

# 用户画像

赵海臣

# 用户画像概述

## 🌀 用户画像概述：

- ★ 用户画像是现实用户的原型抽象技术，根据用户的行为偏好和对事物认知的差异，把用户划分为不同的用户类别,之后从各个类别中提取出典型属性特征，将具体的描述语言或标签赋予姓名、图片和其他自然属性中，这样就构成了一个人物原型。

## 🌀 用户画像作用：

- ★ 精确营销
- ★ 用户分析
- ★ 标签化信息

# 动态画像与静态画像

🌀静态画像：用户固有属性，基本上不会变，或者变化很小。只进行插入操作，或者更新频率很低。

★ 主要来源于用户所填的个人信息，注册信息，登录信息等，或者模型算出来的标签。

- 人口属性
- 商品消费属性
- 如果有空值，可以建立模型来预测概率

🌀动态画像：用户行为产生的数据，会随时间变化。经常需要更新的动态画像表。

★ 主要来源于用户的注册、浏览、点击、购买、签收、评论等在线信息，以及基于行为信息使用算法猜测的标签。

# 静态画像——用户基本属性建模1

🌀 用户基本属性是根据用户注册信息，登陆ip，身份证信息等所推测出来的标签信息。

★ 数据来源：用户信息表，用户登陆信息表

用户基本属性建模——标签1	
标签	反应的问题
用户id	
生日	年龄段偏好
省份	地域偏好习惯
城市等级	大中小城市消费偏好
手机品牌	不同手机品牌的消费偏好
手机型号	
手机操作系统(ios/android)	
注册时间	新/老用户
邮箱运营商	
手机运营商	

# 静态画像——用户基本属性建模2

可以通过消费券奖励，在线用户调查问卷可以获得用户如下信息，在达到足够的数据量后训练分类器模型进行更大范围的预测：

## 用户基本属性建模——标签2

标签	反应的问题
婚姻状况	单身、已婚的偏好
月收入	收入阶层的偏好
是否有车	
是否有婴儿	母婴用户与否
学历	不同学历的偏好
职业	不同职业的偏好

# 用户性别模型

🌀 用户性别是很重要的用户信息，男女购买商品的差异显著，因此是重要的画像信息，用户自己虽然填了性别，仍需要用算法进行确认。

用户性别模型		
标签	值	信息
用户性别	男	购买的商品，通过机器学习分类算法来训练，并应用于识别性别
	女	
	未识别	
婴儿性别	仅有男孩	购买的商品，通过机器学习分类算法来训练，并应用于识别性别
	仅有女孩	
	男女都有	
	无法识别	

# 用户性别模型的后验方法

## 🌀 用户性别模型的后验方法

- ★ 随机抽取几千条用户信息，让客服进行电话确认
- ★ 与用户自己所填写的信息进行对比，确认准确率

# 动态画像——订单数据建模1

根据用户订单消费情况提取用户标签，主要用于了解用户的总体消费情况，以及根据消费习惯与消费能力作定向营销。

★ 数据来源：订单数据、购物车数据、退货数据、用户信息。

订单数据建模——标签列表1

标签	反应的问题
第一次消费时间	什么时候来购物，多久没购物了？
最近一次消费时间	
首单距今时间	
尾单距今时间	
近30天购买次数(不含拒退)	近期的消费能力如何？
近30天购买金额(不含拒退)	
近30天购买次数(含拒退)	
近30天购买金额(含拒退)	



# 动态画像——订单数据建模2

订单数据建模——标签列表2

标签	反应的问题
常用收货地址	定向营销，常用消费属性
支付方式	
最近30天光顾购物车次数	用户的消费情绪，冲动型or理性消费
最近30天购物车商品件数	
最小消费金额	客户总体消费情况怎样？
累计消费金额(不含拒退)	
最大消费金额	
累计消费次数	
累计使用代金券金额	

# 动态画像——购买类目建模

根据用户购买1、2、3级类目提取用户标签，用于了解类目的购买人群和针对某一类目的营销。

★ 数据来源：订单数据、购物车数据、商品类目数据

购买类目建模	
标签	反应的问题
近30天购买1、2、3级类目的次数	用户近期购买的类目
近30天购买1、2、3级类目的金额	
近90天购买1、2、3级类目的次数	用户多久没购买一个类目
近90天购买1、2、3级类目的金额	
近30天购物车1、2、3级类目的次数	用户近期感兴趣、有购买倾向的类目
近30天购物车1、2、3级类目的金额	
近90天购物车1、2、3级类目的次数	用户购物选择比较过程，推断其生活阶段定位
近90天购物车1、2、3级类目的金额	

# 用户忠诚度模型

🌀 忠诚度高的用户越多，对网站发展有直接关系。

用户忠诚度模型		
标签	值	判断依据
用户忠诚度	忠诚型用户	规则+分类算法： 1. 浏览型用户：只有浏览数据 2. 购买天数大于一定天数为忠诚用户 3. 购买天数小于一定天数都是有优惠才买的 4. 其它类型：通过购买天数，最后一次购买时间，购买金额进行分类
	偶尔型用户	
	投资型用户	
	浏览型用户	
	未识别	

# 用户类目偏好模型

通过现有的用户专场标签推荐模型，每个用户提取最高分数的10个标签、品牌作为用户的类目标签。

# 用户族群标签模型

根据用户在不同类目(维度)的消费金额作为不同维度的值，再根据用户在不同维度的值进行聚类，获得用户偏好族群，若有需要，进一步对族群进行打标签，比如按照所有族群用户的购物类目比例进行标签。

用户类目模型		
标签	值	计算方法
用户族群标签	用户族群1、 用户族群2、 用户族群3、 用户族群4、 用户族群5、 ...	根据用户在不同类目(维度)的消费金额作为不同维度的值，再根据用户在不同维度的值进行聚类，获得不同的簇。

# 用户活跃RFM模型

用户活跃模型		
标签	值	计算方法
R(Recency)	最近一次消费时间	
F(Frequency)	近180天消费频率	
M(Monetary)	近180天消费金额	
用户价值	$a \cdot R + b \cdot F + c \cdot M$	基于R、F、M值对用户进行聚类，对每一个簇取平均值RFM，用户价值= $a \cdot R + b \cdot F + c \cdot M$
活跃状态	注册未购买 高频 中频 低频 沉睡 流失	注册未购买: 多是第三方登陆 高频、中频、低频: 最近60天有订单的日期数 沉睡: 近90天有购买，近60天无购买 流失: 近90天无购买，曾经有购买

**THE END**

**THANK YOU!**