

王浩

15821903289 . wanghao.ftd@foxmail.com

个人信息

- 男, 1992 年出生
- 求职意向: 搜索算法工程师
- 工作经验: 3.5 年

教育经历

- 硕士, 上海交通大学**, 计算机科学与技术专业, 2015.9~2018.4
 - 研究方向: 自然语言处理; 导师: 赵海教授
- 学士, 西安交通大学**, 软件工程专业, 2011.9~2015.7

工作经历

- 腾讯科技(深圳)有限公司, 应用研究, 2018.4~2021.10
- 负责看点搜索下的视频搜索体验及视频内容理解能力建设
- 负责CDG FIT虚拟银行开户证件OCR能力建设

项目经历

- 看点搜索: 短视频结构化理解与应用**

项目背景

通过对不同垂类短视频进行结构化内容理解, 从而应用到多个产品形态。

视频理解模块1: 视频结构化DIS标签生成

对多个垂类的短视频标题训练**Prompt tuning + Bert**结合的序列标注模型来生产标签, 既能利用Bert大规模预训练任务的效果, 同时Prompt tuning的思路也解决了扩展领域和嵌套的问题。最终在人工标注的测试集上达到平均F1为83%左右的效果。

视频理解模块2: 多模态视频主题分类

模块1覆盖了多个垂类短视频的大部分结构化标签, 不过在影视领域需要识别出视频主题类型, 这类标签无法通过标题获取, 需要理解视频内容进行识别。

设计了视频多模态理解模型, 这里模态包括: **标题文本Bert embedding**以及**视频 ViT embedding**; 首先基于双流模型进行对比学习作为预训练任务, 再通过单流模型进行分类学习。最终在人工标注的影视类短视频主题分类测试集上F1达到92%。

线上效果

视频理解的两个模块产生了丰富的短视频标签, 直接支持了看点视频搜索中的“视频筛选器”以及“视频合集”两个产品形态。线上小流量实验效果为: 资源点曝光比相对提升123.5% (17.32% --> 38.72%), 有点率相对提升15.3% (21.84% --> 25.19%)。其中筛选器中标签词的CTR为6.17%。

看点搜索: 短视频部分精排特征建设

视频簇紧密度特征

通过挖掘视频搜索精排队列中资源的视觉embedding关系特征, 加入 LTR 模型, 提高视频搜索的 CTR 指标。通过视觉 embedding 紧密度寻找队列中心点, 结合资源队列中的高 CTR 资源为中心点, 计算其他视频与中心点视频的相似度特征, 调高优质资源的排序优先级, 从而提升视频搜索的 CTR。该特征目前在相关性 LTR 实验中, 在验证集ndcg@-3 指标上能够得到 0.7% 的绝对提升。

- 金融场景多证件OCR能力建设**

项目背景

支持腾讯虚拟银行（现富融银行）开户场景下的证件OCR能力；包括：香港身份证，大陆身份证，港澳通行证，护照。

识别系统方案

首先用分类模型判断证件类型，然后使用定位模型对证件区域进行矫正，采用PixelLink进行文字检测。在文字识别部分，模型使用**CNN+CTC loss**进行文字识别。

最终效果

测试集上取得多个证件end to end 96%的准确率。

•

✂ 技能清单

- 自然语言处理
- 视频多模态理解
- C++、Python