

◎模式识别与人工智能◎

领域知识学习中的马尔可夫逻辑网应用研究

于 凤¹, 郑德权^{1,2}, 刘 祥²YU Feng¹, ZHENG Dequan^{1,2}, LIU Xiang²

1. 哈尔滨商业大学 计算机与信息工程学院, 哈尔滨 150028

2. 哈尔滨工业大学 计算机科学与技术学院, 哈尔滨 150001

1. School of Computer and Information Engineering, Harbin University of Commerce, Harbin 150028, China

2. School of Computer Science and Technology, Harbin Institute of Technology, Harbin 150001, China

YU Feng, ZHENG Dequan, LIU Xiang. Markov logic networks technology application for domain knowledge learning. *Computer Engineering and Applications*, 2016, 52(22): 149-154.

Abstract: It has been the focus of academia and application field since knowledge discovery in data is proposed, but facing the increasingly complex application environment, most of the existing research methods are based on statistical learning, which often neglect the relationship between knowledge and the knowledge change over time, leading to that the effect of application is often just passable. How to grasp the statistical relationship between knowledge accurately to make the correct knowledge learning, have become a problem in knowledge research. Recently, with the development of research on statistical relational learning, Markov logic networks, which combines the probability graph and first-order logic theory, is applied successfully to natural language processing, machine learning, social network analysis and so on. In this paper, a Markov Logic Networks is presented for knowledge learning, which brings in first-order logic to learn the relationship between knowledge, and finish logical reasoning. Compared with SVM, this proposal method gets about 10 percent higher effect by classification knowledge learning in the text categorization fields.

Key words: knowledge learning; Markov logic networks; classification; Support Vector Machine(SVM)

摘 要: 现有的知识学习多基于统计方法, 常常忽略了知识间的关系以及随时间的变化情况, 在应用效果方面往往差强人意。如何准确把握知识间的统计关系, 进行正确的知识学习, 成为知识研究的一个重点和难点。近几年, 随着统计关系学习研究的兴起, 结合概率图模型和一阶逻辑理论的马尔可夫逻辑网被成功应用于自然语言处理、机器学习、社会关系分析等领域中。基于马尔可夫逻辑网技术, 提出一种知识学习方法, 在传统知识获取方法的基础上, 引入一阶逻辑来学习知识间的关系, 进行逻辑推理。在文本分类的应用实验中, 通过对分类知识学习, 与传统的SVM相比, 所提出方法的准确率提高10%左右。

关键词: 知识学习; 马尔可夫逻辑网; 分类; 支持向量机

文献标志码: A **中图分类号:** TP391.2 **doi:** 10.3778/j.issn.1002-8331.1601-0204

1 引言

知识库是针对不同领域或者一些具体问题求解的需要, 运用人工智能(AI)、数据库等相关技术, 自动或半自动提取出的知识片段集合, 这些知识片段具有易存储、

易操作、互相联系等特性。知识库作为当今人工智能和数据库技术相结合的产物, 其构建技术也集成相关领域的最新进展(包括知识表示、知识获取和知识推理等)。

当前, 已有的一些知识库主要是面向通用领域的需

基金项目: 国家高技术研究发展计划(863)(No.2015AA015405); 黑龙江省教育厅规划课题(No.GJC1215032)。

作者简介: 于凤(1969—), 女, 副教授, 研究领域为自然语言处理, 文本分类, E-mail: yufeng@hrbcu.edu.cn; 郑德权(1968—), 男, 博士, 副教授, 研究领域为自然语言理解, 数据挖掘; 刘祥(1988—), 男, 硕士研究生, 研究领域为自然语言处理。

收稿日期: 2016-01-14 **修回日期:** 2016-05-16 **文章编号:** 1002-8331(2016)22-0149-06

求研发,因此较少能够满足专业领域自然语言处理问题的求解需求。因此,如何能够高效、准确地学习并获取专业领域知识,并构建自己的领域知识库,使计算机能够“理解”相应领域的问题需求,顺利完成目标,提供必要的知识保障,必然成为自然语言处理系统的关键问题之一。

2 领域知识及其学习

“领域知识”最初来自于相关的人工智能领域,在一些基于知识的专家系统和自然语言处理系统中有广泛的应用。最先给出明确定义的是Maja D'Hondt等几位专家,他们指出:领域知识系指存在于某个领域中的概念集之间的相互关系、概念的约束集合^[1]。

伴随知识工程任务的崛起,“领域知识”一词按照不同的应用需求,被定义不同的范畴。“领域知识”学习,也按照不同的需求具有一定的差异性。例如,在软件工程等任务中,“领域知识”被描述为“与系统环境有关的知识”,相应的领域知识学习目标也变成了系统环境中所定义的知识。而在自然语言处理任务中,“领域知识”是非空的有限集合,用来描述该领域中概念之间的相互关联、概念之间的相互约束、新概念之间的关系规则如何推导或计算。因此,本文定义领域知识为特定领域的语料类别特征,它所描述的是类别之间的联系和约束(区别),提供对新信息推理的依据。此外,本文定义领域知识学习为按照领域类别特征所建立起来的知识库,用以学习各种分类规则。

近几年,随着统计关系学习研究的兴起,结合概率图模型和一阶逻辑理论的马尔可夫逻辑网(Markov Logic Networks, MLNs)^[2]被成功应用于自然语言处理、机器学习、社会关系分析等领域^[3-5]。

本文基于MLNs技术,提出了一种知识学习方法,该方法是在传统知识获取的基础上,引入一阶逻辑对知识间的关系进行学习,从而进行谓词推理,来解决实际问题。此外,本文将文本分类问题作为具体应用,用以验证所提出方法的有效性。

2.1 传统的知识学习方法

作为传统的知识学习方法,支持向量机(Support Vector Machine, SVM)是Cortes和Vapnik在1995年首次提出^[6],其坚守计算学习理论中结构风险最小化原则^[7],结合统计学习理论VC维理论(Vapnik-Chervonenkis dimension)^[8],在有限样本中,对模型复杂性和学习能力寻找最佳折中,以期获得最好的泛化能力。SVM在解决小样本、非线性、高维模式识别等问题时,表现出独特的优势。

文本分类的主要过程是首先将文档表示为算法易

于处理的向量形式,然后通过特征选择和权值计算,构建分类器。可以将文本分类看成一个典型的二值分类任务,且由于文本向量的高维数以及数据稀疏等问题,在SVM出现后,研究者迅速将两者结合起来,并且在各方面均取得了较好的结果^[9]。目前,SVM是文本分类中效果最好的算法之一,因此本文在采用MLNs进行分类知识学习时,与SVM算法的结果进行对比,以期获得更佳的对比效果。

2.2 基于MLNs的知识学习方法

MLNs是将马尔可夫网络与一阶逻辑相结合的一种统计关系学习模型,是一种基于概率图模型的马尔可夫网,它的基本思想是在结合一阶逻辑规则的同时,让那些硬性规则有所松弛,即,针对一个特定问题,如果违反了其中的一条规则时,其存在的可能性将降低,但并非不可能。一个问题违反的规则越少,这个问题存在的可能性就越大。

因此,上述约束程度可以用权值进行表示,给每个规则都加上了一个特定的权值,它反映了对满足该规则的问题的约束力。一个规则的权重越大,对于满足和不满足该规则的不同问题而言,它们之间的差异就越大。在将权值无限增加后,通过MLNs得出的结果会向通过一阶逻辑知识得出的结果靠拢^[10]。

将MLNs与不确定推理中的一阶逻辑理论相结合,便构成了一种统计关系学习方法。在实际问题中,应用一阶逻辑理论,建立带有权值的一阶逻辑知识库,该知识库将包含所有构建马尔可夫网的原子。因此,MLNs可以作为构造马尔可夫网的基础。

文本分类问题中的知识学习技术一般都是基于统计概率来实现的,这种方法没有很好地考虑到知识之间的联系和规则,而MLNs能够比较好地解决此类问题。

MLNs涉及到一阶逻辑,需要进行相关模型的学习与推理,因此,在应用到文本分类知识学习中后,其实验步骤及流程与传统分类方法有很大的区别。

本文提出的基于MLNs的文本分类知识学习过程如算法1所示。

算法1 基于MLNs的文本分类知识学习过程

- 第1步 创建一阶逻辑规则;
- 第2步 数据预处理;
- 第3步 进行特征选择;
- 第4步 构建相关的知识库;
- 第5步 权值学习过程;
- 第6步 进行谓词推理;
- 第7步 分类效果的评价。

传统的文本分类知识学习过程,只考虑了统计学习方法(如SVM),其权值学习部分需要进行事先计算,并

通过建立一个分类器来进行新样本的预测。基于MLNs的文本分类知识学习过程,需要事先设计一阶逻辑规则以进行关系学习,权值学习部分则通过自我学习来得到,文本分类知识学习则是通过生成闭MLNs进行类别谓词推理,以判断其属性。图1给出了二值分类任务生成的闭MLNs示例。

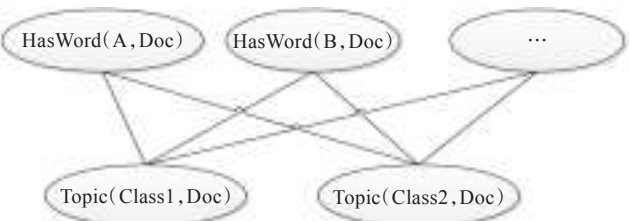


图1 二值分类任务生成的闭MLNs

3 MLNs的技术实现

目前基于传统机器学习方法(如SVM)的文本分类研究颇多,本文将文献[11]中基于SVM的文本分类知识学习技术的实验作为主要参照基线结果,详细介绍基于MLNs的领域知识学习技术的实现过程。

3.1 马尔可夫逻辑表示

在MLNs模型中,构造知识库的首要任务是创建一阶逻辑规则,通过建立相关规则,为后面的权值学习和谓词推理提供支持。因此,为了提高分类的准确率,需要选择正确的领域类别知识以及合理的样本来表示文本。

文本分类任务是在给定的分类体系下,根据其内容对未知类别标号的文本进行归类,该文本可以归为多类,而对于给定的类集合而言也可以不属于任何类。

人们通常采用文本中的“词”这一基本单位来描述文本,在进行文本描述时,也通常“精挑细选”那些最具有代表性的词。在传统的文本分类任务中,上述挑选出来的词通常叫做“特征词”。

在本文中,根据上述描述文本的方法,可以定义两个谓词来表示文本分类,即:

(1) $HasWord(Word, text)$:表示词 $Word$ 包含在文本 $text$ 中;

(2) $Topic(Class, text)$:表示文本 $text$ 属于类别 $Class$ 。

综上,可以构建如图2所示的马尔可夫逻辑表示,包括一阶逻辑规则谓词和公式。

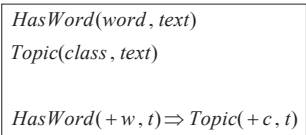


图2 文本分类的马尔可夫逻辑表示

图2中,公式 $HasWord(+w, t) \Rightarrow Topic(+c, t)$ 表示:若文档 t 中有词 w ,则判断该文档属于类别 c ,其中 w

和 c 是变量。其中, $HasWord$ 是作为该马尔可夫逻辑表示的证据谓词, $Topic$ 是作为查询谓词。

上述MLNs表示中,公式中变量之前的符号“+”,表示每个公式会根据变量所代表的不同个体,分别学习获得不同的权值。因此,图2中所构建的公式 $HasWord(+w, t) \Rightarrow Topic(+c, t)$,即使对同一篇文档 t ,在分别对应不同的词 w 和不同的类别 c 时,计算所得的权值也不同。

3.2 实验数据集

本文将文本分类问题作为具体应用,以验证所提出方法的有效性。分类的实验数据集为选自某特定领域2816篇新闻文档,包含了表1中所列出的10个类别,并将其中1882篇作为训练集,934篇作为测试集。以此来校验MLNs在领域知识学习任务中的效果。

表1 分类问题的实验数据集

类别名	缩写	训练集数量	测试集数量
政治	A	134	67
国家社会	B	134	66
经济	C	143	71
国际关系	D	147	73
军事总况	E	217	108
军队建设	F	166	83
武器装备	G	301	149
兵要地志	H	136	68
战争冲突	I	166	82
边海空情	J	338	167

3.3 特征选择

在机器学习和统计学中,特征选择也被称为属性选择、变量选择或变量子集选择。它是指为了构建模型而选择相关特征子集的过程(如属性、指标等),目标是根据某一特定评价标准选择最重要的特征子集。

人们通常采用文本中的“词”这一基本单位来进行文本描述,也通常选择那些最具有代表性的词,构成词的集合,而并非将所有词都作为特征。在传统的文本分类任务中,上述挑选出来的词通常叫做“特征词”。

如果将文本中的所有词都作为特征,会造成特征词的数量极其庞大,并直接导致以下问题:

(1)特征分析及模型训练时间延长,这个问题在MLNs的权值学习时体现的会更加突出。

(2)特征维度过高,直接引起“维度灾难”,会出现模型过拟合现象,实际应用的效果很不理想。

因此,特征选择就是试图通过剔除冗余特征以及不相关特征,去减少特征数据集中的属性(或者称为特征)的数目,能够使得模型得到一定的简化,达到提高模型效果和缩短运行时间的目的。与降维的区别是:特征选择的方法是从原始特征数据集中挑选出子集,是一种包含的关系,不更改原始的特征空间。

目前常用的特征词选择方法很多,如基于文档频率 TF-IDF 方法(Term Frequency-Inverse Document Frequency)、卡方检验法、互信息法、信息增益法(Information Gain, IG)、期望值交叉熵法^[12-13]等。

在信息增益中,特征选择的衡量标准是看该特征能够为分类系统带来多少信息,带来的信息越多,说明该特征越具有类别区分能力。对一个特征 T 而言,系统有 T 和没 T 时所包含的信息量会发生变化,信息量的差值就是这个特征给系统带来的信息量,即信息增益值。鉴于信息增益能够较好地衡量特征 T 与类别 C 之间的联系,本文在进行特征选择时,采用信息增益值作为评价标准。

信息增益的计算方法如式(1)所示,实际计算中,特征词 t 对文本类别 c 的信息增益值,通过文本类别 c 中有特征词 t 和没有特征词 t 的文本数来衡量。

$$IG(t) = - \sum_{i=1}^m P(c_i) \lg P(c_i) + P(t) \sum_{i=1}^m P(c_i|t) \lg P(c_i|t) + P(\bar{t}) \sum_{i=1}^m P(c_i|\bar{t}) \lg P(c_i|\bar{t}) \tag{1}$$

其中, $P(c_i)$ 表示 c_i 类文本在所有文本集中所占的比例; $P(t)$ 表示包含词 t 的那些文本在所有文本集中所占的比例; $P(c_i|t)$ 表示包含词 t 的所有文本属于类 c_i 的条件概率; $P(\bar{t}) = 1 - P(t)$, 表示不包含词 t 的文本在所有文本集中所占的比例; $P(c_i|\bar{t})$ 表示不含有词 t 的所有文本中,属于类 c_i 的条件概率; m 表示文本集类别总数。

本文的特征选择方法如算法 2 所示。

算法 2 基于信息增益方法选择文本特征词的过程

第 1 步 对文本进行分词处理,本文实验中采用了中科院计算所的汉语词法分析系统 ICTCLAS。

第 2 步 采用开源程序 Lucene 的 Java 版建立倒排索引。

第 3 步 建立索引,求出各个词的信息增益值,并根据每个词的信息增益值的大小,除去低于特定阈值的那些词,剩下的词将作为文本数据集的特征候选。

第 4 步 从上述候选特征集中依次选取不同数目的特征量进行实验,选取比例分别是 5%、10%、25%、50%、75% 以及 100%,根据实验结果进行特征的最佳数量选择。

实验结果如表 2 所示。

表 2 不同数量特征词的查准率比较

比例	5%	10%	25%	50%	75%	100%
查准率	68.57%	74.33%	74.16%	74.88%	74.24%	73.76%

由表 2 的结果可以看出,当特征词从 5% 增加到 10% 时,查准率有较明显的提高,并且在选取 50% 时取得最佳结果。对于 MLNs 而言,选取的特征词较多时,高维特征将会构建更为复杂庞大的马尔可夫网,处理起

来耗时越多。但考虑实验效果,仍然选取 50% 的特征数量进行实验。

3.4 领域知识库的建立

知识库是用于知识管理的一种特殊的知识集群,知识库中的信息源于领域专家,是针对某一领域问题求解的需要而建立,包括基本事实、规则和其他相关信息等。

本文提出的基于 MLNs 的知识学习策略,与传统方法进行文本分类的区别之一是需要构建自己的知识库。该知识库中存取的信息是原始数据特征与一阶逻辑规则相结合而生成的谓词原子,需要预先建立相应的一阶逻辑规则。图 3 给出了领域知识库的部分示例。

HasWord("战场", F50)

HasWord("协议", B67)

...

Topic("B", B67)

Topic("F", F50)

...

图 3 领域知识库部分结果示例

图 3 中, *HasWord* 和 *Topic* 是定义的两个谓词;“B67”和“F50”分别表示表 1 中对应类别文本集中的文本序号;“战场”和“协议”是从对应类别中提取的特征词。本文所提出的领域知识库,相当于将特征词和类别作为“变量”,带入一阶逻辑规则中的谓词公式中。在进行新文本推理时,也要转化为图 3 所示的形式。在图 2 的示例中,由于 *Topic* 作为查询谓词出现,其中的类别需要推理才能完成,因此实际操作过程中,仅需将文本转换为 *HasWord* 证据谓词的形式即可。

按照本文的描述,建立领域知识库的过程即为将文本分类任务的训练集中的每个样本,转换成特征词与类别的一阶逻辑谓词形式。建立了相应的领域知识库,即可以进行 MLNs 的学习,其中关键的问题是权值计算和谓词推理两部分。

3.5 权值计算和谓词推理

MLNs 为一阶谓词附加权值,容忍知识库中存在不完整和互相矛盾的知识,具有较好的处理不确定性问题的能力。采用 MLNs 进行知识学习,权值计算和谓词推理也是有别于传统方法的。

传统方法需要建立文本分类器,进行文本类别预测,其权值为每个特征词的权值,一般通过计算训练集中的特征词而获得(例如 TF 或 TF-IDF 方法)。MLNs 计算的权值是闭 MLNs 中每条边的权值指(如图 1 所示),是在知识库构建完毕之后,通过知识库与一阶逻辑规则的结合,自动学习获得。

此外,MLNs 的权值也可以人工设置,只是本文研究的环境中,领域类别之间、特征词之间都有着或多或少的联系,为了不相互干扰,本文仍沿用权值自动学习

的方法。并且实验中采用判别式训练方法进行权值学习,其最终生成一系列由闭谓词到闭谓词的带有权重的公式。图4为MLNs的部分权值学习结果示例,左边的数值代表权重(如:173.793、2.591 5等)。

```
// 173.793 HasWord("会议",p) => Topic("A",p)
173.793 !HasWord("会议",a1) v Topic("A",a1)

// -0.203133 HasWord("军队",p) => Topic("A",p)
-0.203133 !HasWord("军队",a1) v Topic("A",a1)

// 2.5915 HasWord("火力",p) => Topic("A",p)
2.5915 !HasWord("火力",a1) v Topic("A",a1)

// -1.30408 HasWord("多种",p) => Topic("A",p)
-1.30408 !HasWord("多种",a1) v Topic("A",a1)

// -0.338531 HasWord("芯片",p) => Topic("A",p)
-0.338531 !HasWord("芯片",a1) v Topic("A",a1)

// -1.20845 HasWord("宏观",p) => Topic("A",p)
-1.20845 !HasWord("宏观",a1) v Topic("A",a1)
```

图4 MLNs的权值学习结果示例

MLNs的权值计算完毕,就可以进行谓词推理。本文的谓词推理主要是指从 *HasWord* 到 *Topic* 的推理过程,是将训练样本集转化为 *HasWord* 形式后,再结合学习得到的权值以及一阶逻辑规则,通过推理实现类别的预测。谓词推理采用的是融合切片采样的马尔可夫蒙特卡洛(MCMC)方法,即MC-SAT算法^[14]。

MC-SAT算法的主要思路:是在原MCMC算法中加入一个辅助参数,用以捕捉变量之间的某些确定关系,再结合可满足性测试和模拟退火算法进行切片采样^[15]。文献[16]证明了该方法严格遵从MCMC算法的原则,并验证了该算法的良好效果。本文的相关实验采用了由Domingos研究团队开发的平台 Alchemy^[17],该平台的主体采用了MC-SAT算法,并结合了另外一种基于“需要时再激活”的惰性思想的算法——LazySAT算法,从而有效地避免了不必要的原子和从句的展开,提高了实验效率。

3.6 性能评价

传统的评价文本分类的性能指标,单个类别的分类效果通常采用查准率(Precision)、查全率(Recall)和 F1 值评价。其描述形式如公式(2)所示(其中, *P* 表示查准率, *R* 表示查全率)。

$$P = \frac{\text{分类正确的文本数}}{\text{实际分类的文本数}} \times 100\%$$
$$R = \frac{\text{分类正确的文本数}}{\text{应有的文本数}} \times 100\%$$
$$F1 = \frac{P \times R \times 2}{P + R} \times 100\%$$
 (2)

公式(2)描述的评价方法主要针对单个类别的结果,仅是一种局部评价方式。另外一种衡量文本分类器的指标称为宏平均(Macro-average)与微平均(micro-average),该指标是为了评估算法在整个数据集上的分类性能^[18]。

宏平均是每一个类的性能指标的算术平均值,而微平均是每一个文档的性能指标的算术平均。

计算宏平均时,先得到每个类别的查准率和查全率,然后计算所有类别的查准率和查全率的平均值。公式(3)给出了宏平均的查准率(*Macro-P*)、查全率(*Macro-R*)、F1值(*Macro-F1*)的计算方法,其中 *|C|* 表示类别总数。

$$Macro - P = \frac{\sum_{i=1}^{|C|} P_i}{|C|}$$
$$Macro - R = \frac{\sum_{i=1}^{|C|} R_i}{|C|}$$
$$Macro - F1 = \frac{2 \times Macro - P \times Macro - R}{Macro - P + Macro - R}$$
 (3)

微平均是每个文档性能指标的算术平均,公式(4)给出了微平均的查准率(*Micro-P*)、查全率(*Micro-R*)、F1值(*Micro-F1*)的计算方法。

$$Micro - P = \frac{\sum_{i=1}^{|C|} a_i}{\sum_{i=1}^{|C|} a_i + \sum_{i=1}^{|C|} c_i}$$
$$Micro - R = \frac{\sum_{i=1}^{|C|} a_i}{\sum_{i=1}^{|C|} a_i + \sum_{i=1}^{|C|} b_i}$$
$$Micro - F1 = \frac{2 \times Micro - P \times Micro - R}{Micro - P + Micro - R}$$
 (4)

其中,对于给定的某个类别, *a* 表示被正确分到该类的文档个数, *b* 表示被误分到该类的文档个数, *c* 表示属于该类但被误分到其他类别的文档个数。

按照宏平均的公式,文本集中的每个类别重要度均相同,因此较容易受较小类别的影响;而微平均更看重大类的分类效果。在实际应用中,如果类别的结果差异较大,宏平均和微平均计算得到的结果可能存在较大不同。

4 实验结果与分析

表1给出了本文分类问题的实验数据集,baseline实验采用基于SVM的文本分类知识学习技术,采用宏平均和微平均评价实验结果,表3给出了部分对比实验结果。

表3 对比实验结果					%
方法	SVM	MLN15	MLN30	MLN50	MLN100
Macro - P	73.49	69.47	74.25	78.34	80.72
Macro - R	74.76	70.65	74.98	81.35	83.83
Macro - F1	74.12	70.06	74.61	79.82	82.25
Micro - P	75.70	73.61	79.12	81.16	83.47
Micro - R	74.60	73.24	78.21	81.89	84.02
Micro - F1	75.15	73.43	78.66	81.53	83.74

表3列出的方法中,SVM代表的是基于SVM的文本分类知识学习得出的结果,而MLN15、MLN30、MLN50和MLN100分别代表用MLNs进行权值学习时迭代15轮次、30轮次、50轮次和100轮次的结果。由于MLNs生成的是图结构,实验时迭代次数过多时会非常耗时,因此为了提高效率,选取了若干轮次进行对比实验。

由表3的结果可以看出,采用MLNs进行分类知识学习时,分类效果与MLNs权值学习的迭代次数有关。在进行了15轮次权值学习迭代时,仍然是基于SVM的传统知识学习分类技术略高。随着MLNs的权值学习迭代轮次的增加,在达到30轮次权值学习迭代后,分类效果已经超过了基于SVM的文本分类结果。迭代达到100轮次后,分类效果比baseline的结果提高10%左右。

总之,结合了统计学习和关系学习的MLNs应用在领域知识学习中可行,该方法能很好地把握知识之间的联系,挖掘出更加准确的知识。

5 结束语

本文重点研究了领域知识学习技术,基于马尔可夫逻辑网技术,提出了一种文本分类知识学习方法,详细介绍了这种知识学习技术应用在相关领域中的实验设计流程,包括文本分类的马尔可夫逻辑表示、特征选择、领域知识库构建以及MLNs的权值学习和谓词推理等。相应的实验结果表明,与传统基于SVM的文本分类知识学习的结果相比,采用MLNs的文本分类知识学习方法,在权值学习迭代到一定轮数后能获得更好的分类结果,在把握领域知识之间的联系方面显得更突出。

参考文献:

- [1] D'Hondt M, D'Hondt T. Is domain knowledge an aspect? [C]// Proceedings of the ECOOP99 Aspect Oriented Programming Workshop, 1999: 293-294.
- [2] Richardson M, Domingos P. Markov logic networks [J]. Machine Learning, 2006, 62(1): 107-136.
- [3] He H, Li Z G, Yao C C, et al. Sentiment classification technology based on Markov logic networks [J]. New Review of Hypermedia and Multimedia, 2016(4): 1-15.
- [4] Mahajan M J A. Robotic task-planning by using Markov logic networks [J]. International Journal of Trend in Research and Development, 2015, 2(5): 9-12.
- [5] 杨立公, 汤世平, 朱俭. 基于马尔可夫逻辑网的句子情感分析方法 [J]. 北京理工大学学报, 2013, 33(6): 600-604.
- [6] Cortes C, Vapnik V. Support Vector Machine [J]. Machine Learning, 1995, 20(3): 273-297.
- [7] Vapnik V N. The nature of statistical learning theory [M]. New York: Springer, 1995.
- [8] Vapnik V, Levin E, Le Cun Y. Measuring the VC-Dimension of a learning machine [J]. Neural Computation, 1994, 6(5): 851-876.
- [9] Joachims T. Text categorization with support vector machines: Learning with many relevant features [C]// Proceedings of the 10th European Conference on Machine Learning (ECML-98). Chemnitz: Springer-Verlag, 1998: 137-142.
- [10] Getoor L, Taskar B. Markov Logic: A unifying framework for statistical relational learning [M]// Introduction to Statistical Relational Learning, 2007: 339-371.
- [11] Manyika J, Chui M, Brown B, et al. Big data: The next frontier for innovation, competition, and productivity [D]. McKinsey Global Institute, 2011.
- [12] Yang Y M, Pederson J O. A Comparative study on feature selection in text categorization [C]// Proceedings of the 14th International Conference on Machine Learning. San Francisco: Morgan Kaufmann Publishers, 1997: 412-420.
- [13] Mladenic D, Grobelnik M. Feature selection for unbalanced class distribution and Naive Bayes [C]// Proceedings of the Sixteenth International Conference on Machine Learning. Bled: Morgan Kaufmann, 1999: 258-267.
- [14] Poon H, Domingos P. Sound and efficient inference with probabilistic and deterministic dependencies [C]// Proceedings of the Twenty-first National Conference on Artificial Intelligence, Boston, 2006: 458-463.
- [15] 孙舒杨, 刘大有, 孙成敏, 等. 统计关系学习模型 Markov 逻辑网综述 [J]. 计算机应用研究, 2007, 24(2): 1-3.
- [16] Liu D C, Nocedal J. On the limited memory BFGS method for large scale optimization [J]. Mathematical Programming, 1989, 45(1): 503-528.
- [17] Kok S, Sumner M, Richardson M, et al. The alchemy system for statistical relational AI [R/OL]. <http://alchemy.cs.washington.edu/>.
- [18] Kargupta H, Joshi A, Sivakumar K, et al. Data mining: Next generation challenges and future directions [M]. [S.l.]: MIT Press, 2004: 105-125.