

# AI+大数据，为场景重新赋能

张弛

Face++ 旷视

# 大数据的前世今生

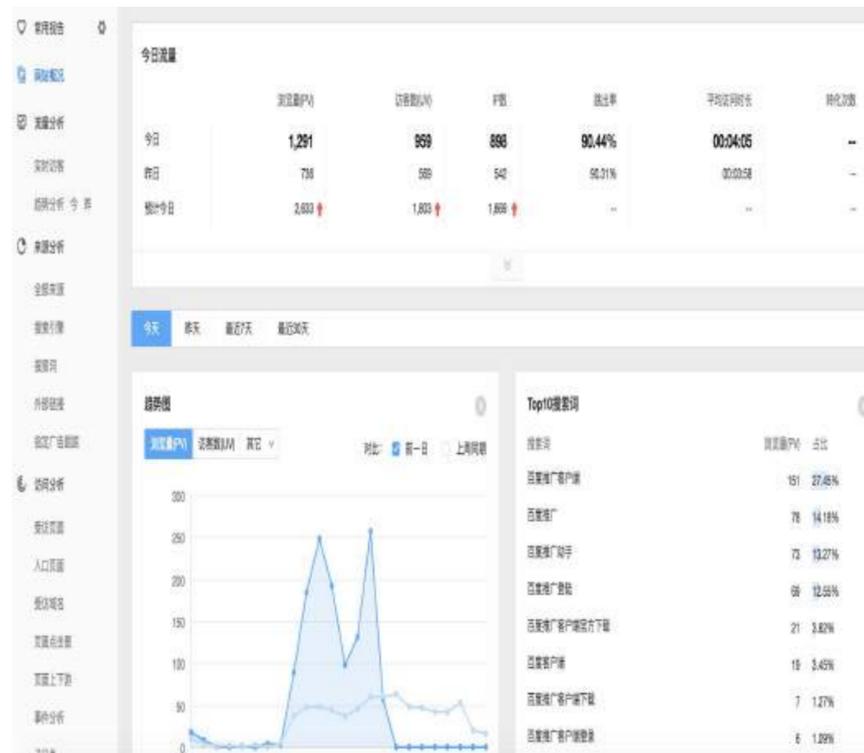


先秦对星象做相关分析



刘伯温对历史知识做机器学习

人肉机器学习，只有少数牛人能做到



网站分析时代



移动互联网时代

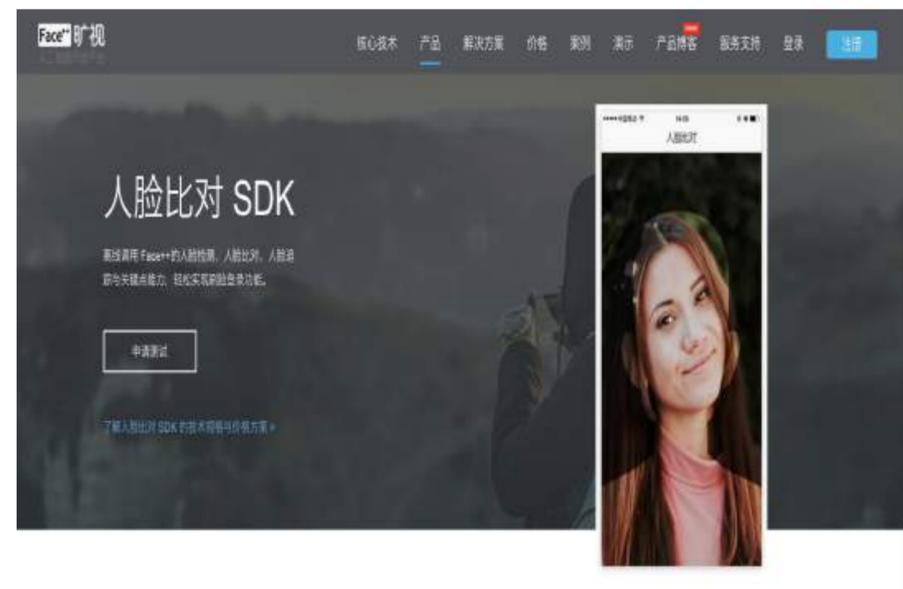


AI时代

数据应用的演变：统计、分析、触达、决策

AI使数据无处不在

# 传统的数据



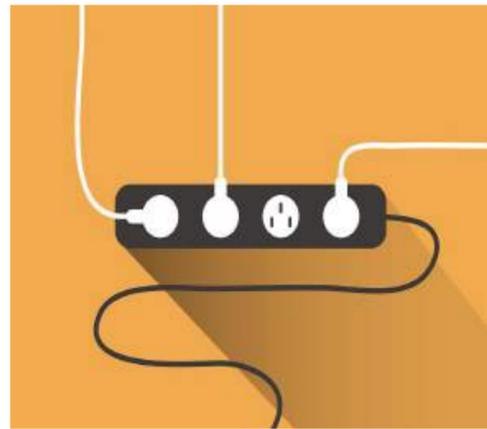
PC



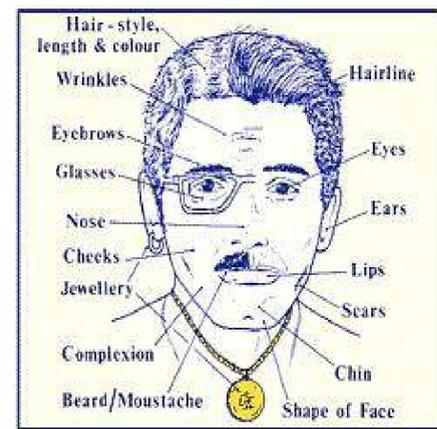
APP



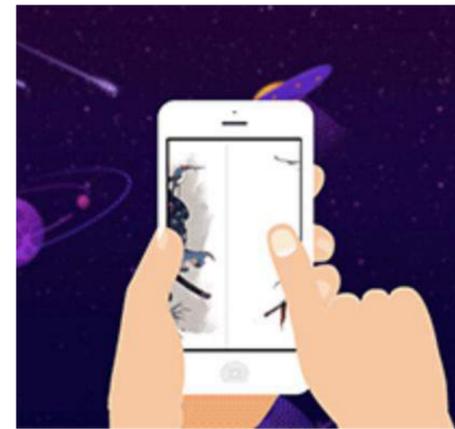
H5



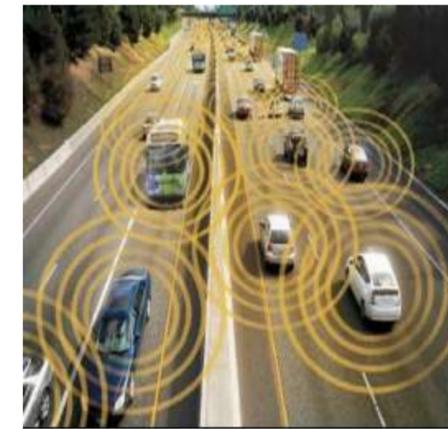
物联网



人体特征



生物探针



车联网



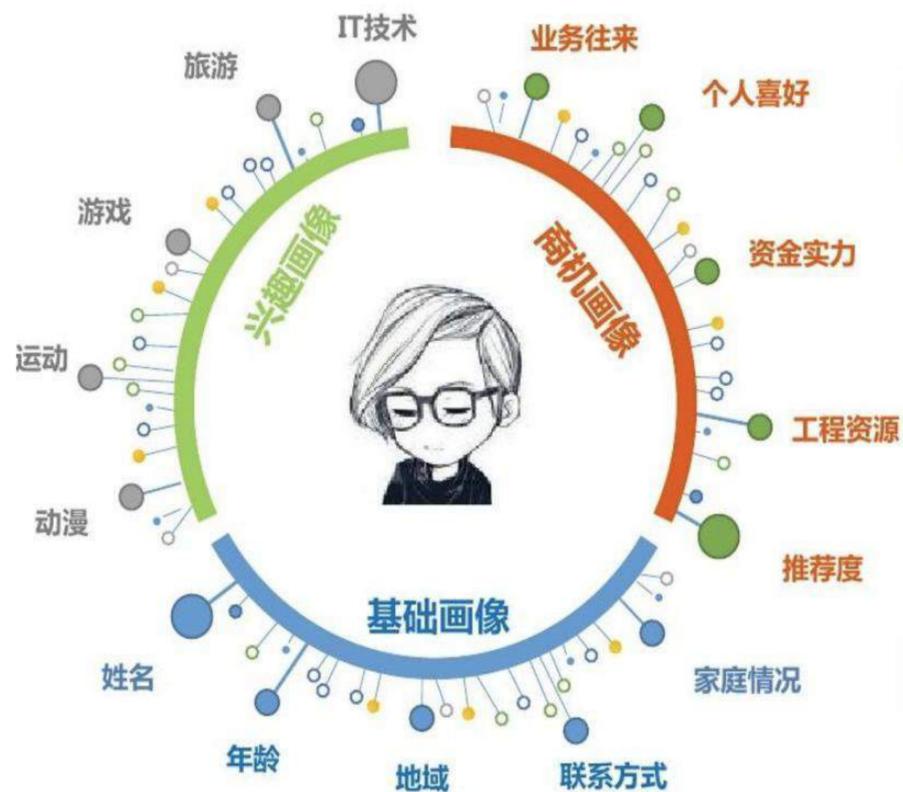
智慧城市

机器所见即为数据

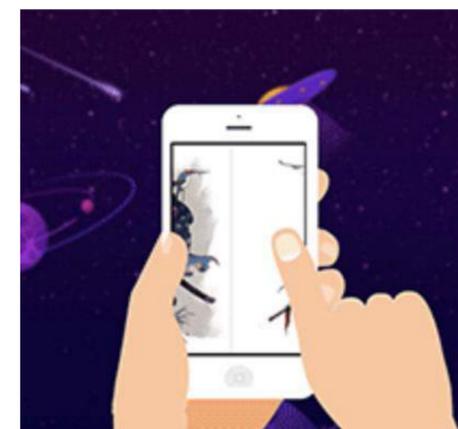
# AI+大数据的化学反应

AI + 大数据 + 电商 = 更舒适的体验

## #1 个性化推荐应用



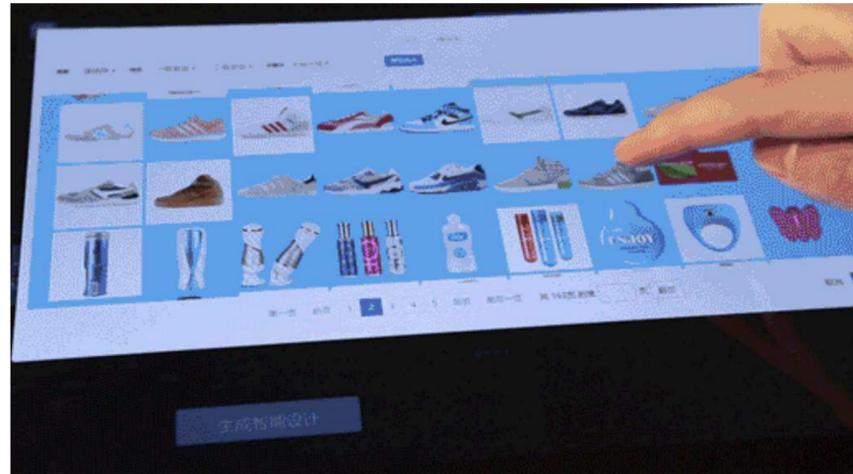
标签实时更新



在算法下，性别可以分为19种

AI + 大数据 + 职业 = 替代人力

## #2 替代人力应用



设计师危险了



编辑危险了



没女朋友没事了？

## #2 替代人力应用

当然，说“危险”有点危言耸听，我们一直致力于用机器替代简单重复的工作，让人有更多的创造时间

## 新闻资讯类数据分析师的痛苦



流量的增长，是产品升级的影响，还是渠道的影响，还是新闻热点的影响？

将分析师的思考逻辑，赋予机器，完成每天固定化的分析项目

AI + 大数据 + 补贴策略 = 营销成本降低

## #3 价值预测应用

普惠·新人专享-安心计划0085期

6.00%+4% 新手专享  
预期年化利率

30 期限(天) | 500 起投金额(元)

已出售44%，可售金额5505600.00元

起始日	到期日	到账日
2017-12-05	2018-01-04	到期日后1~3个工作日

新手专享 新手享10%年化 ?

收益方式 到期后一次性还本付息

产品介绍

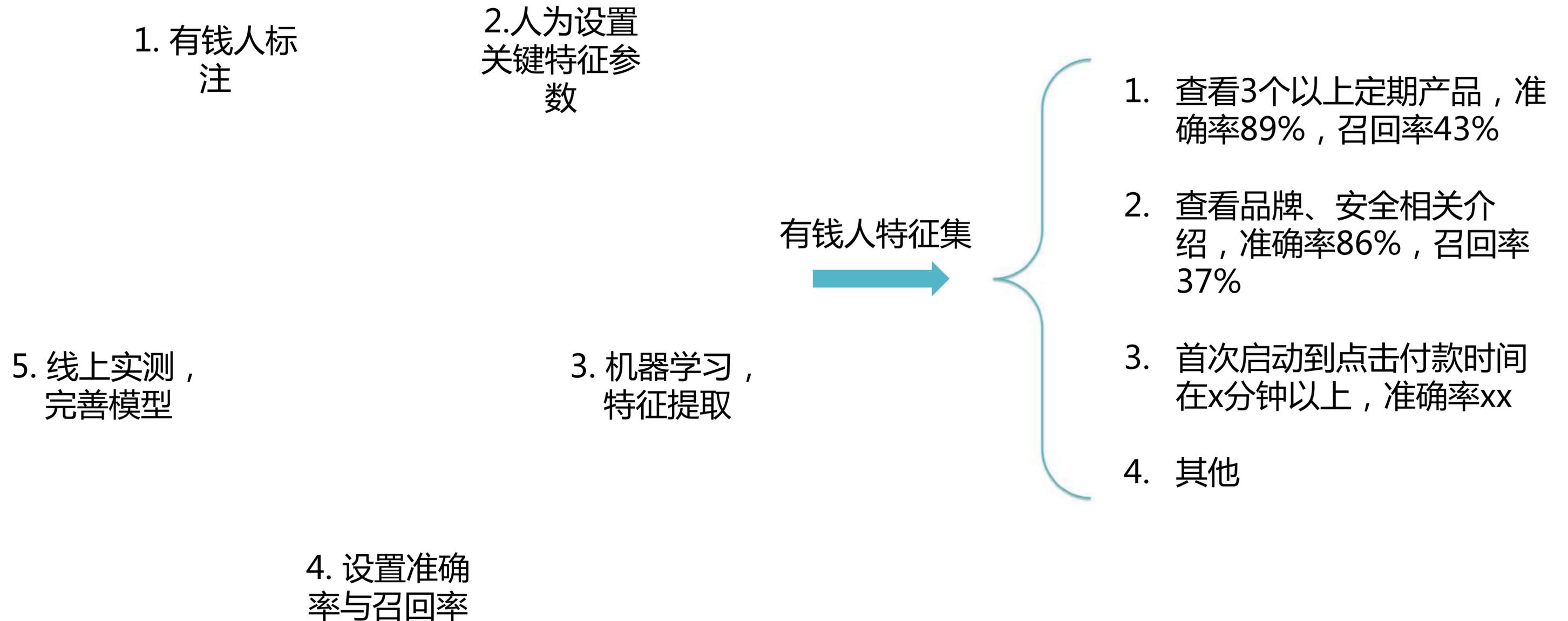
安心计划是本平台为用户提供的本息自动循环复投及到期

 立即加入



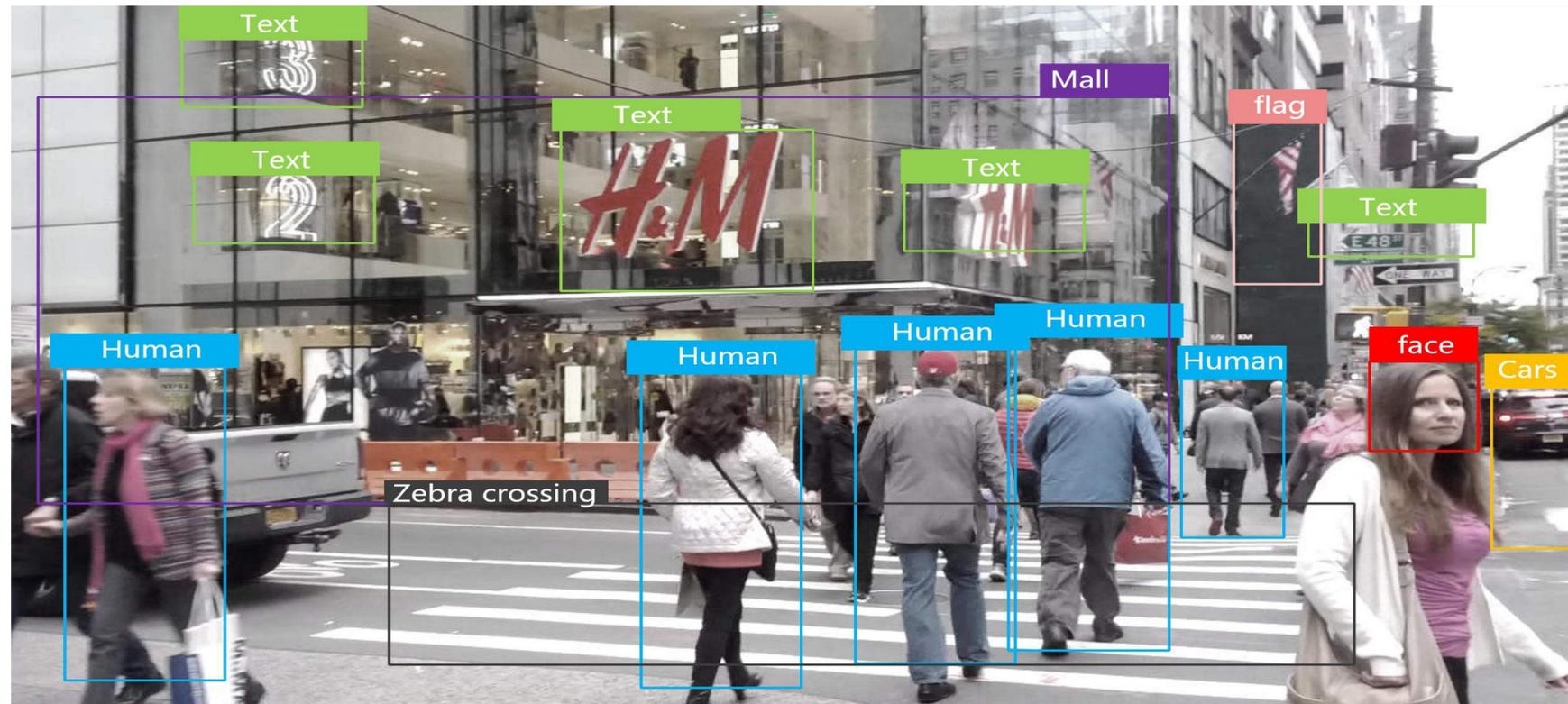
投资理财类产品往往需要利息补贴，  
有限的营销成本下，补贴该给谁？

## #3 价值预测应用



AI + 大数据 + 传统零售 = 精准营销

## #4 新零售应用



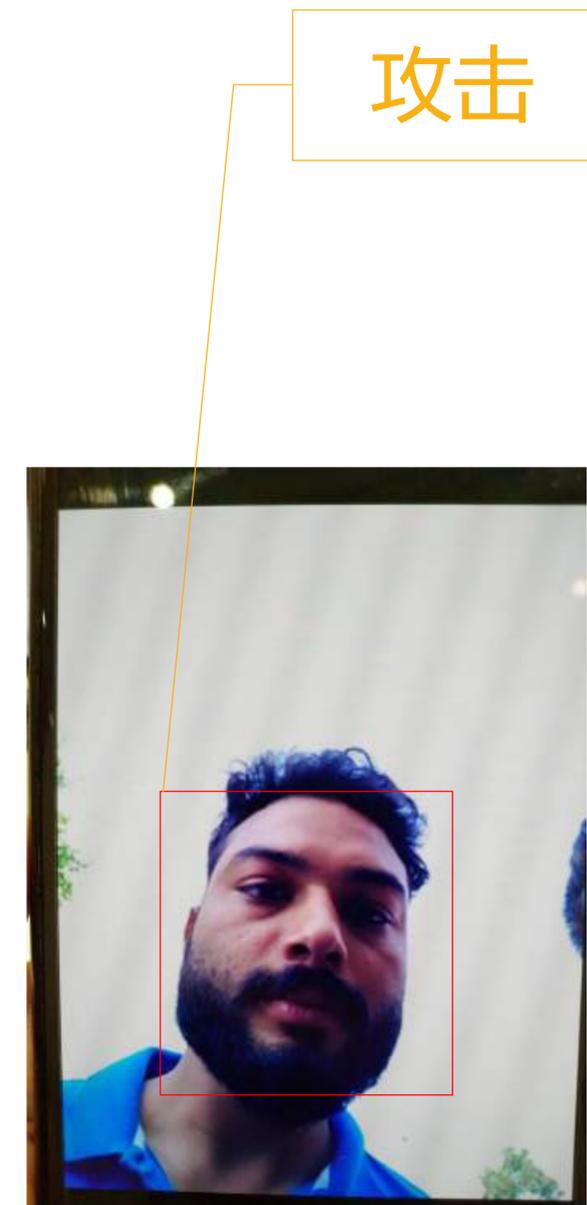
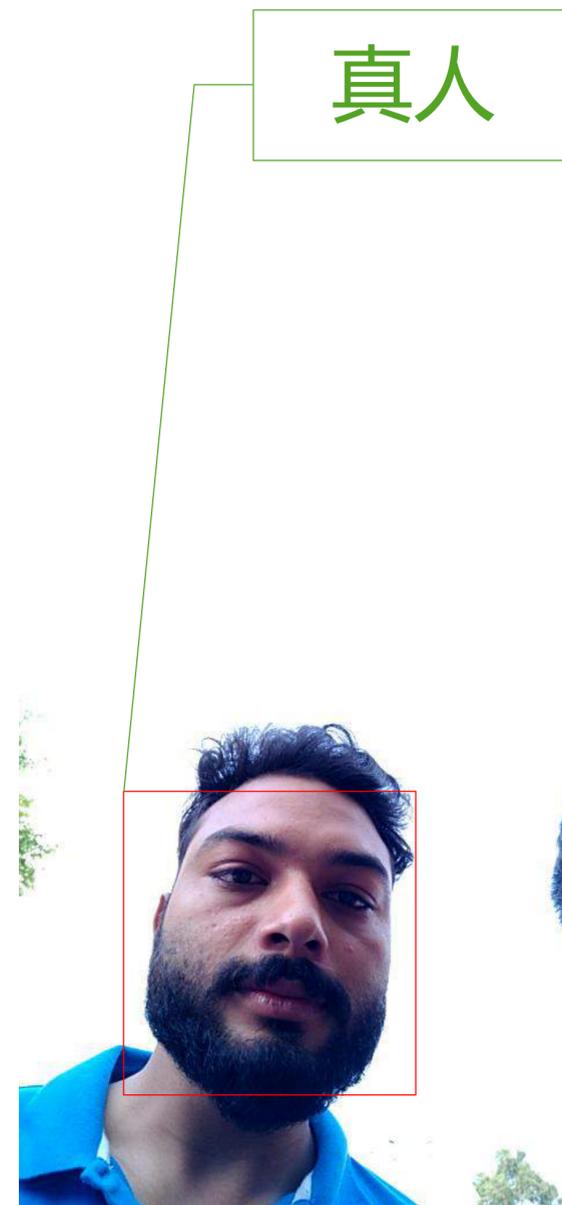
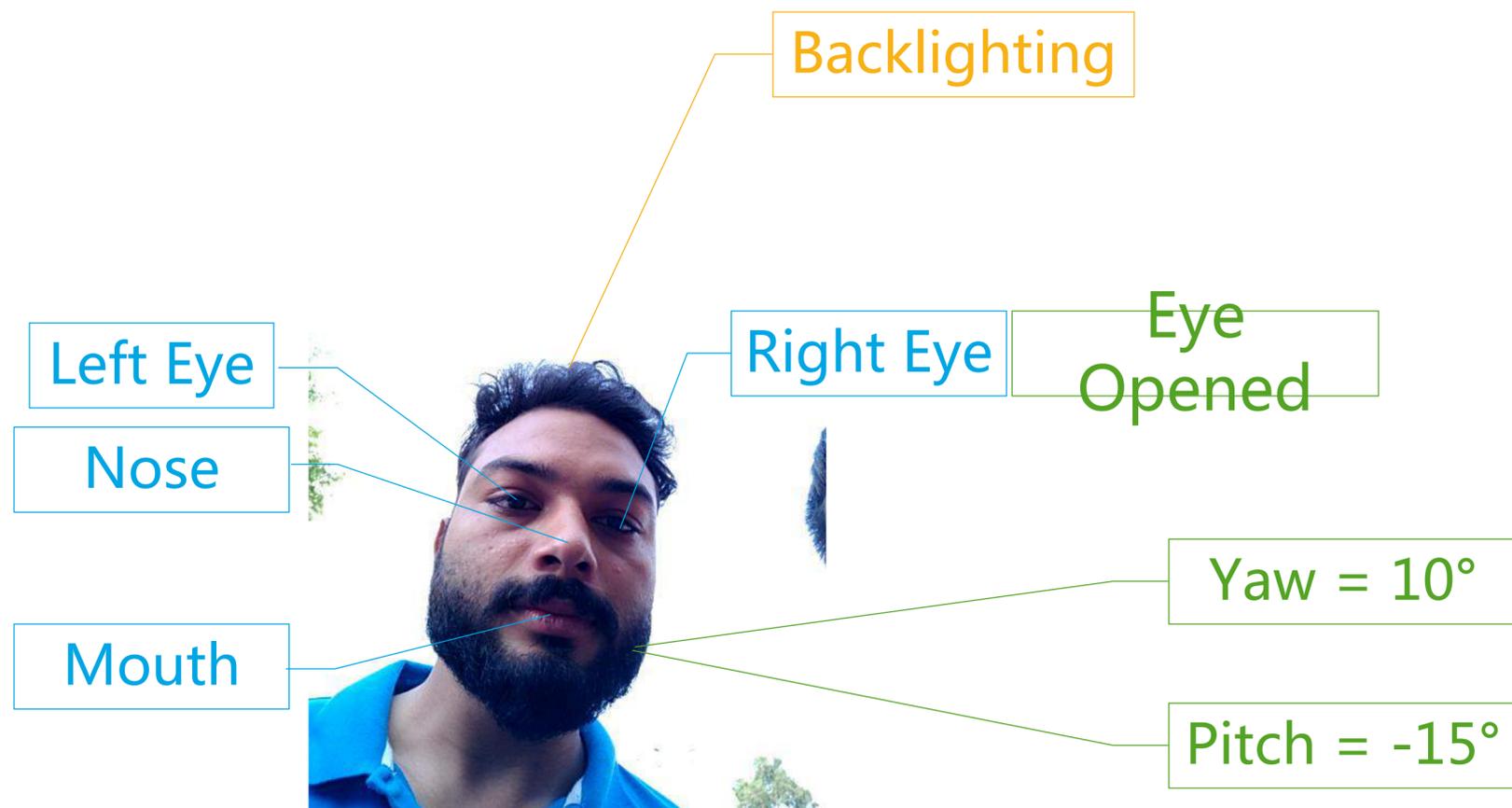
第二次来HM了  
喜欢看裙子  
平均每次逛20分钟  
平均每次试5件衣服  
喜欢与店员交流  
导购知道怎么做了吧！

AI + 大数据 + 金融 = 风控安全

## #5 金融安全的应用



## #5 金融安全的应用



AI + 大数据 + 交通 = 道路安全

## #6 道路安全应用



1. 司机风险系数（疲劳驾驶，行车习惯）
2. 时速控制与道路规划（时速与事故的关系，道路规划规避拥堵）
3. 智慧交通建设（公交建设、智慧红绿灯）

AI与大数据的融合赋能，同样存在着挑战！

## #1 你的数据准确么？

## #1 你的数据准确么？

无埋点  
人工埋点方式



行业平均10-15%的  
数据采集错误率



前两大错误原因：

1. 埋点方案设计问题
2. 事件触发机制设定的不正确

## #1 如何保障数据的准确性

事件组	区域	自定义事件	参考事件id格式	map参数 (涉及到的key)	value
首屏活动组	首屏广告区	点击首屏广告, 进入广告详情	ad_ac1		
	弹窗活动区	首页中出现活动弹窗	ad_ac2		
		首页弹窗中点击进入活动详情 (例如点击拆)	ad_ac3	标题	如新手红包、五月活动等
		首页弹窗中点击“我已领取”按钮或直接关闭弹窗时发送	ad_ac4		
新手体验组	新手弹窗区	新手活动拆红包后点击“马上投资”按钮	new_pop_ac1		
		新手体验标中点击“马上购买”, 提示投资成功时发送	new_pop_ac2		
		新手活动中点击“完善信息, 领取收益”按钮	new_pop_ac3		
	新手福利区	新用户注册成功后出现领取成功提示	new_wf_ac1		
		领取成功后点击“立即查看”按钮	new_wf_ac2		
		新手福利中点击按钮区域某一按钮	new_wf_ac3	标题	
		新手福利518红包中, 点击“查看红包”按钮	new_wf_ac4		
		新手福利518红包中, 点击“马上投资”按钮	new_wf_ac5		
		新手福利中, 点击下方常见问题	new_wf_ac6	标题	
		新手福利中, 点击下方“查看更多问题”	new_wf_ac7		
		新手福利中点击专享复投宝某一产品	new_wf_ac8	类型 标题 利率 起投金额 期限 剩余金额	即将开始、购买、售罄等 产品id
	个人账户升级, 信息完善路径	点击产品时出现“个人账户升级”提示	new_it_ac1		
		在升级提示中点击“立即升级”按钮	new_it_ac2		
		在开通存管账户页面中点击“前往存管页面绑定”按钮	new_it_ac3		
在开通存管账户页面点击“获取验证码”按钮		new_it_ac4			
在开通存管账户页面点击“同意协议确定注册”按钮		new_it_ac5			
提示“存管账户开通成功”时发送		new_it_ac6			
点击投资时出现“完善账户信息”提示		new_it_ac7			
在账户信息中点击“马上完善”按钮		new_it_ac8			
完成风险评估, 点击确定按钮, 完善成功时发送		new_it_ac9			
登录注册组	登录注册流程	老用户输入账号密码, 点击“登录”按钮, 登录成功发送	login_ac1		
		新用户输入手机号密码, 点击“注册”按钮, 注册成功发送	login_ac2	邀请人标识	是否填写了邀请人
		新用户注册时点击“邀请人手机号码”按钮	login_ac3		
	首屏焦点图	点击首屏焦点图区域	sy_banner	标题	如: 银行存管正式上线

我们需要花费更多的时间投入到埋点规划, 且必须要有数据埋点测试的流程

## #2 机器学习对象还是IMEI么？

## #2 数据的关联融合



服务端



pc



app



H5

用户唯一id = ( IMEI , cookieid , 服务端用户账号id , 多账号体系 )

我们需要时刻准备好设备id的找回

## #3 市场多种多样的运营手段会产生巨大噪声

## #3 运营手段带来的机器学习噪声



有钱人和学生同时逛奢侈品店，逛街路线肯定不一样，我们是否可以按照一样的逻辑预测APP上用户的价值呢？



## #3 运营手段带来的机器学习噪声

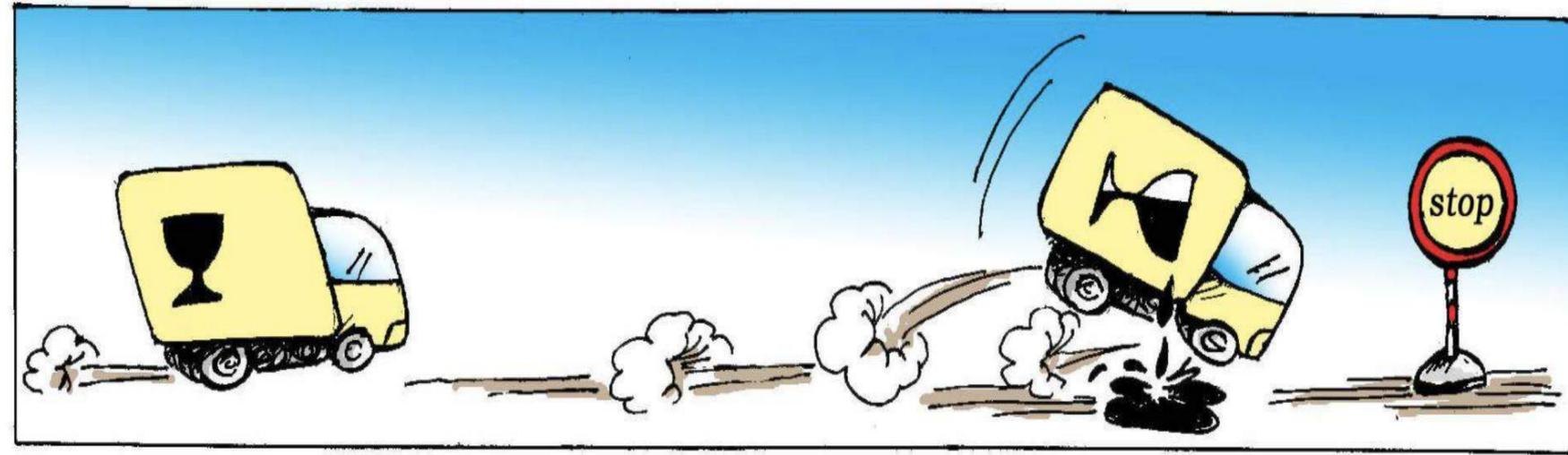


**NO !**

运营手段会干预正常的用户行为，让行为变得具有普遍规律性，对机器学习造成大量噪声和误差

## #4 AI与大数据的融合，依然需要大量的人力标注

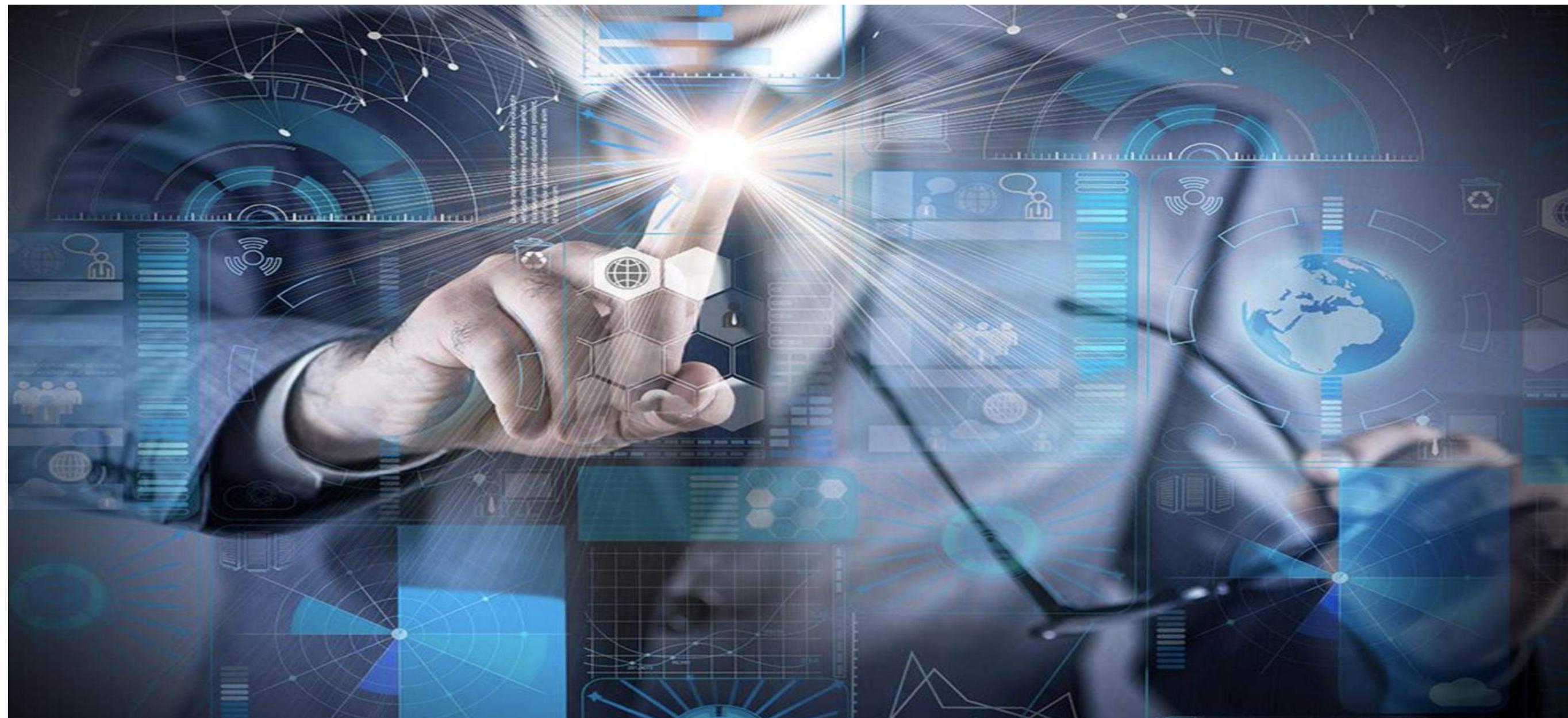
## #4 AI与大数据的融合仍然需要人力标注



急刹车

交通安全的AI大数据应用上，当我们把急刹车纳入风险因子的同时，需要我们来定义，什么是急刹车？

# AI与大数据的展望





Thanks