

基于地统计学与支持向量回归的QSAR建模

陈渊, 袁哲明, 周玮, 熊兴耀

湖南农业大学生物安全科学技术学院, 长沙 410128|湖南农业大学, 湖南省作物种质创新与资源利用重点实验室, 长沙 410128

摘要:

基于主成分分析(PCA)、地统计学(GS)和支持向量回归(SVR), 提出了一种新的定量构效关系(QSAR)个体化预测方法——Weight-PCA-GS-SVR. 其基本思路是: 先以PCA降维并消除自变量间的信息冗余, 继以SVR经非线性主成分筛选去除与因变量无关的主成分, 再以保留主成分计算样本间的加权距离, 然后以高维GS确定公用变程; 每一个待测样本都以自身为中心从训练集中找出加权距离小于公用变程的私有k个近邻, 以SVR训练建模完成个体化预测. Weight-PCA-GS-SVR从行、列两个方向对模型进行了优化, 为自变量提供了一种新的加权方法, 为解决最优k近邻选择难题提供了新的思路, 并具有SVR原来的优点. 经3个化合物活性实例数据集验证, 新方法在所有参比模型中预测精度最高, 且明显优于文献报道结果, Weight-PCA-GS-SVR在QSAR等回归预测领域有较广泛的应用前景.

关键词: 定量构效关系 地统计学 支持向量回归 主成分分析 个体化预测

收稿日期 2009-03-16 修回日期 2009-04-15 网络版发布日期 2009-05-26

通讯作者: 袁哲明 Email: zhmyuan@sina.com

本刊中的类似文章

1. 周原;梅虎;梁桂兆;李志良.取代基物化参数及其在药物定量构效关系中的应用[J]. 物理化学学报, 2006,22(04): 486-491
2. 彭涛;裴剑锋;周家驹.酪氨酸激酶抑制剂的三维定量构效关系研究[J]. 物理化学学报, 2003,19(02): 163-166
3. 王任小;高澹;刘亮;来鲁华.化合物的空间取向对CoMFA结果的影响[J]. 物理化学学报, 1998,14(01): 1-4
4. 梁桂兆;梅虎;周鹏;周原;李志良.三维原子场作用全息矢量用于二氢叶酸还原酶抑制剂及苦味二肽QSAR研究[J]. 物理化学学报, 2006,22(03): 388-390
5. 冯军;周家驹;李仁利.比较分子场分析研究吡嗪酮的体系的三维构效关系[J]. 物理化学学报, 1995,11(03): 206-210
6. 沈斌;陆忠华;迟学斌;吕海峰;任天瑞.GABA受体抑制剂的柔性原子受体模型研究[J]. 物理化学学报, 2005,21(07): 800-803
7. 乔颖欣;周家驹.带有分子轨道能量的3D-QSAR对*N*-氨基咪唑的研[J]. 物理化学学报, 2006,22(02): 209-214
8. 黄钦;侯廷军;徐俊杰.基于遗传算法的Caco-2细胞穿透系数的研究[J]. 物理化学学报, 2005,21(04): 372-377
9. 胡桂香;邹建卫;蒋勇军;王艳花;俞庆森.从药物的三维分子结构预测人体小肠吸收[J]. 物理化学学报, 2004,20(05): 512-517
10. 丁俊杰;丁晓琴;赵立峰;陈冀胜.二氢吡啶类化合物的三维定量构效关系[J]. 物理化学学报, 2003,19(12): 1108-1113
11. 冯长君;沐来龙;杨伟华;蔡可迎.有机污染物的生物富集因子与拓扑指数的数学模型[J]. 物理化学学报, 2008,24(06): 1053-1057
12. 梅虎 刘丽 杨力 李建 闫宁 王琴.原子类型电拓扑状态指数预测咪唑啉啉生物体的抗癌性[J]. 物理化学学报, 2009,25(04): 747-751
13. 骆兆文;王丹丹;来鲁华;徐俊杰;李崇熙.雪花胺类化合物的三维构效关系研究[J]. 物理化学学报, 1995,11(05): 419-423
14. 吴文娟;赖璐;郑康成;云逢存.抗癌性咪唑啉啉生物体的定量构效关系[J]. 物理化学学报, 2005,21(01): 28-32
15. 王瑾玲;孙命;苏华庆;缪方明.咪唑-1-羟酸酯类化合物的构效关系研究[J]. 物理化学学报, 1998,14(05): 444-447
16. 朱丽荔;徐俊杰.褪黑激素受体拮抗剂的三维定量构效关系研究[J]. 物理化学学报, 2002,18(12): 1087-1092
17. 王宝雷;马宁;王建国;马翼;李正名;李永红.新磺酰胺类化合物除草活性的3D-QSAR分析[J]. 物理化学学报,

扩展功能

本文信息

PDF(694KB)

服务与反馈

把本文推荐给朋友

加入我的书架

加入引用管理器

引用本文

Email Alert

文章反馈

浏览反馈信息

本文关键词相关文章

▶ 定量构效关系

▶ 地统计学

▶ 支持向量回归

▶ 主成分分析

▶ 个体化预测

本文作者相关文章

▶ 陈渊

▶ 袁哲明

▶ 周玮

▶ 熊兴耀

18. 王任小;刘亮;来鲁华;唐有祺.凝血酶抑制剂的结构与活性的关系[J]. 物理化学学报, 1998,14(10): 887-892
19. 梅虎;周原;孙立力;李志良.一种新的氨基酸描述子及其在肽QSAR中的应用[J]. 物理化学学报, 2004,20(08): 821-825
20. 王任小;李维忠;来鲁华;唐有祺.酶-配体复合物亲和性的计算[J]. 物理化学学报, 1998,14(09): 826-832
21. 王任小;冯亚彬;来鲁华;唐有祺.磷脂酶A₂ 抑制剂的结构和活性关系[J]. 物理化学学报, 1998,14(10): 893-897
22. 魏卓 张怀 崔巍 计明娟.马来酰胺类糖原合成酶激酶-3 β 抑制剂的分子对接和三维定量构效关系[J]. 物理化学学报, 2009,25(05): 890-896
23. 蒋玉仁;秦伟.苯并噻酮衍生物的3D-QSAR分析[J]. 物理化学学报, 2008,24(10): 1859-1863
24. 黄忠平;潘锦红;蔡国强;俞庆森;林瑞森.方酸衍生物的光敏性与结构关系的理论研究[J]. 物理化学学报, 1998,14(06): 557-561
25. 陈红明;周家驹;谢桂荣;任天瑞.一种基于虚拟受体模型的定量构效关系研究方法[J]. 物理化学学报, 1997,13(07): 626-631
26. 杨光富;刘华银;杨秀凤;杨华铮.1,2,4-三唑并[1, 5-a]嘧啶-2-磺酰胺类除草剂的CoMFA研究[J]. 物理化学学报, 1999,15(02): 190-192
27. 仝建波;周鹏;张生万;梁桂兆;田菲菲;李美萍;李声时.三维全息原子场作用矢量用于HEPT类抗艾滋病药物的QSAR研究[J]. 物理化学学报, 2006,22(06): 721-725
28. 邹霞娟;来鲁华;金桂玉;黄桂琴.新型含噻吩酮基双酰胺类化合物的3D-QSAR研究[J]. 物理化学学报, 2002,18(06): 513-516
29. 张华北;李波;戴梅.[⁹⁹Tc^m(NO)Cl(PL)₂]⁺类配合物的定量构效关系[J]. 物理化学学报, 2003,19(05): 460-463
30. 潘咏梅;计明娟.基于遗传算法的PTP1B抑制剂的二维定量构效关系研究[J]. 物理化学学报, 2003,19(08): 695-700
31. 仝建波;张生万.一种新的三维氨基酸描述子及其在肽类药物QSAR中的应用[J]. 物理化学学报, 2007,23(01): 37-43
32. 宋哲;刘涛;刘伟;朱鸣华;王晓钢.抗原肽与MHC分子相互作用的QSAR模型研究[J]. 物理化学学报, 2007,23(02): 198-205
33. 朱龙观;俞庆森;陈凯先;蔡国强;林瑞森.喹诺酮类N₁位定量构效关系研究[J]. 物理化学学报, 1995,11(10): 925-928