

当A/B实验平台遇上强化学习



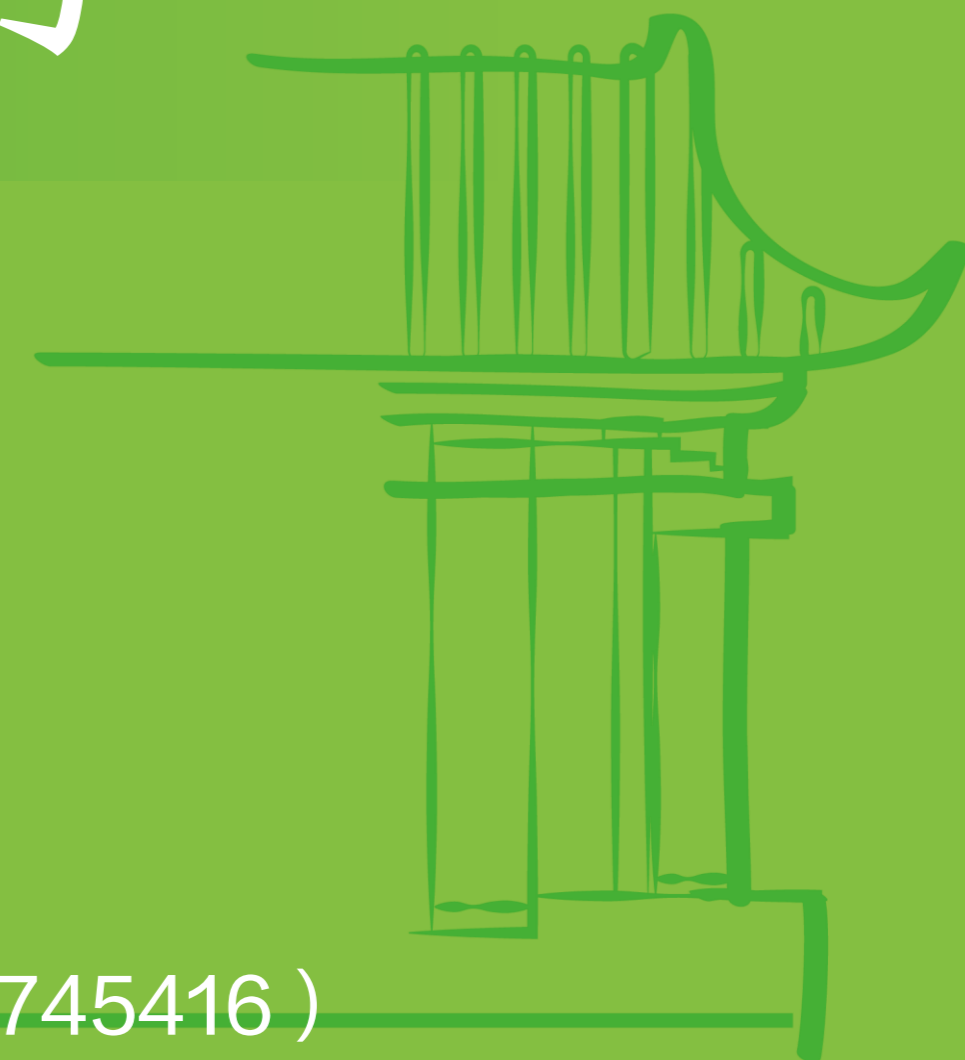
彭天翼 助理教授

哥伦比亚大学
商学院决策风险与运筹系

🎤 Host: 袁晓 助理教授

🕒 2024年7月22日 星期一 10:00am

📍 线上讲座 (Zoom 会议号: 812 5017 9391, 密码: 745416)



Abstract

在当今互联网时代，A/B测试已被奉为评估算法效果的黄金标准。然而，Interference问题（即不同实验单元相互影响的现象）却一直是A/B测试中的一大挑战。这一问题的存在使得业界投入巨大资金构建的实验平台的测试结果变得不再可靠。为了克服Interference问题，我们提出了一种基于强化学习框架的创新方案，用于重新审视A/B测试。该方案通过求解强化学习中的Q值的差分来估计 treatment effect，因此我们将其命名为“Difference-in-Q”（简称DQ）estimator。从理论角度看，我们发现DQ在bias-variance trade-off方面表现出色：一方面，DQ显著减少了传统estimator的bias；另一方面，DQ的variance较任何unbias estimator都可以实现指数级的缩减。我们与抖音公司合作，将DQ应用于实际商业场景，在初步测试中，其均方误差减少了超过99%。同时，在商业级共享汽车simulator中，DQ也展现出了卓越的性能。在本次演讲中，我将介绍DQ的理论与实践，并探讨下一代智能实验平台的展望。

Biography

彭天翼，哥伦比亚大学商学院决策风险与运筹系助理教授。他于2023年取得麻省理工学院博士学位，2017年毕业于清华大学姚班。2023-2024在Cimulate.AI担任人工智能研究员。他的研究兴趣聚焦于生成式人工智能，强化学习，和实验评估。他喜欢关注前沿理论问题在实际问题中的应用，并曾与字节跳动，百威啤酒，Liberty Mutual等公司展开合作。他曾获得INFORMS Daniel H. Wagner实践卓越奖，Applied Probability Society Best Student Prize, RMP Jeff McGill Student Paper Award, 和MSOM Best Student Prize Finalist。