

## 使用 BigQuery ML 构建电子商务推荐系统

## 使用 BigQuery ML 构建图片识别系统

### 1. 使用 BigQuery ML 构建电子商务推荐系统

XX电商平台项目包括用户管理，商家管理，物流跟踪，用户推荐系统等模块，我主要负责其中的用户推荐模块

#### 1.1 Situation（发生了什么情况）

为了针对用户个性化推荐商品，提高用户对推荐商品的购买率，需要分析用户对商品购买意愿，刻画用户画像，推荐相似画像的用户特定商品提高购买率。流程如下：

- 1.提取用户基本信息，提取用户对商品的浏览持续时长并划分等级
- 2.模型选择，模型预测，部署
- 3.整合到源项目

#### 1.2 Task（任务是什么？）--难点

- 需要提取何种信息进行模型训练？如何提取信息？
- 如何刻画用户画像？
- 整合项目到生产环境中 --可作为商城推荐下的一个子模块

#### 1.3 Action（提出了什么方案，如何解决）

- 1.针对用户对商品的购买意愿分析
  - 从数据库或消息队列取出数据--清洗取出关键信息  
userid, itemid, session\_duration
  - session\_duration可以抽象成用户购买意愿等级作为划分依据
  - 创建矩阵分解模型(Create Model) 将模型部署到生产环境
- 2.描绘用户画像&&推荐预热
  - 用户基本信息表和交易表区分不同年龄段，地区用户的购买意愿级别，提取出top20做同地区/同年龄段 预热推荐
  - 根据用户选择，可推荐数据库中预测购买意愿等级高的同类产品

## 1.4 Result (结果如何?)

- 数据仓库选择BigQuery ML, 训练模型利用数据仓库中create model 进行
- 成功解决以上问题

## 2. 使用 Flask,TensorFlow 构建图片识别系统

包括口罩识别提示模块, 垃圾分类模块

### 1.1 Situation (发生了什么情况)

为了提醒大家疫情期间进入公共场所积极佩戴口罩, 加强垃圾分类相关知识, 推出了有红外测温功能的口罩检测系统, 同时该系统支持用户上传图片进行垃圾分类辅助检测

### 1.2 Task (任务是什么?) --难点

- 口罩检测系统? 如何提取信息?
- 记录数据库员工考勤信息
- 整合多类型检测

### 1.3 Action (提出了什么方案, 如何解决)

- 1.不同提示功能, 体温超过阈值, 未佩戴口罩
- 2.考虑到检测仪性能, 需要模型尽可能轻巧
- 3.个性化需求, 引入垃圾分类检测接口 (网站)

## 1.4 Result (结果如何?)

- 人脸识别率98
- 人脸口罩识别95

## 2. 自动驾驶系统

自动驾驶系统中障碍物感知系统模型搭建, KPI统计

### 1.1 Situation (发生了什么情况)

自动驾驶系统中需要实现对障碍物的感知 (距离, 障碍物类型) 以及衡量障碍物感知系统的工作的准确率

- 当摄像头拍摄范围内有行人，车辆出现时可以正确识别，生成3D检测框，距离本车车头最近距离，以及碰撞时间
- 根据后台反馈的信号信息统计感知系统的准确率

## 1.2 Task (任务是什么?) --难点

- 如何对数据进行统一处理，两种不同格式数据如何对应
- 对行人和车辆作出区分，以便计算中心点，碰撞时间
- 在不同场景下障碍物感知系统不同于实验环境，如何根据信号值信息统计KPI

## 1.3 Action (提出了什么方案，如何解决)

- 提出基于3D处理模型+2D特征值提取的模型作为感知算法
- 2D图片不含有深度信息无法准确定位
- 3D点云数据不含有颜色信息且3D逐点检测处理信息量大
- 对人，车的检测框预处理设置不同大小，同时设置不同阈值
- 行人检测信息点少，更依赖图像信息，则2D特征提取权重更高
- 车辆检测 点云数据提供更充分并且误报少，所以车辆检测2D特处理流权重更高
- 源信号和经过系统处理后的信号进行对照并匹配目标物
- 通过对系统输出信号值清洗，选出计算KPI相关信号值
- 通过不同方法计算KPI，距离，IOU
- 逐帧匹配目标物时抽象成矩阵进行处理

## 1.4 Result (结果如何?)

- 多模态检测比纯3D检测快 5fp--16fp
- 准确率2D 76 ---81.04 -----
- KPI统计准确率79