

极客马拉松大赛

【小工具—移动端】

(基于AI技术的移动端违规商品审核)

目录

- 队伍介绍
- 项目简介
- 技术突破点
- 业务贡献/商业价值
- 方案详细介绍

一、队伍介绍

团队风采展示照片
(请在此处提交团队照片)

基本信息

- 队伍名称: 以AI之名
- 队伍队长: 梁延彬 (liangyanbin)
- 队伍成员: 江佳蓓 (jiangjiabei1)、伦宇宸 (lunyuchen1)
林鹏 (linpeng24)、刘倩 (liuqian255)
- 队伍Slogan: Let's hacker together
- 技术方向:
小工具 — 移动端、AI

二、项目简介

项目名称

- 基于AI技术的移动端违规商品审核

一段话简介

- 利用移动端上的AI图像和文本识别技术，将后端的商品审核前置到移动端上进行，通过增加创建商品的一次审核通过率，从而减少外采系统商品审核的次数从而减少审核费用，同时可即时反馈商品审核结果，提高审核效率。

三、技术突破点

行业/京东技术现状

- 商家在平台创建商品提交后，平台多使用外采系统对商品图片、商品（SPU + SKU）名称、商品描述等进行审核。
以京东印尼为例，使用外采系统Basedo审核，主要问题如下：
 - 1、审核费用高，审核按次计费，审核费用支出与平台商家数量、规模成正比。
印尼目前月均大约100万次商品审核，需支付外采系统费用 6K~7K 美元，年均6万~7万美元。
 - 2、审核周期长：商家提交商品信息到后台 → 系统审核 → 审核不通过 → 商家修改后再次提交。
整体审核周期比较长。

三、技术突破点

核心突破点（新方法、新思路）

将原本在后台的图片审核和文本审核环节，前置到移动端进行，使用移动端的开源AI SDK对商品图片、商品名称、商品描述进行违规识别。

1、图像和文本分类模型训练

使用图片和对应的分类标签投喂给模型，模型学会去识别这种图片对应的标签。
通过给予足够多的电商商品图片训练数据，模型就能够学习去预测新的图片是否属于训练数据中的分类。

2、图片和文本的分类识别

商品创建时，通过移动端AI SDK，一张图片被输入模型，模型会输出标签和代表识别概率的数组，如下图：



标签	概率
兔子	0.07
仓鼠	0.02
狗	0.91

通过获取数组概率最大的标签，如果概率值较大超过约定的阈值，如 ≥ 0.8 ，则认为识别成功。

四、业务贡献/商业价值

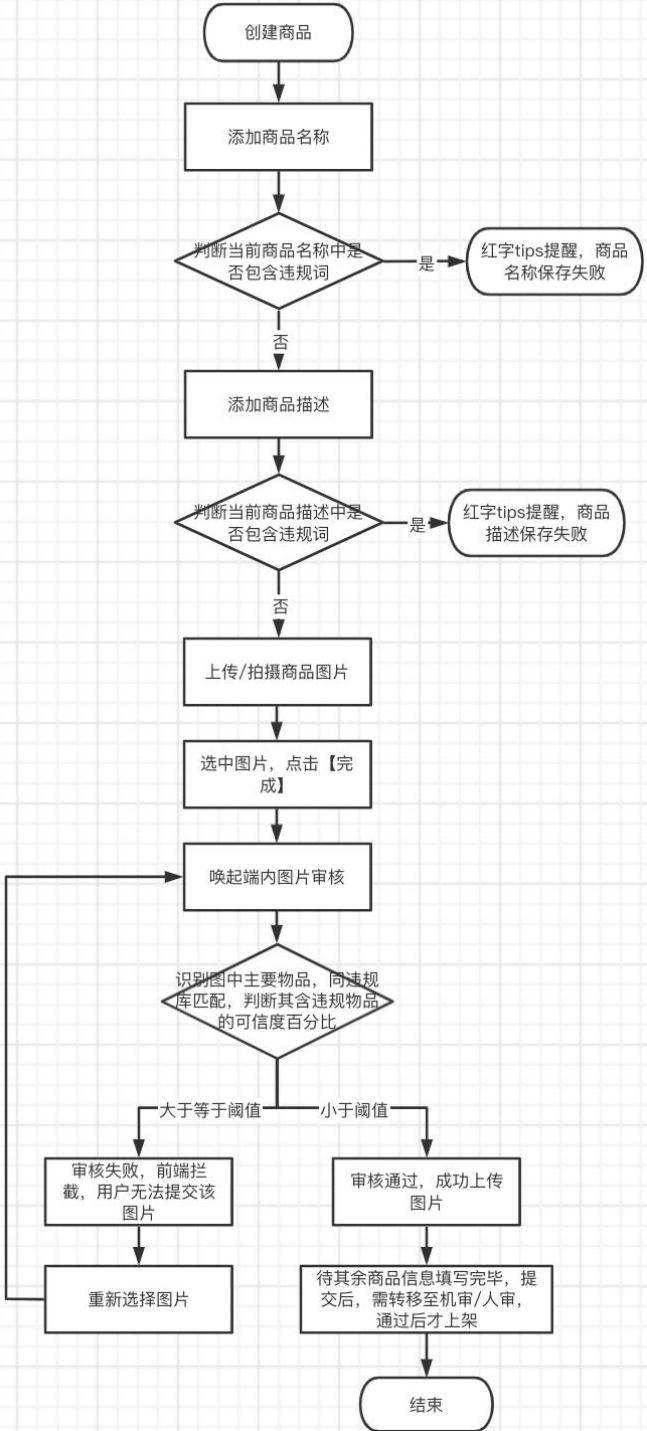
- **成本降低：**当前商品审核多是由外采审核系统实现，业务成本取决于商品提交次数（以印尼站为例，月均提交100万次商品审核，需支付外采系统 6K~7K美元）。在移动端使用AI 技术对商品图片、商品名称及描述进行违规识别，当识别结果高于一定阈值（如80%）时，移动端直接拦截；反之再将商品提交至后台或者人工审核，一定程度上减少审核商品数量，从而降低外采服务器成本及对应的审核人力投入。
- **效率提升：**使用移动端进行审核，将部分服务端审核压力前移到移动端，由之前的后台集中审核改为移动端分布式审核+后端审核的方式。
- **体验优化：**相比起当前用户须填写完成所有商品相关信息并提交后，快则几分钟，慢则几小时到两天才能够得知机审或者人审结果，将商品图片内容、名称及描述违规审核前置用户确认提交之前，即时反馈商品图片及描述初审结论，降低用户商品因含违规内容所导致的失败率，进一步提升用户端上创建商品的效率及体验。

五、方案详细介绍

(一) 产品方案

该小工具为移动端场景，通过AI技术，从文本分类和图片分类识别两方面在端内实现商品审核闭环。

- 用户在移动端创建商品，当在二级页面完成商品名称及商品描述的编辑并点击【保存】时，调用AI SDK，对文本框内的文字信息进行语义分析，并同对应的违规词库进行匹配：若产生命中违规词库的情况，前端toast提示因含敏感词保存失败，引导用户对当前文案进行修改，直至通过审核。
- 用户在上传图片的场景（例如上传商品图片，或商品描述中添加图片等），唤起本地图库或相机选中对应图片，并点击【完成】后，调用SDK，对选中的图片进行AI识别，将出参的字段及可信度与事先维护的违规库中的违规词进行匹配：若出参的违规词可信度 \geq 预设阈值，说明图片违规风险高，即审核失败，前端拦截，toast提示用户更换图片；反之端内审核通过，待其余商品信息填写完毕，提交后，转移至后台机审。



五、方案详细介绍

(二) 技术方案

(1) AI开源平台选型

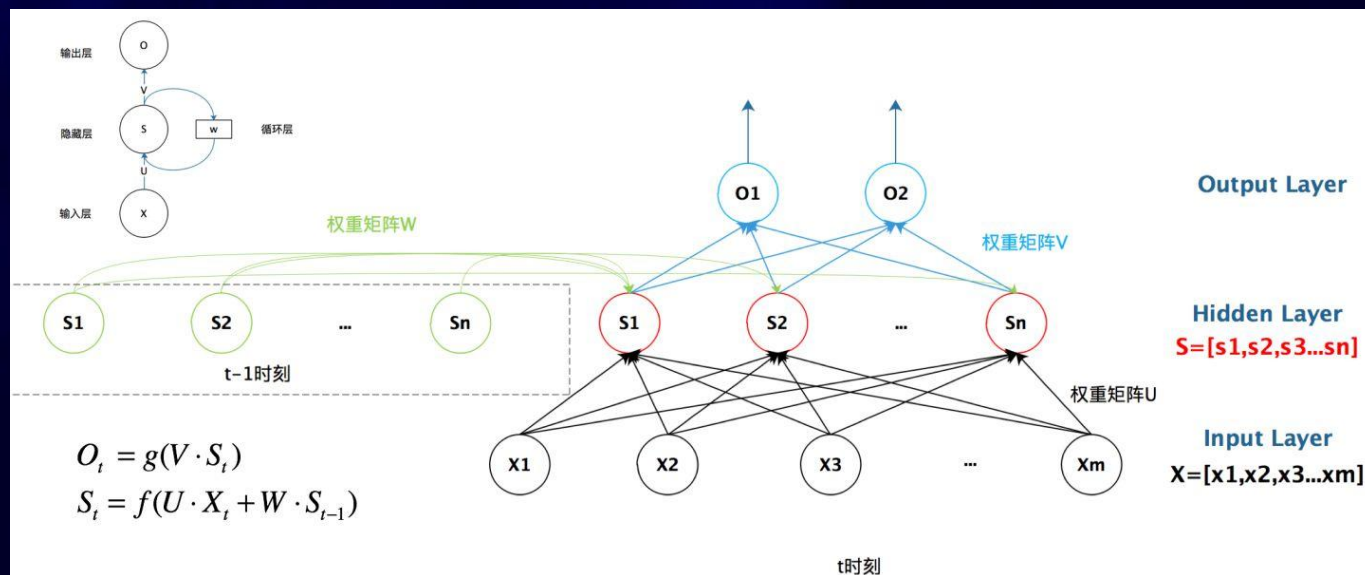
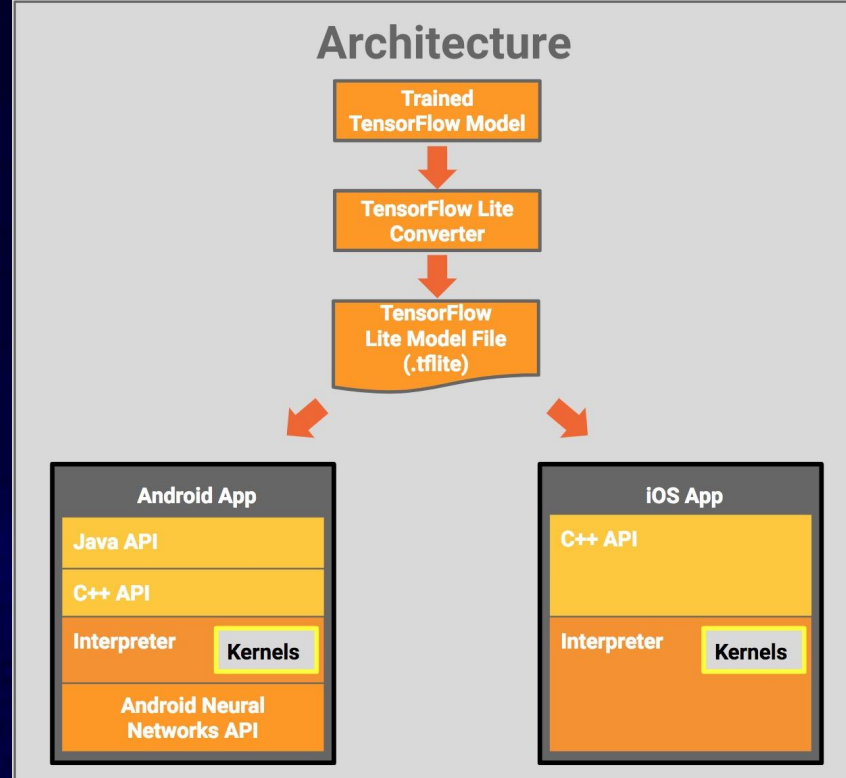
TensorFlow Lite: 移动和嵌入式设备轻量级解决方案

- 1、支持一系列量子数和浮点的核心运算符,并针对移动平台进行了优化。
- 2、基于 FlatBuffers定义了一个新的模型文件格式,代码提交更小,解析速度更快。
- 3、On-device 解释器,内核经过优化,可在移动设备上更快执行。
- 4、针对支持的设备提供了一个利用硬件加速的接口。
- 5、TensorFlow 转换器将 TensorFlow 训练好的模型转换为 TensorFlow Lite 格式 (.tflite)。

总结: 性能更优、体积更小

(2) 神经网络选型

考虑到文本语句分析时,孤立的理解这句话的每个词是不够的,我们需要处理这些词连接起来的整个序列,所以文本分类识别场景选用循环神经网络RNN。



五、方案详细介绍

(二) 技术方案

(1) 图片训练方向

A、合规训练

将平台三级类目下商品批量下载，选取一些数据进行训练，计划每个三级类目10~20个，每个一级类目100~200个，训练总数1000~2000

B、违规训练

与审核平台保持一致：

Middle finger 、 nazi、 nudity、 jd.id placeholder

(2) 文字训练方向

对违规的文字进行训练

<https://joyspace.jd.com/sheets/1A70P5xu51y4OokBvsHq>

四个维度的关键字：

Tobacco、Alcohol、Sensitive、Sex