



(12)发明专利申请

(10)申请公布号 CN 109697255 A

(43)申请公布日 2019.04.30

(21)申请号 201710990806.3

(22)申请日 2017.10.23

(71)申请人 中国科学院沈阳自动化研究所
地址 110016 辽宁省沈阳市沈河区南塔街114号

(72)发明人 张会文 张伟 冷雨泉 陈正仓
付明亮 韩小宁 蔺兆宝

(74)专利代理机构 沈阳科苑专利商标代理有限公司 21002

代理人 李巨智

(51)Int.Cl.
G06F 16/951(2019.01)
G06K 9/62(2006.01)
G06N 20/00(2019.01)

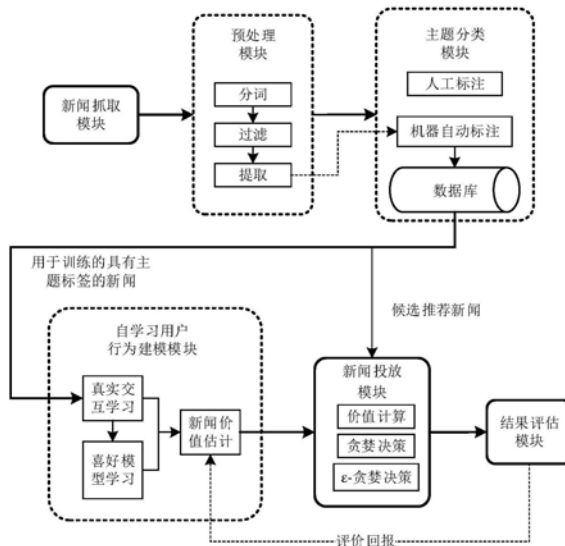
权利要求书2页 说明书6页 附图5页

(54)发明名称

一种基于在线自学习的个性化新闻投放系统及方法

(57)摘要

本发明涉及一种基于在线自学习的个性化新闻投放系统,新闻获取模块用于通过网络爬虫在网络上实时抓取新闻,将抓取到的新闻发送给预处理模块;预处理模块用于对抓取到的新闻进行预处理,并将预处理后的新闻发送到主题分类模块;主题分类模块用于对得到的新闻进行主题分类,并将分类后的新闻存储于数据库中;自学习用户行为建模模块用于通过记录用户读取新闻的经验来估计不同主题新闻对于用户的价值;新闻投放模块用于根据当前候选推荐新闻的价值排序向用户投放用户认为最有价值的新闻。本发明简化了传统推荐方法的复杂计算,在新闻投放上实现了个性化,提升了服务质量,减少了对用户信息的依赖,实现了虚拟的用计算机模拟用户决策的过程。



CN 109697255 A

1. 一种基于在线自学习的个性化新闻投放系统,其特征在於,包括:

新闻获取模块,用于通过网络爬虫在网络上实时抓取新闻,将抓取到的新闻发送给预处理模块;

预处理模块,用于对抓取到的新闻进行预处理,并将预处理后的新闻发送到主题分类模块;

主题分类模块,用于对得到的新闻进行主题分类,并将分类后的新闻存储于数据库中;

自学习用户行为建模模块,用于通过记录用户读取新闻的经验来估计不同主题新闻对于用户的价值;

新闻投放模块,用于根据当前候选推荐新闻的价值排序向用户投放用户认为最有价值的新闻。

2. 根据权利要求1所述的基于在线自学习的个性化新闻投放系统,其特征在於:所述主题分类模块包括人工标注单元和自动标注单元,

人工标注单元用于通过人的主观判断对新闻进行主题标注;

自动标注单元用于根据主题分类模型对新闻进行主题标注。

3. 根据权利要求1所述的基于在线自学习的个性化新闻投放系统,其特征在於:所述自学习用户行为建模模块包括真实经验学习单元和模型学习单元;

真实经验学习单元用于根据用户当前的新闻浏览行为对新闻价值进行估计;

模型学习单元用于根据用户浏览习惯构建的用户喜好模型,产生模拟的浏览行为,并基于模拟的浏览行为进行新闻价值估计。

4. 根据权利要求1或3所述的基于在线自学习的个性化新闻投放系统,其特征在於:所述经验包括用户主动读取主题新闻获得的经验以及初步根据用户行为,向用户推荐新闻获得的经验。

5. 根据权利要求1所述的基于在线自学习的个性化新闻投放系统,其特征在於:所述新闻投放模块包括价值计算单元和决策单元;

价值计算单元用于根据学习到的用户模型对推荐的新闻进行价值计算,即计算用户的喜好程度;

决策单元包括贪婪决策单元和 ϵ -贪婪决策单元,其中

贪婪决策单元用于选择价值最大的新闻主题向用户推荐;

ϵ -贪婪决策单元用于以 ϵ 的概率向用户随机推荐新闻主题,以 $1-\epsilon$ 的概率向用户推荐价值最大的新闻主题。

6. 根据权利要求1所述的基于在线自学习的个性化新闻投放系统,其特征在於:还包括结果评估模块,对新闻投放模块的决策效果进行评估。

7. 根据权利要求1~6任一项所述系统的基于在线自学习的个性化新闻投放方法,其特征在於,包括以下过程:

步骤1:新闻获取模块通过网络爬虫在网络上实时抓取新闻,将抓取到的新闻发送给预处理模块;

步骤2:通过预处理模块对抓取到的新闻进行预处理,并将预处理后的新闻发送到主题分类模块;

步骤3:通过主题分类模块对得到的新闻进行主题分类,并将分类后的新闻存储于数据

库中；

步骤4:自学习用户行为建模模块根据用户当前的新闻浏览行为进行新闻价值估计以及通过用户浏览习惯构建用户喜好模型,产生模拟的浏览行为,并基于模拟的浏览行为进行新闻价值估计；

步骤5:新闻投放模块根据当前候选推荐新闻的价值排序向用户投放用户认为最有价值的新闻。

8.根据权利要求7所述的基于在线自学习的个性化新闻投放方法,其特征在于,所述用户喜好模型包括用户浏览不同主题新闻s的转移概率分布 $P_{ss'}$ 和对某个主题新闻的喜好值R;其中

$$P_{ss'} = \frac{\#(\text{从 } s \text{ 开始转移到 } s')}{\#(\text{出现 } s \text{ 的次数})}$$

对某个主题新闻s的喜好值R由某个主题新闻的浏览次数 $C(s)$ 和该主题新闻的浏览时间 $T(s)$ 决定:

$$R(s) = \frac{C(s)T(s)}{\sum_s C(s)} R_b$$

9.根据权利要求7所述的基于在线自学习的个性化新闻投放方法,其特征在于,所述新闻价值估计包括:

步骤1:初始化所有不同主题新闻的价值V为0;

步骤2:建立一个空的列表Memory,用于保存状态转移序列和喜好值R;

步骤3:从数据库Memory中抽取一个样本,更新相应主题新闻的价值V,更新表达式为:

$$V_s^{t+1} = V_s^t + \alpha [R(s) + \gamma V_s^t - V_s^t];$$

步骤4:循环迭代步骤3,直至前后两次不同主题新闻价值V的差值小于设定阈值。

10.根据权利要求7所述的基于在线自学习的个性化新闻投放方法,其特征在于,所述价值排序为根据不同主题新闻的价值V的由大到小进行排序。

一种基于在线自学习的个性化新闻投放系统及方法

技术领域

[0001] 本发明涉及人工智能、机器学习、自然语言处理、数据挖掘和自适应控制 领域,具体地说是一种基于在线自学习的个性化新闻投放系统及方法。

背景技术

[0002] 经历了20世纪末,从1995年到2001年的互联网泡沫,在21世纪初互 联网重新焕发了生机,实现了快速的发展。随着各种互联网产品和网站的出现,比如淘宝、美团、58同城以及近来逐渐普及开来的网上外卖,可以说人们的衣 食住行都与互联网紧密交织。

[0003] 不同于以前推送式的网络服务,现在的网络服务更多的注重用户体验以及 用户的参与感。这与交互分不开,如何实现更智能的交互方式?当然人与人的 交互是最自然的。如何使互联网也表现出类人的行为,这就需要互联网具有智 能性。

[0004] 这种智能性体现在个性化的推送服务。现有的个性化推送方式需要对个人 喜好建模。这种建模要么通过人为设计的特征来构造,要么需要用户提供大量 的个人信息。人为设计的特征需要耗费大量的人力,另外这些特征往往带有主 观判断,不一定能够准确描述个人的喜好。让用户在注册的时候输入大量个人 信息,然后经过一些算法自动建模的方式虽然避免了人为设计的元素。但是输 入用户信息,可能会涉及个人隐私问题。另外,用户可能也会因为输入信息太 过繁琐而降低用户体验。还有一个问题是,这些模型往往是静态的,并不能动 态的捕捉用户的兴趣点,从而更新原来的模型。

发明内容

[0005] 针对现有技术的不足,本发明提供一种基于在线自学习的个性化新闻投放 系统及方法。

[0006] 本发明为实现上述目的所采用的技术方案是:

[0007] 一种基于在线自学习的个性化新闻投放系统,包括:

[0008] 新闻获取模块,用于通过网络爬虫在网络上实时抓取新闻,将抓取到的新 闻发送给预处理模块;

[0009] 预处理模块,用于对抓取到的新闻进行预处理,并将预处理后的新闻发送 到主题分类模块;

[0010] 主题分类模块,用于对得到的新闻进行主题分类,并将分类后的新闻存储 于数据库中;

[0011] 自学习用户行为建模模块,用于通过记录用户读取新闻的经验来估计不同 主题新闻对于用户的价值;

[0012] 新闻投放模块,用于根据当前候选推荐新闻的价值排序向用户投放用户认 为最有价值的新闻。

[0013] 所述主题分类模块包括人工标注单元和自动标注单元,

[0014] 人工标注单元用于通过人的主观判断对新闻进行主题标注;

- [0015] 自动标注单元用于根据主题分类模型对新闻进行主题标注。
- [0016] 所述自学习用户行为建模模块包括真实经验学习单元和模型学习单元；
- [0017] 真实经验学习单元用于根据用户当前的新闻浏览行为对新闻价值进行估计；
- [0018] 模型学习单元用于根据用户浏览习惯构建的用户喜好模型，产生模拟的浏览行为，并基于模拟的浏览行为进行新闻价值估计。
- [0019] 所述经验包括用户主动读取主题新闻获得的经验以及初步根据用户行为，向用户推荐新闻获得的经验。
- [0020] 所述新闻投放模块包括价值计算单元和决策单元；
- [0021] 价值计算单元用于根据学习到的用户模型对推荐的新闻进行价值计算，即计算用户的喜好程度；
- [0022] 决策单元包括贪婪决策单元和 ϵ -贪婪决策单元，其中
- [0023] 贪婪决策单元用于选择价值最大的新闻主题向用户推荐；
- [0024] ϵ -贪婪决策单元用于以 ϵ 的概率向用户随机推荐新闻主题，以 $1-\epsilon$ 的概率向用户推荐价值最大的新闻主题。
- [0025] 还包括结果评估模块，对新闻投放模块的决策效果进行评估。
- [0026] 一种基于在线自学习的个性化新闻投放方法，包括以下过程：
- [0027] 步骤1：新闻获取模块通过网络爬虫在网络上实时抓取新闻，将抓取到的新闻发送给预处理模块；
- [0028] 步骤2：通过预处理模块对抓取到的新闻进行预处理，并将预处理后的新闻发送到主题分类模块；
- [0029] 步骤3：通过主题分类模块对得到的新闻进行主题分类，并将分类后的新闻存储于数据库中；
- [0030] 步骤4：自学习用户行为建模模块根据用户当前的新闻浏览行为进行新闻价值估计以及通过用户浏览习惯构建用户喜好模型，产生模拟的浏览行为，并基于模拟的浏览行为进行新闻价值估计；
- [0031] 步骤5：新闻投放模块根据当前候选推荐新闻的价值排序向用户投放用户认为最有价值的新闻。
- [0032] 所述用户喜好模型包括用户浏览不同主题新闻 s 的转移概率分布 $P_{ss'}$ 和对某个主题新闻的喜好值 R ；其中

$$[0033] \quad P_{ss'} = \frac{\#(\text{从 } s \text{ 开始转移到 } s')}{\#(\text{出现 } s \text{ 的次数})}$$

- [0034] 对某个主题新闻 s 的喜好值 R 由某个主题新闻的浏览次数 $C(s)$ 和该主题新闻的浏览时间 $T(s)$ 决定：

$$[0035] \quad R(s) = \frac{C(s)T(s)}{\sum_s C(s)} R_b$$

- [0036] 所述新闻价值估计包括：
- [0037] 步骤1：初始化所有不同主题新闻的价值 V 为0；
- [0038] 步骤2：建立一个空的列表Memory，用于保存状态转移序列和喜好值 R ；

[0039] 步骤3:从数据库Memory中抽取一个样本,更新相应主题新闻的价值V,更新表达式为: $V_s^{t+1} = V_s^t + \alpha[R(s) + \gamma V_s^t - V_s^t]$;

[0040] 步骤4:循环迭代步骤3,直至前后两次不同主题新闻价值V的差值小于设定阈值。

[0041] 所述价值排序为根据不同主题新闻的价值V的由大到小进行排序。

[0042] 本发明具有以下有益效果及优点:

[0043] 1.本发明中用户行为建模中不直接构建用户的偏好特征向量,间接的通过新闻价值函数反映用户爱好;

[0044] 2.由于不直接对用户行为建模,不需要用户的个人资料,比如注册信息等,保护用户个人隐私;

[0045] 3.本发明中的学习是在线的学习,能够主动发觉用户的兴趣变化,更新模型;

[0046] 4.本发明可以在行为学习中同时利用实际经验和虚拟经验,可以快速的学习到用户的兴趣;

附图说明

[0047] 图1是本发明的系统架构示意图;

[0048] 图2是本发明的新闻预处理过程图;

[0049] 图3是本发明的主题层次聚类示意图;

[0050] 图4是本发明的学习模型框架图;

[0051] 图5是本发明的自学习用户行为建模系统组成图;

[0052] 图6是本发明的计算V的记忆库Memory;

[0053] 图7是本发明的用户喜好模型图;

[0054] 图8是本发明的新闻投放过程图。

具体实施方式

[0055] 下面结合附图及实施例对本发明做进一步的详细说明。

[0056] 如图1所示为本发明的系统架构示意图。

[0057] 该系统设有以下组成部件:新闻获取模块、前期预处理模块、主题分类模块、基于学习的用户行为建模模块、新闻投放模块以及用户体验评估,其中:

[0058] 新闻获取模块,负责采用网络爬虫在网络上实时抓取新闻,包括新闻的标题、新闻内容、新闻链接甚至是相关评论。

[0059] 预处理模块,用于对抓取的新闻初步处理,具体包括分词过滤、词性标注、特征提取以及文章的特征向量表示等。

[0060] 主题分类模块,主要用于对所获得的新闻进行主题的划分,划分方法包括人工标注法和机器自动标注法。人工标注法根据个人的主观判断和先验知识对新闻进行主题划分。机器自动标注法通过对预处理得到的文本特征向量进行聚类实现主题标注。聚类方式包括硬聚类和软聚类两种。在产生训练样本时本发明采用硬聚类,在对候选推荐样本进行聚类时采用软聚类。经过聚类后的文本被分割成不同主题的子集,存储于数据库。用 D_{ver}^x 表示第x个主题包含所有文本的平均特征向量。具体包括下列操作内容:

[0061] 聚类产生K个主题类；

[0062] 利用公式 $D_{ver}^x = \frac{1}{m_x} \sum_{i=1}^{m_x} D_i$ 计算每个主题类的平均特征向量；

[0063] 经过聚类后的文本被分割成不同主题的子集，分别存储于数据库。

[0064] 自学习用户行为建模模块，主要通过记录用户读取新闻的经验来估计不同主题新闻对于用户的价值。经验包括2种，一种为用户浏览新闻获得的经验，另一种是由用户模型向用户推荐新闻获得的虚拟经验。两种经验都用于学习不同主题新闻的价值函数。具体包括真实经验学习单元、模型学习单元和价值估计单元。

[0065] 一个主题的新闻的价值定义为： $V^n(s) = R(s) + \gamma \sum_{s' \in S} P_{sa}(s') V^n(s')$ ， s' 代表下一个浏览的新闻的主题类型。 $R(s)$ 是某个用户对于某主题新闻的喜好值。第二项表示下个主题新闻的期望喜好值。可见某个主题新闻的价值是累积的，下个主题的新闻会影响前面的价值。这反映了新闻的引导作用。

[0066] 新闻投放模块，根据获得主题新闻价值和当前候选推荐新闻的价值排序向用户投放用户最感兴趣的新闻。包括价值计算单元、贪婪决策单元和 ϵ -贪婪决策单元。从数据库中得到一个候选的新闻集合。经过初步过滤筛选，再进行软聚类处理，得到候选新闻和每个主题类的相似度 w_{ix} ， x 表示第 x 个主题。利用自学习用户行为建模模块的价值估计单元得到 V ，计算出文本 i 的综合价值 V_i 。根据 V_i 的大小有两种策略向用户投放新闻：贪婪策略和 ϵ -贪婪策略。

[0067] 结果评估模块，对新闻投放模块的决策效果进行评估。从投放给用户，到用户的反馈构成了一次经验，这个经验被添加到记忆库中，用于新闻价值的评估。

[0068] 如图2所示为本发明的新闻预处理流程图。

[0069] 首先去除标记词、虚词和标点符号，保留具有实际意义的名词以及一些特殊意义的实体名词。然后通过词典构造实体名词序列，下一步根据停用词表，过滤掉停用词。至此获得一个完成的名词序列，通过特征选择，简化文档的表示，把简化的特征序列存储起来。

[0070] 对上一步获得新闻进行预处理操作，包括分词过滤、词性标注、特征提取以及文档的特征向量表示。经过分词和词性标注后得到新闻文本的名词序列，在经过特征提取得到文本的特征向量表示。比如假设总共有 m 个文本，那么文本的向量空间表示可写为： $D = \{D_1, D_2, \dots, D_m\}$ ， $D_i = (d_{i1}, d_{i2}, \dots, d_{in})$ ， $i = 1, 2, \dots, m$ 。 n 表示特征向量的维度， d_{ij} 表示文档 D_i 中第 j 个词条 t_j 的特征值。

[0071] 如图3所示为本发明的主题层次聚类示意图。

[0072] 所述主题分类模块包括以下组成单元：

[0073] 人工标注单元，通过人为的给文本划定主题词，比如这篇文章是娱乐方面的，涉及到篮球。那么主题词可以给定为娱乐、体育等，当然进一步还可添加体育项目主题词，这里对应的是篮球。划分越精细，越能实现准确的个性化投放。

[0074] 聚类处理单元，能实现主题聚类，常用的方法有K-means聚类、层次聚类等。经过聚类或者人工标注的新闻存储在数据库。作为后一步的训练样本和投放新闻的语料库。

[0075] 数据库，用于存储经过带有主题标签的新闻数据。

[0076] 如图4所示是本发明的学习模型框架图。

[0077] 框架包括新闻投放系统和用户。投放系统通过观察用户的行为以及用户反

馈——当前的新闻是否是用户感兴趣的,来学习用户的行为策略,然后利用学习到的策略来模仿用户决策。具体的要素包括三个:当期状态 s ,投放系统得到用户的反馈信息 R ,用户做出的行为 a 。

[0078] s 表示投放系统面临的状态,具体到本方法指的是不同主题的新闻,一个主题表示一个状态;

[0079] R 表示喜好值,为用户对推荐新闻的兴趣。如果用户不感兴趣或者根本没看投放的新闻, R 取一个负值,表示对错误投放的惩罚。如果用户感兴趣给 R 一个正值,具体大小跟用户兴趣度成正比;

[0080] a 表示用户看新闻时的选择动作,比如当前用户看了一篇关于娱乐的新闻,下一刻阅读了一篇关于体育新闻,这个过程称为一次行为动作。

[0081] 如图5是本发明的自学习用户行为建模组成图。

[0082] 自学习用户行为建模模块,主要通过记录用户读取新闻的经验来估计不同主题新闻对于用户的价值。经验包括2种,一种是用户当前浏览的主题新闻获得的经验,另一种是由用户喜好模型产生的模拟经验。产生方法是:1)随机的从以训练的状态中选择一个主题 s ;2)以喜好模型Model中的转移概率 $P_{ss'}$ 获得下一个状态 s' ;3)依次按同样的方法,就生成了一次浏览经验。

[0083] 两种经验都用于对不同主题新闻价值的学习。具体包括真实经验学习单元、模型学习单元和新闻文档价值估计单元。

[0084] 自学习用户行为建模:基于一个学习模型,利用训练样本,假设用户随机的获得一个初始训练样本,通过用户和训练样本的交互获得不同新闻主题的价值 V 。 V 的值可以利用蒙特卡罗方法更新,随着训练次数的增加,每个主题的价值 V 越来越稳定,最终收敛到一个值 V^* 。 V 的大小蕴含了用户的喜好信息,可以用作投放新闻的依据。其次模型学习可以建立用户的喜好模型,用Model表示,Model随着训练进行越来越准确。

[0085] 实施例1:

[0086] 该步骤包括下列操作内容:

[0087] 首先用户的一系列浏览点击行为形成一次经验,也就是一个训练样本序列。

[0088] 举例说明:用户在某天晚上依次看了以下新闻:娱乐的(黄晓明大婚),NBA(科比复出)……。这就可以看成一次经验,如果用 s 表示每次看的新闻主题,那么这个过程可以

描述为一个序列: $s_0 \xrightarrow{a_0} s_1 \xrightarrow{a_1} s_2 \xrightarrow{a_2} s_3 \xrightarrow{a_3} \dots$

[0089] 用户一段时间内的浏览行为构成了多次经验,根据这些经验更新价值 V 和用户喜好模型,具体过程是:

[0090] 1)初始化 V 全为0;

[0091] 2)初始化一个空的记忆库Memory,用于保存经验,如图6所示;初始化一个新的列表Model用于建立用户喜好模型,如图7所示;

[0092] 3)重复以下步骤:

[0093] 从数据库Memory中随机抽取一个样本,更新相应主题新闻的价值 V ,更新表达式为: $V_s^{t+1} = V_s^t + \alpha[R(s) + \gamma V_s^t - V_s^t]$;利用经验更新Model。

[0094] 4)当步骤3)中前后两次 V 的差值小于设定的阈值,结束迭代。得到所有主题新闻

的最优价值 V^* 和用户喜好模型Model。

[0095] 如图6所示为本发明的计算V的记忆库Memory。

[0096] 记忆库抽象的存储了用户的浏览经验。每一个经验数据是一次状态转移。比如用户依次看了新闻:娱乐的(黄晓明大婚),对应主题 s_3 ,NBA(科比复出),对应主题 s_5 。改用户在娱乐主题新闻上停留了30分钟,并且之前有50次浏览娱乐新闻的经历。因此根据

$$[0097] \quad R(s) = \frac{C(s)T(s)}{\sum_s C(s)} R_b$$

[0098] 得到 s_3 对应的喜好值 $R(s_3) = 20$ 。那么这样的一次转移在Memory中表示为: $\langle s_3, 20, s_5 \rangle$ 。Memory由很多个这样的经验转移构成。

[0099] 如图7所示为本发明的用户喜好模型示意图。

[0100] 用户喜好模型Model根据Memory中的数据更新某用户的喜好模型参数P和R。Memory中每增加一个数据,Model更新一次。比如某用户依次看了新闻:娱乐的(黄晓明大婚),对应主题 s_3 ,NBA(科比复出),对应主题 s_5 。在这之前该用户总共浏览娱乐新闻的次数为50,其中有5次看完娱乐新闻后立马切换到了NBA新闻。这50次的平均喜好值 $R(s_3) = 20$,这一次得到了立即回报 $R = 40$ 。那么 $P_{s_3s_5}$ 更新为:

$$[0101] \quad P_{s_3s_5} = \frac{\#(\text{从 } s_3 \text{ 开始转移到 } s_5)}{\#(\text{出现 } s_3 \text{ 的次数})} = \frac{5 + 1}{50} = 0.12$$

[0102] $R(s_3)$ 更新为: $R(s_3) = 50 \times 20 + 40 / 51 = 20.4$ 。

[0103] 如图8所示为本发明的新闻投放过程图。

[0104] 从新闻集合中产生获选推荐新闻,候选推荐新闻已经经过预处理,并得到每个新闻文本的特征向量。

[0105] 根据聚类结果,通过余弦相似公式 $w_{ik} = \frac{\langle D_{ave}^x, D_i \rangle}{\|D_{ave}^x\| \cdot \|D_i\|}$ 计算第i个新闻文本和聚类产生的第x个主题的相似度。

[0106] 借助训练得到的各个主题新闻的价值 V^* ,根据公式 $V_i = \sum_{x=1}^k w_{ix} V_x^*$ 计算文本i的综合价值 V_i 。K表示所有主题的总数,也是状态集合{S}元素的个数。

[0107] 根据求得的价值,进行降序排列,并提取需要投放的前n个新闻。

[0108] 新闻投放:根据 V_i 的大小有两种策略向用户投放新闻:贪婪策略和 ϵ -贪婪策略。

[0109] 效果评估:完成一次投放决策后,用户得到一个投放的新闻集合,用户在该文本集合上的表现会反馈到模型学习模块,进一步改善用户行为模型。

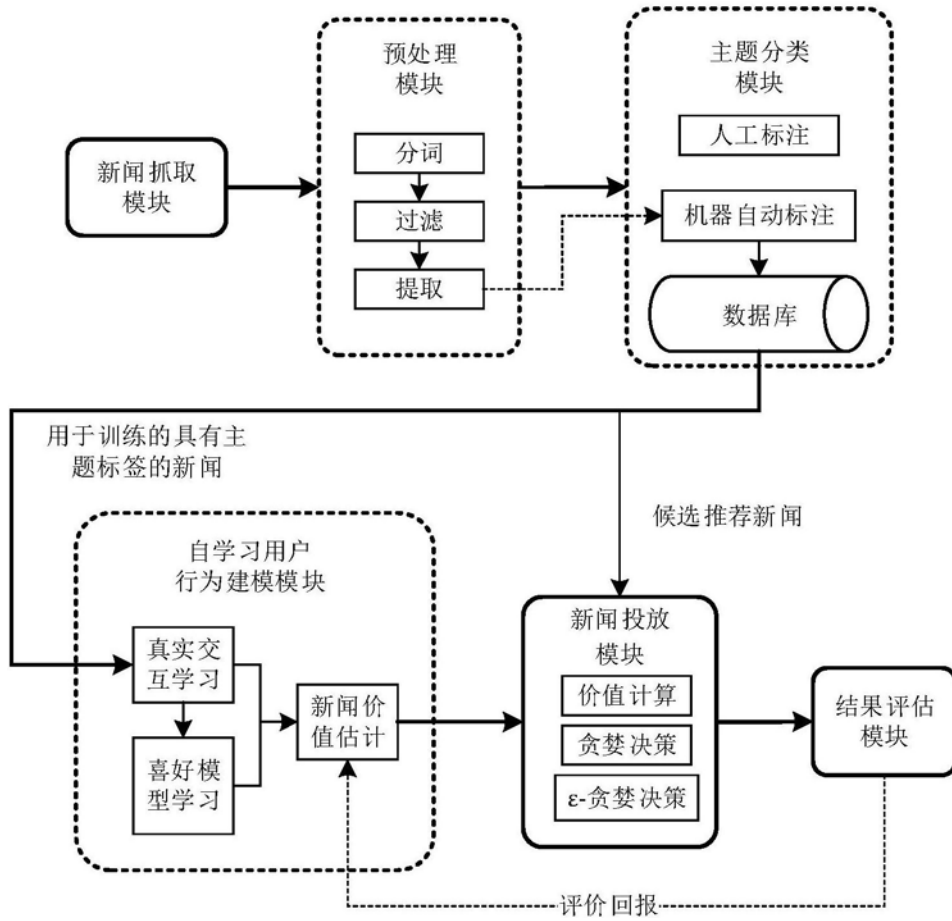


图1

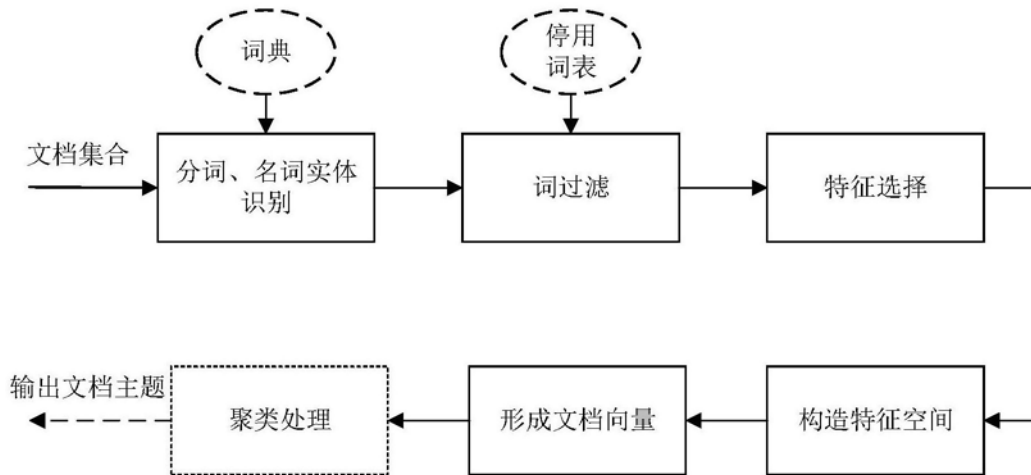


图2

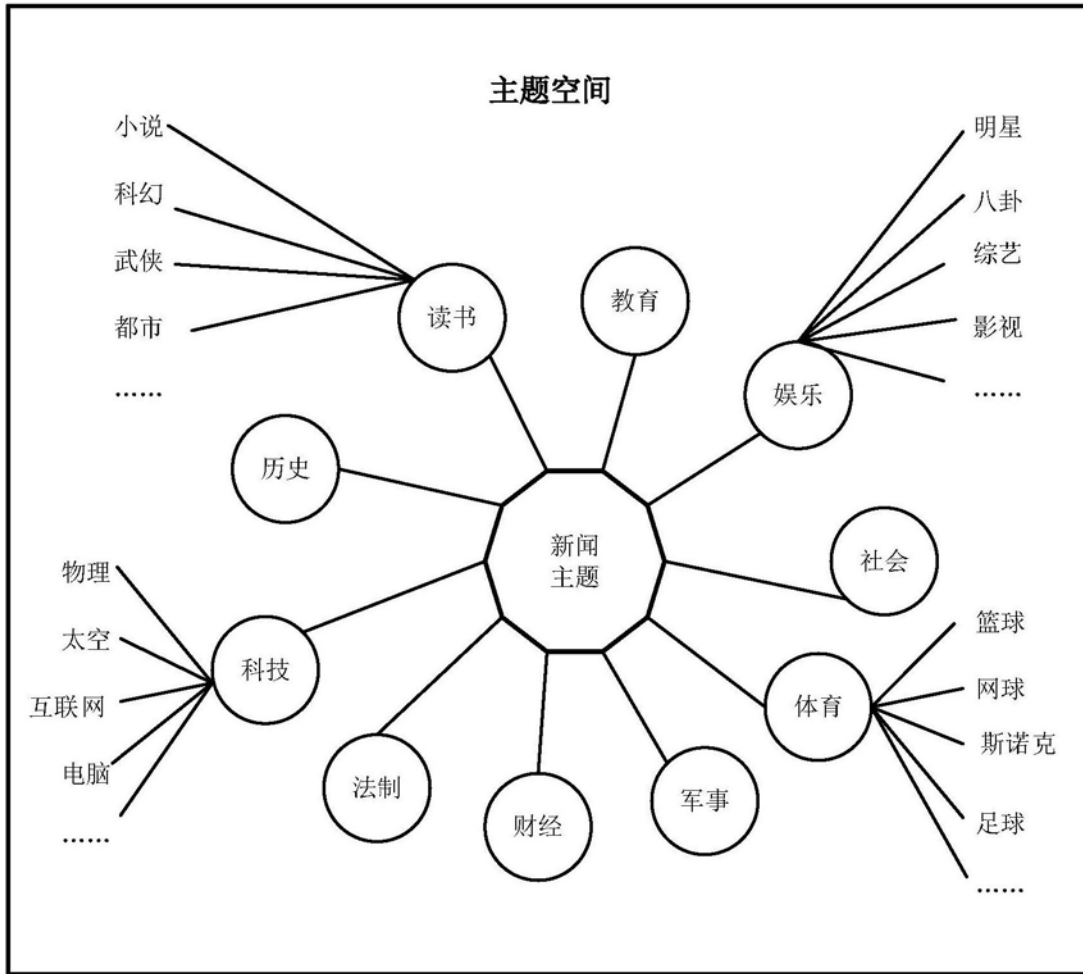


图3

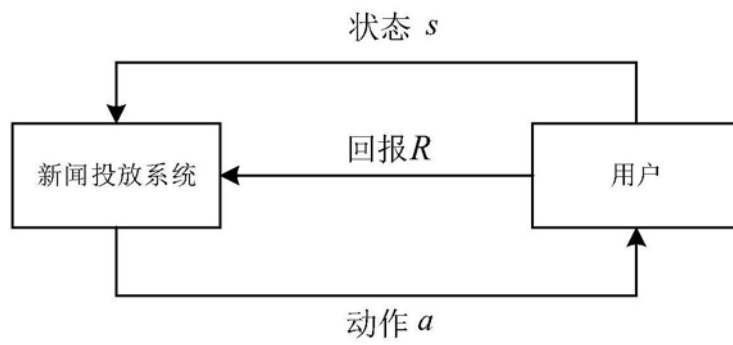


图4

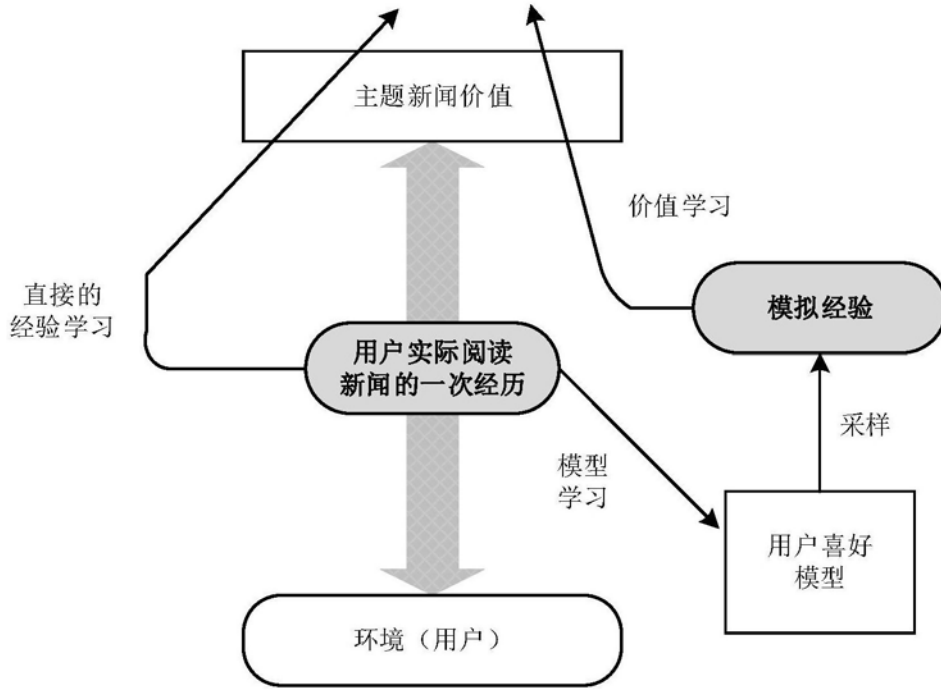


图5

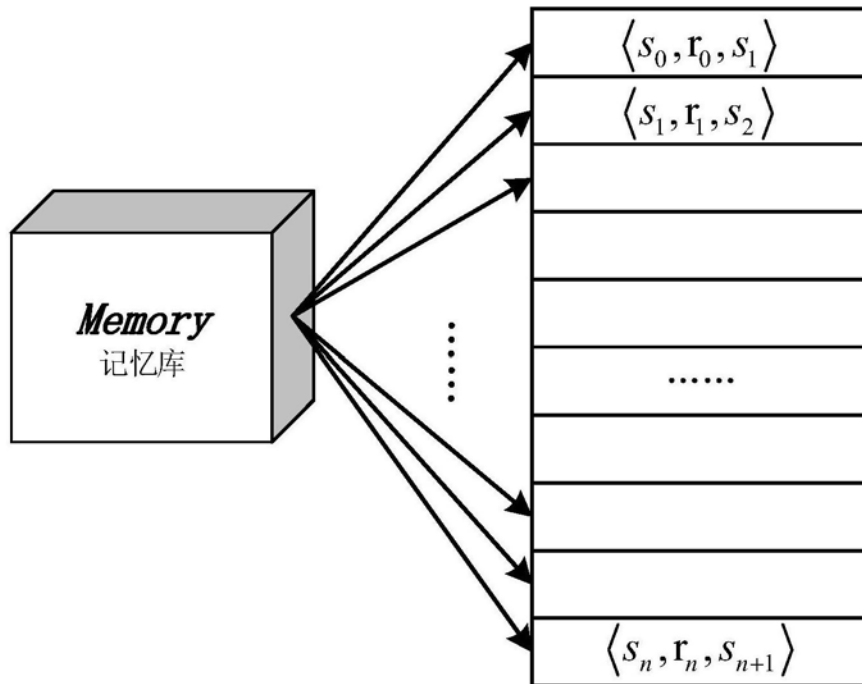


图6

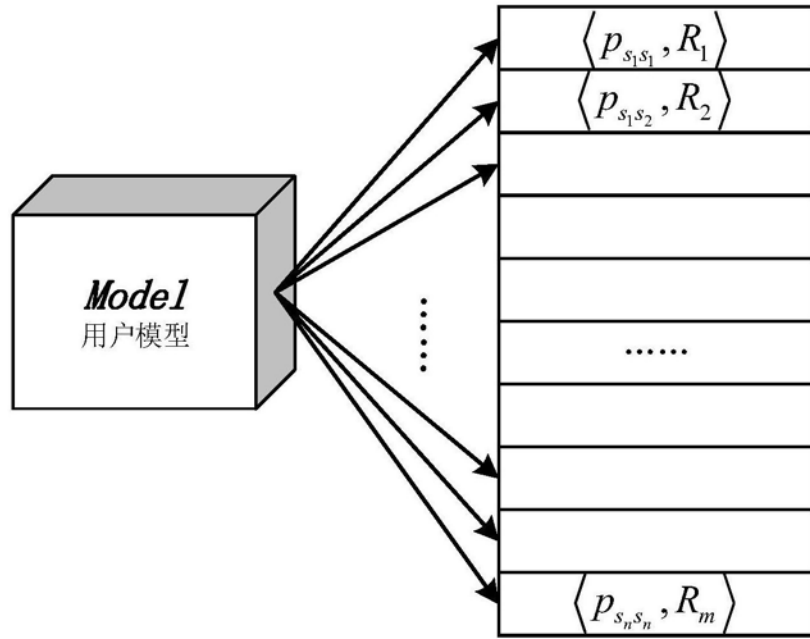


图7

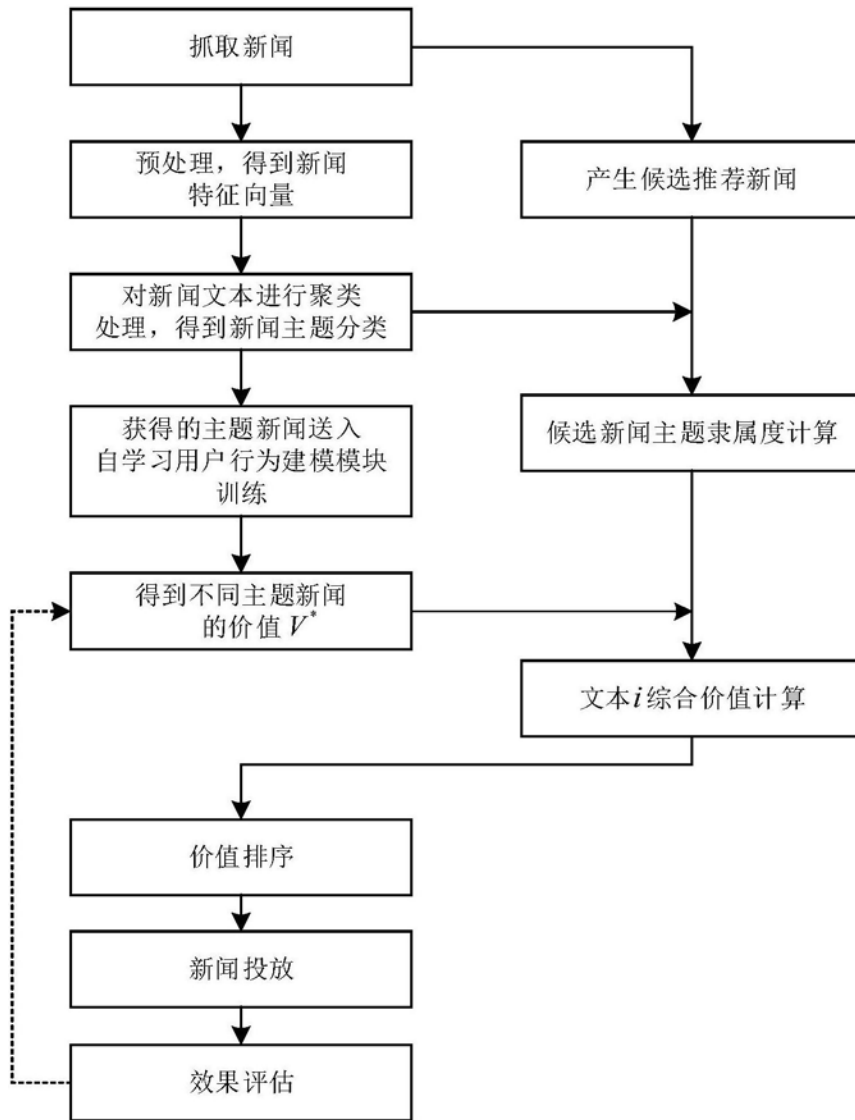


图8