

# 宁文韬

个人主页: stevenn9981.github.io/

邮箱: stevenn9981@gmail.com

电话: (+86) 13979805143 / (+852) 96718094

研究兴趣: 推荐系统, 信息检索, 大语言模型

## 教育背景

### 香港大学

博士 - 计算机科学专业

香港

2020 – 2024

- 论文: 在信息检索/数据库领域顶会 (WWW, CIKM, VLDB, ICDE 等) 上共发表 9 篇论文, 其中 7 篇为第一作者。

### 南方科技大学

本科 - 计算机科学与技术专业, GPA: 3.73/4.00

深圳

2016 – 2020

## 工作 / 实习经历

### 腾讯, 广告营销服务线

算法工程师 (青云计划校招生)

深圳

2024 年 12 月 – 至今

- 基础模型升级 (2025 年 8 月 – 至今) :

- \* 公众号 pCTR 模型及 PCAD pCVR 模型升级 Unimixer-百 M 模型 (主要负责人): 参与广告推荐模型的参数 Scaling Up 升级, 构建广告领域推荐大模型。将推荐模型中 CPU 时代常用的 DCNv2, FwFFM 等结构, 替换成更加 GPU 亲和的 Rankmixer 结构, 并在其中加入各种创新点 (例如 field 拆分, task token, 增加异源异构 token, 更换 pre-norm 等), 升级为 Unimixer 架构。目前 pCVR 模型已在小流量实验中初步验证有效, GMV 提升 3%。
- \* 视频号 pCVR 模型升级 Unimixer-SparseMoE-1B 模型 (主要负责人): 视频号 pCVR 模型的 Unimixer-百 M 升级已在小流量实验下验证有效。准备在推全的同时, 探索更大参数的模型, 找出参数 Scaling Up 天花板, 目前正在探索 SparseMoE-1B 模型的效果。

- 出价算法升级 (2024 年 12 月 – 2025 年 7 月) :

- \* 智投 \* 付费 ROI 广告的出价调控升级 (主要负责人): 针对付费 ROI 广告特点, 设计特定的出价调控方案, 做到成本持平情况下, 跑量提升。目前该算法在智投付费 ROI 广告已全量上线。
- \* 探索基于强化学习的出价调控方案 (主要负责人): 采用 USCB 算法, 期望做到成本持平下的跑量最大化, 以取代线上的 MPC 算法。目前离线模拟可做到在成本打平情况下消耗显著上涨, 并打通线上实验链路。

### 字节跳动, TikTok 部门

推荐算法实习生

上海

2022 年 11 月 - 2023 年 3 月

- 探索不同的推荐策略, 提高 TikTok 创作者的投稿率。

- \* **投稿偏好模型**: 使用用户投稿视频数据流训练模型。通过向作者推荐更多模型认为他们可以创作的视频, 来激发创作灵感, 促进创作者的投稿。该策略获得了创作者投稿率 0.5% 的提升。
- \* **流量激励策略**: 使用 Uplift 模型来寻找对低热度视频不敏感的用户, 向他们更多推荐这些视频。该策略可使得视频创作者获得流量激励, 促进作者发布。同时定向对低敏感度用户推荐视频, 可防止用户流失。
- \* **评论激励策略**: 探究了创作者收到的评论数量与其投稿作品收到的评论数量之间的相关性。我们发现视频评论数从 0 到 1 的过程最能激发作者创作欲, 因此我们增加低评论视频的推荐, 从而激励创作者投稿。

### TCL 香港研究院

研究实习生

香港

2021 年 6 月 – 2021 年 8 月

- 针对现有的基于元路径的推荐算法, 我们提出了一种自动有效的基于强化学习的元路径搜索框架。
- 提出了一种新的基于元路径和图神经网络 (GNN) 的一种推荐算法。

## 论文发表

1. Wentao Ning, Reynold Cheng, Xiao Yan, Nur Al Hasan Haldar, Ben Kao, Nan Huo, Bo Tang. *Towards Efficient and Effective Recommendation Unlearning*. In **Guide-AI @ VLDB 2025**.
2. Wentao Ning, Reynold Cheng, Xiao Yan, Ben Kao, Nan Huo, Nur Al Hasan Haldar, Bo Tang. *Debiasing Recommendation with Personal Popularity*. In **WWW 2024** (被选中为 Oral Presentation, 接收率 189 / 2008 = 9.4%).

3. Wentao Ning, Xiao Yan, Weiwen Liu, Reynold Cheng, Rui Zhang, Bo Tang. *Multi-domain Recommendation with Embedding Disentangling and Domain Alignment*. In **CIKM** 2023.
4. Wentao Ning, Reynold Cheng, Jiajun Shen, Nur Al Hasan Haldar, Ben Kao, Xiao Yan, Nan Huo, Tian Li, Wai Kit Lam, Bo Tang. *Automatic Meta-Path Discovery for Effective Graph-Based Recommendation*. In **CIKM** 2022.
5. Wentao Ning, Xiao Yan, and Bo Tang. *Towards Efficient MaxBRNN Computation for Streaming Updates*. In **ICDE** 2021.
6. Wentao Ning, Qiandong Tang, Yi Zhao, Chuan Yang, Xiaofeng Wang, Teng Wang, Haotian Liu, Chaozu Zhang, Zhiyuan Zhou, Qiaomu Shen, and Bo Tang. *CheetahVIS: a visual analytical system for large urban bus data*. In **VLDB** 2020.
7. Carrie Wang, Wentao Ning, Xiaoman Wu, Reynold Cheng. *HINCare: An Intelligent Helper Recommender System for Elderly Care*. In **WWW** 2024.
8. Nan Huo, Reynold Cheng, Ben Kao, Wentao Ning, Nur Al Hasan Haldar, Xiaodong Li, Jinyang Li, Tian Li, Mohammad Matin Najafi, Ge Qu. *ZeroEA: A Zero-Training Entity Alignment Framework via Pre-Trained Language Model*. In **VLDB** 2024.
9. Reynold Cheng, Chenhao Ma, Xiaodong Li, Yixiang Fang, Ye Liu, Victor Y.L. Wong, Esther Lee, Tai Hing Lam, Sai Yin Ho, Man Ping Wang, Weijie Gong, Wentao Ning, Ben Kao. *The Social Technology and Research (STAR) Lab in the University of Hong Kong*. ACM **SIGMOD Record** 2022.

## 其他项目

---

- 腾讯广告算法大赛（全模态生成式推荐）

- 作为算法大赛内测队伍参与赛题评测，并取得内测团队第一名；之后作为核心队伍被选中参与赛题的评测优化，基线代码优化等工作。
- 在比赛中，我们使用生成式模型架构，端到端生成推荐的物品。同时，采用 RQ-Kmeans 等方法，将物品多模态信息融入，增强推荐效果。

- 高效、有效的推荐模型遗忘（发表于 **GuideAI@VLDB'2025**, 第一作者）

- 提出了一种新的推荐遗忘框架，可以让模型遗忘指定的一部分训练数据。该框架可适用于任意推荐模型，由两个关键设计组成，whitening module 和 phantom users。
- 该框架只使用需要被遗忘的数据进行训练，避免访问整个训练数据集，实现了快速遗忘。另外，通过匿名保存协同信息，也保留了较好的推荐准确性。

- 消除推荐算法中的流行度偏差（发表与 **WWW'2024**, 第一作者）

- 本文分析了现有的消除推荐中流行度偏差（popularity bias）方法的局限性，并提出一个新的用户感知版本的 personal popularity 来解决现有方法的局限性。
- 利用反事实推理技术来估计和控制 personal popularity 和 global popularity 在推荐中的因果效应，以减轻流行度偏差。该框架可以应用于不同的推荐模型，用以消除其流行度偏差。

- 提升多领域推荐效果（发表于 **CIKM'2023**, 第一作者）

- 本文提出了一种应用于多领域推荐的嵌入解纠缠架构（embedding disentangling），该架构可以在 embedding 层面明确地解纠缠了域间知识（inter-domain knowledge）和域内知识（intra-domain knowledge）。
- 同时也提出了一种基于随机行走（random walk）的领域对齐策略，以识别来自不同领域的相似用户/项目，有助于知识共享和避免过拟合。

- 基于强化学习的元路径挖掘技术（发表于 **CIKM'2022**, 第一作者）

- 本文提出了一个通用的基于强化学习的元路径选择框架 RMS，该框架可以被应用于任何基于元路径的推荐模型中，增强推荐模型的效果。
- 同时，本文也提出了一种基于元路径和图神经网络（GNN）的推荐算法 RMS-HRec，并设计了对应的训练策略，充分挖掘元路径在推荐任务中的潜力。

## 获奖情况

---

• 香港大学研究生奖学金	2020–2024
• 南方科技大学计算机系优秀毕业生	2020
• 南方科技大学树礼书院优秀毕业生	2020