

谭昌炼


个人信息

- 2020届
- 手机: 18818264891
- 专业: 计算机技术

城市: 现居上海

邮箱: 18818264891

岗位: 算法工程师



工作及教育经历

英语流利说	2020.12-至今	算法部	算法工程师	
富途证券	2020.07-2020.12	大数据应用组	算法研究岗	
复旦大学	2017.09~2020.06	计算机学院	研究生/计算机技术专业	GPA:
				3.31/4.00
复旦大学	2012.09~2017.06	药学院	本科/药学专业	GPA:
				3.14/4.00 (专业前25%)

项目经历

1. 英语流利说 - Darwin成人英语推题 - 算法开发 - 2021.07 - 2021.12
- 英语流利说是一款英语学习App。根据用户的英语水平和学习历史量身定制学习课程，使用AI技术实时调整学习内容。旨在颠覆传统知识图谱体系，构建全新教学模型。深度自适应学习系统，不断自我学习，自我进化。让「千人千面」的学习系统，成为可能！
 - 我主要负责了在Darwin成人英语和Bell发音两条业务线的推题工作。两条业务线采用，新用户低价首购，7天到期后再复购的运营模式，核心业务指标为首购用户的复购率。
 - 当前的推课最小粒度为一个session，结合Session内跳转规则来跳过部分session内的内容。session为单次请求内容的最小单元，用户无法根据session内用户的答题情况做即时的调整和反馈，限制了算法侧的操作性能。
 - 尝试将session级别的推荐逐步向activity级别推荐迁移，收益点在于：1、及时响应调整推送内容，处理一些典型问题（用户重复感、学习区）2、支持逐题推送后，打破原有multiple session页推送的定势，即时响应，更为灵活，可以支持更多的模式，比如新的推课模式，如题目feed流，用户可自由把握学习结束的时机。
 - 受开发资源的限制，计划分两步走：1. 通过预测用户答题动态组装Provider，使用 IRT, KT 等模型，预测用户答题情况，来“近似”实现逐题推送。所有的题目，仍组装成一个session推送，沿用现有的推送逻辑。2、未来实现逐题推荐。彻底改变现有以Activity为维度的推送方式，用户每回答一道题目会提交答案请求，并即时请求下一道题目。后者挑战点在于：1、涉及到后端、客户端全链路的RPC改造，投入大。2、调整后QPS显著上升，工程上需保障压力上升后的服务质量。
 - 在算法上，原有模型使用RNN在用户完整学习历史上对用户状态进行建模，使用了RL中(Top-K off-policy correction)来进行排序，reward则是用户未来的复购。我的工作是在原有的模型基础上，新加session内activity index做为特征信息，添加用户的掌握度提升、留存作为reward，阻断NEXT_SESSION类型的activity的梯度回传，复购率提升了+3.27%，z检验置信度95%。工作集中在离线指标设计、GPU训练速度优化、样本筛选、强化学习Reward设计、多目标优化。

2. 英语流利说 - bell发音推题 - 算法开发 - 2020.12 - 2021.06

- 当前Bell中的组课是根据用户对于知识点的掌握情况和之前学习记录，利用Bell题库重新组织的一种课程形式，在课程推送的时候，课程中的所有题目就已经确定。为了能够在更细粒度上(单题)级别上对用户做到自适应的推荐，需要对组课做逐题推送。推送的目标则是用户未来的掌握度提升情况。
- 在工程上，这个项目在组课上支持了逐题推送。在策略上，将组课的推送条件放开，用户能够更多地被推送到组课。在算法上，使用了RNN在session级别对用户状态进行建模，reward则是用户留存情况。通过AB test实验，复购率相对规则推题提升了+2.96%，z检验置信度95%。工作集中在离线指标设计、GPU训练速度优化。

3. 富途证券 - 社交内容推荐 - 算法开发 - 2020.07 - 2020.12

内容推荐系统

- 为用户推荐内容，目标是推荐列表feeds有效推荐率达到5%，有效推荐率= $U(\text{关注}+\text{点击}+\text{互动})/\text{feeds曝光量}$ 。
- 推荐系统分为召回、排序、重排序等阶段。负责多路召回中的最新、最热、向量召回，以及推荐结果的实时性、多样性等指标监测。
- 向量召回采用DSSM双塔结构，其中内容塔用HAN提取文本稠密特征，用户塔用CARN捕捉用户长短期兴趣。对于每天用户生成的新内容，可以用内容塔实时产生内容向量，解决内容冷启动的问题。DSSM双塔结构集成了用户行为、内容的语义信息及其他稀疏特征。
- 向量召回在测试集上做点击/不点击二分类任务，auc为0.794。选取合适的分类阈值后，在测试集(正负样本1:10)正样本上召回率为61.2%，精准率为29.8%。

4. 学校 - 电商推荐 - 算法研究 2019.03- 2020.03

- 项目以经过修改过的中文亚马逊电商数据集作为依托，提供了从前端应用、后台服务、算法设计实现、平台部署等多方位的闭环的业务实现。
- 改进了混合推荐中的实时推荐模块，提出了基于会话的神经网络模型CARN，以用户一个Session中的历史点击物品作为输入，预测用户下一最可能点击的K个物品。模型采用了最后点击物品与历史点击物品序列协作的注意力机制，考虑了用户Session内长时、短时兴趣的相互影响。在RSC15数据集上，CARN较同类注意力机制算法STAMP，MRR@10 (0.3508) 提升4.7%，HiteRate@10 (0.5901) 提升1.3%。

专业技能

- 推荐算法：熟悉常见的推荐算法，如ItemCF、UserCF、LFM等，熟悉Session-based的推荐，如GRU4REC、STAMP等，了解基于内容、UGC、社交的推荐。
- 深度学习：熟悉CNN、LSTM、GRU等神经网络，熟悉Attention机制，熟悉序列建模方法熟悉TensorFlow。了解强化学习。
- 传统机器学习：熟悉常见的机器学习算法，如朴素贝叶斯、逻辑回归等分类算法，k-means、DBSCAN等聚类算法，PCA等降维方法，AdaBoost等集成学习方法。
- 语言：熟练掌握Python，熟悉go、Java。
- 大数据：了解Spark、Kafka、Hive、Hadoop、Hbase等大数据工具，用Spark写过离线计算任务。

获奖经历

- 复旦大学 学业奖学金 2019-10
- 复旦大学 国家励志奖学金 2015-10
- 复旦大学 赛默飞奖学金 2013-10

其他信息

- 语言：CET6英语（CET-6）
- 喜欢钻研技术 等等