

支持向量机规则提取*

王 强,沈永平,陈英武

(国防科技大学 信息系统与管理学院,湖南 长沙 410073)

摘 要 :支持向量机是一种黑箱模型,其学习到的知识蕴含在决策函数中,不仅影响了用户对利用支持向量机技术构建智能系统的信心,还阻碍了支持向量机技术在数据挖掘领域的应用。由于对支持向量机规则提取进行研究有助于解决上述问题,因此该领域正成为机器学习和智能计算界的研究热点。分析了具有代表性的支持向量机规则提取算法,并提出该领域未来的研究重点。

关键词 :支持向量机;机器学习;规则提取;知识获取;数据挖掘

中图分类号 :TP18 文献标识码 :A

Rule Extraction from Support Vector Machines

WANG Qiang, SHEN Yong-ping, CHEN Ying-wu

(College of Information System and Management, National Univ. of Defense Technology, Changsha 410073, China)

Abstract Support vector machines is a blackbox model whose knowledge is concealed in the decision function. This has not only weakened the confidence of users in building intelligent systems using support vector machines techniques, but also hindered the application of support vector machines to data mining. Since extracting rules from support vector machines help to solve those problems, this area is becoming a hot topic in both machine learning and intelligent computing communities. In this paper, the typical algorithms for rule extraction from support vector machines are introduced, and some issues valuable for future exploration in this area are indicated.

Key words :support vector machines; machine learning; rule extraction; knowledge acquisition; data mining

支持向量机(support vector machine, SVM)由于其具有较为严格的统计学习理论基础,较好地解决了其它机器学习方法存在的过学习问题、局部最小问题和维数灾问题,已经被成功地应用到很多不同领域的数据挖掘中,并取得了大量成果。然而,其进一步的发展却受到了一个固有缺陷的限制,即 SVM 具有“黑箱性”。SVM 学习到的知识蕴藏在由少量支持向量样本“加权”得到的决策函数中,用户无法了解 SVM 到底学到了什么、能处理什么样的任务,也无从知道 SVM 如何进行分类和预测、为什么得出这样或那样的推理结论。一般来说,“可解释性”是可靠系统的必备特性,由于通常的 SVM 模型都是“不可解释性”的,这在一定程度影响了用户通过 SVM 技术构建智能决策支持系统的信心。SVM 方法虽然可以给出所建模型的推广能力的严格的界,从而保证对未知示例产生正确的决策结果,但这并没有消除用户对可理解性的偏好。此外,训练好的 SVM 学习到的知识不能以容易理解的形式提交给决策者,这也是 SVM 方法难以用于管理决策领域数据挖掘的主要原因之一。

规则提取是建立具有“黑箱性”机器学习系统解释机制的一种途径,其目的是将学习系统中隐含的知识以一种易于理解的方式明确地表达出来。一方面,以规则的形式表示学习系统获取的知识可以方便使用者和设计者理解学习系统的推理过程;另一方面,通过提取出来的规则,用户可以发现输入数据中一些以前被忽略的重要关系,从而帮助用户进行决策处理。此外,这些规则知识可直接用于构造决策支持系统的知识库。

正因为规则提取具有上述重要意义,这方面的工作受到越来越多的关注,规则提取研究目前已成为机器学习领域一个重要的研究方向。近年来,对于神经网络规则提取的研究取得了一定的成果^[1],而对

* 收稿日期 2005 - 09 - 20

作者简介:王强(1971—),男,博士生。

SVM 规则提取的研究工作正引起研究人员的重视。

SVM 学习到的知识蕴藏在由少量支持向量和其相关参数构成的决策函数中,规则提取的任务就是将这些知识用易于理解的方式表达出来。根据提取规则算法设计思想的不同,目前的 SVM 规则提取方法大致可以分成两大类,即基于结构分析的方法和基于学习的方法。本文将对这两大类中的典型算法进行介绍和分析。

1 基于结构分析的规则提取算法

基于结构分析的 SVM 规则提取算法把规则提取过程看成是一个搜索过程。其基本思想是以支持向量为单位将 SVM 分解为若干单个支持向量的集合,对每一支持向量搜索和提取规则,最后对这些规则进行组合以描述整个 SVM 的特性,从而把已训练好的 SVM 映射成对应的规则。目前比较典型的算法有以下几种:

(1) SVM + prototypes 算法

SVM + prototypes 算法是由 Núñez 等人^[2-3]提出的一种基于结构分析的规则提取方法。该算法的基本思想是在 SVM 得到决策函数后,采用 k -均值聚类算法确定每一类别的典型向量。应用几何分析方法,由典型向量和支撑向量在输入空间中定义一个椭圆体的区域,进而将该区域映射为相应的“IF... THEN...”规则。获得的规则可表示为方程式规则或区间规则(如图 1 所示)。

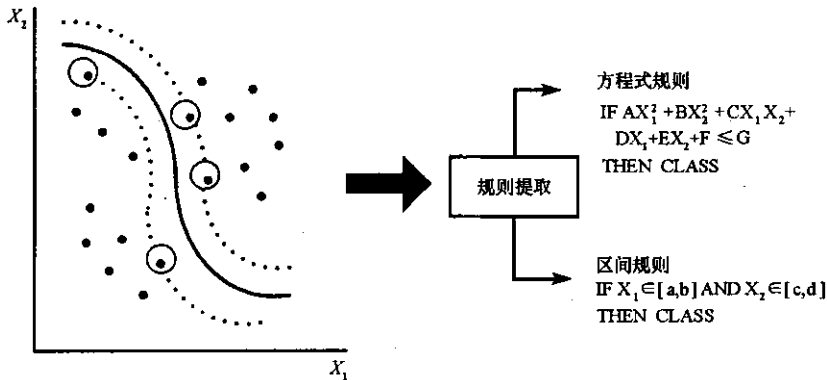


图 1 SVM + prototypes 算法产生的规则

Fig. 1 Rules generated by SVM + prototypes algorithm

算法首先由典型向量和离其最远的支持向量决定椭圆体的一条轴线,再通过几何分析的方法确定椭圆体的另一条轴线及其顶点。算法中采用三类数据来确定这些顶点:支持向量本身或从支持向量派生出来的数据点,亦或是离典型向量最远的数据点。算法采用递增策略决定每一类别椭圆体的数量,从而控制产生规则的数量。算法从初始典型向量开始,生成相关的椭圆体区域。对该椭圆体区域进行分割测试,如果结果表明不存在类别之间的交叉重叠,则将此区域转换为一条规则。否则,产生新的椭圆体区域。重复这个过程直到分割测试表明新的椭圆体区域不存在类别之间的交叉重叠或达到规定的最大迭代次数,将分割得到的所有椭圆体区域转换为相应的规则。

SVM + prototypes 算法的主要缺陷是由确定典型向量的聚类算法引起的。 k -均值聚类算法要求用户必须事先给出聚类数 k ,并且初始聚类中心的选择是随机的。这些因素都会影响聚类算法的有效性,进而影响产生规则的数量及其准确性。另外,当数据维数增加时,用简单的几何分析方法确定高维空间区域的边界将是一件非常复杂的工作。

(2) RuExSVM 算法

RuExSVM 算法是由 Fu 等人^[4]提出的一种从 SVM 中提取规则三个阶段算法。该算法的基本思想是在 SVM 得到决策函数后,由支持向量和决策超曲线在输入空间中定义一个超矩形区域,进而将该区域映射为相应的“IF... THEN...”规则。

在算法的初始阶段,沿着输入空间的每一条轴线,对给定类别的某一支持向量进行直线延伸,这些

直线与 SVM 决策超曲线相交。由这些交点可以确定该支持向量的一个初始超矩形区域。如图 2 所示为在二维的输入空间中,支持向量 A 和 C 各自的超矩形区域。算法中直线与决策超曲线的交点由牛顿法进行求解。获取初始超矩形区域后,算法进入调整阶段,其目的是提高规则对数据分类的准确性。在某一支持向量决定的初始超矩形区域内,搜索其它类别的数据点。通过收缩超矩形区域来除去这些点,得到新的超矩形区域。重复这个过程直到超矩形区域内不存在任何其它类别的数据点。此时就可以将此超矩形区域转换为相应的规则。最后,算法通过规则间交叉重叠的判断对冗余规则进行剪枝,获得一个简洁的规则集。

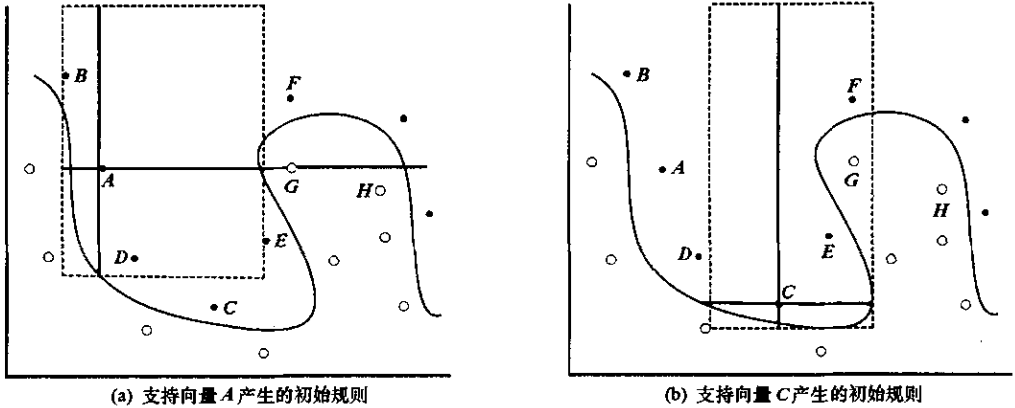


图 2 RulExSVM 算法产生的规则
Fig.2 Rules generated by RulExSVM algorithm

RulExSVM 算法充分利用了 SVM 决策边界所蕴含的信息,所提取的规则具有较高的准确性,同时规则的前提数较少。但是算法需要对每一个支持向量在输入空间的所有维度中进行规则搜索,其复杂度为 $O(mnt)$,其中 m 为输入空间的维数, n 为支持向量的个数, t 为重复迭代次数。这使其适用范围受到很大的限制,不适合在线数据挖掘过程和高维数据情况下的规则提取。

(3) 属性组合算法

魏玲等人^[5-6]针对决策系统知识发现,提出了一种从 SVM 中提取决策系统决策规则的属性组合算法。该算法的主要思想是首先对决策系统进行属性约简,然后应用 SVM 得到决策函数,再对每一个支持向量的输入空间属性进行任意组合,从而提取规则。算法从第一个支持向量开始,组合属性形成规则。用获得的每条规则检验原始训练样本集 U ,正确规则集合记为 G_k ,能被正确分类的样本集合记为 U_k 。如果 $\cup U_k = U$,则结束,获得的规则为 $\cup G_k$;否则,对样本集 $U - U_k$ 进行训练,得到新支持向量集,再对每一个支持向量进行属性组合形成规则,并用获得的每条规则检验样本集 $U - U_k$ 。重复上述过程,直到 $\cup U_k = U$ 获得的规则为 $\cup G_k$ 。

该算法通过属性约简改善了规则获取的难度和速度,适用于决策系统的规则提取。然而,决策系统的缩小可能会丢失一些信息,使提取到的规则不完善。另外,该算法的复杂度为 $O(k^m nt)$,其中 m 为输入空间的维数, n 为支持向量的个数, k 为属性取值数, t 为重复迭代次数。当属性取值数 k 和属性维数 m 较高时,会出现属性组合爆炸,使算法的时间代价相当高。

2 基于学习的规则提取算法

另一类 SVM 规则提取方法是基于学习的规则提取算法,这类方法和基于结构分析的方法不同之处在于使用该方法进行提取规则时不考虑 SVM 的类型和结构,只关心 SVM 的输入与输出之间的映射关系。基于学习的 SVM 规则提取算法将规则提取视为一个归纳学习问题。其基本思想是对给定一个训练好的 SVM 以及用于对其进行训练的数据,将 SVM 视为一个产生学习样本的“黑箱”,由能够产生规则的第二种学习算法来提取规则,这些规则将以与 SVM 相同的方式对示例进行划分。

Barakat 和 Diederich^[7]提出了一种基于学习的 SVM 规则提取算法。该算法使用两个有差异而又有

相互联系的数据集,分别用于训练 SVM 和第二种学习算法,实现决策规则的提取。首先,应用 SVM,由带类别标识的训练数据集 A 建立目标问题的模型。然后产生一个与训练数据集 A 具有相同属性而属性取值不同的不带类别标识的数据集 B 。将数据集 B 作为测试集,用训练好的 SVM 为这些数据产生类别标识,以验证 SVM 的推广能力。由此得到一个综合数据集 C 。最后使用各种具有解释能力的机器学习方法从综合数据集 C 中获取反映 SVM 推广能力的规则。Barakat 和 Diederich 采用决策树学习的 C5 算法进行规则提取。

该算法最大的特点是其具有一定的通用性,不受 SVM 类型和结构的限制。该算法的主要缺陷在于决策树的可理解性较差,且数据集 B 的产生没有严格的依据。同时,由于采用了决策树算法,其算法的复杂度为 $O(nmk)$,其中 n 为学习样本集大小, m 为输入空间的维数, k 为属性取值数。因此,算法的效率与属性维数密切相关。为此,Barakat 和 Diederich^[8]提出先应用回归模型进行数据约简,以降低算法的时间开销。

综合基于结构分析和基于学习两种规则提取算法的基本思想,Barakat 和 Diederich^[9]又提出一种“折衷的”规则提取方法。该算法的基本思路与基于学习的规则提取算法一致,只是在数据集 B 的构造上有所改变。算法中数据集 B 由训练数据集 A 中的支持向量除去类别标识后组成。他们认为,这样构造数据集 B ,充分利用了学习模型中支持向量提供的信息,从而实现基于结构分析和基于学习两种规则提取算法的结合。但从算法的本质来看,其仍然是一种基于学习的规则提取算法。

3 进一步研究的内容

作为一种获取 SVM 中隐含知识的手段,SVM 规则提取算法提供了一把打开蕴藏丰富知识的“黑盒”的钥匙。现有的 SVM 规则提取算法在应用中仍有很多不足之处。对这些不足之处的改进将成为新的研究目标。以下几个方面的问题可望成为该领域的重要研究内容。

(1)目前的 SVM 规则提取算法,无论是基于结构分析还是基于学习的,都有很高的计算复杂度,这就使得在数据规模较大时难以奏效。而在现实世界问题中真实数据的规模通常都比较大。因此,如果不能大幅度降低现有方法的计算复杂度或者提出新的计算复杂度小的方法,SVM 规则提取领域研究成果的实用性将很难得到保证。

(2)目前 SVM 规则提取的工作主要集中在输出为离散值的 SVM 分类器上,对于输出为连续值的回归估计型 SVM 研究很少。而后者在时序预测、信号处理等领域有广泛的应用。因此,从回归估计型 SVM 中提取规则将是一项重要的研究课题。

(3)现有的 SVM 规则提取方法几乎都力图从 SVM 中提取产生式规则或类似形式的知识。但在某些场合,非规则形式的知识往往具有更好的表达能力。因此,从 SVM 中提取非规则形式的知识有可能成为今后的一大研究重点。在这方面,Melnik 和 Pollack 的工作有一定的启发性^[10]。他们提出了一种新颖的、从神经网络分类器中提取定性知识的方法。该方法不依赖于输入空间的维数,可以提取出分类决策区域图,不仅适用于神经网络,还可用于 SVM 等其它类型的分类器。

(4)从 SVM 与模糊推理模型等价转换的角度,将 SVM 决策函数蕴含的决策知识用模糊规则进行解释。目前关于 SVM 与模糊推理系统关系的研究已经证明,在一定条件下,SVM 和模糊推理模型可以相互转换^[11-14]。SVM 模型可以解释成取加、乘运算的模糊推理模型,每个作为支持向量的样本数据及其系数对应一条模糊推理的规则。某些模糊推理模型也可以看成是由 SVM 方法得到的一种支持向量机,一条模糊推理规则对应一个“支持向量”。将 SVM 转换为等价模糊推理模型,可以很容易地获取一些模糊规则。

4 结语

缺乏对隐含知识的解释能力是阻碍 SVM 技术被广泛应用的重要因素之一。从 SVM 中提取规则的研究工作,除了可以用易于理解的方式解释 SVM 学习到的知识外,对于知识发现的研究,专家系统中的知识获取工作的研究,以及如何融合 SVM 技术和符号学习技术的研究都很有参考价值。

现有的各种 SVM 规则提取算法都还存在着不同的缺陷,如算法的复杂度较高,知识的表示形式单一等。我们相信,随着规则提取研究的不断深入,SVM 技术将会得到更为广泛的应用。

参考文献:

- [1] Tickle A , Andrews R , Mostefa G , et al. The Truth Will Come to Light : Directions and Challenges in Extracting the Knowledge Embedded within Trained Artificial Neural Networks[J] . IEEE Transactions on Neural Networks ,1998 ,9(6) :1057 - 1068 .
- [2] Núñez H , Angulo C , Català A. Rule Extraction from Support Vector Machines[A] . In : Proceedings of the 2002 European Symposium on Artificial Neural Networks[C] ,2002 :107 - 112 .
- [3] Núñez H , Angulo C , Català A. Support Vector Machines with Symbolic Interpretation[A] . In : Proceedings of the VII Brazilian Symposium on Neural Networks[C] ,2002 :142 - 147 .
- [4] Fu X J , Ong C J , Keerthi S. Extracting the Knowledge Embedded in Support Vector Machine[A] . In : Proceedings of 2004 IEEE International Joint Conference on Neural Networks[C] ,2004 :291 - 296 .
- [5] Wei L , Qi J J , Zhang W X. Knowledge Discovery of Decision Table Based on Support Vector Machine[C] . In : Proceedings of 2003 IEEE International Conference on Machine Learning and Cybernetics[C] ,2003 :1195 - 1200 .
- [6] 魏玲 祁建军 张文修. 基于支持向量机的决策系统知识发现[J] . 西安交通大学学报 2003 ,37(10) :995 - 998 .
- [7] Barakat N , Diederich J. Learning-based Rule-extraction from Support Vector Machines[A] . In : Proceedings of the 14th International Conference on Computer Theory and Applications[C] ,2004 :247 - 252 .
- [8] Diederich J , Barakat N. Hybrid Rule-extraction from Support Vector Machines[A] . In : Proceedings of 2004 IEEE Conference on Cybernetics and Intelligent Systems[C] ,2004 :1270 - 1275 .
- [9] Barakat N , Diederich J. Eclectic Rule-extraction from Support Vector Machines[J] . International Journal of Computational Intelligence ,2005 ,9(1) :59 - 62 .
- [10] Melnik O , Pollack J. Using Graphs to Analyze High-dimensional Classifiers[A] . In : Proceedings of the IEEE-INNS-ENNS International Joint Conference on Neural Networks[C] ,2000 :425 - 430 .
- [11] Chen Y X , Wang J Z. Kernel Machines and Additive Fuzzy Systems : Classification and Function Approximation[A] . In : Proceedings of the IEEE International Conference on Fuzzy Systems[C] ,2003 :789 - 795 .
- [12] Chen Y X , Wang J Z. Support Vector Learning for Fuzzy Rule-based Classification Systems[J] . IEEE Transactions on Fuzzy Systems ,2003 ,11(6) :716 - 728 .
- [13] Chen Y Y , Feng H Z. A Kernel Method for Fuzzy Systems Modeling and Approximate Reasoning[A] . In : Proceedings of the IEEE 22nd International Conference of the North American Fuzzy Information Processing Society[C] ,2003 :307 - 311 .
- [14] 陈永义. 支持向量机方法与模糊系统[J] . 模糊系统与数学 ,2005 ,19(1) :1 - 11 .

(上接第 87 页)

参考文献:

- [1] 何峰 梁甸农 刘建平. 星载寄生式 SAR 系统干涉信号模型与相对高程测量性能分析[J] . 国防科技大学学报 ,2005 ,27(3) :72 - 76 .
- [2] Homer J , Kubik K , Mojarrabi B , et al. Passive Bistatic Radar Sensing with LEOS Based Transmitters[A] . IGARSS IEEE International Conference [C] ,2002 :438 - 440 .
- [3] Antonio M. A Bistatic SAR for Earth Observation[A] . Proceedings of SPIE[C] ,2000 :4173 .
- [4] Massonnet D. Capabilities and Limitations of the Interferometric Cartwheel[J] . IEEE Transactions on Geoscience and Remote Sensing ,2001 :506 - 520 .
- [5] 汤子跃 张守融. 双站合成孔径雷达系统原理[M] .北京 :科学出版社 ,2003 .
- [6] Aguasca A , Broquetas A , Fdez de Muniain J , et al. Remote Bistatic Receiver Synchronization Using DLL Techniques[A] . International Radar Conference[C] ,1992 :489 - 492 .
- [7] 杨振起 张永顺 骆永军. 双基地雷达系统[M] .北京 :国防工业出版社 ,1998 .

