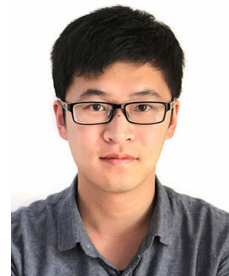


罗少泽

LUO Shaoze



✉ luoshao23@foxmail.com

☎ (+86) 152-1671-6974

🌐 github.com/luoshao23

📍 上海市徐汇区文定路 468 弄 1 号

📅 1990 年 8 月 25 日

🔗 luoshao23.github.io

👤 工作经历

算法工程师

2024.4 至今

主要负责视频推荐，改善用户使用体验，进一步提升留存、拉新。

- * 多目标优化：完善精排多目标，提升排序能力。先后构建或优化了“退出内流”、“个人主页时长停留”和“有效播放”等目标，累计提升留存 0.12%，时长 +0.57%，并显著提升了用户互动行为。探索多目标 LTR 优化新方向，通过提升融合打分个性化，最终提升用户消费满意度。
- * 召回链路优化：对主召回模型优化特征、结构、样本进行优化，留存 +0.0133%；负采样系统框架重构，重构后新系统可承接 30+ 个批式任务 qps，特征调研更灵活，支持自定义范围负样本采样。
- * 帮助建设自建视频能力，优化自建内容分发，自建视频播放占比 23%→71.3%，大盘总播放 +3%、留存 +0.27%

小红书科技有限公司

算法工程师

2020.3 – 2024.3

主要负责小红书 APP 社区搜索召回相关性方面工作，通过机器学习、深度学习等算法策略提升搜索相关性、个性化以及时效性，保证社区内容质量，提升用户留存。

- * 维护搭建召回架构，开发多通道融合机制，利用离线深度模型，对所有通道召回笔记进行统一排序打分，并结合线上强化学习策略根据用户反馈即时更新模型。最终提升用户消费指标，搜索次 3 日留存 +0.028pt，query 有点比 +0.261pt，query 前四有点比 +0.206pt。
- * 基于多任务学习范式，在 MMoE 模型框架下，搭建以点击、相关性、消费等目标的统一召回模型，提升了线上召回相关性 (auc+2.2pt)，query 有点比 +0.23pt，query 前四有点比 +0.15pt，L7+0.056%，次 3 留 +0.055%，相关性在人工测评结果集上均有正向提升，同时节省了离线任务耗时和线上资源。
- * 构建全局/长尾向量召回模型，融合预训练 Bert 向量、个性化、上下文特征，考虑字词多粒度语义，选用多层级负样本，构建基于对比学习的深度模型。配合总量 7 亿 +ANN 索引库，累计提升次 3 留 0.06%、query 有点比 +0.35pt，同时显著下降长尾词上的无少结果率。
- * 利用多模态技术，结合文本与图像等多模样本，构建多模态召回模型；利用内外流信息，构建 id 嵌入模型。两个模型累计提升有效视频播放数 2.2%，提升 ces 约 1.34%，改善搜索下视频消费环境。
- * 基于互动点击行为，利用 pagerank、LPA、swing 等图算法，挖掘高质量笔记，在此基础上构建 youtubeDNN 和 DIN 深度模型，提升 query 前四有点比 0.07pt，笔记有效点击率 1.1%。
- * 新增时效性召回通道，通过搜索热词的挖掘与提权，保证笔记供应的即时性。同时在此基础上新增了“小红书搜索发现”功能，通过强化学习为用户提供粗粒度的个性化热词推荐。

百丽电子商务（上海）有限公司

算法建模经理

2018.3 – 2020.3

高瓴资本集团投资，以 BeLLE 鞋业集团为蓝本，搭建以“人-货-场”为基本元素，选款、定量、补货、调价等功能模块于一体的零售业鞋品管理优化平台。主要负责算法设计、数据挖掘、模型搭建。

- * 为帮助业务优化库存结构，进行智能化补货，及时发现畅销款，运用自主研发的基于生命周期理论的销售速率模型、xgboost 等模型 ensemble 为多个女鞋品牌提供销量预估，销量误差中远期可低于 30%，短期预测误差可保持低于 20%，比原业务预估方法降低 8%-10%。
- * 帮助进行鞋款相似性识别，优化货品分类管理结构，通过深度学习、迁移学习等技术训练 CNN 深度神经网络，最终可以获得较好相似度，其中推荐的鞋款超过 80% 获得业务认可。
- * 结合量价关系，运用图算法，给出基于市场事实的最优调价策略，提高期末销售量，减轻期末约 20% 的库存；考虑安全库存及售罄率等指标，利用带约束非线性规划 SLSQP 算法解决利润最优化问题。
- * 借鉴 assortment optimization 中的 EM 算法评估 SKU 效用，提供细至系列 SKU 的补货计划，合理分配补货空间，减少无效补量。

奥浦诺管理咨询（上海）有限公司

Analyst Specialist

2017.1 – 2018.2

为多行业领域客户提供全面有效的商业解决方案。

- * 为全美 TOP3 医药零售商提供基于用户消费行为进行顾客分群识别，基于 HMM 理论进行消费行为预测、并给出合理个性化推荐，提高客户参活，提高当季业务收入额约 1.2million；

- * 借助深度学习算法优化某大型外企银行信用卡评分模型，运用 CNN 衍生有效变量，强化模型输入，提高风险识别能力，减少坏账损失；
- * 为某欧洲银行提供高净值用户识别及信用卡流失率预测。

中国船舶及海洋工程设计研究院助理工程师2016.4 – 2016.11

智能制造赋能，高新船舶与结构物设计研发，搭建多目标优化平台辅助完成总体方案设计与详细设计。运用 CFD 数值离散算法、各类优化算法如 NSGA-II、NLPQL 提高船舶设计自动化、智能化程度，最终精度可用于实际工业制造。

🎓 教育背景

上海交通大学	船舶与海洋工程学院	硕士	GPA:2.76/3.3, Top 3/79	2013.9 – 2016.3
上海交通大学	工商管理	学士	二专	2011.3 – 2013.7
上海交通大学	船舶与海洋工程学院	学士	GPA:89.4/100, 8/81	2009.9 – 2013.7

🏠 社会实践

上海交通大学船舶海洋与建筑工程学院研究生会	主席	2014.9 – 2015.7
-----------------------	----	-----------------

🏆 奖项荣誉

Hackthon@Tacos’ 17, 3rd place	2017.10	上海交通大学优秀毕业生	2016.3
国家奖学金	2015.10	上海交通大学三好学生	2014.10
新科工程奖学金（续评）	2012.11	新科工程奖学金（续评）	2011.11
上海市奖学金	2011.11	新科工程奖学金（初评）	2010.11
全国大学生数学建模竞赛三等奖	2011.11		

👤+ 资格认证

CFA level II	CFA Institute	2019.8
--------------	---------------	--------

🏢 实习交流

横滨国立大学（日本横滨）	短期交流生	2014.12.2–2015.2.28
• 基于 NSGA-II 算法对船舶性能进行多目标协同优化及风阻研究，基于粘性流理论建立 CFD 模型，并通过小尺度风洞试验验证模型精度，在此基础上通过调整集装箱排列分布以减少船舶风阻		
ST Engineering(Marine)（新加坡）	助理工程师	2012.7.23 – 2012.8.31
• 根据目前市场情况，进行两栖登陆艇设计制造项目规划的可行性分析，评估预期收益。两栖登陆艇初步设计、详细设计，考虑稳性、破仓稳性的基础上设计分仓。		

📑 主要项目

视频推荐优化	2024.4 至今
• 将“退出内流”作为负反馈信号引入到排序公式中，改善用户消费满意度。留存 LT30 +0.0538%, HLT30 +0.2347%, 全局核心 ST/U +0.831%, SD/U +0.782%, 视频内流活跃天 +0.156%, 人均时长 +3.028%；	
• 修复原先个人主页时长预估的不合理问题，通过调整为分类目标。其中个人主页的渗透率 +2.1393%，人均次数 +2.7238%，人均时长 +2.7238%，留存 HLT7+0.025%，90D 拉活老用户 LT30+0.0771%，90D 主启老用户 LT7+0.0286%，HLT30+0.0404%，全局核心 st+0.1811%，sd+0.1549%，视频内流人均时长 +0.3537%，互动显著提升。	
• 主召回模型优化：通过对主召回模型特征、结构、损失函数等方面的优化。离线模型批式训练评估 uauc+3.29%，recall+0.01 有提升。协助上线实验，留存 LT30 +0.0133%，HLT30 +0.033%，高活用户 LT30 +0.03%，全局 st/u +0.10%，sd/u +0.12%。	
• 负采样系统框架重构：重构后新系统可承接 30+ 个批式任务 qps，特征调研更灵活，支持自定义范围负样本采样。在主召回优化中引入精排难负例，提升与后链路的一致性，离线评估指标 hitrate+17%(0.54->0.63)。	

- 为统一召回侧排序融合逻辑，减少通道间偏差。搭建多通道融合机制，离线训练考虑上下文信息、个性化信息的深度模型，并结合线上强化学习策略（CMA-ES），分钟级收集实时反馈，更新模型权重，最终输出 agg 排序分，对所有通道召回笔记进行统一排序打分，最终提升了用户消费指标，搜索次 3 日留存 +0.028pt，query 有点比 +0.261pt，query 前四有点比 +0.206pt。得益于框架优化，显著减少链路整体耗时；
- 为提升召回效果，减少了资源开销。通过多任务学习和 MMoE 模型框架，以点击、相关性、消费等目标，训练基于对比学习，选取多级负样本，离线 auc 和 recall 提升显著，线上 query 有点比 +0.23pt，query 前四有点比 +0.15pt，L7+0.056%，次 3 留 +0.055%，相关性在人工测评结果集上均有正向提升。
- 为了提升搜索体验，减少整体无少结果，开发搜索向量召回模型，融合预训练 Bert 向量、个性化、上下文特征，考虑字词多粒度语义，构建基于对比学习的深度模型，选用多层级负样本。配合总库 6 亿 ANN 索引，累计提升次 3 留 0.06%、query 有点比 +0.35pt，同时显著下降长尾词上的无少结果率；
- 为优化社区视频消费环境，利用笔记文本和视频多帧图像信息，开发了基于多模态技术、迁移学习的向量召回双塔模型，通过文本和图像的 self-attention 以及 cross-attention，充分提取笔记信息，通过基于 hnswlib 的 ANN 索引，并配合搜索词的词中心匹配机制，召回与搜索词最为相关的视频笔记，提升了有效视频播放数 2.2%，社区消费指标约 1.34%；
- 为提升搜索时效性，考虑搜索词在不同时间窗口下的每日搜索量变化程度，挖掘搜索侧热词，并对相应热词下的新发布笔记进行不同程度提权，保证笔记供应的即时性，提升了用户 query 满意度（有笔记点击的 query 点击率）。此外通过 contextual UCB 算法，累计用户在小红书热搜中的数据，进而提供粗粒度的个性化热词，提升了用户 query 点击率与笔记消费指标；
- 为优化男性搜索场景冷启动，基于互动点击行为，利用 pagerank、LPA、swing 等图算法，挖掘高质量笔记，在此基础上构建 youtubeDNN 和 DIN 等深度模型，提升 query 前四有点比 0.07pt，笔记有效点击率 1.1%。

BeLLE 快时尚销量预测与库存优化

百丽鞋业

2018.3 – 2020.2

- 为业务提供每季度畅销款建议，帮助掌握销售趋势。结合线下线上数据、智能门店特征因子及外部数据，自主研发了销售速率模型预测短期及长期销量。通过结合 Multiplicative regression 销售速率模型、XGBoost 模型构建 Ensemble 模型为集团多个品牌提供从 sku、大类到总量的短期（每周）与长期（类别季度）销量预估，短期与长期的预测销量，MAPE 可以分别低至 8% 及 25%；
- 借助商品中的文本属性及图像，提出并实现使用 CNN 深度学习网络及 transfer learning 进行模型和参数的迁移学习，得到的图像的 embedding 特征，借此为新款 SKU 提供往年相似鞋款，鞋款相似度得到专业设计师认可。另一方面，将图像向量特征与生命周期曲线相关联，构建 LSTM 预测模型预测新上市鞋款的生命周期，为销量预测提供辅助信息，其预测效果会比基于时序的 prophet 模型在 MAPE 上优 12%；
- 为提高货品当季销售上的单位利润，结合销售速率模型得到价格销量关系，运用 DAG 最长路径算法为货品运营提供最优调价策略；在多约束条件下，运用非线性规划 SLSQP 算法解决最优化问题保证在以毛利率、售罄率、单位利润等关键指标下商品达到当期最优折扣；
- 借鉴 assortment optimization 中的 EM 算法计算 SKU 效用值，再求解最优化问题，得到每个风格内的 SKU 预测销量分配方案；

技能

- 编程语言：Java, Python, Golang, Shell, Matlab, SQL
- 框架：Spark, TensorFlow, Pytorch, Keras
- 外语：<English> 流利，CET6 优秀；<日本語> N2，读写熟练，听说基础
- 证书：国际日语能力测试 N2 级，特许金融分析师 CFA Level 3 candidate，基金从业资格证书
- 特长：绘画，唱歌，篮球，轮滑，二胡

自我评价

- | | |
|--------|---|
| 可靠的硬实力 | 数理基础扎实，掌握多种优化算法、机器学习算法及深度学习算法，部分可独立实现；大数据量化分析挖掘实践经验；Github 代码贡献者（lda-project）；精通多门开发语言 Java, Python, Golang |
| 可信的软实力 | 主观能动性、学习能力及抗压能力强；善于沟通，具有团队合作精神；乐观幽默 |