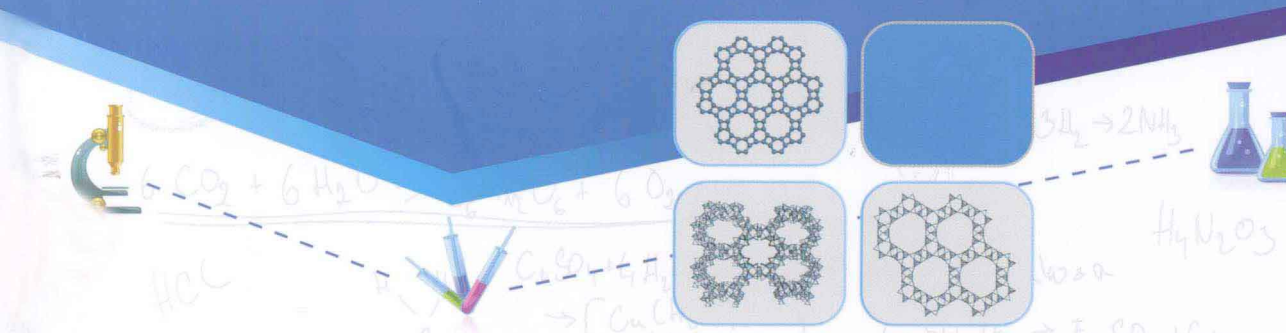
JIQI XUEXI FANGFA

ZAI LINSUANLV FENZISHAI DINGXIANG HECHENG ZHONG DE YINGYONG

机器学习方法

在磷酸铝分子筛定向合成中的应用

齐妙◎著



清华大学出版社

JIQI XUEXI FANGFA

ZAI LINSUANLV FENZISHAI DINGXIANG HECHENG ZHONG DE YINGYONG

机器学习方法

在磷酸铝分子筛定向合成中的应用

齐妙◎著

清华大学出版社
北京

内 容 简 介

本书采用基于统计的机器学习理论和方法对磷酸铝分子筛进行了大量的数据挖掘工作，主要介绍了一些经典的机器学习方法，并在磷酸铝合成数据库上进行了一系列的应用研究，①估计缺失的合成参数，完善磷酸铝合成数据库；②挖掘合成参数对合成产物某一特定结构的影响程度，为定向合成实验提供合理的解释；③处理类不平问题对预测模型的性能影响，提高定向合成实验的成功率。

本书不仅对理论方法进行了详细的介绍，还对其应用进行了具体的描述与解析，不局限于对化学定向合成的研究，可扩展到其他领域的数据分析与建模研究，以期对计算机和化学研究人员进行交叉研究起到抛砖引玉的作用。

图书在版编目(CIP)数据

机器学习方法在磷酸铝分子筛定向合成中的应用 / 齐妙 著. —北京：清华大学出版社，2013

ISBN 978-7-302-34354-7

Ⅰ. ①机… Ⅱ. ①齐… Ⅲ. ①机器学习—应用—磷酸—磷化铝—分子筛—定向分子—合成—研究 Ⅳ. ①O614.3

中国版本图书馆 CIP 数据核字(2013)第 257331 号

责任编辑：王桑娉　胡花蕾
封面设计：张　媚　张嫚丽
版式设计：方加青
责任校对：邱晓玉
责任印制：沈　露

出版发行：清华大学出版社
网　址：http://www.tup.com.cn，http://www.wqbook.com
地　址：北京清华大学学研大厦 A 座　**邮　编**：100084
社总机：010-62770175　**邮　购**：010-62786544
投稿与读者服务：010-62776969，c-service@tup.tsinghua.edu.cn
质　量　反　馈：010-62772015，zhiliang@tup.tsinghua.edu.cn
印 装 者：三河市春园印刷有限公司
经　销：全国新华书店
开　本：169mm×230mm　**印　张**：9.25　**字　数**：107 千字
版　次：2013 年 11 月第 1 版　**印　次**：2013 年 11 月第1 次印刷
定　价：58.00 元

产品编号：056174-01

前　言

随着计算机科学技术的飞速发展，社会各个领域的数据资源与日俱增，利用机器学习分析数据中所隐藏的重要信息并挖掘数据中所蕴涵的规律，已经成为人们关注的热点。目前，机器学习已经有了十分广泛的应用，例如数据挖掘、计算机视觉、自然语言处理、生物特征识别等。除此之外，机器学习还是很多交叉学科的重要支撑技术，如生物信息学、医学、化学计量学等。美国NASA-JPL实验室的科学家2001年9月在《Science》上指出，机器学习对科学研究的整个过程正起到越来越大的支持作用，并认为该领域稳定而快速的发展将对科学技术的发展起到更大的促进作用。

利用机器学习方法进行化学数据分析可优化实验过程，并从化学测量数据中最大限度地提取有用的化学信息，为化学实验提供良好的指导。由于微孔晶体材料在吸附性、离子交换性、催化活性和主客体组装化学等方面体现出优秀的性能，深入研究合成反应条件与产物之间的关系和规律对定向合成具有重要的理论意义与实际应用价值。吉林大学徐如人院士等已经初步开展了利用机器学习方法进行磷酸铝分子筛结构预测与设计等工作，并取得了一定的研

究成果。本书在磷酸铝分子筛合成反应数据库基础上，采用机器学习算法从数据中自动分析、获得规律，并利用规律对未知数据进行预测。重点将机器学习理论与磷酸铝分子筛定向合成实际问题相结合，集中讨论如何将机器学习方法有效地应用到磷酸铝合成数据分析与预测中，特别对结构与合成规律进行了深入的挖掘与研究。由于这是一个交叉学科的研究与应用，希望本书不仅为计算机科学与化学交叉学科的研究者起到抛砖引玉的作用，还希望本书中的结论能为研究磷酸铝合成的学者提供一定的指导与启发，为进一步促进计算机科学与化学交叉学科的发展贡献一点微薄之力。

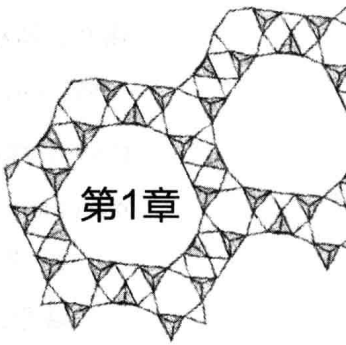
在本书的撰写过程中，非常感谢东北师范大学吕英华教授、孔俊教授、苏忠民教授和马志强教授的指导。特别感谢吉林大学“无机合成与制备化学国家重点实验室”徐如人院士、于吉红教授和李激扬教授的支持。另外，要感谢李劲松、王建中、张明和吴雪茵等人对本书的帮助。同时，本书的顺利出版得到了清华大学出版社编辑老师的认真审校。在此，对他们致以诚挚的谢意。

由于本书涉及计算机科学和化学交叉的知识，加之作者知识水平和能力有限，书中难免出现错误和不足之处，敬请广大读者批评指正，并提出宝贵意见，以使本书更加完善。

齐 妙

2013年8月

目　录



绪 论

多孔化合物与以多孔化合物为主体的多孔材料的共同特征是具有规则、均匀的孔道结构，多孔材料的特征主要包括：孔道与窗口的大小、尺寸、形状、孔道的维数、孔道的走向、孔壁的组成和性质。孔道大小、尺寸是多孔结构材料中最重要的特征，目前通行的分类标准[1, 2]如下：

孔道尺寸在2nm以下的物质称为微孔(micropore)，具有规则的微孔孔道结构的物质称为微孔化合物(microporous compounds)或分子筛(molecule sieve)；孔道尺寸范围在2～50nm之间的物质称介孔(mesopore)，具有有序介孔孔道结构的物质称为介孔材料(mesporous maerials)；孔道尺寸大于50nm的属于大孔(macropore)范围。

据国际分子筛学会(International Zeolite Association，IZA)官方统计，截至2003年，分子筛结构总数已达145种之多，骨架组成元素也发生了巨大的扩展，从沸石的组成元素Si与Al扩展到包括大量过渡元素在内的几十种元素都可作为微孔骨架的组成元素；骨架的调变与二次合成方法的进步，使具有上述一百多种独立的微孔骨架结构的各类微孔化合物数量急剧增长。从微孔、介孔直至大孔，所有的分子筛与多孔材料，其规整孔道骨架的组成全是纯无机化合物，近年来大量兴起以配位聚合物、无机有机杂化物质为主体的有序多孔骨架(porous metal-organic frameworks，MOFs)，并且在结构与功能上都显示出MOFs的特色，这就为多孔材料的多样化与组成的复杂性增添了新的领域，也为多孔材料的进一步发展拓宽了视野。

1.1 沸石分子筛

最早被人们发现的微孔材料是沸石分子筛。根据其有效孔径，可用来筛分大小不同的流体分子，这种作用叫做分子筛作用。沸石分子筛是一种孔径为0.1～1.5nm，内表面积>300m^2/g，空旷体积>0.1cm^3/g的规则微孔孔道结构的微孔化合物，通常被分为天然沸石和人工合成沸石两种。瑞典科学家A.F.Cronstedt在1756年将一种矿物Stibite进行焙烧时发现有气泡产生，类似液体的沸腾现象，将其命名为“沸石”[3]。随着地质勘探、矿物研究工作的逐步展开，越来越多的天然沸石被人们发现。目前为止，已发现的天然沸石有50余种[4]。由于天然沸石的产量已不能满足工业上的大规模需要，因此，用人工合成沸石代替天然沸石已成为生产实践中的迫切要求，Milton和Breck[5]等人丰富并发展了沸石合成方法，在温和的水热(反应温度为25℃～150℃，通常为100℃)条件下进行沸石的合成。从20世纪40年代开始，就已合成出首批低硅沸石，低温水热合成技术为大规模的工业生产提供了有利的条件。到1955年，A型分子筛和X型分子筛已经开始工业性生产。随后，Linde公司、U.C.C.公司、Mobil公司与Exxon等公司模拟天然沸石的类型与生成条件，开发出一系列低硅铝比与中硅铝比的人工合成沸石分子筛，如NaY型沸石、毛沸石、菱沸石、斜发沸石、大孔丝光沸石、L型沸石等，且在气体的吸附、分离、净化、石油炼制与石油化工中众多的催化过程以及离子交换等领域得到了广泛应用。

从1954年至20世纪80年代初，沸石分子筛的发展达到全盛时期，低硅与中硅铝比以至高硅与全硅沸石的全面开发，极大地推动

了分子筛的应用与产业的发展。我国也于1959年成功合成了A型分子筛和X型分子筛。由于沸石领域中从低硅(硅铝比为1.0～1.5)、中硅(硅铝比为2.0～5.0)直至富硅(硅铝比为10～100)与全硅等一大批沸石分子筛的出现，促进了分子筛与微孔化合物结构与性质的研究，推动了应用方面的全面进步。

1.2 磷酸铝分子筛

在分子筛合成技术的发展过程中，大量硅铝以外的其他元素也被用作分子筛的骨架构成元素，促使一大批具有新颖结构和化学计量比的无机微孔物质被合成出来，极大地丰富了无机微孔化合物的结构和组成化学。1982年，U.C.C.公司科学家S.T.Wilson与E.M.Flanigen[6]成功合成、开发出一个全新的磷酸铝分子筛家族$AlPO_4$-*n*(*n*为编号)，其丰富的吸附、催化和组装等性能使得多微孔磷酸铝合成备受关注。这一全新的磷酸铝分子筛家族不仅包括大孔、中孔与小孔的$AlPO_4$-*n*分子筛，而且可以将13种主族金属、过渡金属以及非金属元素——Li、Be、B、Mg、Si、Ga、Ge、As、Ti、Mn、Fe、Co、Zn——引入微孔骨架，生成具有24种独立开放骨架的6大类微孔化合物：$AlPO_4$-*n*、SAPO-*n*、MeASO-*n*(S=Si)、MeAPO-*n*(Me=Fe、Mg、Mn、Zn、Co等)、ElAPO-*n*(El=Ba、Ga、Ge、Li、As等)与ElAPSO-*n*，在后4类中还可生成多个元素的衍生物，这个大家族的微孔化合物成员总数已达200多种。与硅铝沸石分子筛不同的是，所有合成体系中必须有模板剂或结构导向剂的参

与，同时与硅铝酸盐沸石相比，磷酸铝分子筛也很容易形成大孔和超大孔分子筛，例如具有十四元环的$AlPO_4$-8(AET)[7]、十八元环的VPI-5[8](如图1-1所示)、二十元环的JDF-20[9](如图1-2所示)等。这类大孔分子筛的合成成功打破了以往分子筛主孔道不能超过十二元环的界限，极大促进了分子筛合成化学的发展。

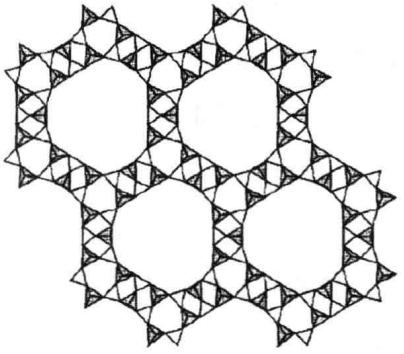


图1-1 十八元环的VPI-5结构

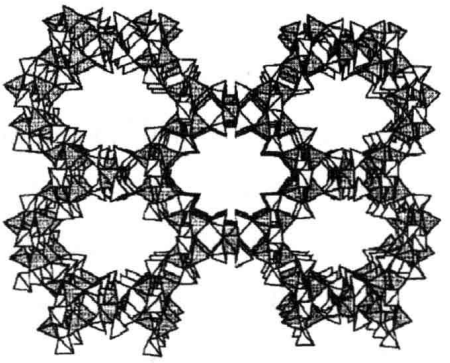


图1-2 二十元环的JDF-20结构

磷酸铝分子筛的骨架结构是由AlO_4四面体和PO_4四面体通过氧桥严格交替连接构成，其中，铝磷比为1/1[10]。Al或P原子可以部分被Si或其他元素所取代，形成具有特殊性质的杂原子磷酸铝分子筛。分子筛的合成主要采用引入不同的有机胺模板剂和水热合成方法，自1982年至今，在磷酸铝分子筛家族的基础上发展了众多铝磷比<1的具有阴离子骨架的磷酸铝的微孔化合物[11]，其中，铝磷比可为1/1、1/2、2/3、3/4、4/5、5/6、6/7、11/12和12/13等。由于磷酸铝及其衍生物的分子筛和微孔金属磷酸盐具有骨架元素种类与孔道结构多样化的特性，因此在吸附分离、催化与先进材料等多方面得到广泛应用，在氧化还原催化、手性催化与大分子催化反应等方面都显示出重要的应用前景。

1.3 分子筛的应用与发展

自20世纪50年代A型分子筛和X型分子筛开始被大规模工业生产和应用以来，分子筛作为主要的离子交换材料、催化材料和吸附分离材料，其合成和工业应用研究始终受到工业界、产业界和学术界的广泛关注。以石油工业为例，基于Y型分子筛催化剂的重油催化裂化制汽油反应工艺，使汽油产率提高了20%。60年来，由于分子筛在石油化工、日用化工等工业领域中起着越来越重要的作用，使得多孔材料的合成成为无机合成的一大热门。

20世纪60年代初，中国科学院大连化学物理研究所就有一个小组开始了分子筛的研究工作。该小组主要从事A型分子筛的合成和研究。经过国内多家大学、科研院所的不断努力，国际同行逐渐认可了我国在此领域的学术地位。其中，徐如人院士负责的吉林大学“无机合成与制备国家重点实验室”在多孔材料的结构分析和合成、分子筛在催化领域的应用等方面取得了优异的成绩，推进了无机合成化学与材料创新的进程。

据Christian Marcilly[12]在2001年的统计，由于上述领域的需要，目前人工合成分子筛全世界的年产量已超过160余万吨，而天然沸石的年产量由于离子交换与吸附材料的需要，产量也已升至每年30万吨以上(约为总量的18%)。合成分子筛的年产总值据统计已超过20亿美元，而与分子筛有关的催化、吸附等材料其年产总值已大大超过分子筛本身的价值，与20世纪60年代相比有了巨大的增长[13]。尽管如此，分子筛在上述传统领域中的作用仍有很大的发展空间与前景，至今已知结构的分子筛多达230多种，从组成元素与骨架结构的

多样性来看，仍有很大的发展空间。半个多世纪以来，分子筛主要应用于石油炼制、石油化工以及70年代后期发展起来的某些精细化工与中间体化工。据推测，未来十年，精细化工、中间体化工大量发展的需要以及石油加工、石油化工传统应用领域的更新与发展，将进一步推动分子筛在催化与吸附分离应用等领域的大踏步发展。

分子筛与多孔物质的传统应用领域主要包括：①作为吸附材料用于分离、净化干燥等领域；②作为催化材料用于需要工业催化过程的石油化工、加工等领域；③作为离子交换材料用于需要进行废料、废液处理的领域。这些领域的需求使得对分子筛与多孔物质的研究持续发展，并不断深入。目前人工合成分子筛的年产量超过160余万吨，年产值超过20亿美元。对分子筛与多孔材料的研究经历了从天然沸石到人工合成沸石，从硅铝分子筛到磷酸铝分子筛，从超大微孔到大孔等阶段。微孔分子筛与介孔分子筛在传统领域起着巨大的作用，有着很大的发展空间。同时，在一些高新技术先进材料应用领域也有着广阔与诱人的前景。

1.4 研究意义与研究内容

基于统计的机器学习方法已经成功地应用于材料科学领域中，典型的例子有沸石合成中模板剂的设计[14]、无机开放式骨架材料模板剂的预测[15, 16]、沸石和微孔磷酸铝合成因素的研究[17, 18]。神经网络已经成功运用于建模和预测各种反应的催化性能，如乙烷氧化脱氢作用[19]、水煤气漂移[20]和甲醇合成[21, 22]。然而，虽然神经网络

方法得到了广泛的应用，但是神经网络方法也存在过拟合和容易陷入局部最优解等缺点。支持向量机方法始于20世纪70年代，由于其能够很好地解决非线性、高维数、局部极小点等问题，因而受到了广泛的关注，在药物发现和医学化学[23, 24]、工艺学[25]和物理化学的定量构效关系(QSAR)[26-28]中得到了广泛的应用。最近，文献[29]将支持向量机运用于各种催化剂(heterogeneous catalysts)的催化性能预测，实验结果证明支持向量机的分类性能优于各种决策树。文献[30]运用支持向量机、神经网络、决策树、聚类分析和主成分分析方法预测沸石的合成，实验结果表明：在相同的合成参数下，支持向量机得到了较好的预测结果。文献[17]利用各种基于参数和非参数统计方法来预测沸石材料ITQ-21的结晶相图。据我们所知，机器学习方法在磷酸铝合成的数据中的应用研究相对较少。徐文国等人[31]采用一元线性回归、多元线性回归分析和BP神经网络预测了磷酸铝分子筛的骨架晶格能，预测结果与计算晶格能吻合，表明晶格能与配位序($\overline{N_2} \sim \overline{N_3}$)具有良好的多元线性关系。刘晓东等人[32]采用决策树方法利用模板剂的特征预测磷酸铝和硅铝酸盐分子筛的环结构，预测结果表明，采用二乙醇胺做模板剂，成功地晶化出了具有十二元环孔道的$AlPO_4$-5。

无机微孔晶体由于其独特的规则孔道结构而被广泛地应用于催化、吸附、分离和离子交换等领域，因而具有新颖结构的微孔晶体的设计、合成以及新合成路线的开发一直备受关注。其中，开放骨架结构的金属磷酸盐化合物由于其结构的多样性和潜在的应用价值，国内外很多学者已经对其开展了广泛而深入的研究。无机微孔晶体化合物的合成十分复杂，材料结晶受诸多因素的影响，例如原