

一种用于病案相似性度量的弱监督学习算法

张振宇¹, 朱培栋^{1,2}, 赵东升³

- (1. 国防科技大学 计算机学院, 湖南 长沙 410073;
2. 长沙学院 电子信息与电气工程学院, 湖南 长沙 410022;
3. 军事医学科学院 网络信息中心, 北京 100039)

摘要:弱监督机器学习算法解决标签模糊类的问题具有更好的优势,该类算法缓解了数据标签的精度要求。病案的相似性度量就是这类问题,其对医疗数据的应用有着极其重要的基础性作用。鉴于现有病案相似性度量算法通常只基于病理关系的理论规则模型提出,忽略了数据本身包含的信息,文中提出一种弱监督机器学习算法应用于病案相似性度量。该算法首先基于多指标概率分配的方法进行病案组的构建,避免陷入局部最优的情况;然后根据理论模型进行标签赋值,充分利用理论信息;最后通过输入、损失函数、学习模型的分析,从机器学习的角度进行病案的相似性度量。与经典病案相似性度量算法相比,该算法提高了病案相似性度量的准确性,解决了高成本标签的问题。

关键词:弱监督;机器学习;病案相似性;理论模型

中图分类号:TP301.6

文献标识码:A

文章编号:1673-629X(2019)09-0001-06

doi:10.3969/j.issn.1673-629X.2019.09.001

A Weakly Supervised Machine Learning Algorithm Applied to Similarity Measure of Medical Records

ZHANG Zhen-yu¹, ZHU Pei-dong^{1,2}, ZHAO Dong-sheng³

- (1. School of Computer, National University of Defense Technology, Changsha 410073, China;
2. School of Electronic Information and Electrical Engineering, Changsha University, Changsha 410022, China;
3. Network Information Center, Academy of Military Medical Sciences, Beijing 100039, China)

Abstract: The weakly supervised machine learning algorithm has a better advantage in solving the label fuzzy class problem, which alleviates the accuracy requirements of data labels. The similarity measure of medical records is such a problem, which plays a significant role in medical applications. Given that existing medical records similarity measurement algorithm is usually based on the theoretical rule model under pathological relationship, this method ignores the information of the data itself. We propose a weakly supervised machine learning algorithm applied to the similarity measure of medical records. To start with, a medical record group is constructed based on a multi-index probability allocation method to avoid local optimal problems. In addition, the label assignment is conducted according to the theoretical model, which makes full use of the theoretical information. At last, through the analysis of input, loss function and learning model, the similarity measure of medical records is carried out from the perspective of machine learning. Compared with classical medical records similarity measurement algorithm, the algorithm proposed improves the accuracy and solves the problem of high cost labels.

Key words: weak supervision; machine learning; medical records similarity; theoretical model

0 引言

随着数字技术在医学领域的应用, PACS 和 HIS 系统的广泛采用, 医院和医学中心每天都产生和处理大量的医学数据。数据的爆炸式增长给人们的生活带来了多方面影响, 一方面伴随着数据的急剧增多, 给数

据的存储、传输、应用等带来了技术难题和挑战; 另一方面, 数据的大量生产为多种应用服务的出现、人们生活质量的提高提供了可能。病案相似性度量就是这样的应用, 其对于理解病案间的关系、识别病案的聚类情况及预测病人病情的发展趋势有着极其重要的基础性

收稿日期: 2018-10-24

修回日期: 2019-02-26

网络出版时间: 2019-04-24

基金项目: 国家自然科学基金(61572514); 长沙市科技计划重点项目(K1705007)

作者简介: 张振宇(1990-), 男, 硕士研究生, 研究方向为机器学习; 朱培栋, 教授, 博导, 研究方向为网络科学; 赵东升, 研究员, 博导, 研究方向为医疗大数据。

网络出版地址: <http://kns.cnki.net/kcms/detail/61.1450.TP.20190424.1047.040.html>

作用。

病案相似性的度量方法有两类:基于理论知识的传统理论模型^[1]和基于病案数据的机器学习模型^[2]。传统理论模型从医学领域知识出发,通过病理分析判断病案之间的相似性大小关系,这种模型的优点在于解释性好,少量疾病之间的相似性度量精度高,缺点在于对专业知识要求较高,同时受到专业领域知识的限制,模型精度提升难度大。基于病案数据的机器学习模型从病案数据本身出发,通过对大量已形成关系的病案数据进行分析学习,进而学习到其中的相似性关系。该模型的优点在于精度可能与数据量正相关,不受限于领域知识,缺点在于解释性不好。

现有的一些病案相似性度量算法大都是基于理论知识的理论模型^[3]。针对该类算法的不足,文中从病案数据的实际情况出发,综合传统理论模型和机器学习模型,传统理论模型完成弱标签的工作,机器学习模型进行病案相似性的学习,充分利用各模型的优势,对病案的相似性进行度量学习。

1 问题描述及建模

1.1 问题描述

对病案进行精准的描述是度量相似性的前提。在文中,病案被描述为由属性特征构成的向量: $\mathbf{b}_i = (a_1^i, a_2^i, \dots, a_n^i, s_i)$ 。其中, \mathbf{b}_i 表示病案 i ; a_1^i 表示病案的属性特征; s_i 表示病案的特殊特征—诊断信息。

病案数据的病案属性特征具有大量性、异源性和动态权的特点。其中,大量性表现在病案数据属性特征多;异源性表现在病案数据属性特征来源不同,使得特征权重未知、共线性大、特征量纲不统一等;动态权表现在病案数据的属性特征在不同类别的比较中有着不同的权重。

两病案之间的相似性其实是病案对应的向量之间的距离。刻画两向量之间距离的方法有很多,比如:欧氏距离、曼哈顿距离、汉明距离、信息熵、相关系数等^[4]。这些距离计算方法都需要向量各维度等权或者权重已知,然而,基于上述病案数据的特点,这恰恰是病案数据无法满足的。因此,文中从机器学习的角度对病案向量之间的距离进行学习表示。其模型表示如下:

$$S_{ij} = f(\mathbf{b}_i, \mathbf{b}_j) \quad (1)$$

其中, S_{ij} 表示病案 i 和病案 j 之间的距离; \mathbf{b}_i 表示病案 i , \mathbf{b}_j 表示病案 j ,函数 f 表示从病案到病案相似性值的映射,函数 f 即为所求。

1.2 问题的模型方案设计

根据前文对问题模型化的描述,文中模型的输入为病案组 \mathbf{b}_i 和 \mathbf{b}_j ,经过模型的映射,得到模型输出为

病案的相似性值 S_{ij} 。采用传统理论模型和机器学习模型相结合的方法:传统理论模型对病案数据进行相似性标签的工作,将病案无标签数据转化为弱标签数据;机器学习模型在弱标签病案数据的基础上进行学习^[5],通过训练得到模型,如图1所示。

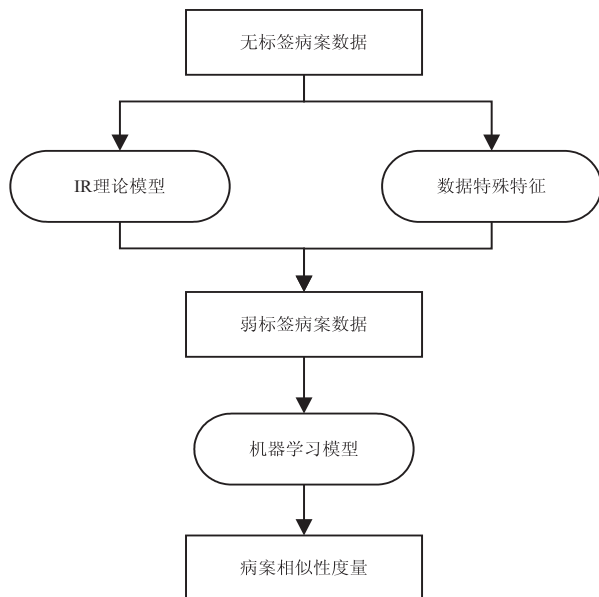


图1 病案相似性模型流程

病案数据一般不存在大量相似性标签,使得用机器学习算法进行病案相似性学习时无法使用监督学习方法,这对提高算法上限有一定的限制。因此,文中通过机器学习与传统理论模型相结合的方式进行。如图1所示,首先,将无标签病案数据转化为有标签病案数据,提出“伪标记”的思想,利用现有的理论模型方法以及数据特征的标记创建一个弱注释的训练数据集,传统理论模型是基于病理推导进行,数据特征标记是对特殊的病案特征属性的标签化,文中采用诊断数据;然后,对弱标签病案数据进行机器学习算法的训练,进而得到病案相似性度量的方法。

2 病案相似性度量

2.1 病案标签

病案标签的设计其本质是将无标签病案数据转化为弱标签病案数据,原始病案数据只包含病案属性特征以及特征对应的特征值,标签病案数据应该增加病案间相似性特征及特征值。由此,病案标签设计就分为两个步骤:病案组的设计和病案组的标签。病案组的设计主要讨论如何合理且均匀地将病案两两成组;病案组的标签主要讨论如何对所有的病案组进行标签值的赋予。

2.1.1 病案组的设计

病案组的设计是将病案两两成组的过程,确定哪些病案组成病案组对病案相似性学习有重要影响。从

逆向的角度考虑:假设病案主要分为两类,即 A 和 B,其中同处于 A 类或 B 类中的病案相似性较高,分别处于 A 类和 B 类中的病案相似性较低,病案组集合需要同时有都处于 A 类的病案组、都处于 B 类的病案组、分别处于 A 类和 B 类的病案组。同时,需要考虑数据的均衡性问题。

病案真正的分类情况是未知的,很难准确判定所有病案组的分布情况,如何避免该问题是病案组设计需要重点考虑的。文中设计一种选排的病案数据集设计方法。病案的选排组合方法需要选择不同的病案组,所以该方法的关键问题就是选择病案组的标准,病案组的选择标准在病案的分布上体现在三个方面:病案组分布的完整性,病案的分布具有社区性,表示各种病案类别之间需要一定数量的病案成组;病案组分布的均匀性,表示各种病案类别之间病案成组的数量与病案数量正相关;病案组分布的边界性,体现在病案形成的社区的边界上,病案社区靠近边界的病案形成的病案组具有更好的边界性。

病案组分布的性质都是通过病案之间的距离体现的。由于没有精确度量病案之间距离的方法,文中提出类似于遗传算法中的“遗传选择”的思想—多指标概率分配的方法进行文档对的选择^[6]。具体就是,从数据的角度进行病案距离的量化,由于距离量化的低精度性,选取多种量化方法^[7],对每种量化方法都进行选择,这样可以一定程度上避免最优解消失的问题,具体设计如表 1 所示。

表 1 病案组选择

距离度量	选择概率
欧氏距离	a_1
余弦距离	a_2
杰卡德距离	a_3
调整余弦距离	a_4
随机生成	a_5

其中,每种距离的度量方式不一样,选择概率表示病案组集合中对应的病案组数量,所有选择概率之和为 1,即有 $a_1 + a_2 + a_3 + a_4 + a_5 = 1$ 。

选择概率是病案组数据集分布情况的关键参数。基于 K-means 算法中初始聚类中心选择的思想,文中提出一种迭代优化的操作对选择概率参数进行优化^[8]。随机初始选择概率,然后进行训练学习,通过每次学习后的网络模型进行病案相似性测试,通过测试损失函数对选择概率参数进行优化,进而迭代求解。其实质就是把选择概率作为学习模型的中间参数进行学习求解。

该过程得到的病案数据如下:

$$b_i = (a_1^i, a_2^i, \cdots, a_n^i, s_i) \tag{2}$$

$$Z' = \{(b_1^1, b_2^1), (b_1^2, b_2^2), \cdots, (b_1^t, b_2^t)\} \tag{3}$$

其中, b_i 表示病案 i ; a_j^i 表示病案 i 的第 j 个特征; s_i 表示病案 i 的诊断特征; Z' 表示不带标签的病案组集合,共有 t 个病案组,每个病案组包含两个病案。

2.1.2 病案组的标签

病案组的标签是对该病案组中两个病案之间的相似性的度量,用相似性值表示。文中将病案的相似性值限定在 0 和 1 之间,当病案组对应的标签为 1 时,表示两病案完全一样;当病案组标签为 0 时,表示两病案完全不同。病案组标签越大,表明两病案的相似性越高。

从传统 IR 理论模型—BM25 算法的思想出发^[9],结合病案数据中的特殊属性,设计一种病案组相似性弱标签的生成方法。BM25 算法是一种用于文本相似度分析的检索算法:有一个查询文档 Q 和一批文档 D,先对文档 Q 进行切分,得到单词集合 q,然后每个单词的分数由三部分组成:单词 q 与文档 D 之间的相关性、单词 q 与查询 Q 之间的相关性、每个单词的权重。最后对于每个单词的分数求和,就得到了查询 Q 和文档 D 之间的相似性分数,该算法很好地解决了文本相似性的度量问题。对文中病案,病案特征可类比于 BM25 算法中的单词;特征值可类比于 BM25 算法中单词间的相关性。由前文对 BM25 算法的描述,需要得到特征的权重,才能通过该模型进行病案标签的设计。

文中设计一种基于数据稳定性和特殊特征相关性的权值拟合方法^[10]。其中,数据稳定性方面,基于信息理论中后验概率的思想提出,这是一种通过结果讨论条件的分布情况的思想,结果集在一定程度上相对稳定,故对结果形成具有影响的因素的突变性应该受到限制,所以,对结果具有较大影响的因素相对稳定;特殊特征相关性方面,病案中诊断数据属于特殊特征,该特征和病案之间具有极大的相关性,因此,所有病案的诊断值向量与特征向量之间的夹角大小是对特征权重的部分表征。

根据该原理,病案数据表示如下:其中有 m 个病案,每个病案有 $n + 1$ 个特征,最后一个特征 s 为特殊特征—诊断信息。

$$B = \begin{bmatrix} a_1^1 & a_2^1 & \cdots & a_n^1 & s_1 \\ a_1^2 & a_2^2 & \cdots & a_n^2 & s_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_1^m & a_2^m & \cdots & a_n^m & s_m \end{bmatrix} \tag{4}$$

每个特征对应的权重为:

$$\hat{y}_i = \frac{\vec{A}_i \cdot \vec{S}}{|\vec{A}_i| \cdot |\vec{S}|} \cdot \sigma_i \quad (5)$$

$$y_i = \frac{\hat{y}_i}{\sum_{j=1}^n \hat{y}_j} \quad (6)$$

其中, \vec{A}_i 表示第 i 个特征向量; \vec{S} 表示诊断信息特征向量; σ_i 表示第 i 个特征向量的方差; y_i 为归一化之后第 i 个特征的权重。

综上所述, 基于 BM25 算法的思想, 文中分别通过特征之间的相似性以及特征对应的权重来表征病案之间的相似性:

$$s_{ij} = 1 - \sum_{x=1}^n \left(\frac{|a_x^i - a_x^j|}{a_x^{\max} - a_x^{\min}} \cdot y_x \right) \quad (7)$$

其中, s_{ij} 表示病案 i 与病案 j 之间的相似性; y_x 为归一化之后的第 x 个特征的权重; a_x^{\max} 、 a_x^{\min} 表示第 x 个特征的最大值与最小值。

病案组数据集进行标签化, 该过程得到的病案数据如下:

$$Z = \{ (b_1^1, b_2^1, s(b_1^1, b_2^1)), (b_1^2, b_2^2, s(b_1^2, b_2^2)), \dots, (b_1^t, b_2^t, s(b_1^t, b_2^t)) \} \quad (8)$$

其中, Z 表示带标签的病案组集合, 共有 t 个病案组, 每个病案组包含两个病案和标签值。

2.2 病案相似性学习

前文对病案数据进行了分析处理, 得到了带标签的病案组的数据集, 其中标签表示病案组对应病案之间的相似性。文中通过机器学习的方法对病案间的相似性进行训练学习, 主要通过损失函数的分析、学习算法的选择进行设计^[11]。

2.2.1 模型的损失函数

机器学习中模型的损失函数是训练过程中对误差

的表示, 准确的损失函数能够精确反映学习模型的不足, 进而通过对应的参数调整对误差进行修正, 因此, 损失函数的选择对学习模型的效率具有很大的影响。文分别设计三种损失函数, 对每种损失函数的效果进行验证分析。

(1) 评分损失函数。

评分损失函数是最直接的损失函数, 模型输入为两病案数据, 输出为相似性评分值, 损失函数通过预测评分值与标签评分值之间的绝对值表示, 模型结构如图 2 所示。

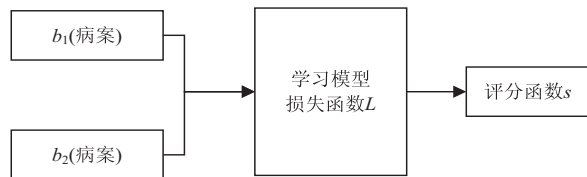


图 2 评分损失函数结构

这种损失函数的设计实际上是一个预测“病案组”的检索分数的逐点排序模型, 这种结构的目标就是掌握一个“分数功能”, 这一功能能够决定一个“病案组”的检索分数。损失函数的表达式如下:

$$L(\theta) = \frac{1}{t} \sqrt{\sum_{i=1}^t [s(b_1^i, b_2^i, \theta) - s(b_1^i, b_2^i)]^2} \quad (9)$$

其中, $L(\theta)$ 表示损失值; t 表示病案组的数量; $s(b_1^i, b_2^i, \theta)$ 表示病案 b_1^i 与病案 b_2^i 的模型预测评分值; $s(b_1^i, b_2^i)$ 表示病案 b_1^i 与病案 b_2^i 的标签评分值。

(2) 排序损失函数。

排序损失函数是比较性的损失函数, 模型输入为三个病案数据, 其中一个查询病案, 两个排序病案, 输出为相似性评分值, 损失函数通过查询病案分别与两排序病案之间预测评分值的高低与标签评分值的高低之间的差异表示。模型结构如图 3 所示。

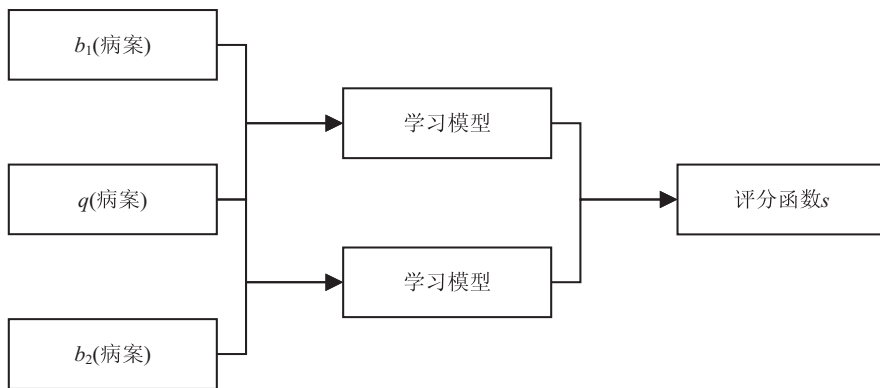


图 3 排序损失函数结构

这种损失函数的设计并不是为了使评分标准化, 而是关注于病案相似性的相对大小, 具体来说就是在训练中使用了两个参数相同的学习模型, 通过两个排序病案与查询病案之间相似性大小的差异表示损失函

数, 表达式如下:

$$L(\theta) = \frac{1}{t} \sum_{i=1}^t \max \{ 0, 1 - \text{sign} [(s(q_i, b_1^i, \theta) - s(q_i, b_2^i, \theta)) \cdot (s(q_i, b_2^i) - s(q_i, b_3^i))] \} \quad (10)$$

其中, $L(\theta)$ 表示损失值; t 表示病案组的数量; $s(q_i, b_1^i, \theta)$ 表示病案 q_i 与病案 b_1^i 的模型预测评分值; $s(q_i, b_1^i)$ 表示病案 q_i 与病案 b_1^i 的标签评分值。

(3) 排序概率损失函数。

排序概率损失函数是概率比较性的损失函数, 模型输入为三个病案数据, 其中一个查询病案, 两个排序病案, 输出为评分比较概率值, 损失函数通过查询病案分别与两排序病案之间预测评分值的高低与标签评分值的高低之间的差异概率表示。模型结构如图 4 所示。

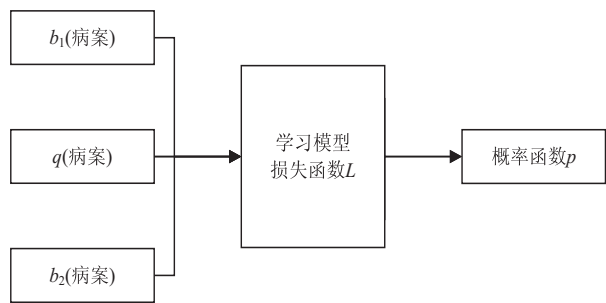


图4 排序概率损失函数结构

这种损失函数的设计是在排序损失函数的基础上考虑概率浮动得到的, 具体来说就是在训练中使用了两个参数相同的学习模型, 通过两个排序病案与查询病案之间相似性大小的差异概率表示损失函数, 表达式如下:

$$L(\theta) = -\frac{1}{t} \sum_{i=1}^t p(q_i, b_1^i, b_2^i) \cdot \log(p(q_i, b_1^i, b_2^i, \theta)) + (1 - p(q_i, b_1^i, b_2^i)) \cdot (1 - \log(p(q_i, b_1^i, b_2^i, \theta))) \tag{11}$$

$$p(q_i, b_1^i, b_2^i) = \frac{s(q_i, b_1^i)}{s(q_i, b_1^i) + s(q_i, b_2^i)} \tag{12}$$

其中, $L(\theta)$ 表示损失值; t 表示病案组的数量; $p(q_i, b_1^i, b_2^i)$ 表示标签评分值下病案 q_i 和病案 b_1^i 间相似性大于病案 q_i 和病案 b_2^i 间相似性的概率; $p(q_i, b_1^i, b_2^i, \theta)$ 表示模型预测评分值下病案 q_i 和病案 b_1^i 间相似性大于病案 q_i 和病案 b_2^i 间相似性的概率。

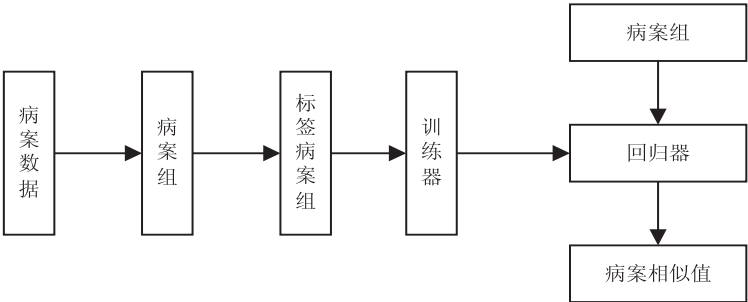


图6 算法流程

从病案信息出发, 得到病案信息中的结构型病案数据; 通过设计的距离度量方法和选择概率从病案数

2.2.2 模型的学习算法

病案数据具有数据维度大、数据数量多的特点, 鉴于病案数据的特征复杂, 文中采用 BP 神经网络算法进行病案相似性度量的训练学习^[12]。

该网络模型由输入层、隐藏层和输出层组成。输入层将病案所有维度完整输入网络; 隐藏层是一个完全连接层网络, 用于病案的特征处理; 输出层的情况取决于模型的损失函数。激活函数根据损失函数的不同选用不同的函数: 评分损失函数模型用线性激活函数^[13]; 排序损失函数用 tanh 函数作为激活函数; 排序概率损失函数用 sigmoid 函数作为激活函数。

学习模型的结构如图 5 所示。

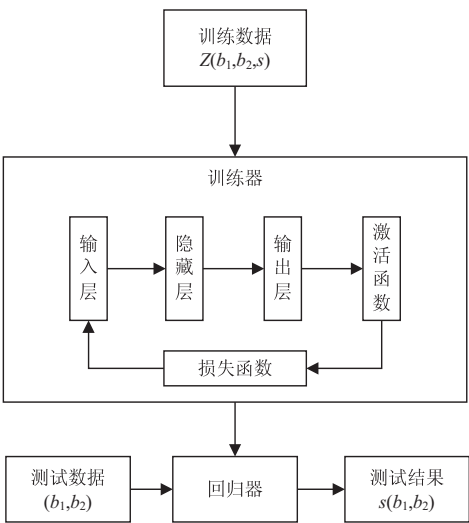


图5 学习模型结构

3 算法描述与分析

3.1 算法描述

基于上文的一系列工作, 从病案数据出发, 对病案进行病案组生成、病案标签设计得到大量带标签的训练数据, 然后通过弱监督学习器的设计对病案相似性进行训练学习得到一个回归器, 该回归器对于输入的两病案能够输出其对应的相似性值。文中给出一种全新的病案相似性度量方法, 算法流程如图 6 所示。

据中得到病案组集合; 基于理论模型的思想结合特殊特征对病案组数据进行标签化工作, 得到带有标签的

病案组集合;将带标签的病案组数据作为训练数据集对设计的机器学习训练器进行训练学习,得到回归器;回归器可以完成病案相似性度量的工作,对输入的病案组计算得到病案相似性值的输出。由此,综合理论模型和机器学习模型得到一种弱监督病案相似性度量方法。

3.2 算法分析

对提出的病案相似性度量算法进行编程实现,并基于病案数据集与各算法进行比较分析。采用医疗数

据中心的实际病案数据和公共数据集 Robust04;评价指标^[14]采用 MAP、P@ 20 和 nDCG@ 20,其中 MAP 表示所有检索病案的平均精度,P@ 20 表示检索的前 20 个病案的平均精度,nDCG@ 20 表示检索的前 20 个病案的累计折扣精度,即其每个病案精度对应权重不同,靠前的病案权重较大;对两类病案数据集进行实验;不同数据集下各算法的 MAP、P@ 20 和 nDCG@ 20 的分布如表 2 所示。

表 2 病案相似性度量精度

方法	Robust04			医疗数据中心		
	MAP	P@ 20	nDCG@ 20	MAP	P@ 20	nDCG@ 20
BM25	0.250 3	0.356 9	0.410 2	0.234 2	0.283 7	0.370 2
RanksSVM	0.242 3	0.350 1	0.399 9	0.240 2	0.293 1	0.389 9
弱监督算法	0.283 7	0.380 2	0.438 9	0.238 7	0.326 7	0.433 0

其中,BM25 为基于理论的模型;RanksSVM 为基于 SVM 的弱监督学习算法;弱监督算法为文中设计的算法,采用排序概率损失函数。通过实验结果可以看到,在不同的数据集中,弱监督算法在各个评价指标下均有较大的优势;在 nDCG@ 20 指标下文中算法优势最大,MAP 指标下文中算法优势相对较小,表明该算法对具有较大相似性的病案组敏感。

4 结束语

随着标签数据的获取成本越来越高,弱监督机器学习算法的应用也越来越广。病案相似性度量就是这样的情况,其对医学研究和医疗应用均有重要作用,而病案相似性标签的获取成本非常高。文中从理论模型和弱标签机器学习模型着手,通过理论模型进行病案数据的弱标签设计,进而将带有弱标签的病案数据作为训练数据对设计的机器学习模型进行训练学习,得到能够度量病案相似性的回归器,该回归器对输入的两病案进行处理得到病案相似性值的输出。通过对比文中算法和理论算法 BM25 以及 RanksSVM 在不同数据集上不同评价指标下的实验结果,表明文中设计的弱标签机器学习算法应用于病案相似性度量时具有较大优势,其对高相似度的病案敏感。

参考文献:

[1] GOPAL A,GORDON J J,KIM J,et al. An electronic medical record search engine for clinical outcome research[J]. International Journal of Radiation Oncology, Biology, Physics,2014,90(1):S817-S818.

[2] PHAM T,TRAN T,PHUNG D,et al. Predicting healthcare trajectories from medical records; a deep learning approach [J]. Journal of Biomedical Informatics,2017,69:218-229.

[3] ZHAO Yawei,MING Yuewei,LIU Xinwang,et al. Large-scale k-means clustering via variance reduction[J]. Neurocomputing,2018,307:184-194.

[4] HURLIN C, LEYMARIE J, PATIN A. Loss functions for loss given default model comparison[J]. European Journal of Operational Research,2018,268(1):348-360.

[5] MOON B. Weighted Hamming metric structures[J]. Discrete Mathematics,2018,341(11):3174-3181.

[6] 辛宇,童孟军,华宇婷. 一种基于最优特征选择改进的遗传算法[J]. 传感技术学报,2018,31(11):1747-1752.

[7] 郑梅. 基于常数复杂度距离函数的推荐算法研究[D]. 成都:西南石油大学,2017.

[8] 徐森. 弱监督多标记学习[D]. 南京:南京大学,2017.

[9] 贾峥,宗瑞杰,段会龙,等. 基于电子病历的患者相似性分析综述[J]. 中国生物医学工程学报,2018,37(3):353-366.

[10] 亢华爱,彭新光. 一种降低误分类代价的权值分布优化算法[J]. 太原理工大学学报,2005,36(4):398-400.

[11] JUNG C,SHEN Yanbo,JIAO Licheng. Learning to rank with ensemble ranking SVM[J]. Neural Processing Letters,2015,42(3):703-714.

[12] GÉRY M,LARGERON C. BM25t;a BM25 extension for focused information retrieval[J]. Knowledge and Information Systems,2012,32(1):217-241.

[13] 关成立,杨岳. PSO 算法优化 BP 神经网络的应用研究[J]. 信息技术与信息化,2018(8):66-68.

[14] 王静疆. 搜索引擎评价指标体系比较研究[J]. 图书情报工作,2008,52(10):136-138.