

机器学习预测食品重金属检测中铜离子 对汞离子荧光信号的干扰

Machine learning prediction of copper ion interference with mercury ion
fluorescence signals in food heavy metal detection

宋方亮^{1,2} 梁 盈^{1,2} 董 界³ 王雪洁^{1,2} 钱 洁^{1,2}

SONG Fangliang^{1,2} LIANG Ying^{1,2} DONG Jie³ WANG Xuejie^{1,2} QIAN Jie^{1,2}

(1. 中南林业科技大学食品科学与工程学院, 湖南 长沙 410004; 2. 水稻及副产物深加工国家
工程研究中心分子营养分中心, 湖南 长沙 410004; 3. 中南大学湘雅药学院, 湖南 长沙 410013)
(1. College of Food Science and Engineering, Central South University of Forestry and Technology,
Changsha, Hunan 410004, China; 2. Molecular Nutrition Branch, National Engineering Research
Center of Rice and By-Product Deep Processing, Changsha, Hunan 410004, China; 3. Xiangya
School of Pharmaceutical Sciences, Central South University, Changsha, Hunan 410013, China)

摘要:目的: 构建一个人工智能预测模型, 在存在 Cu^{2+} 干扰的复杂食品检测环境下预测荧光探针对于 Hg^{2+} 的选择性。方法: 采用荧光探针技术结合 7 种先进经典的机器学习模型, 预测分析存在 Cu^{2+} 干扰时探针对于 Hg^{2+} 的选择性, 并比较各模型的预测效果, 选择最优模型。结果: 基于分子二维描述符 (molecular 2D descriptors, Mol2D) 和极端梯度提升算法成功建立了在交叉验证和测试集中准确度为 0.786 和 0.810 的高效模型, 在 Cu^{2+} 干扰下准确预判 Hg^{2+} 的探针选择性。结论: 该模型通过选择性预判对 Hg^{2+} 荧光分子探针的设计进行改进, 使 Hg^{2+} 荧光探针的设计更加高效可靠。

关键词: 汞离子检测; 荧光分子探针; 探针选择性; 机器学习; 化学信息学

Abstract: **Objective:** To construct an artificial intelligence prediction model to predict the selectivity of fluorescent probes for Hg^{2+} in a complex food testing environment in the presence of Cu^{2+} interference. **Methods:** Fluorescent probe technology combined with seven advanced classical machine learning models was used to predict and analyze the selectivity of the probe for

Hg^{2+} in the presence of Cu^{2+} interference, and to compare the prediction effect of each model and select the optimal model.

Results: Efficient models with accuracies of 0.786 and 0.810 in the cross-validation and test sets were successfully established based on Molecular 2D Descriptors (Mol2D) and extreme gradient boosting algorithms to accurately predict the probe selectivity of Hg^{2+} under Cu^{2+} interference. **Conclusion:** The model is improved for the design of Hg^{2+} fluorescent molecular probes by selective prediction, which makes the design of Hg^{2+} fluorescent probes more efficient and reliable.

Keywords: mercury ion detection; fluorescent molecular probes; probe selectivity; machine learning; cheminformatics

Hg^{2+} 是食品中广泛存在的有毒重金属离子。富集了来自海洋中污染物的鱼类、贝类等海产品和受到工业或农药污染土壤中的水果与蔬菜, 被捕捞或采摘, 再端上餐桌以供食用, 处于这种食物加工链下, Hg^{2+} 可能会被人体吸收并在体内蓄积, 导致神经系统、肾脏等身体器官受损^[1-3]。中国国家标准对于食品中汞 (Hg) 的含量有特定的限制, 根据《食品安全国家标准 食品中汞的限量》(GB 2762—2017) 规定通过肉类、禽蛋类、水产品类以及水果、蔬菜中汞的限量不得大于 0.05 mg/kg。显然, 加工食品中出现汞离子含量超标的情况是非常严重的, 因此进行准确高效的检测是必要的^[4]。目前, 电感耦合等离子体质谱法 (ICP-MS)^[5] 和原子荧光 (AFS)^[6] 等方法已被用于测定食品中的汞, 虽然这些传统方法在稳定、准确性检测 Hg^{2+} 上各有优势, 但它们通常需要专业的检测设备

基金项目: 国家自然科学基金 (编号: 32372349); 国家重点研发计划 (编号: 2022YFF1100203); 湖南省科技创新人才项目 (编号: 2022RC3056)

作者简介: 宋方亮, 女, 中南林业科技大学在读硕士研究生。

通信作者: 梁盈 (1981—), 女, 中南林业科技大学教授, 博士。

E-mail: liangying498@163.com

董界 (1989—), 男, 中南大学副教授, 博士。

E-mail: jiedong@csu.edu.cn

收稿日期: 2024-02-11 **改回日期:** 2024-04-26

和经验丰富的技术人员且操作复杂耗时^[7]。

荧光分子探针方法因其快速、简单、灵敏、可视化等优势在化学检测和分析、医学、食品科学等领域得到广泛应用^[8-9]。近几年,有大量研究强调了荧光分子探针检测 Hg^{2+} 的优势。探针对于目标离子的检测是通过特定的反应基团产生荧光以得到其可视化的结果,在合成过程中,进行对特定基团的选择性实验验证往往是最关键的,但也是最耗时、耗材的^[10-11]。并且在 Cu^{2+} 存在的复杂环境下,用于检测 Hg^{2+} 的探针也会受到 Cu^{2+} 的干扰,不便于准确区分目标离子^[12-15]。

目前,关于荧光探针检测 Hg^{2+} 的选择性预测方法未见报道。研究拟提出一种分子结构表征与经典机器学习算法相结合的选择性验证新策略,在非试剂试验条件下通过高精度模型在探针合成前预先评估出对 Hg^{2+} 的选择性或者 Cu^{2+} 的干扰可能。以期为基于荧光探针高效检测食品中 Hg^{2+} 的发展提供参考。

1 材料与方法

研究基于多种机器学习方法预测荧光分子探针在 Cu^{2+} 存在环境下对 Hg^{2+} 的选择性检测。具体的操作流程见图 1。

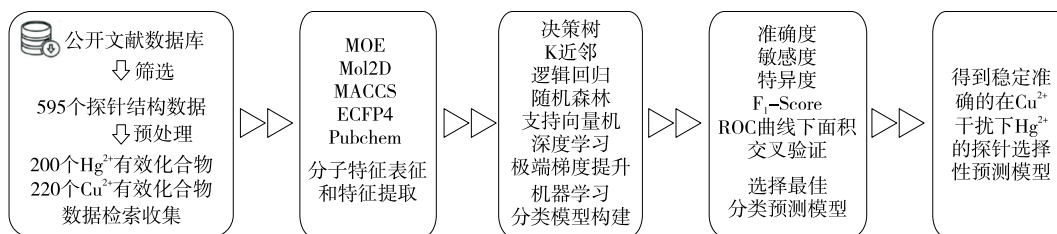


图 1 研究流程

Figure 1 Research procedure

1.1 汞离子探针选择性数据收集

完整数据是从公开文献数据库 Web of Science、PubMed 和 Google 学术中收集的。通过使用“Fluorescent molecular probes + $\text{Hg}^{2+}/\text{Cu}^{2+}$ ”“Fluorescent molecular probes + Mercury ion / Copper ion”关键词进行检索,从近 20 年已发表的文献中检索并筛选得到 595 篇相关文献,对每一篇文献内容进行概读,删去非响应型、非荧光型探针等无效文献,只考虑荧光探针选择性检测的汞离子和干扰性铜离子,从剩下的文献中得到 216 和 226 个有效的汞离子和铜离子荧光探针结构数据,其中探针 SMILES(simplified molecular input line entry system)结构是通过开源软件 StoneMIND Collector(Win V1.5.1)转换获取的。

1.2 数据预处理

首先,对得到的探针 SMILES 数据进行了手动检查,修正了存在问题的 6 个结构,对不平衡电荷结构添加了氢原子。其次,使用基于计算机的分子建模软件 MOE (molecular operating environment, 版本 2019) 中的 wash 功能保留每个结构的最大分子片段,基于化学特征提取的计算机辅助工具 ChemDes 将所有 SMILES 数据转换为 InChIKey 标识符进行去重,InChIKey 相同的数据只保留其中之一。最后,得到 Hg^{2+} 和 Cu^{2+} 有效探针数据 200 和 220 个,数据的整理结果总结在表 1 中。主成分分析(principal component analysis, PCA)是一种降维统计方法,在很小信息损失的情况下可将多个变量转化为多个主成分^[16]。基于 PCA 对数据进行聚类,以观察 Hg^{2+} 和 Cu^{2+} 探针选择性数据在二维空间的聚集情况。

表 1 数据处理结果

Table 1 Results of data processing

组别	数据类别	原始数量	预处理后数量
阳性(正)类	Hg^{2+}	216	200
阴性(负)类	Cu^{2+}	226	220
总计	—	442	420

1.3 分子结构表征

结构表征通过量化物理和拓扑结构,将二维或三维分子结构转换为分子描述符或分子指纹。研究中综合了 5 种结构表征方式来弥补它们在化学空间中的局限性。使用 RDKit (<http://www.rdkit.org/>) 和基于 Python 的用于生物医学信息学和化学信息学领域的计算工具 PyBioMed^[17] 计算分子二维描述符(molecular 2D descriptors, Mol2D)、MACCS 指纹、Pubchem 指纹和 ECFP4 指纹,通过 MOE 软件计算得到了二维和三维描述符(研究中统一命名为 MOE 描述符)。这些描述符和指纹涵盖了大量的物理化学和拓扑信息,包括原子、键、环、模式等。为了减少噪声和提高建模效率,对 Mol2D 和 MOE 描述符进行了数据预处理。① 去掉方差为零的描述符;② 剔除相关性高(>0.95)的描述符;③ 使用基于随机森林的递归特征消除和交叉验证方法(RF-RFECV)选择最佳特征子集。

1.4 机器学习算法

选择决策树(decision tree, DT)^[18]、K-近邻(K-nearest neighbors, KNN)^[19]、逻辑回归(logistic regression, LR)^[20]、随机森林(random forest, RF)^[21]、支

持向量机(support vector machine,SVM)^[22]、极端梯度提升(eXtreme gradient boosting,XGBoost)^[23]和深度神经网络(deep neural network,DNN)^[24]算法进行建模。DT是基于树状结构的模型,适用于分类和回归问题,简单易解释。RF是一种集成学习模型,由多个决策树组成,进一步提高预测准确性。XGBoost是梯度提升树算法的变体,在结构化数据和大规模数据集上表现卓越。KNN是基于邻居关系的算法,适用于简单问题和小规模数据。逻辑回归是处理二分类问题的线性模型,简单高效。支持向量机通过找到最优超平面进行分类或回归,在高维空间中表现良好。DNN包括多个隐藏层,通过学习更复杂的模式和关系来处理复杂的任务,对大规模数据和图像处理具有强大表现力。

1.5 模型评估指标

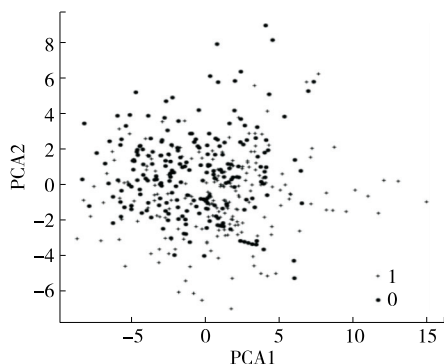
在机器学习领域,分类模型的评估指标是衡量模型性能的关键工具。这些指标帮助研究者了解模型在未知数据上的预测能力。常见的分类模型评估指标包括准确率(Accuracy)、精确率(Precision)、召回率(Recall)、 F_1 分数(F_1 Score)以及 AUC-ROC (area under the receiver operating characteristic curve)等。准确率提供了模型正确分类样本的总体比例,而精确率和召回率则分别衡量了模型预测为正样本的实例中真正为正样本的比例,以及真正为正样本的实例中被模型正确预测的比例^[25]。 F_1 分数是精确率和召回率的调和平均数,用于在两者间取得平衡,体现出模型的综合预测性能^[25]。AUC-ROC 则通过绘制不同阈值下的真正例率(true positive rate)和假正例率(false positive rate)来评估模型的整体性能,数值越大表明模型的性能越好^[26]。选择适当的评估指标对于理解模型性能、优化模型参数以及选择最佳模型至关重要。

2 结果与分析

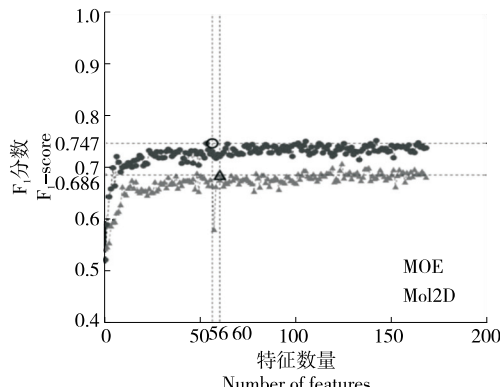
图 2(a)为 Hg^{2+} 和 Cu^{2+} 终点分类数据集的 PCA 结

果。结果发现,数据集中的分子都聚集在一起,并没有明显地分成两个或多个聚类。PCA 无法直接聚类出 Hg^{2+} 和 Cu^{2+} 的选择性终点。因此,使用更多更复杂的模型。从图 2(b)可以看出,通过 RF-RFECV 对 Mol2D 和 MOE 描述符特征逐个筛选移除对模型影响小的特征,在模型 F_1 -Score 结果达到最高 0.747 和 0.686 时分别得到 56 和 60 个最佳的特征子集。整个数据集中 420 个化合物被通过 4 : 1 的比例重复随机劈分为包含 336 个化合物的训练集包含 84 种化合物的测试集。使用选择后的这两种描述符和 3 种指纹结合 7 种机器学习算法构建了 35 种 Hg^{2+} 探针选择性模型,模型评估结果见表 2。总体而言,模型性能良好,五折交叉验证的结果与测试集的结果相似,表明数据集的划分是合理的。基于几种不同评估指标性能最好的是 Mol2D-XGBoost 模型,其在测试集上准确率为 0.810,AUC 值为 0.881。

图 3(a)显示了基于 Mol2D 描述因子的最佳结果。在假正率为 0.15 时,真正率已经达到 0.8,ROC 曲线的整体幅度偏向左上,使用 XGBoost 算法的优势是显而易见的,AUC 值为 0.881,表示其在训练数据上能够更好地区分正样本和负样本,对潜在的新数据具有更强的预测能力。MOE 描述符的最佳结果在图 3(b)中呈现,依旧使用 XGBoost 算法达到最佳结果,AUC 数值为 0.824,与 Mol2D-XGBoost 模型的曲线结果相比略差一些。在 3 种不同的指纹中,Pubchem 指纹显示出比 ECFP4 指纹和 MACCS 指纹更优异的表现,图 4(a)中最佳模型 Pubchem-XGBoost 的 ROC 曲线在假正率为 0.3 时真正率的提升幅度最高,AUC 值为 0.844。结合图 4 中 3 种指纹的最佳模型 ROC 曲线,它们均结合 XGBoost 和 RF 实现了最高的 Hg^{2+} 选择性预测结果。综合考虑了基于不同算法和分子表征方法构建模型的 AUC 值和 ACC 值,XGBoost 算法和 RF 算法始终表现出更平衡更好的模型性能,既有较高的分类准确性,又保持了良好的概率估计能力。



(a) 数据PCA分布图



(b) 描述符特征选择图

图 2 整体数据 PCA 分布图和描述符的 RF-RFECV 特征选择过程

Figure 2 Overall data PCA distribution and descriptor RF-RFECV feature selection process

表 2 最佳分子表征 Mol2D 描述符的模型性能

Table 2 Model performance results for the best molecularly characterized Mol2D descriptors

分类模型	五折交叉验证(5-CV)				测试集			
	敏感度	特异度	准确度	AUC 值	敏感度	特异度	准确度	AUC 值
DT	0.669	0.737	0.705	0.726	0.765	0.720	0.738	0.749
KNN	0.656	0.739	0.700	0.698	0.700	0.773	0.738	0.736
LR	0.690	0.758	0.726	0.784	0.700	0.841	0.774	0.802
RF	0.714	0.786	0.753	0.837	0.791	0.829	0.810	0.877
SVM	0.785	0.708	0.744	0.815	0.816	0.717	0.762	0.864
XGBoost	0.791	0.781	0.786	0.848	0.805	0.814	0.810	0.881
DNN	0.690	0.740	0.717	0.779	0.750	0.771	0.762	0.840

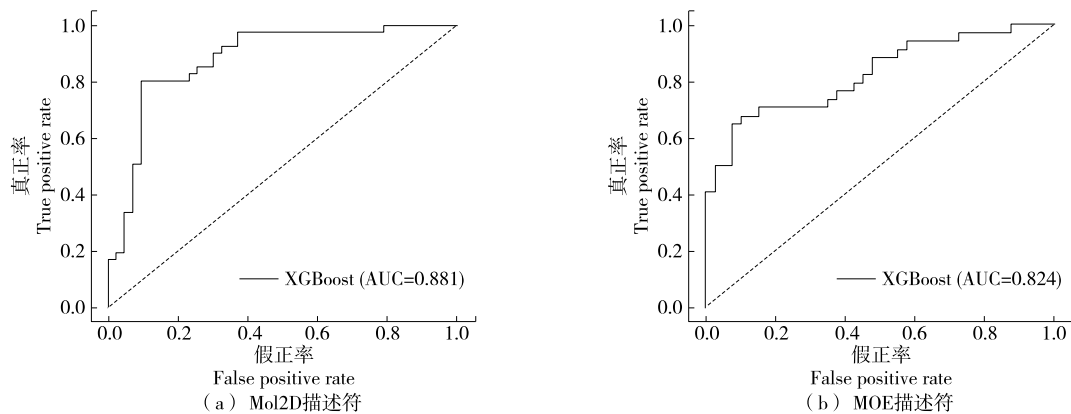


图 3 基于不同分子描述符得到的 Hg^{2+} 最佳预测模型的 ROC 曲线结果图

Figure 3 ROC curve results for the best prediction model of Hg^{2+} based on different molecular descriptors

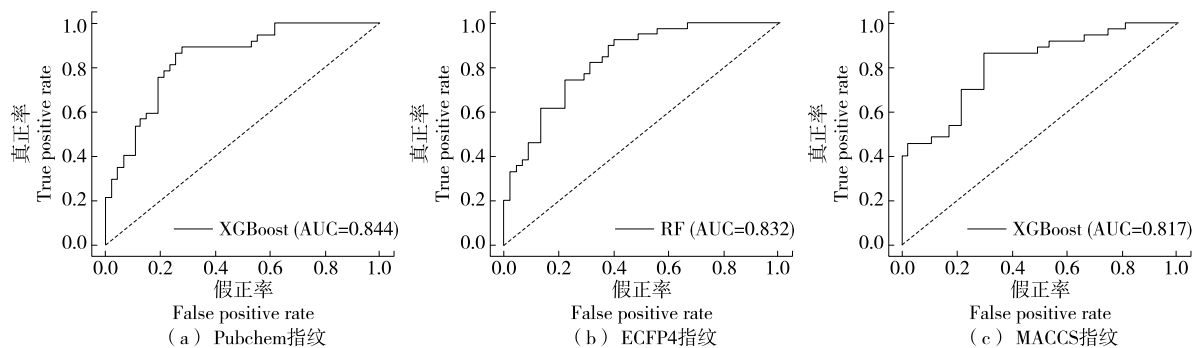


图 4 基于不同分子指纹得到的 Hg^{2+} 最佳预测模型的 ROC 曲线结果图

Figure 4 ROC curve results for the best prediction model of Hg^{2+} based on different molecular fingerprints

图 5 完整表示了所有分子表征和算法组合的模型结果,每一个箱线图三角符号都表示在该表征或算法上的准确度。考虑在不同算法的结果中,大部分模型组合能在不同的数据集随机划分下稳定实现一定的性能。明显可以看出,基于 XGBoost 算法和 RF 算法的模型比其他表现优秀,而基于 DT 算法和 KNN 算法的几个模型表现相对较差。对于 DT 算法,MOE 描述符表征的模型表现最好(ACC 值为 0.789,AUC 值为 0.736),最差的是 ECFP4,ACC 值为 0.726。在 KNN 算法上,ECFP4 指纹依旧表现最差。基于 LR 算法、SVM 算法和 DNN 算法,

它们的模型结果一直都是比较稳定的,均结合 Mol2D 模型精度最高,ACC 值 ≥ 0.762 ,结合 MACCS 指纹效果较差。在 XGBoost 算法和 RF 算法中,Mol2D-XGBoost 和 Mol2D-RF 都达到了 0.810 的模型预测精度,AUC 值均 ≥ 0.877 ,同样在 MACCS 中表现是最差的。总体而言,大部分组合模型表现良好。Mol2D 指纹似乎更适合表征该数据集,而 MACCS 指纹是最不适合的。因为许多 Hg^{2+} 选择性的探针具有共同的基团结构。在这种情况下,指纹描述符的区分能力受到了限制,而使用具有一些连续属性的描述符特征可能适合该数据集。

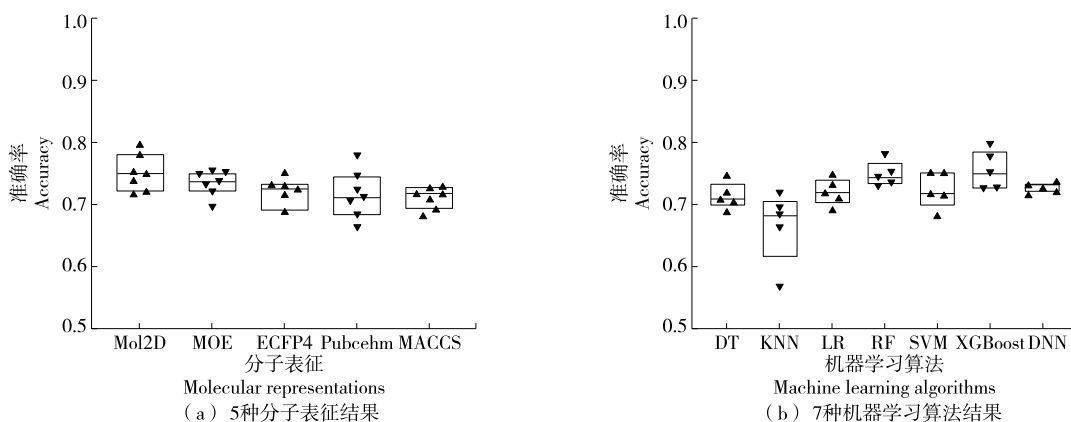


图 5 所有模型结果可视化箱线图

Figure 5 Boxplot of all model results visualization

3 结论

研究了一种基于先进的机器学习算法的 Hg^{2+} 探针选择性预判方法。结果表明,使用包含化合物分子量、氢原子数量、单键数量、双键数量的二维描述符 Mol2D 与极端梯度提升算法组合模型在性能上均优于其他模型,测试集上的预测准确率达到 0.810,ROC 曲线下面积达到 0.881。该方法达到研究的预期目标,最终模型在交叉验证和测试集的结果相差不大,结果是可靠的。对新数据进行模型应用域适配后,可以用来快速准确地预先判别食品中检测 Hg^{2+} 的探针选择性以及 Cu^{2+} 的干扰。研究构建的 Hg^{2+} 探针数据集还有一定的数据增长空间,下一步会进一步扩大数据收集范围,考虑在更多复杂食品加工环境下影响 Hg^{2+} 选择性的因素,扩大模型应用范围,使 Hg^{2+} 的选择性预评估更加全面。

参考文献

- [1] PRAJAPATI S, CHALLIS J K, JARDINE T D, et al. Levels of pesticides and trace metals in water, sediment, and fish of a large, agriculturally-dominated river[J]. *Chemosphere*, 2022, 308: 136236.
- [2] ULLAH N, UR REHMAN M, AHMAD B, et al. Assessment of heavy metals accumulation in agricultural soil, vegetables and associated health risks[J]. *PLoS One*, 2022, 17(6): e0267719.
- [3] 王丽, 许华杰, 李海丽, 等. 茂名地区主要水产品汞含量分析与风险评估[J]. *食品与机械*, 2021, 37(9): 73-77.
WANG L, XU H J, LI H L, et al. Mercury content analysis and risk assessment of the major aquatic products in Maoming[J]. *Food & Machinery*, 2021, 37(9): 73-77.
- [4] LIAQUAT H, IMRAN M, LATIF S, et al. Multifunctional nanomaterials and nanocomposites for sensing and monitoring of environmentally hazardous heavy metal contaminants [J]. *Environmental Research*, 2022, 214: 113795.
- [5] MICHALSKI R, JABŁONSKA M, SZOPA S, et al. Application of ion chromatography with ICP-MS or MS detection to the determination of selected halides and metal/metalloids species[J]. *Critical Reviews in Analytical Chemistry*, 2011, 41(2): 133-150.
- [6] BOLSHOV M A, KOLOSHNIKOV V G, RUDNEV S N, et al. Detection of trace amounts of toxic metals in environmental samples by laser-excited atomic fluorescence spectrometry [J]. *Journal of Analytical Atomic Spectrometry*, 1992, 7(2): 99-104.
- [7] DONG H, GIRMATSION M, WANG R Y, et al. Construction of fluorescent logic gates for the detection of mercury (II) and ciprofloxacin based on phycocyanin[J]. *Methods and Applications in Fluorescence*, 2022, 10(3): 035008.
- [8] GONG L D, ZHANG L X, SHEN R, et al. A novel conjugated polymer fluorescence probe for the detection of copper ions in aqueous media[J]. *Methods and Applications in Fluorescence*, 2022, 11(1): 015001.
- [9] 武玲, 韦庆益, 蒲洪彬. 黄绿色碳量子点的热合成荧光探针及其在白皮杉醇促氧化作用中的研究[J]. *现代食品科技*, 2019, 35(10): 174-181.
WU L, WEI Q Y, PU H B. A thermal synthesis of yellow-green carbon dots as a fluorescent probe for the study of the pro-oxidation of piceatannol[J]. *Modern Food Science and Technology*, 2019, 35(10): 174-181.
- [10] LI G C, GAO G Q, CHENG J Y, et al. Two new reversible naphthalimide-based fluorescent chemosensors for Hg^{2+} [J]. *Luminescence*, 2016, 31(4): 992-996.
- [11] HUANG M X, LU C H, HUANG Q D, et al. A novel and fast responsive turn-on fluorescent probe for the highly selective detection of Cd^{2+} based on photo-induced electron transfer[J]. *RSC Advances*, 2019, 9(62): 36 011-36 019.
- [12] SURESH M, SHRIVASTAV A, MISHRA S, et al. A rhodamine-based chemosensor that works in the biological system[J]. *Organic Letters*, 2008, 10(14): 3 013-3 016.
- [13] KAUR P, KAUR S, SINGH K. Bis (N-methylindolyl) methane-based chemical probes for Hg^{2+} and Cu^{2+} and molecular IMPLICATION gate operating in fluorescence mode[J]. *Organic & Biomolecular Chemistry*, 2012, 10(8): 1 497-1 501.

(下转第 153 页)

- effect of volatile oil of *Acorus tatarinowii* on anti-Alzheimer's disease based on GC-MS combined with network pharmacology[J]. Chinese Pharmaceutical Affairs, 2023, 37(10): 1 129-1 139.
- [24] 钱一帆, 朱韵韵, 姬珊珊, 等. 石菖蒲挥发油超临界 CO₂ 萃取及胶体磨法包合工艺的研究[J]. 药学服务与研究, 2016, 16(2): 121-125.
- QIAN Y F, ZHU Y Y, JI S S, et al. Supercritical CO₂ fluid extraction and colloid mill inclusion methodology of essential oil from *Acorus tatarinowii* Schott [J]. Pharmaceutical Care and Research, 2016, 16(2): 121-125.
- [25] 唐一梅, 张博, 房海娜, 等. 离子液体辅助提取石菖蒲挥发油的工艺研究[J]. 西北大学学报(自然科学版), 2018, 48(3): 393-398.
- TANG Y M, ZHANG B, FANG H N, et al. Study of ionic liquids-assisted extraction process of *Acorus tatarinowii* Schott. volatile oil[J]. Journal of Northwest University (Natural Science Edition), 2018, 48(3): 393-398.
- [26] 罗球珠, 杨棣华, 巫资葵, 等. 石菖蒲挥发油成分的气相色谱-质谱分析[J]. 中国药物经济学, 2021, 16(8): 116-120.
- LUO Q Z, YANG D H, WU Z L, et al. Analysis of the constituents of volatile oil in *Rhizoma Acorus tatarinowii* by using GC-MS[J]. China Journal of Pharmaceutical Economics, 2021, 16(8): 116-120.
- [27] 张莉, 纪娟, 刘艺, 等. 不同方法提取石菖蒲挥发油气质联用成分分析[J]. 辽宁中医药大学学报, 2016, 18(1): 53-55.
- ZHANG L, JI J, LIU Y, et al. GC-MS analysis the constituents of volatile oils from *Acorus tatarinowii* Schott extracted by different methods[J]. Journal of Liaoning University of Traditional Chinese Medicine, 2016, 18(1): 53-55.
- [28] 韩沅沅, 方洪, 谷丽华, 等. 基于网络药理学石菖蒲挥发油干预中枢神经系统疾病共性物质基础和作用机制探讨[J]. 上海中医药大学学报, 2022, 36(3): 52-60, 88.
- HAN Y Y, FANG H, GU L H, et al. Exploring common material basis and mechanism of volatile oil from *Acori tatarinowii* Rhizoma in intervening central nervous system diseases based on network pharmacology [J]. Academic Journal of Shanghai University of Traditional Chinese Medicine, 2022, 36(3): 52-60, 88.
- [29] 侯宇, 郑娇, 屠鹏飞, 等. 基于中空纤维吸附靶酶的流体灌注系统用于苦丁茶总皂苷活性成分筛选[J]. 药科学报, 2020, 55(3): 495-500.
- HOU Y, ZHENG J, TU P F, et al. Perfusion enzyme affinity selection system based on hollow fiber for screening active compounds from total saponins of *Kudingcha* [J]. Acta Pharmaceutica Sinica, 2020, 55(3): 495-500.
- [30] 苗静琨, 陈启雄, 吴小玫, 等. 石菖蒲 α -细辛醚对 Lithium-Pilocarpine 癫痫模型 GABA 系统的调控作用[J]. 中国药理学通报, 2011, 27(8): 1 067-1 071.
- MIAO J K, CHEN Q X, WU X M, et al. Effect of α -asarone on regulation of GABA system in the Lithium-Pilocarpine model of epilepsy[J]. Chinese Pharmacological Bulletin, 2011, 27(8): 1 067-1 071.
- [31] 邢倩倩, 刘振民, 洪青, 等. 气相色谱-质谱联用技术在食品分析方面的应用进展[J]. 食品与机械, 2023, 39(6): 234-240.
- XING Q Q, LIU Z M, HONG Q, et al. Application progress of gaschromatography-mass spectrometry in food analysis[J]. Food & Machinery, 2023, 39(6): 234-240.
- (上接第 66 页)
- [14] KAUR M, AHN Y H, CHOI K, et al. A bifunctional colorimetric fluorescent probe for Hg²⁺ and Cu²⁺ based on a carbazole-pyrimidine conjugate: Chromogenic and fluorogenic recognition on TLC, silica-gel and filter paper[J]. Organic & Biomolecular Chemistry, 2015, 13(26): 7 149-7 153.
- [15] HE T, LIN C L, GU Z Y, et al. Sensing behavior and logic operation of a colorimetric fluorescence sensor for Hg²⁺ /Cu²⁺ ions[J]. Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy, 2016, 167: 66-71.
- [16] JOLLIFFE I T, CADIMA J. Principal component analysis: A review and recent developments[J]. Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 2016, 374(2 065): 20150202.
- [17] DONG J, YAO Z J, ZHANG L, et al. PyBioMed: A python library for various molecular representations of chemicals, proteins and DNAs and their interactions[J]. Journal of Cheminformatics, 2018, 10: 1-11.
- [18] QUINLAN J R. Induction of decision trees[J]. Machine Learning, 1986, 1(1): 81-106.
- [19] PETERSON L. K-nearest neighbor[J]. Scholarpedia, 2009, 4(2): 1 883.
- [20] HESS A S, HESS J R. Logistic regression[J]. Transfusion, 2019, 59(7): 2 197-2 198.
- [21] BREIMAN L. Random forests[J]. Machine Learning, 2001, 45: 5-32.
- [22] MAMMONE A, TURCHI M, CRISTIANINI N. Support vector machines[J]. WIREs Computational Statistics, 2009, 1(3): 283-289.
- [23] SHERIDAN R P, WANG W M, LIAW A, et al. Extreme gradient boosting as a method for quantitative structure-activity relationships[J]. Journal of Chemical Information and Modeling, 2016, 56(12): 2 353-2 360.
- [24] GARDNER M, DORLING S. Artificial neural networks (the multilayer perceptron): A review of applications in the atmospheric sciences[J]. Atmospheric Environment, 1998, 32(14/15): 2 627-2 636.
- [25] CHICCO D, JURMAN G. The advantages of the Matthews correlation coefficient (MCC) over F₁ score and accuracy in binary classification evaluation[J]. BMC Genomics, 2020, 21: 1-13.
- [26] NARKHEDE S. Understanding auc-roc curve[J]. Towards data science, 2018, 26(1): 220-227.