

黄英凡

(+86) 188-1072-6027 · huangyingfan@outlook.com · <https://github.com/huang-xx>

教育背景

华北电力大学, 信息与计算科学, GPA 1/46, 本科 2013.9 - 2017.6
中国科学院计算技术研究所, 计算机应用技术, 硕士 2017.9 - 2020.7

工作经历

2020 年校招入职京东, 担任【京东零售-技术与数据中心-数据与智能部】算法工程师, 主要参与多模态基础能力建设, 以及探索多模态及大语言模型在京东业务上的使用。在最近 6 个季度的绩效评估中, 本人共获得了 3 个 A 以及 1 个 A+ 的评级, 是近 2 年组内平均绩效最高的员工 (本人绩效成绩图)。

项目经历

1. 京东商城多模态预训练模型 (独立完成)

- **背景:** 以 CLIP、BLIP 为代表的多模态模型展示出强大的理解和生成能力, 在电商数据上构建多模态模型具有很强的业务价值;
- **难点:** 模型需要高可扩展性, 支持多种下游任务; 电商数据特性对于数据增强的限制; 商品库中存在低质量样本; 模型训练需要千万量级别图文对, 资源消耗大, 训练耗时长;
- **方案:** 选用 BLIP 作为预训练模型, 同时支持下游检索及生成任务; 数据增强时不使用颜色变换, 控制随机裁剪比例; 数据选用高 GMV 及活跃商品, 使用 EMA 建立软标签; 使用 LMDB 格式存储数据加速 IO, 使用混合精度加速训练;
- **效果:** 当前在阿里天池的中文多模态榜单 MUGE 上综合排名第 3 (组织名 JD_DK_Pretrain_Group, 点击“测试模型 v4”查看本人京东邮箱)。

2. 多模态检索应用于京东选品业务支持 (合作完成)

- **背景:** 京东内部基于“关键字匹配”的选品方案, 不具备深层次语义理解能力, 在处理“风格、场景”等抽象概念时, 严重受到标题中蹭词的影响, 使得选品效果不佳;
- **难点:** 理解“颜色、风格、场景、样式”等抽象概念, 降低“蹭词”的影响, 提升选品准确率;
- **方案:** 基于多模态预训练模型建立了一套“跨模态检索”选品方案 (使用“文本召回图片”), 利用多模态模型对于抽象概念的深层次理解能力, 精准召回符合抽象概念的商品;
- **本人工作:** 将预训练模型在业务类目数据上针对“跨模态检索”任务微调, 将 Text Encoder 和 Vision Encoder 分别转化为 TensorRT 格式, 后者由本人建立周级调度程序, 每周产出约 2000 万商品表征 (前者由同事搭建在线推理服务);
- **效果:** 覆盖了 219 个二级类目, 平均准确率相比文本提升 12.27% (80.58% → 92.85%) (图示);

3. 京东站内同品、相似品识别及百亿补贴业务支持 (京东零售集团重点战略项目, 部分参与)

- **背景:** 同品项目将同款商品进行聚合展示, 提升用户购买效率, 推动商家良性竞争; 相似品用于加速检索从其他平台搬家的非标品, 以建立跨平台的比价能力; 百亿补贴用于提升用户黏性;
- **难点:** 建立一套全面、准确的商品表征, 使得相似商品的 embedding 表征相近。为下游的同品/相似品聚合奠定基础, 为百亿补贴的商品审核提效;
- **方案:** 经过实验, 选择预训练多模态模型的单模态表征经过后期加权融合作为商品的全局表征; 对于同品/百亿补贴业务, 图片和文本比重为 1:1; 对于相似品业务, 图片与文本比重为 7:3;
- **本人工作:**
 - 基于预训练的基础模型, 在不同业务的商品底池上进行“多模态召回”任务的微调;
 - 针对同品/相似品业务, 搭建从【商品信息表】到【多模态召回表】数据流: 天级活跃商品表 -> 增量商品信息表 -> 增量商品 embedding 表 -> 全量商品 embedding 表 -> Faiss 召回; (业务天级新增商品数约为 350 万, 全量商品约为 1 亿)
 - 建立商品表征效果评估系统, 周期性从【多模态召回表】中取样 anchor 商品及其召回的相似商品交由运营同事评估结果, 运营同事将每个 (anchor、召回商品) pair 的二分类结果放入指定路径, 评估系统自动生成 Precision、mAP 等指标;
 - 针对百亿补贴项目, 搭建多模态在线推理服务 (基于 Triton 框架), 实时得出 2 个商品相似度得分;
- **效果:**

- 同品识别项目,运营同事在 40 个二级类目上的评估显示,多模态方案初始聚合准确率 54.52%,显著高于【基于图片特征】的 16.14%;其中,高销准确率平均准确率 94.35%,高销召回率 88.7% (图示);
 - 百亿补贴项目,线上服务覆盖 336 个重点三级类目,运营同事审核效率提升 30% (图示);
4. 多模态特征用于广告推荐 (仅作为特征提供方)
- 母婴场景标签小批量实验,孩子类标签预测,准确率提升5.66%;
 - 首页推荐离线 $auc + 0.2\%$,线上收益点消 2 – 3 个点;
5. 多模态生成模型探索 (独立完成)
- 基于预训练的基础模型,在“图片描述生成”(image caption)任务上进行微调,用于商家发品阶段时,辅助生成营销文案。目前主要针对衣物类别,以线上服务的形式提供调用 (图示);
 - 基于 Stable Diffusion 的文本生成图片探索:在 Stable-Diffusion-2-1 的 checkpoint 基础上,利用开源的 Chinese-CLIP 替换英文的 Text Encoder,在中文数据集 TaiSu (过滤后数据量为 1.2 亿)上微调一个 epoch (只训练 Text Encoder 参数)。当前模型以 WebUI 的方式部署在京东内部的 AI Hub 上 (图示);

前沿技术

- NLP 模型:
 1. Encoder: BERT、ALBERT、SpanBERT、RoBERTa、ELECTRA、DistilBERT、TinyBERT 等;
 2. Decoder: XLNet、GPT 系列 (含 InstructGPT)、LLaMA 等;
 3. Encoder-Decoder 系列: UniLM、BART、T5、GLM 等;
- 多模态模型: ALBEF、VLMO、BLIP、CoCa、BEIT-V3、Flamingo、BLIP2 等;
- CV 自监督模型: MoCo、BEIT、MAE 等;
- 大模型训练技巧: 梯度累加、激活重计算、混合精度训练、数据并行、Tensor 并行、流水线并行、Zero 技术等 (个人笔记);

科研经历

以第一作者完成论文 **STGAT: Modeling Spatial-Temporal Interactions for Human Trajectory Prediction**, 被 CCF A 类会议 ICCV 2019 接受为 **Oral** 论文 (录用率 4.3%), 代码链接 **Code**。当前 Google Scholar 引用次数 360 余次。

主要获奖情况

- 北纬优秀科研硕士生奖 (计算所全所每年 3 个名额)

个人总结

- 具备良好的沟通和协作能力,能快速适应与不同风格的团队合作完成项目;
- 具有较强的学习能力和定位问题的能力,能独立处理具有一定难度的问题;
- 工作认真负责、善于协作、乐于分享。