适用于连续数值标签的兴趣漂移增量学习方法

张 品 蒲菊华 刘永利 熊 璋

(北京航空航天大学 计算机学院, 北京 100191)

摘 要: 针对现有兴趣漂移增量学习方法大多针对包含二值数据标签的用户反馈进行学习的不足,提出了一种适用于连续数值标签反馈的兴趣漂移增量学习方法,以或然概念集的形式描述用户兴趣,将用户反馈中包含的数据标签视为用户对该实例的喜好概率,并采用基于指数近度加权平均的方法对兴趣模型进行增量学习. 在不同学习任务下的实验结果表明,该方法能够在反馈数据标签为连续数值的条件下达到比现有方法更好的学习效果.

关 键 词: 学习算法; 强化学习; 模糊集

中图分类号: TP 391

文献标识码: A 文 章 编 号: 1001-5965(2009) 09-1057-05

Learning drifting user interest incrementally from numerically labeled feedbacks

Zhang Pin Pu Juhua Liu Yongli Xiong Zhang

(School of Computer Science and Technology, Beijing University of Aeronautics and Astronautics, Beijing 100191, China)

Abstract: Most incremental approaches for learning drifting user interests assume that data instances in user feedbacks are binary labeled. A novel incremental learning approach was presented which learns drifting user interests from numerically labeled feedbacks instead of binary labeled ones. User interests were modeled as a set of probabilistic concepts. Numerical instance labels were considered as probabilities that the user likes those instances. Feedbacks were used to update user interest models incrementally based on an exponential recency-weighted average algorithm. Experimental results on different learning tasks showed that the approach outperforms existing approaches in numerically labeled feedback environment.

Key words: learning algorithms; reinforcement learning; fuzzy sets

对用户兴趣的学习是提供个性化服务的Web系统中必不可少的知识获取环节. 由于用户的兴趣会随时间的推移发生变化[1],因此如何以最快的速度和最低的用户参与程度对用户的兴趣漂移进行跟踪学习,就成为了上述Web系统需解决的关键问题. 增量学习方法是被研究者广泛采用的一种适用于兴趣漂移环境的用户兴趣学习方法,这类方法主要来源于概念漂移学习领域,通过利用个性化服务的用户反馈作为带有数据标签的训练数据,跟踪用户兴趣变化,并不断对现有用户兴趣模型进行更新.

在个性化服务应用中,表示用户对个性化内容喜好程度或评分的反馈权值(即数据标签)可能以连续数值的形式存在(如NewsDude[2]),这种连续数值形式的用户反馈通常比传统概念漂移学习场景中的二值反馈(标识实例为正相关或负相关)所反映的兴趣信息更精确. 但现有的绝大部分针对兴趣漂移或概念漂移的增量学习方法(如TIX[3], MTDR[4], FLORA[5]等)均没有在学习过程中考虑数据标签为连续数值的情况,因此当反馈中包含比二值标签更多的信息时,上述方法并不能获得比二值反馈条件下更优的性能.

根据以上分析,本文提出了一种兴趣漂移增量学习方法 PALDI(Probabilistic Approach for

收稿日期: 2008-09-18

基金项目: "十一五"科技支撑计划重大项目(2006BAB04A13)

作者简介: 张 品(1983-),男,内蒙古包头人,硕士生, zhangpin@cse.buaa.edu.cn

Learning Drifting Interest), 该方法针对连续数值形式的用户反馈进行设计, 能够在反馈权值为连续数值的情况下达到比现有方法更好的学习效果; 同时, 该方法的学习效果受到方法参数选取的影响较现有方法更小, 具有更好的适应性.

1 用户兴趣建模

定义 **1** 兴趣概念.

兴趣概念是一个或然概念[6], 对个性化内容实例集合的一个子集 X, 一个兴趣概念可表示为从集合 X 到 $[0, 1]$ 范围内实数集的映射 $c: X \rightarrow [0, 1]$, 即对任一实例 $x \in X$, $c(x)$ 表示 x 与兴趣概念 c 的关联概率或关联程度. 本文中称 X 为 c 的实例域. 一个兴趣概念是对一组个性化内容实例共有特征的抽象表示, $c(x)$ 越接近 1, 表明 x 与 c 表示的抽象特征越接近. 因此可将兴趣概念视为一个分类, 其依一定概率"包含"若干内容实例; 一个内容实例依一定概率"隶属"于一个或多个兴趣概念.

对个性化内容实例集合中的任一实例 x' 及 $c: X \rightarrow [0, 1]$, 若 $x' \in X$, 则称 x' 与 c 相关.

定义 **2** 用户兴趣模型.

设 $C = \{c_1, c_2, \cdots, c_n\}$ 是描述特定用户兴趣的兴趣概念的集合, 令 X_i 表示集合 C 中兴趣概念 c_i 的实例域, 令 $X_C = \bigcup_i X_i$. 对给定 $x \in X_C$, 令 $p_x = \{c_i \mid x \in X_i, c_i \in C\}$ 表示 C 中所有实例域包含 x 的兴趣概念的集合, 则用户对特定个性化内容实例的兴趣可表示为条件概率 $\Pr_{p_x}(I \mid x)$, 其中 I 表示随机事件: "感兴趣", 由于对任一 $c_i \in p_x$, 都有 $x \in X_i$ 成立, 由全概率公式可以得出

$$\Pr_{p_x}(I \mid x) = \sum_{c_i \in p_x} \Pr(I \mid c_i) \times \Pr(c_i \mid x) \tag{1}$$

从式(1)可以看出, 用户对 x 感兴趣的概率取决于用户对 c_i 的喜好概率 $\Pr(I \mid c_i)$, 以及 x 与 c_i 的关联概率 $\Pr(c_i \mid x)$ (即 $c_i(x)$). 由于 $\Pr(c_i \mid x)$ 的大小取决于 c_i 与 x 的特征相似程度, 因此只要记录了每个兴趣概念的特征和喜好概率, 即可通过计算得出用户对具体内容实例的兴趣. 由此可以得出如下定义: 用户兴趣模型 UP 是一个由多个兴趣概念构成的集合, c_i 包含有一个可计算的概念特征 D_i, 以及一个表示用户对该概念喜好概率的兴趣权值 w_i, 即

$$\mathrm{UP} = \{\langle D_1, w_1\rangle, \langle D_2, w_2\rangle, \cdots, \langle D_n, w_n\rangle\}$$

定义 **3** 兴趣学习.

基于上述兴趣模型定义, 本文所指的兴趣漂移增量学习, 是指在兴趣漂移条件下, 利用用户反馈对现有用户兴趣模型中的 w_i 及 D_i ($i = 1, 2, \cdots, n$) 进行增量式的更新. 前述用户反馈是一个由 x 和反馈权值 d 构成的二元组 $\mathrm{UF} = \langle x, d\rangle$, 该二元组表示用户对 x 的喜好概率为 d ($d \in [0, 1]$), 由此可将用户兴趣增量学习过程表示为

$$w_i^{\mathrm{new}} = \delta(w_i^{\mathrm{old}}, \mathrm{UF})$$

及

$$D_i^{\mathrm{new}} = \varphi(D_i^{\mathrm{old}}, \mathrm{UF}) \qquad i = 1, 2, \cdots, n$$

其中, w_i^{old} 与 w_i^{new} 分别为更新前后的条件概率 w_i, 函数 δ 和函数 φ 分别为针对 w_i 和 D_i 的增量学习算法的计算表达式.

定义 **4** 兴趣预测.

在给定用户兴趣模型 UP 的条件下, 根据式(1), 用户对新的个性化内容实例 x'' 的兴趣 $\Pr(I \mid x'')$ 可由下式得出

$$\Pr(I \mid x'') = \sum_{c_i \in p_{x''}} \Pr(I \mid c_i) \times \Pr(c_i \mid x'') = \sum_{i \in [1, n]} w_i \times \mathrm{sim}(D_i, x'') \tag{2}$$

其中, $\mathrm{sim}(D_i, x'')$ 为该内容实例与 c_i 之间的特征相似度.

考虑到个性化服务应用中的大部分内容实例 (Web 页面[7]、电视节目[8]、电影[9]新闻[2, 10]等) 都带有一定数量的文本信息, 因此本文将以文本特征作为个性化内容实例和兴趣概念特征的可计算描述形式. 为便于进行对比实验, 本文采用在其他研究中广泛使用的基于 TF-IDF (Term Frequency-Inverse Document Frequency) 权重的特征词向量[4]作为内容实例和兴趣概念的特征表示方式, 并以余弦相似度[4]作为兴趣概念和内容实例之间特征相似度的度量方法.

2 PALDI学习方法

本文提出的 PALDI 方法的学习算法流程如下:

输入: model 待学习的目标用户模型;

c_i model中的第 i 个兴趣概念;

D_i c_i 的概念特征;

$\{\langle x_1, d_1\rangle, \langle x_2, d_2\rangle, \cdots, \langle x_m, d_m\rangle\}$ 反馈特征序列;

θ 特征相似度阈值, 取值范围为 $(0, 1)$;

M model中的兴趣概念数目阈值.

输出: 更新后的 model

函数: Size(model) 返回 model 中的兴趣概念

数目;

$\delta(w_i^{old}, <x_j, d_j>)$　返回利用反馈 $<x_j, d_j>$ 更新后的 w_i^{old} 的值;

$\Phi(D_i^{old}, <x_j, d_j>)$　返回利用反馈 $<x_j, d_j>$ 更新后的 D_i^{old} 的值.

步骤:

1) For each x_j in $\{<x_1, d_1>, <x_2, d_2>, \cdots, <x_m, d_m>\}$ do

2) let $s = \mathrm{sim}(D_i, x_j)$ such that $\mathrm{sim}(D_k, x_j) = \max_k\{\mathrm{sim}(D_k, x_j)\}$

3) if $s < \theta$ and $\mathrm{Size}(\mathrm{model}) < M$

4) then create a new interest concept c' with $D' = x_j, w' = d_j$

5) For each c_i in model do

6) $w_i^{new} = \delta(w_i^{old}, <x_j, d_j>)$

7) $D_i^{new} = \Phi(D_i^{old}, <x_j, d_j>)$

8) End for

End for

对个性化服务系统获得的每一条用户反馈 $<x_j, d_j>$, 首先计算内容实例 x_j 与模型中所有兴趣概念的特征相似度, 若相似度最大值小于相似度阈值 θ, 即该实例与模型中现有兴趣概念的关联程度均较低, 则建立一个新的兴趣概念以对该实例的特征进行描述; 然后利用该反馈对模型中所有兴趣概念的特征和兴趣权值进行更新. 为避免因内容实例特征分散而导致的兴趣概念数目无限增长[4], 本文引入了 M 以对兴趣概念的最大数目进行约束, 在兴趣概念数目达到该阈值后, 将不再建立新的兴趣概念, 而是用现有兴趣概念对新学习到的内容实例的特征进行描述.

下面详细描述针对 w_i 和 D_i 的增量学习算法(即上文中步骤 6)和步骤 7)的具体计算过程).

2 1　兴趣权值的学习

基于指数近度加权平均方法[11], 本文采用下述表达式对 w_i 进行增量计算:

$$w_i^{new} = \delta(w_i^{old}, <x_j, d_j>) = w_i^{old} + \beta_i^j (d_j - w_i^{old}) \quad (3)$$

其中, d_j 为反馈信息序列中的第 j 条反馈信息中包含的反馈权值; β_i^j 为第 j 条反馈信息对应于条件概率 w_i 的更新率, 取值范围为 $[0, 1]$.

在决定每一内容实例的反馈权值在概念兴趣权值 w_i 中所占的权重时, 不仅应考虑反馈信息的先后顺序, 还应考虑该实例与兴趣概念的关联程度. 同时, 针对不同用户个体的兴趣变化速度和幅度各不相同的情况, 在更新兴趣权值时还应体现对应于不同用户的更新幅度的差别. 本文在确定更新率时, 兼顾了实例-概念关联程度的影响和针对不同用户的更新幅度灵活性要求, 提出了更新率的实例关联程度因子 r_i^x 及更新幅度偏好因子 q, 令 $\beta_i^j = q r_i^{x_j}$. r_i^x 即式 (1)中的条件概率 $\Pr(c_i|x)$, 表示 c_i 与 x 的关联程度. q 记录了用户对其兴趣模型更新幅度的设置, 其取值范围为 $[0, 1]$, 越接近 1的数值表示兴趣模型的增量更新幅度越大, 模型对兴趣漂移的响应越快. 引入 r_i^x 及 q 后的增量学习表达式为

$$w_i^{new} = w_i^{old} + q r_i^{x_j} (d_j - w_i^{old}) \quad (4)$$

其中 x_j 为反馈信息序列中的第 j 条反馈信息中包含的实例.

由于在更新兴趣权值的计算表达式中已经考虑了实例-概念关联程度的影响, 因此在学习过程中可以用上述表达式对兴趣模型中的每个兴趣概念的兴趣权值进行统一方式的更新, 而不用预先设定兴趣概念的更新范围, 从而减少了学习方法中的参数数量, 提高了方法的适应性.

2 2　兴趣概念特征的学习

根据上文定义, 一个兴趣概念可以视为一个模糊集[12], 因此可以采用针对模糊集的类别特征计算方式对兴趣概念特征进行学习. 具体方法如下.

对 c_i: $X_i \rightarrow [0, 1]$, 其概念特征 D_i 可表示为[12]

$$D_i = \frac{\sum_{x_j \in X_i} [c_i(x_j)]^{\gamma} x_j}{\sum_{x_j \in X_i} [c_i(x_j)]^{\gamma}} \quad (5)$$

其中 $\gamma > 1$ 为一个控制模糊集模糊程度的常数.

由式 (5)可以看出, D_i 也可以表示为加权平均的形式, 平均值的分量是内容实例 x_j 各分量的权值是该实例与该兴趣概念的关联程度的幂. 与更新兴趣权值的学习算法相对应, 为了突出最新学习到的实例在决定兴趣概念特征时的作用, 本文将式 (5)中的计算表达式替换为指数近度加权平均方法的表达式:

$$D_i^{new} = \Phi(D_i^{old}, <x_j, d_j>) = D_i^{old} + q(r_i^{x_j})^{\gamma}(x_j - D_i^{old}) \quad (6)$$

3　实验评估

本文实验的目的在于检验和比较 PALDI方法在兴趣漂移条件下对用户兴趣描述的准确程度和对兴趣漂移的响应速度. 本文实验选择 MTDR[4]方法作为对照方法, MTDR 方法是针对多兴

趣类别的兴趣漂移学习方法, 该方法在数据标签为二值标签的条件下能够达到较高的兴趣预测精度, 根据文献[13]中的实验数据, 该方法的学习效果优于 Rocchio及 FLORA等其他兴趣漂移学习方法. 在本文实验中, MTDR方法的参数选取与文献[4]一致. 对于 PALDI方法的参数, 实验中设置 $q=1$, 即使兴趣模型以最快速度跟踪用户兴趣变化; 设置 $\gamma=1.1$, 使兴趣权值和兴趣概念特征的更新速度基本保持一致; 由于参数 θ和 M 在学习过程中的作用与 MTDR方法中的对应参数类似, 因此为提高实验结果的可比性, 故将 θ和 M 设置为与 MTDR方法相同, 即 $\theta=0.175$, $M=8$.

为使 MTDR方法能够处理连续数值标签, 在将反馈数据用于 MTDR方法前, 首先对数据标签 L_j 以 0.5为阈值进行二值化处理, 以得到二值标签 B_j.

3.1 实验数据

本文采用 Reuters-21578数据集的子集 ModApte split作为实验数据集. 为使实验结果具有较好的可比性, 本文按照文献[4]选择的 5个主题(Trade, Coffee, Crude, Sugar, Acq)分别从数据集中抽取与该主题对应的文档各 100篇(80篇训练文档 + 20篇测试文档), 抽取过程保证属于同主题的文档之间具有较高的相似度. 抽取出的文档在实验前都进行了去除中止字、抽取词干、识别词组、生成特征词向量的预处理操作. 抽取出的 500篇文档共同构成实验场景的个性化内容实例集合.

为验证本文方法在连续数值型数据标签条件下的实际效果, 本文采用如下方式生成数据标签:

假设用户对属于特定主题的内容实例的兴趣服从正态分布[14], 对于实验中假设用户“喜欢”的文档主题, 依照分布 $N(0.9, 0.0036)$生成属于该主题的文档的数据标签(即兴趣权值均值为 0.9, 标准差为 0.06); 对于实验中假设用户“不喜欢”的文档主题, 依照分布 $N(0.1, 0.0036)$生成属于该主题的文档的数据标签(即兴趣权值均值为 0.1, 标准差为 0.06). 标签的生成过程保证标签数值位于[0, 1]范围内.

3.2 学习任务

实验选取了难度不同的 3个学习任务对 2种方法的学习效果进行测试, 每个学习任务都被划分为若干学习阶段, 每学习阶段包含 10个评估周期. 每一学习阶段代表用户兴趣的一个稳定阶段, 从一个学习阶段到下一个学习阶段的转换表示一次兴趣漂移的发生. 在每一学习阶段开始前, 都根据该阶段的目标兴趣概念, 采用前述方法为数据集(包括训练集和测试集)中的所有文档生成标签. 在每个评估周期中, 首先采用从训练集中随机选取的 10篇文档(对应于各个主题的文档各 2篇), 用于进行增量学习; 然后在固定的测试集(共 100篇测试文档)中对兴趣模型的准确程度进行检验. 每个学习任务的最终实验数据取 10次学习过程的平均值. 实验采用的学习任务以及每个任务中各学习阶段的目标兴趣概念如表 1所示.

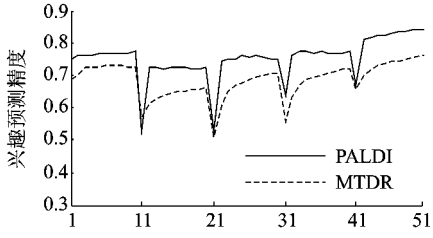
表 1 学习任务及任务中各学习阶段的目标概念

学习任务	学习阶段(评估周期)				
	1(1~10)	2(11~20)	3(21~30)	4(31~40)	5(41~50)
1	Trade	Coffee	Crude	Sugar	Acq
2	Trade, Coffee	Coffee, Crude	Crude, Sugar	Sugar, Acq	
3	Trade, Coffee, Crude	Coffee, Crude, Sugar	Crude, Sugar, Acq		

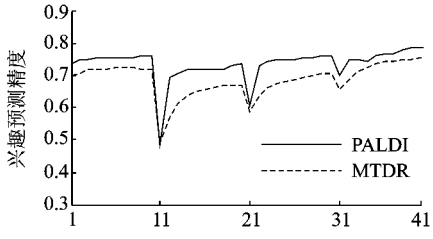
由于引入了连续数值标签, 本文采用标准化平均绝对误差, 即通过比较预测兴趣概率 $\Pr(I|x_j)$和实际反馈权值 d_j, 来评价 2个学习方法的兴趣预测精度.

3.3 实验结果

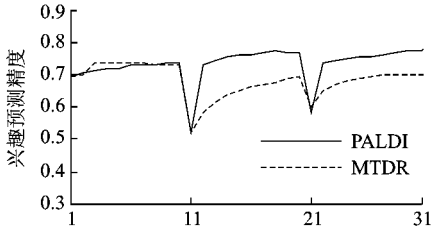
图 1展示了 PALDI方法在不同学习任务中与现有 MTDR方法的兴趣预测精度对比, 从图中可以看出, 在连续数值型反馈条件下, PALDI方法在各个学习任务上都达到了优于 MTDR方法的预测精度和兴趣漂移响应速度, 这说明 PALDI针对连续数值反馈进行的学习方法设计是有效的. 对比图 1a~图 1c, 可以看出 PALDI方法的预测精度在学习任务 1, 2, 3中依次下降, 这也体现出了学习任务的不同难度对学习效果的影响; 但在同一学习任务中, PALDI方法的预测精度并没有随着兴趣漂移次数的累积而出现明显的下降, 一定程度上说明该方法能够在用户兴趣不断变化的环境中持续为个性化服务提供兴趣学习支持, 从而使个性化服务系统在兴趣漂移条件下能够获得相对稳定的兴趣预测精度.



a 学习任务 1



b 学习任务 2



c 学习任务 3

图 1 兴趣预测精度对比

4 结束语

本文提出了一种适用于连续数值型反馈数据的兴趣漂移增量学习方法 PALDI,其以或然概念集的形式描述用户兴趣,将用户反馈中包含的数据标签视为用户对该实例的喜好概率,并采用基于指数近度加权平均的方法对兴趣模型进行增量学习.实验表明,PALDI方法在各个学习任务上都达到了优于现有方法(MTDR)的预测精度和兴趣漂移响应速度.

参考文献 (**References**)

[1] Koychev I, Schwab I. Adaptation to drifting user's interest [C]//ECML Workshop: Machine Learning in New Information Age. Barcelona, Spain [s.n.], 2000: 39-45

[2] Billsus D, Pazzani M J. A hybrid user model for news classification[C]//The Seventh International Conference on User Modeling. Wein, New York: Springer-Verlag, 1999: 99-108

[3] Forman G. Tackling concept drift by temporal inductive transfer [C]//The 29th Annual International ACM SIGIR Conference. Seattle, Washington: Springer-Verlag, 2006: 252-259

[4] Widyantoro D H, Ioerger T R, Yen J. Learning user interest dynamics with a three-descriptor representation[J]. J Am Soc Information Science, 2001, 52(3): 212-225

[5] Widmer G, Kubat M. Learning in the presence of concept drift and hidden contexts[J]. Machine Learning, 1996, 23(1): 69-101

[6] Kearns M J, Schapire R E. Efficient distribution-free learning of probabilistic concepts[C]//31st Annual Symposium on Foundations of Computer Science. St Louis, MO: IEEE Computer Society Press, 1990(1): 382-391

[7] Ji J, Liu C, Yan J, et al. Bayesian network structure learning and its application to personalized recommendation in a B2C portal[C]//International Conference on Web Intelligence. Beijing: Computer Society, 2004: 179-184

[8] Kamahara J, Asakawa T, Shimojo S, et al. A community-based recommendation system to reveal unexpected interests[C]//11th International Multimedia Modelling Conference. Melbourne, Australia: IEEE Computer Society, 2005: 433-438

[9] Good N, Schafer J, Konstan J, et al. Combining collaborative filtering with personal agents for better recommendations[C]//Sixteenth National Conferenceon Artificial Intelligence. Orlando, Florida: MIT Press, 1999: 439-446

[10] Singh S, Shepherd M, Duffy J, et al. An adaptive user profile for filtering news based on a user interest hierarchy[C]//69th Annual Meeting of the American Society for Information Science and Technology. Austin, Texas: Wiley, 2006(43): 1-27

[11] Kim C O, Kwon I H, Baek J G.. Asynchronous action-reward learning for nonstationary serial supply chain inventory control [J]. Applied Intelligence, 2008, 28(1): 1-16

[12] Mika S, Jain L C. Introduction to fuzzy clustering [J]. Innovations in Fuzzy Clustering, 2006, 205: 1-8

[13] Hendrato W D. Concept drift learning and its application to adaptive information filtering[D]. Texas: Texas A&M University, 2003

[14] Hofmann T. Collaborative filtering via gaussian probabilistic latent semantic analysis[C]//26th Annual International ACM SIGIR. New York: Association for Computing Machinery, 2003: 259-266

(责任编辑:赵海容)