基于项目分类的协同过滤改进算法 *

熊忠阳, 刘 芹, 张玉芳, 李文田
(重庆大学 计算机学院, 重庆 400044)

摘 要: 为了解决用户评分数据稀疏性和用户最近邻寻找的准确性问题,提出了一种基于项目分类的协同过滤推荐改进算法。该算法首先利用项目分类信息为类内未评分项目预测评分值;然后通过计算类内用户间的相似度得到目标用户的最近邻居;最后进行推荐。实验结果表明,该算法可以准确地获取用户兴趣最近邻,有效地解决数据稀疏性问题;同时,该算法还极大地提高了系统的工作效率及可扩展性。

关键词: 项目分类; 协同过滤; 评分预测; 兴趣最近邻; 推荐系统
中图分类号: TP311 **文献标志码**: A **文章编号**: 1001-3695(2012)02-0493-04
doi:10.3969/j.issn.1001-3695.2012.02.025

Improved algorithm of collaborative filtering based on item classification

XIONG Zhong-yang, LIU Qin, ZHANG Yu-fang, LI Wen-tian
(*College of Computer Science*, *Chongqing University*, *Chongqing* 400044, *China*)

Abstract: To overcome the drawbacks caused by the data sparseness and inaccurate of the user neighbors, this paper came up with an improved collaborative filtering recommendation algorithm, basing on the technique of item classification. The algorithm first rated the unrated items by applying the item classification, and then calculated the user similarity within classes for nearest-neighbors, after which it could recommend the items based on the final prediction. Experimental results show that this algorithm can not only improve the accuracy of nearest neighbor search, but also increase the efficiency and scalability of the system.

Key words: item classification; collaborative filtering; rating predication; interest nearest neighbors; recommendation systems

0 引言

随着互联网和电子商务系统的迅猛发展,个性化推荐系统越来越重要。个性化推荐系统不仅能够快速地发现用户兴趣并作出推荐,还能发现用户潜在兴趣,提高电子商务网站的销售量。而最近邻协同过滤(collaborative filtering)是目前应用最广泛最成功的推荐技术之一[13]。

另一方面,随着用户数目和商品数量的不断增加,协同过滤算法的数据稀疏性问题及其带来的用户相似性度量不准确问题也越来越突出,这直接导致了系统的推荐质量迅速下降。针对数据稀疏性问题,提高系统推荐效率,研究者们提出了基于项目[47](包括填充均值[6]和评分预测[5,7])、基于用户[1,2,8,9]、结合用户和项目[3,7,10,11]、基于聚类[1,10,12]、基于维度化简[13]等多种协同过滤推荐算法。同时,为提高系统的相似性度量精度,Pearson 相关系数法[14]、Cosine 相似度法[6]、LHN-I 指标法[15]、基于权重的相似性[16]等计算方法也陆续被提出。另外,利用修正的条件概率方法计算项目相似性[4];采用基于用户兴趣度的相似性计算方法[8];基于项目类别的相似性方法[2]和结合共同邻居及用户评分的相似性度量[17]方法也被研究者们提出以改进相似度计算的准确性。

以上方法虽从不同的角度克服了数据稀疏性及传统相似度计算带来的弊端,但并没有准确地获取用户兴趣,从而导致推荐效果仍不理想。若想提高系统的推荐质量,推荐系统所采用的推荐算法必须能够准确获取用户的兴趣邻居。如何准确获取用户的兴趣邻居就成为提高推荐系统质量的关键。为此,本文提出了一种利用项目分类信息对类内项目进行评分预测,然后在类内寻找用户最近邻居的协同过滤改进算法。

1 传统的协同过滤算法

传统的协同过滤算法一般分为三个步骤:a)数据表述。通常是获得一个 $m \times n$ 的用户—项目评分矩阵,m 行代表用户数,n 列代表项目数,矩阵元素 $R_{i,j}$ 表示用户 i 对项目 j 的评分值。b)发现 k 最近邻。根据用户—项目评分矩阵计算用户或项目的相似度,按照相似度从大到小为当前用户或项目求得一个最近邻集合。c)产生推荐数据集。用户获得 k 最近邻后,可得到用户对任意项目的兴趣度及 Top-N 推荐集。所以要得到用户或项目评分最相似的 k 近邻,必须计算用户或项目之间的相似度。

1.1 传统相似性度量方法

传统的相似性度量方法[6,18]一般主要为余弦相似性(cosine)、修正的余弦相似性(adjusted cosine)和相关相似性(cor-

收稿日期: 2011-07-28; **修回日期**: 2011-08-30 **基金项目**: 中央高校研究生科技创新基金资助项目(CDJXS11180012)

作者简介: 熊忠阳(1962-),男,重庆人,教授,博导,主要研究方向为数据挖掘与应用、互联网应用关键技术;刘芹(1987-),女,山东潍坊人,硕士研究生,主要研究方向为个性化推荐、云模型(tianshilq@126.com);张玉芳(1965-),女,教授,硕导,主要研究方向为数据挖掘、网络入侵检测等;李文田(1987-),男,山东潍坊人,硕士研究生,主要研究方向为物联网技术.

relation)。

1)余弦相似性

用户评分被看做是 n 维空间向量,如果用户对项目没有进行评分,则用户对该项目的评分设为 0,用户之间的相似度通过向量间的余弦夹角度量:

$$\mathrm{sim}(u_i,u_j)=\cos(u_i,u_j)=\frac{u_i\cdot u_j}{|u_i|\times|u_j|}=\frac{\sum_{c=1}^{n}R_{i,c}\cdot R_{j,c}}{\sqrt{\sum_{c=1}^{n}R_{i,c}^2}\sqrt{\sum_{c=1}^{n}R_{j,c}^2}} \tag{1}$$

其中:$\mathrm{sim}(u_i,u_j)$ 为用户 u_i 与 u_j 之间的相似度;$R_{i,c}$、$R_{j,c}$ 分别表示用户 u_i、u_j 对项目 c 的评分。

2)修正的余弦相似性

修正的余弦相似性度量方法通过减去用户对项目的平均评分来弥补余弦相似性度量方法中没有考虑不同用户评分尺度不同的缺陷。设用户 u_i 和 u_j 评过分的项目集合分别为 I_i 和 I_j,而 I_{ij} 表示用户 u_i 和 u_j 共同评分的项目集合。故用户 u_i 和 u_j 的相似度公式为

$$\mathrm{sim}(u_i,u_j)=\frac{\sum_{c\in I_{ij}}(R_{i,c}-\overline{R_i})(R_{j,c}-\overline{R_j})}{\sqrt{\sum_{c\in I_i}(R_{i,c}-\overline{R_i})^2}\sqrt{\sum_{c\in I_j}(R_{j,c}-\overline{R_j})^2}} \tag{2}$$

其中:$R_{i,c}$、$R_{j,c}$ 分别表示用户 u_i 和 u_j 对项目 c 的评分;$\overline{R_i}$、$\overline{R_j}$ 分别表示用户 u_i 和 u_j 对已评分项目的平均评分。

3)相关相似性

用户 u_i 和 u_j 的相似性可以通过计算 Pearson 相关性得到。假设 I_{ij} 表示用户 u_i 和 u_j 共同评分的项目集合,那么用户 u_i 和 u_j 的相似度可以由式(3)得到

$$\mathrm{sim}(u_i,u_j)=\frac{\sum_{c\in I_{ij}}(R_{i,c}-\overline{R_i})(R_{j,c}-\overline{R_j})}{\sqrt{\sum_{c\in I_{ij}}(R_{i,c}-\overline{R_i})^2}\sqrt{\sum_{c\in I_{ij}}(R_{j,c}-\overline{R_j})^2}} \tag{3}$$

其中:$R_{i,c}$、$R_{j,c}$ 分别表示用户 u_i 和 u_j 对项目 c 的评分;$\overline{R_i}$、$\overline{R_j}$ 分别表示用户 u_i 和 u_j 对已评分项目的平均评分。

1.2　现有协同过滤算法问题分析

协同过滤算法存在的缺点主要有稀疏性、可扩展性和冷启动问题。事实上,随着大型电子商务站点用户数目和商品数量的不断增加,使得用户评分矩阵成为高维矩阵;用户评分项目一般不会超过项目总数的 1%[2],因此数据稀疏性成为影响推荐质量的最主要因素。目前的协同过滤算法也主要是针对这三种弊端提出改进以提高推荐质量。虽然研究者们提出的各种填充算法在一定程度上能弥补数据稀疏性带来的问题,但上述算法并没有考虑到用户兴趣与项目分类之间的相关性,从而导致得到的用户兴趣邻居与最佳邻居不一致,影响推荐质量与系统效率。针对该问题,本文提出了一种基于项目分类的协同过滤改进算法。

2　基于项目分类的协同过滤改进算法

一般用户的兴趣只有特定的几类,而且用户只浏览或购买感兴趣的特定类别商品,并给自己关注的商品类别打分,所以关注共同类别的用户可以认为兴趣是相似的。当类内(相同类别内)项目相似度与类间(不同类别间)项目相似度相同时,往往类内商品(项目)的相似性会更高。另外,用户兴趣不同时,用户兴趣最近邻也不同,故在类内查找用户兴趣最近邻会更准确。基于上述考虑,本文提出的基于项目分类的协同过滤算法首先计算类内项目相似度对未评分项目进行评分预测填充,弥补数据稀疏性;然后计算类内用户相似度来获取针对用户某个兴趣的兴趣最近邻,提高搜寻用户最近邻的准确度;最终产生推荐集。基于项目分类的协同过滤改进算法,在评分矩阵更新时只计算新增项目所在的类别即可,从而在实际应用中可以极大地提高系统的效率及可扩展性。

2.1　未评分项目的预测评分

在实际的电子商务网站中,所有的商品项都被划分到若干不同的商品类别中,如果没有类别划分,可以通过聚类算法对项目生成 k 个分类。假设项目类别 $C=C_1\cup C_2\cup\cdots\cup C_k$ 表示所有项目组成的集合,而 $C_j=\{I_{j,1},I_{j,2},\cdots,I_{j,k}\}$ 表示其中的第 j 类。当类内项目相似度与类间项目相似度相同时,往往类内商品的相似性会更高。基于此,本文提出在类内计算项目的相似度并得到项目的最近项目邻居,然后对未评分项目进行预测填充。详细步骤见算法 1。

算法 1　未评分项目的预测评分算法

输入:用户—项目矩阵 R_{mxn} 及项目类别信息矩阵 $C_{i,j}$。

输出:预测用户 c 对未评分项目 j 的预测评分值 $P_{c,j}$。

a)根据用户—项目矩阵 R_{mxn} 及项目类别信息矩阵 $C_{i,j}$ 将评分记录分为 k 个类别矩阵,即 $R_{mxn}=R_1\cup R_2\cup\cdots\cup R_k$,其中 R_j 为第 j 类的评分矩阵,且一个项目可能属于多个类别。

b)根据分类后的用户—项目评分矩阵,采用式(1)分别计算未评分项目 j 与 j 所在类内其他项目之间的相似度:

$$\mathrm{sim}(i,j)=\cos(i,j)=\frac{i,j}{|i|,|j|}=\frac{\sum_{c=1}^{m}R_{c,i}\cdot R_{c,j}}{\sqrt{\sum_{c=1}^{m}R_{c,i}^2}\sqrt{\sum_{c=1}^{m}R_{c,j}^2}}$$

其中:$R_{c,i}$、$R_{c,j}$ 表示用户 c 分别对类内项目 i 和 j 的评分。

c)在类内空间查找项目 j 的最近邻项目集合 $M_j=\{I_1,I_2,\cdots,I_v\}$,使得 $j\notin M_j$,且项目 I_1 与 j 的相似度最高,项目 I_2 其次,依此类推。

d)得到未评分项目 j 的最近邻项目集合 M_j 后,根据文献[5]中提出的方法预测用户 c 对项目 j 的评分:

$$P_{c,j}=\frac{\sum_{n\in M_j}\mathrm{sim}_{j,n}\cdot R_{c,n}}{\sum_{n\in M_j}(|\mathrm{sim}_{j,n}|)} \tag{4}$$

其中:$\mathrm{sim}_{j,n}$ 是项目 j 与 n 的相似度;$R_{c,n}$ 是用户 c 对项目 n 的评分;n 是最近邻 M_j 内任意一个相似项目。

e)当项目 j 属于多个类别时,重复步骤 b)d),分别计算每个类内用户 c 对项目 j 的评分;然后将所有类的预测值取平均值作为用户 c 对项目 j 的预测评分 $P_{c,j}$;最后对评分 $P_{c,j}$ 采用“四舍五入”[3]存入评分矩阵 R_{mxn}。

2.2　推荐集计算

关注共同项目类别的用户兴趣爱好是相似的,并且当用户的兴趣不同时,兴趣最近邻也不同。假设针对电影项目类别划分,用户的兴趣如表 1 所示。由表 1 可知,针对兴趣类别“喜剧片”,用户 A 的兴趣邻居是用户 B 和 D;针对兴趣类别“爱情片”,用户 A 的兴趣邻居是用户 C 和 E。即针对用户的

不同兴趣,用户的兴趣邻居也不同,而且这也比较符合实际情况。基于此思想,本文提出的个性化推荐算法如下:首先通过类内项目最近邻居来预测未评分项目的评分值;然后根据用户兴趣在类内计算用户相似度获得兴趣最近邻;最终产生推荐。

表 1　用户—兴趣表

用户	用户兴趣类别	
	兴趣 1	兴趣 2
A	喜剧片	爱情片
B	喜剧片	武打片
C	悬疑片	爱情片
D	喜剧片	悬疑片
E	动画片	爱情片

算法 2　基于项目分类的协同过滤改进算法

输入:用户评分表及项目类别矩阵 $C_{i,j}$。

输出:目标用户 UID 对项目 IID 的预测评分 $P_{\mathrm{UID,IID}}$。

a)根据用户评分表,计算用户—项目矩阵 R_{mxn},其中,m 为用户数;n 为项目数。然后根据项目类别矩阵 $C_{i,j}$,将 R_{mxn} 分为 k 个类别矩阵,即

$$R_{mxn} = R_1 \cup R_2 \cup \cdots \cup R_k$$

其中:R_j 为第 j 类的评分矩阵。用户 c 对项目 j 的评分 $R_{c,j}$ 为 $r_{c,j}$,而用户 c 对未评分项目 j 的评分 $R_{c,j}$ 则由算法 1 计算得到,即对任意的项目 $j \in U_{cj}$,用户 c 对项目 j 的评分为

$$R_{c,j} = \begin{cases} r_{c,j} & \text{if User } c \text{ Rated Item } j \\ P_{c,j} & \text{if User } c \text{ Not Rated Item } j \end{cases}$$

b)根据分类后的用户—项目评分矩阵,采用式(1)计算类内用户之间的相似度,并组成类内用户相似度矩阵。

c)计算目标项目 IID 所在的类内,与用户 UID 相似度最高的若干项作为用户 UID 相应兴趣的最近邻,即在目标项目 IID 所在的类内,寻找与用户 UID 相似度最高的邻居集 $C_{\mathrm{UID}} = \{C_1, C_2, \cdots, C_k\}$,并且用户 C_1 与用户 UID 的相似度最高;C_2 次之,依此类推。

d)根据兴趣最近邻 C_{UID} 中每个用户对项目 IID 评分的加权平均值,得到 $P_{\mathrm{UID,IID}}$,计算方法如下:

$$P_{\mathrm{UID,IID}} = \overline{P_{\mathrm{UID}}} + \frac{\sum_{u \in C_{\mathrm{UID}}} \mathrm{sim}(\mathrm{UID}, u) \times (R_{u,\mathrm{IID}} - \overline{R_u})}{\sum_{u \in C_{\mathrm{UID}}} |\mathrm{sim}(\mathrm{UID}, u)|} \tag{5}$$

其中:$R_{u,\mathrm{IID}}$ 为用户 u 对项目 IID 的评分;$\mathrm{sim}(\mathrm{UID}, u)$ 为用户 UID 与 u 的相似度;$\overline{R_{\mathrm{UID}}}$、$\overline{R_u}$ 分别表示用户 UID 和 u 对项目的平均评分。

e)当项目属于多个类别时,分别计算类内用户 UID 对项目 IID 的预测评分 $P_{\mathrm{UID,IID}}$,然后取所有类预测评分的平均值作为用户 UID 对项目 IID 的最终预测评分结果。

3　实验结果及分析

3.1　数据集

实验采用的数据集是 Movielens 站点(http://movielens.umn.edu)提供的 100 k 的公开数据集(1997 年 9 月 19 日—1998 年 4 月 22 日的数据集)。该数据集由美国 Minnesota 大学的 GroupLens 研究小组创建并维护,包括 943 个用户对 1 682 部电影的评分记录,其中注册用户必须至少对它所拥有的电影中的 20 部进行评价才可以使用该系统,所以每个用户至少对 20 部电影进行过评价。

该数据集包含 19(018)类不同的电影类别,每一部电影至少属于一个类别且可以同时属于多个类别,实验时只利用 118 类进行测试(0 类为未知类,少数异常数据,故不用)。用户评分的范围是 15,评分数值越大,表示用户对电影的兴趣就越大,即 1 表示最不喜欢,5 表示最喜欢。采用的数据集是按照 80% 和 20% 的比例划分的 base 训练集和 test 测试集,然后进行 5-折交叉实验并取五组数据的平均值进行验证。用户评分数据集的稀疏等级为 1 - 100000/(943 × 1682) = 0.9370。

3.2　评价标准

平均绝对偏差 MAE(mean absolute error)是目前使用最广泛的评价推荐系统精确度的评价标准。平均绝对偏差 MAE 主要是计算测试集中用户实际评分与利用推荐算法预测出来的评估值之间的绝对差,平均绝对偏差 MAE 值越小,系统的推荐质量越高。

假设预测的用户评价集表示为$\{p_1, p_2, \cdots, p_N\}$,而相应的实际的用户评价集为$\{r_1, r_2, \cdots, r_N\}$,则平均绝对偏差 MAE 的定义为

$$\mathrm{MAE} = \frac{\sum_{i=1}^{N} |p_i - r_i|}{N} \tag{6}$$

3.3　实验结果及分析

计算项目间的相似性时,由于余弦相似性度量方法比较容易实现,预测速度比较快,而且预测精度也比较高[5],计算用户间的相似性方法也同理,故本文选择余弦相似性作为项目间相似性的度量标准。

实验 1

实验内容:a)当项目邻居不变时,用户邻居从 10 增加到 50 的 MAE 值变化情况;b)对比项目邻居个数从 10 增加到 50 的 MAE 值变化情况。

实验结果如图 1 所示。

由图 1 可知:a)项目邻居个数 ItemNum 固定时,MAE 值随用户邻居个数 UserNum 的增加而降低,当 UserNum ≥ 40 时,MAE 值趋于稳定;b)当用户邻居个数 UserNum 固定时,MAE 值随项目邻居个数 ItemNum 的增加而降低。结果表明在一定范围内,项目邻居个数和用户邻居个数的变化均可对系统的推荐质量产生较大影响。

实验 2

实验内容:在同等数据集及相同比例训练集和测试集的基础上,比较本文算法与相似算法的推荐效果。

实验结果如图 2 所示。

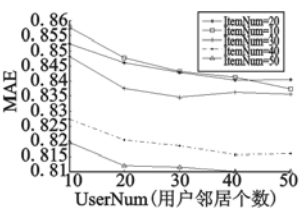

图1　不同邻居的协同过滤算法比较

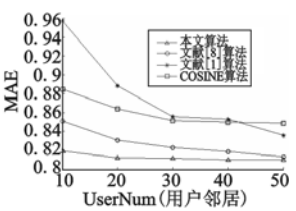

图2　推荐算法的MAE比较

实验结果表明,本文算法对 MAE 控制在[0.80, 0.82]之间,而其他三种相似算法的 MAE 值均大于等于 0.82,且波动范围较大。

4　结束语

由实验结果可知,本算法在优选项目邻居个数和用户邻居个数的情况下,可将 MAE 值始终控制在[0.80,0.82]内。改进的协同过滤算法不仅降低了数据稀疏性对推荐系统的影响,提高了系统效率及可扩展性,而且还能准确找到用户的兴趣最近邻,减小推荐误差。

为进一步提高本算法的推荐质量,在下一步研究中考虑增加项目类别的相似度计算和解决冷启动问题。

参考文献:

[1] 杨芳, 潘一飞, 李杰. 一种改进的协同过滤推荐算法[J]. 河北工业大学学报, 2010,39(3):82-87.

[2] 李聪, 梁昌勇, 董珂. 基于项目类别相似性的协同过滤推荐算法[J]. 合肥工业大学学报:自然科学版, 2008,31(3):360-363.

[3] 李春, 朱珍敏, 高晓芳. 基于邻居决策的协同过滤推荐算法[J]. 计算机工程, 2010,36(13):34-36.

[4] 周军锋, 汤显, 郭景峰. 一种优化的协同过滤推荐算法[J]. 计算机研究与发展, 2004,41(10):1842-1847.

[5] 邓爱林, 朱扬勇, 施伯乐. 基于项目评分预测的协同过滤算法推荐算法[J]. 软件学报, 2003,14(9):1621-1628.

[6] SARWAR B, KARYPIS G, KONSTAN J, *et al*. Item-based collaborative filtering recommendation algorithm[C]//Proc of the 10 th Intering World Wide Web Conference. New York: ACM Press, 2001: 285-295.

[7] 张海鹏, 离烈彪, 李仙,等. 基于项目分类预测的协同过滤推荐算法[J]. 情报学报, 2009,19(6):218-223.

[8] 嵇晓声, 刘宴兵, 罗来明. 协同过滤中基于用户兴趣度的相似性度量方法[J]. 计算机应用, 2010,30(10):2618-2610.

[9] 李大学, 谢名亮, 赵学斌. 结合项目类别信息的协同过滤推荐算法[J]. 重庆邮电大学学报:自然科学版,2010,22(6):823-827.

[10] GONG Song-jie, YE Hong-wu. Joining user clustering and item based collaborative filtering in personalized recommendation services[C]// Proc of International Conference on Industrial and Information Systems. Washington DC: IEEE Computer Society,2009:149-151.

[11] 黄裕洋, 金远平. 一种综合用户和项目因素的协同过滤推荐算法[J]. 东南大学学报:自然科学版,2010,40(5):917-921.

[12] THIESSON B, MEEK C, CHICKERING D M, *et al*. Learning mixtures of DAG models[C]//Proc of the 14 th Conference on Uncertainty in Artificial Intelligence. San Francisco, CA: Morgan Kaufmann, 1998:504-513.

[13] SARWAR B M, KARYPIS G, KONSTAN J A, *et al*. Application of dimensionality reduction in recommender system: a case study[C]// Proc of ACM WebKDD 2000 Workshop. [S.l.]: ACM Press,2000: 264-268.

[14] RESNICK P, IACOVOU N, SUCHAK M, *et al*. GroupLens: an open architecture for collaborative filtering of Netnews[C]//Proc of ACM Conference on Computer Supported Cooperative Work. New York: ACM Press,1994:175-186.

[15] LEICHT E A, HOLME P, NEWMAN M E J. Vertex similarity in networks[J]. Physical Review E,2006,73(2):026120.

[16] SHEN Lei, ZHOU Yi-ming. A new user similarity measure for collaborative filtering algorithm[C]//Proc of the 2nd International Conference on Computer Modeling and Simulation. Washington, DC:IEEE Computer Society, 2010:375-379.

[17] 贺银慧, 陈端兵, 陈勇,等. 一种结合共同邻居和用户评分信息的相似度算法[J]. 计算机科学, 2010, 37(9):184-186, 204.

[18] ADOMAVICIUS G, TUZHILIN A. Toward the next generation of recommender systems: a survey of the state-of-the-art and possible extensions[J]. IEEE Trans on Knowledge and Data Engineering, 2005,17(6):734-749.

(上接第 486 页)态的叠加,在较小的种群规模情况下,仍然保持种群中个体的多样性。同时,HCQGS 算法采用克隆算子,在进化中的过程中,根据亲合度的大小,在候选解的附近产生一个变异解的群体,扩大解的搜索范围,有助于防止进化早熟和陷于局部极小值的问题。从图中还可以看出,采用量子比特进行编码的 QGA 要优于传统的 GA 算法。

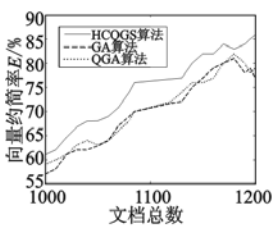

图1　向量约简率对比

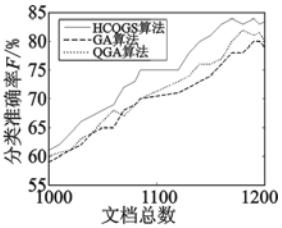

图2　分类准确率对比

4　结束语

文本的特征抽取是一个重要的研究方向。如何把表示文本的高维数特征向量转换为合适的特征子集是实现文本正确分类的关键。本文采用量子比特进行编码,引入人工免疫中的克隆选择策略,提出一种基于混合克隆量子遗传退火策略的文本特征选择方法。实验结果表明,该方法比传统的 GA 和 QGA 算法更能有效地降低文本特征向量的维度,所提取的特征子集能有效提高文本分类的精度和效率,具有一定的应用前景。

参考文献:

[1] DOLOCA A. Feature selection for texture analysis using genetic algorithms[J]. International Journal of Computer Mathematics, 2000,74(3):279-292.

[2] 刘勇国,李学明,张伟,等. 基于遗传算法的特征子集选择[J]. 计算机工程,2003,29(6):19-21.

[3] 赵丽娜,刘培玉,朱振方. 自适应遗传算法在特征选择中的改进及应用[J]. 计算机工程与应用,2009,45(7):39-41.

[4] 张昊,陶然,李志勇,等. 基于自适应模拟退火遗传算法的特征选择方法[J]. 兵工学报,2009,30(1):82-85.

[5] 邱烨,刘培玉. 基于量子遗传算法的文本特征选择方法研究[J]. 计算机工程与应用,2008,44(25):140-142.

[6] 陈绯,郑华. 一种免疫克隆特征选择算法在文本分类中的应用[J]. 计算机工程与科学,2009,31(9):119-121.

[7] XIONG Yan, CHEN Huan-huan, MIAO Fu-you, *et al*. A quantum genetic algorithm to solve combinatorial optimization problem[J]. Acta Electronica Sinica,2004,32(11):1855-1858.

[8] SHANG Rong-hua, JIAO Li-cheng. An immune clonal algorithm for dynamic multi-objective optimization[J]. Journal of Software, 2007,18(11):2700-2711.

[9] 杨咚咚,焦李成,公茂果,等. 求解偏好多目标优化的克隆选择算法[J]. 软件学报,2010,21(1):14-33.