

——萝卜网数据分析培训系列之

用户画像数据建模方法

主讲人：萝卜网特聘讲师

大纲

一、什么是用户画像？

二、为什么需要用户画像

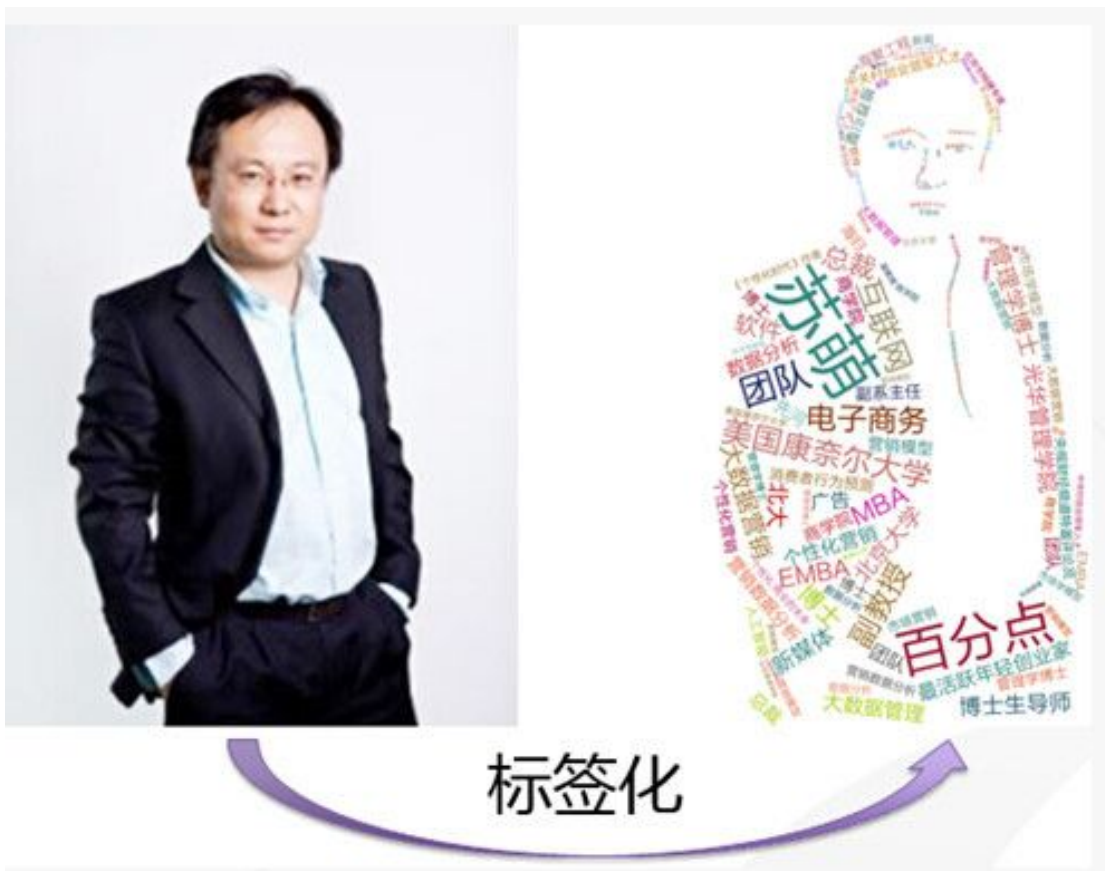
三、如何构建用户画像

四、总结

一、什么是用户画像？

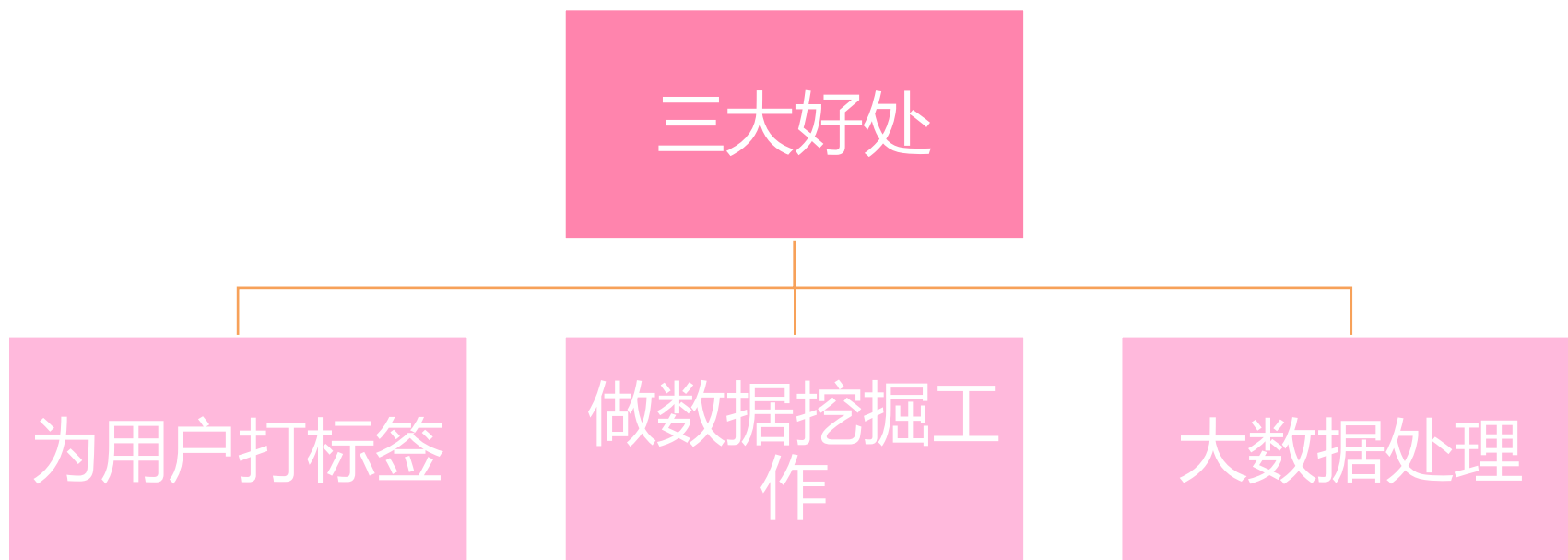
一、什么是用户画像？

用户信息标签化



二、为什么需要用户画像

二、为什么需要用户画像

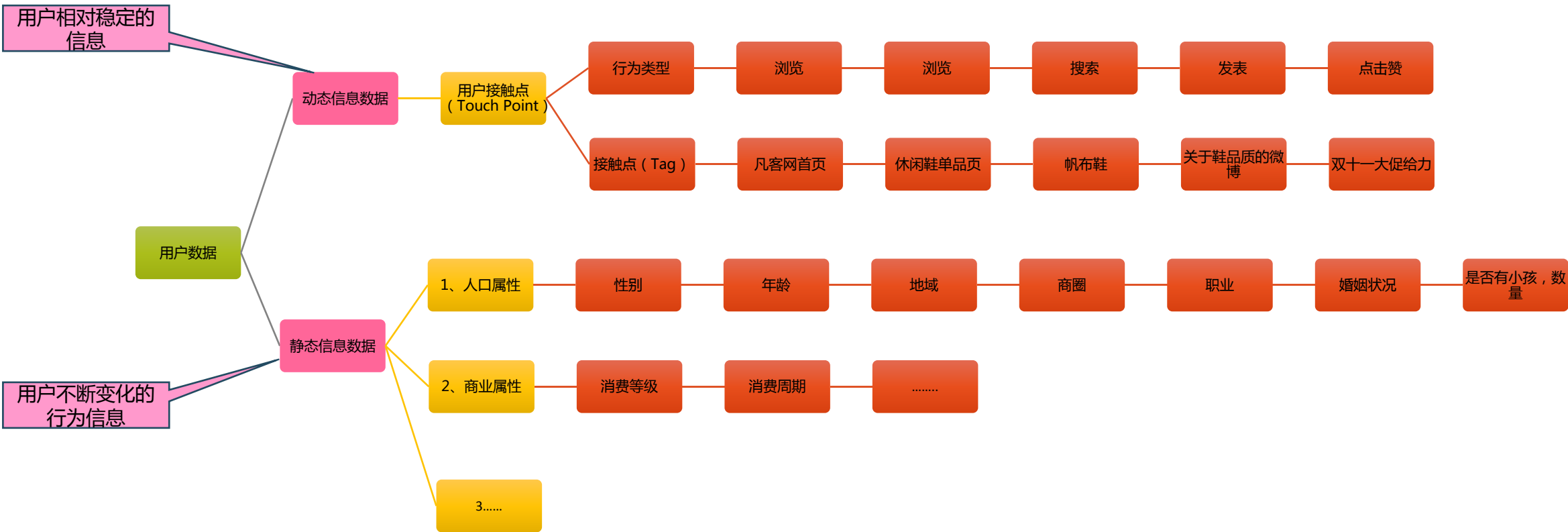


三、如何构建用户画像

三、如何构建用户画像

3.1 数据源分析

本文将用户数据划分为静态信息数据、动态信息数据两大类。



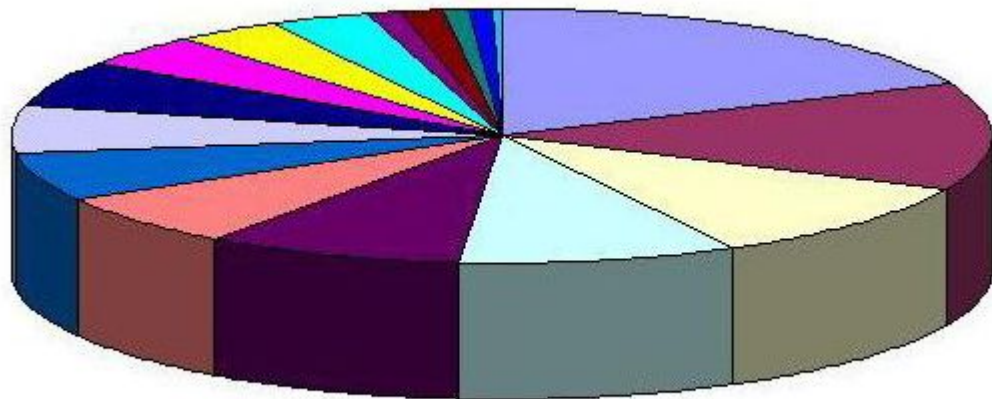
三、如何构建用户画像

3.2 目标分析

用户画像的目标是通过分析用户行为，最终为每个用户打上标签，以及该标签的权重。

标签：表征了内容，用户对该内容有兴趣、偏好、需求等等。

权重：表征了指数，用户的兴趣、偏好指数，也可能表征用户的需求度，可以简单的理解为可信度，概率。



三、如何构建用户画像

3.3 数据建模方法

用户标识方式	效果	备注（局限性）
Cookie	互联网使用最为广泛的方式，能够标识匿名、未注册用户。	通常有一定的有效期，不易跨浏览器、设备。
注册ID	各家网站的用户标识，最常见的互联网会员管理方式。	用户注册意愿越来越低，需要投入大量推广运营成本。
Email	互联网早期较为常用的用户标识方式。目前依然有一定的占有率。	一人有多个email很常见。因此标识会损失些准确性。
微博、微信、QQ	当下业内共识的第三方登录ID，提供OAuth授权机制	标识准确性，持久性上是个较好的折中方案。
手机号	移动端最精准的标识	较难获取到，视产品激励用户填写意愿。
身份证	最官方的标识	难获取到，视产品激励用户填写意愿。

以上列举了互联网主要的用户标识方法，获取方式由易到难。视企业的用户粘性，可以获取的标识信息有所差异。

三、如何构建用户画像

3.3 数据建模方法

什么时间：时间包括两个重要信息，时间戳+时间长度。

什么地点：用户接触点，Touch Point。

内容：每个url网址（页面/屏幕）中的内容。

什么事：用户行为类型

标签权重=衰减因子×行为权重×网址子权重

四、总结

四、总结

本文并未涉及具体算法，更多的是阐述了一种分析思想，在计划构建用户画像时，能够给您提供一个系统性、框架性的思维指导。

核心在于对用户接触点的理解，接触点内容直接决定了标签信息。



THANKS FOR WATCHING

进入萝卜官网观看更多牛逼课程

<http://www.luobo360.com/>