

学校编码: 10384

分类号 _____

学号: X2006223013

UDC _____

厦 门 大 学

硕 士 学 位 论 文

灰色支持向量机在小样本预测中的应用研究

Application Research of Grey Support Vector Machines
on Small Sample Set

林耀进

指导教师姓名: 吴顺祥 教授

陈宝兴 教授

专 业 名 称: 控制理论与控制工程

论文提交日期: 2009 年 5 月

论文答辩时间: 2009 年 6 月

学位授予日期: 年 月

答辩委员会主席: _____

评 阅 人: _____

2009 年 6 月

厦门大学博硕士学位论文摘要库

厦门大学学位论文原创性声明

兹提交的学位论文，是本人在导师指导下独立完成的研究成果。本人在论文写作中参考的其他个人或集体的研究成果，均在文中以明确方式标明。本人依法享有和承担由此论文产生的权利和责任。

声明人（签名）：

年 月 日

厦门大学博硕士学位论文摘要库

厦门大学学位论文著作权使用声明

本人完全了解厦门大学有关保留、使用学位论文的规定。厦门大学有权保留并向国家主管部门或其它指定机构送交论文的纸质版和电子版，有权将学位论文用于非营利目的的少量复制并允许论文进入学校图书馆被查阅，有权将学位论文的内容编入有关数据库进行检索，有权将学位论文的标题和摘要汇编出版。保密的学位论文在解密后适应本规定。

本学位论文属于

- 1、保密（ ），在 年解密后适用本授权书。
- 2、不保密（ ）

（请在以上相应括号内打“√”）

作者签名：

日期： 年 月 日

导师签名：

日期： 年 月 日

厦门大学博硕士学位论文摘要库

摘 要

灰色系统理论的研究对象是“部分信息已知，部分信息未知”的“小样本”、“贫”信息不确定系统，它通过对“部分”已知信息的生成、开发了解、认识现实世界，实现对系统运行行为演化规律的正确把握和描述。灰色动态建模是灰色系统理论的核心，也是灰色系统理论与世界相结合的桥梁。

支持向量机是建立在统计学习理论的VC维理论和结构风险最小化原则基础上的新型机器学习方法。它根据有限样本信息在模型的复杂度和期望风险之间寻求最佳折衷，能够获得更好的泛化性能。与传统的神经网络等学习方法相比，支持向量机具有泛化能力强、维数不敏感、收敛到全局最优等优点，很好地解决了前者容易出现的过学习、局部极值、维数灾难等棘手问题。

论文主要研究工作包括以下3个方面：

1. 针对灰色关联因子分析和支持向量机的分类机理，提出了一种混合算法。用灰色关联分析方法作为属性预处理器，依据属性关联度改变每个属性的权重，然后基于支持向量机进行回归建模和预测，并通过实例证明该模型的有效性。

2. 阐述预测技术中灰色预测模型GM(1, 1)，通过分析其建模机制找出影响模型精度的各种因素（背景值、初值、光滑度），并针对各种影响因素分别提出背景值预测模型(BGM模型)、初值预测模型(IGM模型)、光滑度预测模型(SGM模型)，最后分别通过实例证明其有效性。

3. 将背景值预测模型(BGM模型)、初值预测模型(IGM模型)、光滑度预测模型(SGM模型)的预测数据作为输入因子，实际数据作为输出因子，然后基于支持向量机进行回归建模和预测，并通过实例证明该模型的有效性。

关键词：灰色系统理论；支持向量机；预测

厦门大学博硕士学位论文摘要库

Abstract

Grey System Theory (GST) studies on the indeterminate system with “a few samples” and “poor” information, which is in the situation of “part of information known, part of information unknown”. By generating and developing the “part of information known”, GST can help us understand and recognize the real world, and help us rightly master and describe the operational behavior and evolutionary law of the investigated system. Grey Dynamic Modeling Technique is the core of GST, and is also the bridge between the GST and practice.

Support vector machines (SVM) is a novel and powerful machine learning approach developed in the framework of statistical learning theory, which bases on the VC theory and the principle of structural risk minimization. It always performs well in many practical applications with high generalization because of its better traditional learning approaches, such as Neural Network, SVM holds the advantages of good generalization, being insensitive to high dimension data and convergence to global optimum, so it solves the intractable problems of the former, such as over-learning, local minima, dimension curse etc.

The main works of this paper include the following three parts:

1. A admixture algorithm is presented base on grey relational analysis and support vector machines. Pretreatment module which grey relational analysis attribution reduction algorithm course endow different weight to each influencing factors, At last the predictive performance is checked.

2. Introducing Grey prediction model GM(1,1), and analyzing the mechanism of GM(1,1) to identify the impact of model accuracy of a variety of factors (Background Value, Initial Value, Smoothness), then three prediction models are presented based on each impact factors, that is, Background value prediction model (BGM model);, Initial value prediction model (IGM model) and Smoothness prediction model (SGM model).

3. Treating the prediction data of BGM model, IGM model, SGM model as input factors, the actual data as output factor, At last the predictive performance is checked.

Key words: grey system theory; support vector machines; prediction

厦门大学博硕士学位论文摘要库

目 录

摘 要.....	I
第一章 绪论	1
1.1 引言	1
1.2 支持向量机的研究现状	1
1.3 灰色理论的研究现状	2
1.4 本文的主要工作	3
1.5 本章小结	3
第二章 支持向量回归机的基本理论	4
2.1 统计学习理论	4
2.2 支持向量回归机原理	5
2.2.1 最优分类面	5
2.2.2 高维空间中的最优分类面	8
2.2.3 核函数	9
2.3 支持向量回归机的几种算法	10
2.3.1 ε -支持向量机方法	10
2.3.2 最小二乘支持向量机方法	11
2.4 本章小结	13
第三章 灰色系统理论	14
3.1 灰色系统的产生	14
3.2 灰色关联分析	15
3.2.1 灰色关联分析	15
3.2.2 灰色关联度	15
3.3 灰色 GM(1, 1) 模型	16
3.3.1 GM(1, 1) 模型	16
3.3.2 改进 GM(1, 1) 模型的研究现状	17
3.4 本章小结	20
第四章 灰关联分析支持向量机模型 (GASVM)	21
4.1 灰色关联分析支持向量机模型	21
4.2 GASVM 模型的应用	24

4.3 本章小结	27
第五章 灰预测支持向量机模型 (GMSVM)	28
5.1 GM(1, 1)模型的分析与改进	28
5.1.1 提高数据序列的光滑度模型 (SGM 模型)	28
5.1.2 预测公式中初值的选取 (IGM 模型)	33
5.1.3 模型中背景值的影响 (BGM 模型)	35
5.2 GMSVM 模型的建立与应用	40
5.2.1 灰色支持向量机预测模型的建立	40
5.2.2 仿真实验	40
5.3 本章小结	42
第六章 工作总结与展望	43
6.1 工作总结	43
6.2 工作展望	44
参考文献.....	45
致 谢.....	48
附录 攻读硕士期间发表的论文	49

CONTENTS

Abstract	I
Chapter 1 Exordium	1
1.1 Introduction.....	1
1.2 Research Progress of Support Vector Machines	1
1.3 Research Progress of Grey Theory	2
1.4 Main Studies	3
1.5 Summary.....	3
Chapter 2 Foundation of Support Regression Vector Machines	4
2.1 Statistical Learning Theory.....	4
2.2 Support Vector Regression Machines Principle.....	5
2.2.1 Optimal Separating Hyperplane	5
2.2.2 High dimension Space of Optimal Separating Hyperplane.....	8
2.2.3 Kernel Function Kernel Function	9
2.3 Some Algorithms of Support Vector Regression Machines.....	10
2.3.1 ε -Support Vector Regression.....	10
2.3.2 Least Square Support Vector Machines Method	11
2.4 Summary.....	13
Chapter 3 Grey System Theory	14
3.1 Grey System Publication	14
3.2 Grey Relational Analysis	15
3.2.1 Grey Relational Analysis	15
3.2.2 Grey Relational Degree	15
3.3 Grey GM(1,1) Model.....	16
3.3.1 GM(1,1) Model.....	16
3.3.2 Improved GM(1,1) Model Research Status.....	17
3.4 Summary.....	20
Chapter 4 Grey Analysis Support Vector Machines Model(GASVM)	21
4.1 Grey Analysis Support Vector Machines Model	21
4.2 GASVM Model Application	24
4.3 Summary.....	27
Chapter 5 Grey Prediction Support Vector Machines Model (GMSVM)	28

5.1 The Analysis and improve of GM(1,1) Model	28
5.1.1 Enhancing Data Sequence Smooth Degree Model(SGM model)	28
5.1.2 Selection of Prediction formula's Initial Value(IGM model)	33
5.1.3 Effect of Background Value(BGM model)	35
5.2 Foundation and Application of GMSVM Model.....	40
5.2.1 Foundation of Grey Support Vector Machines Model.....	40
5.2.2 Application of GASVM Model	40
5.3 Summary.....	42
Chapter 6 Conclusions and Expectation.....	43
6.1 Conclusions.....	43
6.2 Expaction	44
References	45
Acknowledgement	48
Appendix The Published Paper of Studing for Master Degree	49

第一章 绪论

1.1 引言

基于数据的机器学习是现代人工智能技术中的一个重要研究内容和方向,其主要研究是从观测数据(样本)出发寻找规律,并利用这些规律对未来数据或无法观测的数据进行预测。基于数据的机器学习可大致分为三种,它们分别为模式分类、函数回归与预测以及概率密度估计^[1-3]。

V.Vapnik 等人^[1,2]从六、七十年代开始致力于统计学习理论的研究。同时,在统计学习理论基础上发展了一种新的通用学习方法——支持向量机(Support Vector Machine, SVM),它在解决小样本、非线性以及高维模式识别问题中表现出许多特有的优势。SVM 克服了神经网络分类和传统统计分类法的很多缺点,具有较高的泛化能力。与传统的分类算法相比,SVM 在运算速度、结果精度等方面都有着明显的优越性。

灰色系统是八十年代发展起来的一门新学科,灰色系统着重外延明确,内涵不明确的对象^[16]。灰色系统建模方法是着重系统行为数据间、内在关系间挖掘量化的方法,是内涵外延的方法,是外延内涵均取得方法。灰色系统建模实际上是一种以数找数的方法,从系统的一个或几个离散数列中找出系统的变化关系,建立系统的连续变化模型。

灰色支持向量机是将灰色系统与支持向量机有机结合、取长补短,来提高系统的并行计算能力和系统的可用信息的利用率,提高系统建模的效率与模型的精度。

1.2 支持向量机的研究现状

20 世纪六、七十年代,Vapnik 开始致力于统计学学习理论研究。1981 年,他与 Chervonenkis 联合提出了 SVM 的重要基础——VC 维。1982 年,Vapnik 进一步提出了具有划时代意义的结构风险最小化(Structural Risk Minimization SRM)原则,它被称为 SVM 算法的基石。1993 年,Cortes 和 Vapnik 在《The Soft Margin Classifier》一书中,进一步探讨了非线性情况下的最优边界分类问题^[4]。1997 年,Vapnik,Gokowich 和 Smola 发表的《Support Vector Method for Function Approximation, Regression Estimation, and Signal Processing》一

书中，详细介绍了基于支持向量机方法的回归估计方法和信号处理^[5]。

近几年涌现出大量的理论研究成果，使支持向量机理论不断得到完善和发展，更为其应用研究奠定了坚实的基础。主要的研究成果分为以下几个方面。

1、支持向量预选取和简化。观测样本中占极少数的支持向量决定了 SVM 的解，因此，如果能够依据样本的集合分布信息或者密度信息直接估计出支持向量的近似集合，并在该近似集上进行训练，就能够显著加快 SVM 的学习速度，因此，把此类方法称为“支持向量预选取方法”。目前已经提出了的方法包括：类中心距离比方法^[6]，Guard 向量方法^[7]，模糊预选取方法等^[8,9]。

2、泛化性能分析。泛化性能又称为推广能力，是指经训练后的学习机器能利用已得到的规律，对测试样本作出正确的预测和判断。SVM 的泛化性能一般不能准确计算，只能对其上界作出估计，它的泛化性能的上界基于 VC 维。研究表明由于上界过于宽松，没有实际意义。为此，一些学者提出改进方法。比如 Taylor 等人给出了软边界的上界^[10]，Zhou 定量研究了核空间的容量问题^[11]。

3、核函数的研究。支持向量机的泛化能力与核函数密切相关，如何对核函数进行改进既是重点又是难点。Amari 提出了利用实验数据修正改进已有核函数，使其泛化能力加强^[12]。

4、SVM 的改进。为了扩展支持向量机的应用范围或者克服它在某些方面的局限，出现大量的改进算法。Vapnik 提出了可调罚参数的 C-SVM 系列^[1]。Lin 等提出了 FSVM(Fuzzy SVM) 事先给每个样本指定一个隶属度因子，用于度量样本在参与构建分类面时的贡献^[13]。

5、处理多分类。由于支持向量机是针对两分类问题提出，而现实应用中有许多是多分类问题。因此，如何有效地将其扩展到多类问题是当前研究的一个热点。第一种方法是“一对一(1-against-1)”^[14]；第二种方法是“一对多(1-against-rest)”^[15]。

以上简单介绍了有关支持向量机及其相关的一些研究。

1.3 灰色理论的研究现状

1982 年，华中科技大学邓聚龙教授提出了灰色系统的概念，引起国内外学者的关注和重视。灰色系统理论主要研究内容包括：灰色关联分析、灰色聚类评估、灰色系统模型、灰色系统预测等^[16-19]。

1、灰色关联分析。灰色关联分析的基本思想是根据序列曲线几何形状的相

Degree papers are in the "[Xiamen University Electronic Theses and Dissertations Database](#)". Full texts are available in the following ways:

1. If your library is a CALIS member libraries, please log on <http://etd.calis.edu.cn/> and submit requests online, or consult the interlibrary loan department in your library.
2. For users of non-CALIS member libraries, please mail to etd@xmu.edu.cn for delivery details.

厦门大学博硕士论文摘要库