结果（MAE）

Fig. 7 Experimental results with different proportions of ordered information (MAE)

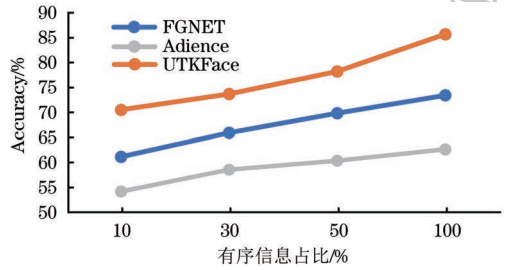

图 8 不同有序信息占比下的实验结果（Accuracy）

Fig. 8 Experimental results with different proportions of ordered information (Accuracy)

从图 7 和图 8 可以看出，在 3 个数据集中，α=10%时，模型表现最差，随着有序信息比例的增加，模型性能也在提高，这与实验预期结果一致。值得注意的是，当 α=50%时，模型效果已经优于 DRFs 或其他基线模型，这意味着在特定情况下，模型性能并不完全依赖于有序信息，偏序信息在一定程度上可以替代有序信息。

3.8 实验总结分析

本实验旨在探究偏序信息在半监督有序分类中的应用。本文假设通过引入偏序信息，可以在有序信息稀少的情况下改善有序分类的性能。

在实验设置中，将数据集拆分成 3 个部分。具体来说，将不同的数据集拆分为仅包含少量有序信息的数据集、包含少量有序信息和大量偏序信息的数据集以及包含大量有序信息的数据集，并进行了一系列实验以验证偏序信息在有序分类中的应用效果。

在各个数据集上的结果均验证了本文的假设，但不同数据集间也有差异性：FGNET 数据集因其规模小、边界不够明显导致其聚类效果差；Adience、UTKFace 数据集在聚类和其他指标中表现良好，体现出了其不同类别间差异性较大、易于区分的特点。

在整个实验过程中，大量的偏序信息可能存在着冗余的情况，导致本文框架性能的下降，并且实验环境的控制条件有限，存在着局限性，可能会影响结果的普遍性。

4 结束语

本文提出了一种融合偏序信息的半监督有序分类框架。该框架使用偏序信息进行排名，在得到全局排名后，利用其排名为数据特征进行加权，接着对其聚类，同时融合有序信息，以此为新数据集进行有序分类。实验结果证明了本文方法的有效性和合理性。下一步工作将引入主动学习方法以探究其在偏序信息中的应用，同时对偏序信息进行数据增强，进一步完善有序分类框架，提高有序分类性能。

参考文献

[1] ZHU F, CHEN X C, CHEN S, et al. Relative margin induced support vector ordinal regression[J]. Expert Systems with Applications, 2023, 231: 120766.

[2] CHU W, KEERTHI S S. Support vector ordinal regression [J]. Neural Computation, 2007, 19(3): 792-815.

[3] BELLMANN P, SCHWENKER F. Ordinal classification: working definition and detection of ordinal structures [J]. IEEE Access, 2020, 8: 164380-164391.

[4] 王雅辉, 钱宇华, 刘郭庆. 基于模糊优势互补互信息的有序决策树算法[J]. 计算机应用, 2021, 41(10): 2785.
WANG Y H, QIAN Y H, LIU G Q. Ordered decision tree algorithm based on fuzzy dominance complementarity mutual information [J]. Journal of Computer Applications, 2021, 41(10): 2785. (in Chinese)

[5] ZHU H P, SHAN H M, ZHANG Y H, et al. Convolutional ordinal regression forest for image ordinal estimation[J]. IEEE Transactions on Neural Networks and Learning Systems, 2021, 33(8): 4084-4095.

[6] WANG J H, CHENG Y, CHEN J T, et al. Ord2Seq: regarding ordinal regression as label sequence prediction[C]// Proceedings of the IEEE/CVF International Conference on Computer Vision. Washington D. C., USA: IEEE Press, 2023: 5865-5875.

[7] CRAMMER K, SINGER Y. PRanking with ranking [J]. Advances in Neural Information Processing Systems, 2001, 14: 641-647.

[8] SHEN L B, JOSHI A K. Ranking and reranking with

perceptron[J]. Machine Learning, 2005, 60: 73-96.
[9] VARGAS V M, GUTIÉRREZ P A, HERVAS-MARTINEZ C. Cumulative link models for deep ordinal classification[J]. Neurocomputing, 2020, 401: 48-58.
[10] SHI W L, GU B, LI X, et al. Quadruply stochastic gradient method for large scale nonlinear semi-supervised ordinal regression AUC optimization[C]//Proceedings of the AAAI Conference on Artificial Intelligence. [S. l.]: AAAI Press, 2020: 5734-5741.
[11] TSUCHIYA T, CHAROENPHAKDEE N, SATO I, et al. Semisupervised ordinal regression based on empirical risk minimization[J]. Neural Computation, 2021, 33(12): 3361-3412.
[12] CHEN H Y, JIA Y Z, GE J M, et al. Incremental learning algorithm for large-scale semi-supervised ordinal regression [J]. Neural Networks, 2022, 149: 124-136.
[13] STEWART N, BROWN G D A, CHATER N. Absolute identification by relative judgment[J]. Psychological Review, 2005, 112(4): 881.
[14] SADER M, VERWAEREN J, PEREZ-FERNANDEZ R, et al. Integrating expert and novice evaluations for augmenting ordinal regression models [J]. Information Fusion, 2019, 51: 1-9.
[15] TANG M Z, PEREZ-FERNANDEZ R, BAETS B. A comparative study of machine learning methods for ordinal classification with absolute and relative information[J]. Knowledge-Based Systems, 2021, 230: 107358.
[16] LIU T Y. Learning to rank for information retrieval[J]. Foundations and Trends® in Information Retrieval, 2009, 3(3): 225-331.
[17] FOGEL F, D'ASPREMONT A, VOJNOVIC M. Spectral ranking using seriation[J]. Journal of Machine Learning Research, 2016, 17(88): 1-45.
[18] 张会影, 圣文顺. 基于标记适应的人脸年龄识别优化算法[J]. 计算机工程, 2025, 51(1): 174-181.
ZHANG H Y, SHENG W S. Improved algorithm for facial age recognition based on label adaptation[J]. Computer Engineering, 2025, 51(1): 174-181. (in Chinese)
[19] TANG M, PÉREZ-FERNÁNDEZ R, BAETS B. Combining absolute and relative information with frequency distributions for ordinal classification [C] // Proceedings of the 18th International Conference on Information Processing and Management of Uncertainty in Knowledge-Based Systems. Lisbon, Portugal: Springer International Publishing, 2020: 594-602.
[20] TANG M, PÉREZ-FERNÁNDEZ R, BAETS B. Ordinal classification with a spectrum of information sources[J]. Expert Systems with Applications, 2022, 208: 118163.
[21] TANG M, PÉREZ-FERNÁNDEZ R, BAETS B. Fusing absolute and relative information for augmenting the method of nearest neighbors for ordinal classification[J]. Information Fusion, 2020, 56: 128-140.
[22] TANG M, PÉREZ-FERNÁNDEZ R, BAETS B. Distance metric learning for augmenting the method of nearest neighbors for ordinal classification with absolute and relative information[J]. Information Fusion, 2021, 65, 72-83.
[23] ARLEGI R, DIMITROV D. Fair elimination-type competitions[J]. European Journal of Operational Research, 2020, 287(2): 528-535.
[24] BRADLEY R A, TERRY M E. Rank analysis of incomplete block designs: I. the method of paired comparisons[J]. Biometrika, 1952, 39(3/4): 324-345.
[25] LUCE R D. Individual choice behavior: a theoretical analysis [M]. [S. l.]: Courier Corporation, 2005.
[26] CARON F, DOUCET A. Efficient Bayesian inference for generalized Bradley–Terry models[J]. Journal of Computational and Graphical Statistics, 2012, 21(1): 174-196.
[27] KENETT D Y, PERC M, BOCCALETTI S. Networks of networks–an introduction[J]. Chaos, Solitons & Fractals, 2015, 80: 1-6.
[28] CUCURINGU M. Sync-rank: robust ranking, constrained ranking and rank aggregation via eigenvector and SDP synchronization[J]. IEEE Transactions on Network Science and Engineering, 2016, 3(1): 58-79.
[29] BACCO C, LARREMORE D B, MOORE C. A physical model for efficient ranking in networks [J]. Science Advances, 2018, 4(7): eaar8260.
[30] D'ASPREMONT A, CUCURINGU M, TYAGI H. Ranking and synchronization from pairwise measurements via SVD[J]. Journal of Machine Learning Research, 2021, 22(19): 1-63.
[31] RIGUTINI L, PAPINI T, MAGGINI M, et al. SortNet: learning to rank by a neural preference function[J]. IEEE Transactions on Neural Networks, 2011, 22(9): 1368-1380.
[32] KÖPPEL M, SEGNER A, WAGENER M, et al. Pairwise learning to rank by neural networks revisited: Reconstruction, theoretical analysis and practical performance[C]//Proceedings of European Conference on Machine Learning and Knowledge Discovery in Databases. Würzburg, Germany: Springer International Publishing, 2020: 237-252.
[33] BURGES C, SHAKED T, RENSHAW E, et al. Learning to rank using gradient descent[C]//Proceedings of the 22nd International Conference on Machine Learning. New York, USA: ACM Press, 2005: 89-96.
[34] MAURYA S K, LIU X, MURATA T. Graph neural networks for fast node ranking approximation[J]. ACM Transactions on Knowledge Discovery from Data, 2021, 15(5): 1-32.
[35] HE Y X, GAN Q, WIPF D, et al. GNNRank: learning global rankings from pairwise comparisons via directed graph neural networks[C]//Proceedings of International Conference on Machine Learning. [S. l.]: PMLR, 2022: 8581-8612.
[36] TONG Z, LIANG Y, SUN C, et al. Digraph inception convolutional networks[J]. Advances in Neural Information Processing Systems, 2020, 33: 17907-17918.
[37] WERRIJ P, KAPTEIN R. Clustering ordinal survey data in a highly structured ranking[C]//Proceedings of Dutch-Belgian Information Retrieval Workshop 2016. New York, USA: ACM, 2016: 1-5.
[38] ALDAHDOOH R T, ASHOUR W. DIMK-means—"distance-based initialization method for k-means clustering algorithm"[J]. International Journal of Intelligent Systems and Applications, 2013, 5(2): 41.
[39] SHIN N H, LEE S H, KIM C S. Moving window regression: a novel approach to ordinal regression [C] // Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Washington D. C., USA: IEEE Press 2022: 18760-18769.
[40] SHI X T, CAO W Z, RASCHKA S. Deep neural networks for rank-consistent ordinal regression based on conditional probabilities[J]. Pattern Analysis and Applications, 2023, 26(3): 941-955.
[41] SHEN W, GUO Y L, WANG Y, et al. Deep regression forests for age estimation [C] // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Washington D. C., USA: IEEE Press, 2018: 2304-2313.
[42] HU Z J, YANG Z Y, HU X F, et al. SimPLE: similar pseudo label exploitation for semi-supervised classification [C] // Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Washington D.C., USA: IEEE Press, 2021: 15099-15108.
[43] SOHN K, BERTHELOT D, CARLINI N, et al. FixMatch: Simplifying semi-supervised learning with consistency and confidence [J]. Advances in Neural Information Processing Systems, 2020, 33: 596-608.

文字编辑 金胡考
栏目编辑 赖玉玲