
上海药物所通过多任务深度神经网络建立药物调控激酶谱的预测分析方法

作者：writer 来源：中国科学院

本文原地址：<https://www.iikx.com/news/progress/6605.html>

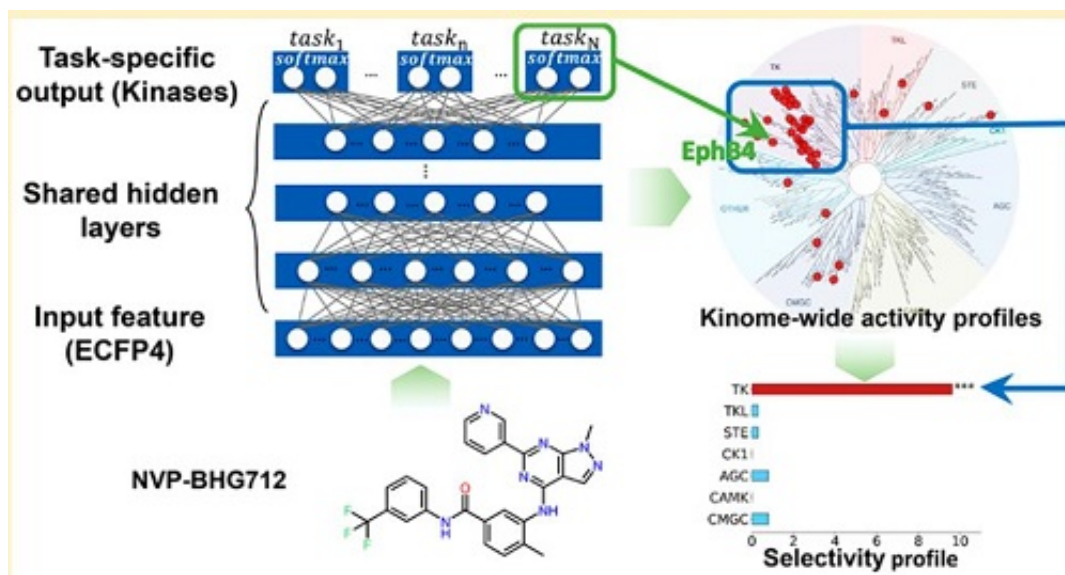
本文仅供学习交流之用，版权归原作者所有，请勿用于商业用途！

上海药物所通过多任务深度神经网络建立药物调控激酶谱的预测分析方法。蛋白激酶(protein kinases)是细胞功能的关键调节分子，是生物体内最大且功能最多样的基因家族之一。因此，激酶是开发治疗癌症、炎症、糖尿病、心血管疾病和阿尔兹海默症等相关疾病药物的重要靶标。然而，由于激酶家族蛋白质(特别是催化域)结构的高度保守性，给高效选择性激酶抑制剂的开发带来了巨大挑战。

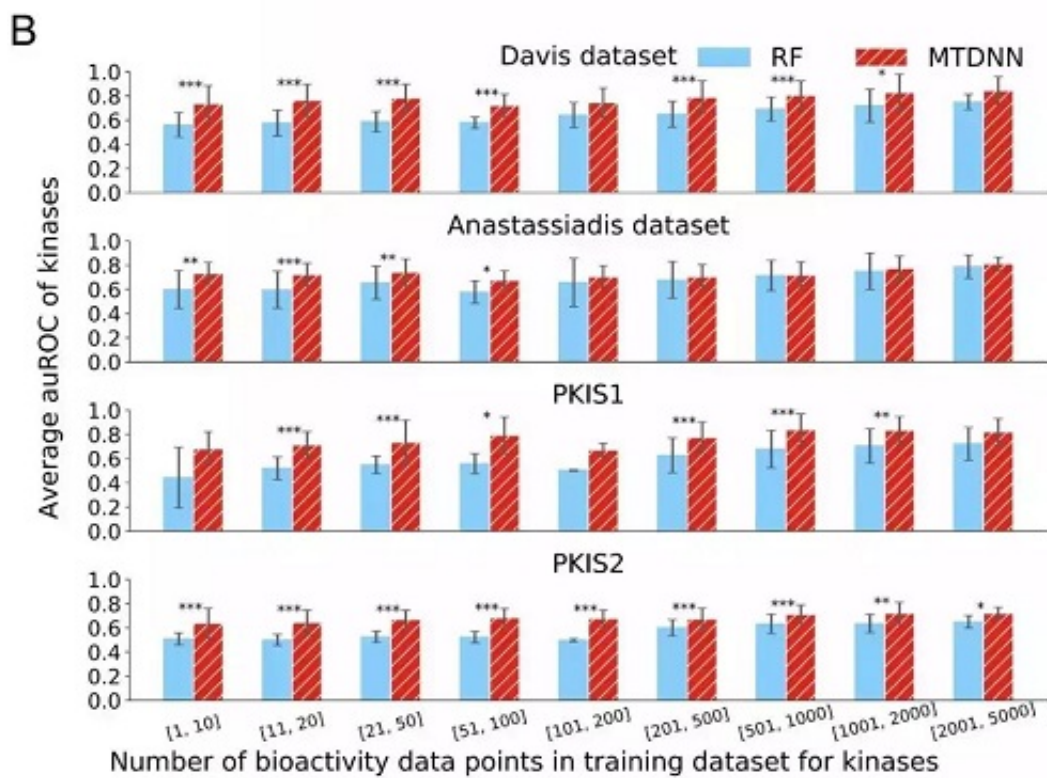
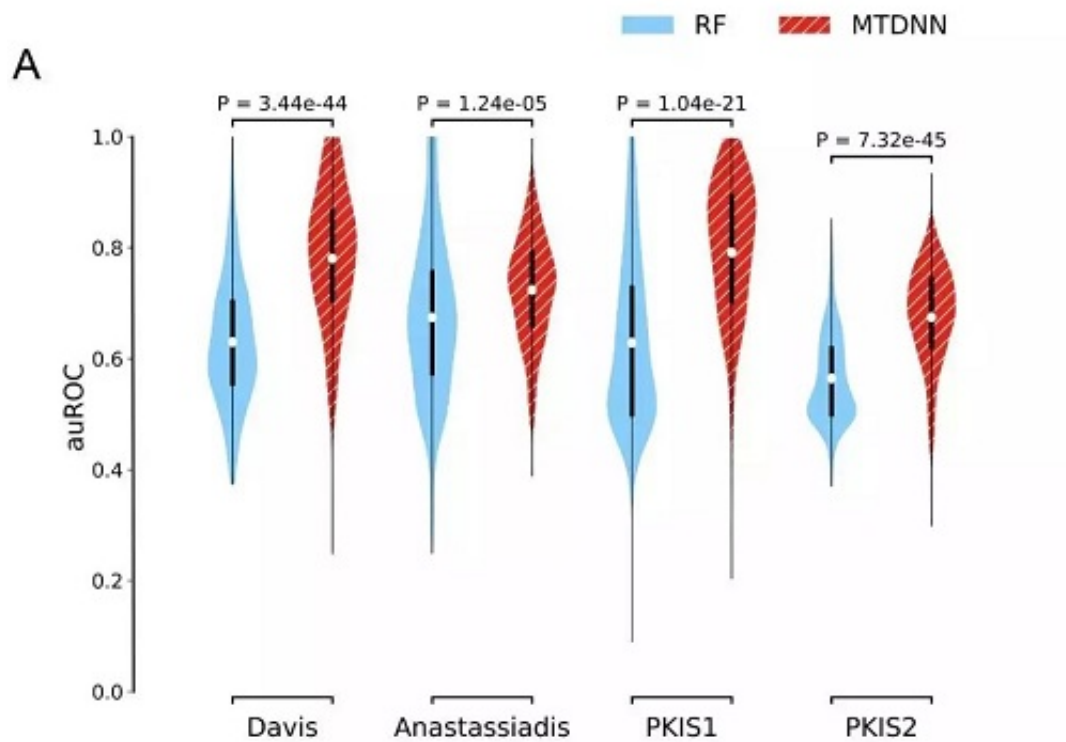
二十一世纪以来，随着计算机计算能力的迅猛提升和大数据的涌现，深度学习在机器学习算法的基础上快速崛起，并在药物研发领域得到广泛应用。然而，对许多尚未经过充分深入研究的激酶靶标，目前已有的数据还远远达不到训练中等规模神经网络所需的量级。因此，传统的单任务神经网络模型通常难以取得较好的泛化性能。

为了解决这一问题，中国科学院上海药物研究所蒋华良、郑明月团队采用了多任务深度神经网络(multitask deep neural network)建立分类模型解决化合物的激酶谱预测问题。多任务深度神经网络通过任务间的迁移学习，可以高效解决具有相关性的多类别分类问题，对于众多激酶靶标，共享的保守催化域使得多重活性预测任务紧密相关。因此，利用多任务深度神经网络可以有效减少特定激酶数据不足对模型泛化性能的限制。此外，多任务神经网络只需建立一个模型即可对整体激酶谱进行预测，无需多次建模，并且通过使用共享表示减少了模型参数的规模，可以使模型的训练学习过程更加高效。

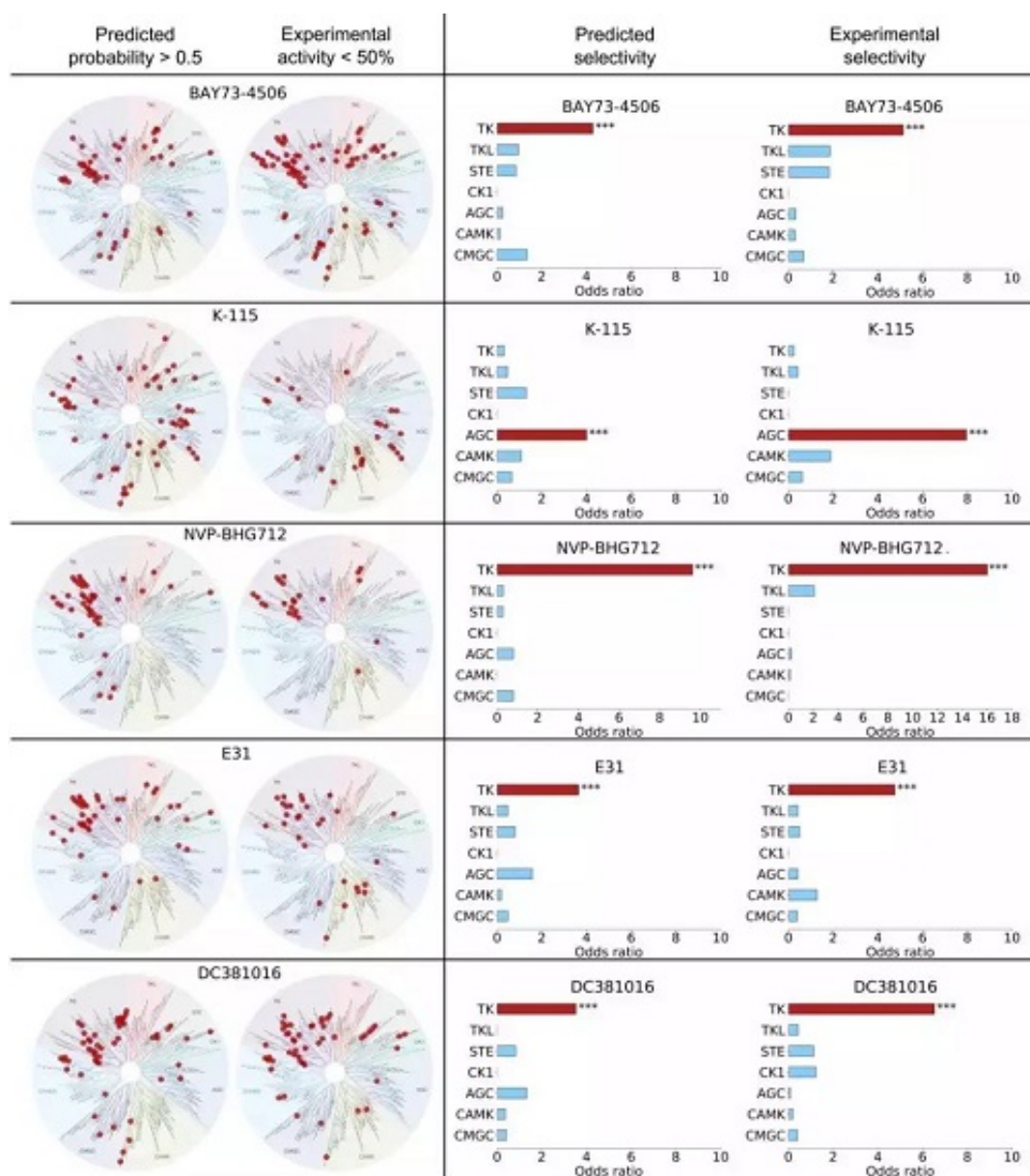
上述研究结果近期在线发表于Journal of Medicinal Chemistry，题为Deep Learning Enhancing Kinome-Wide Polypharmacology Profiling: Model Construction and Experiment Validation，并被选为封面论文。上海药物所药物设计与发现中心(DDDC)郑明月为论文通讯作者，第一作者是药物发现与设计中心博士研究生李叙潼。



基于多任务神经网络的药物激酶谱预测流程



多任务神经网络与随机森林预测表现对比



(左)对五个化合物预测与实验得到的激酶图谱，标记为红色的激酶表示预测或实验测定为活性。
(右)预测与实验得到的激酶家族选择性，当一个家族的odds ratio显著大于1(红色)，认为化合物对该激酶家族具有选择性。

更多 科学进展 请访问 <https://www.iikx.com/news/progress/>

本文版权归原作者所有，请勿用于商业用途，[爱科学iikx.com](https://www.iikx.com)转发