121 | 如何利用机器学习技术来检测广告欺诈?

2018-07-11 洪亮劼

AI技术内参 进入课程 >

讲述：初明明
时长 06:21 大小 2.91M

在上一期的内容中，我们聊了如何帮助广告商扩大受众群这个话题，也就是受众扩展技术。受众扩展的目的是让广告商投放的广告能够接触到更广泛的受众，甚至有可能提高广告效果。

在计算广告高级话题的最后一篇分享，同时也是整个广告模块的最后一篇分享里，我想来聊一聊广告中一个非常棘手，同时也是一个非常实际的问题：欺诈检测（Fraud Detection）。

什么是广告欺诈

广告欺诈是一个多大规模的问题呢?

根据一个统计数字 [1]，到 2015 年的时候，就因为广告欺诈，全美的市场营销和媒体业每年的耗费约为 82 亿美元。这个数字中大约有 56%，也就是 46 亿多美元的耗费来自于“**非法流量**”（Invalid Traffic）。我们把这个数字和全美每年 596 亿的广告支出进行对比，就可以看出这是一个惊人的数字。当然，因为各种欺诈手段层出不穷，并不是所有的欺诈都能够被甄别出来。因此，我们其实有理由相信真实的数字会更高。

那么，怎么来定义广告欺诈呢？什么样的行为算是广告欺诈呢？

我们这里主要讨论三种形式的广告欺诈。这三种广告欺诈模式其实对应着三种流行的广告计费模式。

第一种欺诈叫“**展示欺诈**”（Impression Fraud），也就是造假者产生虚假的竞价信息，然后把这些竞价展示放到广告交易平台上去贩卖，并且在广告商购买了这些展示后获利。

第二种欺诈叫“**点击欺诈**”（Click Fraud），也就是造假者在广告商产生虚假的点击行为。

第三种欺诈叫“**转化欺诈**”（Conversion Fraud），也就是造假者完成某种虚假的动作，例如填写表格，下载某个应用等来虚拟真实的转化事件。

在真实的场景中，这三种欺诈手段经常混合出现。例如点击欺诈和展示欺诈可能同时出现，这样就能在报表中展示一个看似合理的点击率。

广告欺诈的产生源

了解了广告欺诈的基本形式之后，我们来看一下这些欺诈产生的源头都在什么地方。因为广告产业的有利可图，产生欺诈的途径也是多种多样的，我们这里就看一些经典的形式。

首先，有一种欺诈来源途径叫**PPV**（Pay-Per-View）网络。

利用 PPV 进行欺诈的主要流程就是尝试通过购买流量，然后在一些合法的展示机会中插入用户肉眼看不见的 0 像素的标签（Tag），诱导广告商，让广告商以为产生了更多的合法流量。

对于这样的欺诈，一般来说，广告商必须去检测展示机会用户是不是看不见，或者是否是由 0 像素产生的。然后还可以采用黑名单的方式，对屡次利用 PPV 来进行欺诈的 IP 地址进行

屏蔽。

另外一种欺诈手段是通过“**僵尸网络**”（Botnets）。

这种方法主要是试图直接控制用户的终端电脑或者其他的移动设备，从而进行很多方面的攻击。在过去，僵尸网络的一大应用主要是来产生拒接服务的 DDoS（Distributed Denial of Service）攻击和发送垃圾信息。

近年来，因为其灵活性，很多僵尸网络也被用于广告欺诈。僵尸网络的一大作用就是产生浏览信息，而这些浏览的行为是宿主电脑的用户所无法得知的。因此，对付僵尸网络的一大方法，就是检测从某些 IP 地址或者 DNS 产生的流量行为是否发生了突然的根本性的变化。

第三类欺诈手段是“**竞者攻击**”（Competitor Attack）。

正常的广告商设立预算参与竞价购买广告位。而竞争对手可以利用“点击欺诈”的方式产生虚假无效的点击信息，从而消耗广告商的预算。当把竞争对手的预算消耗光以后，攻击者反而可以用比较小的成本拿到这些广告位，因为竞争减少了。

另外，还有一种情况是仅仅大量调入竞争对手的广告而不点击。在这样的情况下，就容易产生非常低的点击率。而很多广告平台依赖点击率来进行排序，因此，如果点击率很低，那代价就是难以赢得竞价，通过这种方式也就间接打压了竞争对手。

欺诈检测

了解了什么是广告欺诈以及不同的广告欺诈来源之后，我们来看一看如何利用机器学习技术，来对各种不同的欺诈行为进行检测和挖掘。

首先介绍一个研究 [2]，作者们提出了一种技术，**利用“同访问”图来分析异常的浏览行为**。这里面有一个最基本的假设：对于大多数用户来说，对两个不同的网站并不具有相同的喜好程度，除非这些网站非常流行。也就是说，对于绝大多数的网站来说，其用户群体是不一样的。

如果用户和这些网站的相互关系发生了变化，那可能就是出现了一些异常的情况。当然，利用图分析的方法，就是把异常发掘当成了一种无监督学习的任务，自然也就会有无标签的困难。

还有一个研究 [3]，作者们提出了一种方法，来**分析用户到底需要花多少时间来浏览显示的像素**。这个方法其实就是来检测是否是 0 像素的展示欺诈。作者们通过研究发现，对于 50% 以上的像素，绝大多数用户至少需要 1~3 秒时间来观看。于是，广告商或者平台就可以用这种停留时间来作为一个最基本的检测手段。

当然，一种最普遍的做法就是把广告欺诈当做一个**监督学习任务**。通过产生各种格样的特性以及把过去已知的欺诈数据当做训练数据来进行学习。这种做法的难点是，欺诈数据在真实世界中毕竟是少数。于是，我们就有了数据不足以及需要训练和不平衡的分类问题。正是因为存在这些问题，欺诈检测依然是一个非常前沿的研究领域。

总结

今天我为你介绍了在线计算广告的最后一个高级话题：欺诈检测。

一起来回顾下要点：第一，我们讲了三种形式的广告欺诈，分别是展示欺诈、点击欺诈和转化欺诈，在真实场景中，这三种欺诈手段经常混合出现；第二，产生欺诈的源头很多，我们简单介绍了三种不同类型的广告欺诈来源，分别是 PPV 网络、僵尸网络和“竞者攻击；第三，我们讨论了欺诈检测的一些基本思路，比如利用图分析、利用停留时间的方法等等。

最后，给你留一个思考题，如何来检测转化欺诈，也就是我们怎么知道广告转化中哪些是虚假的呢?

欢迎你给我留言，和我一起讨论。

参考文献

1. Interactive Advertising Bureau (2015). **What is an untrustworthy supply chain costing the us digital advertising industry**?

2. Stitelman, O., Perlich, C., Dalessandro, B., Hook, R., Raeder, T., and Provost, F. **Using co-visitation networks for detecting large scale online display advertising exchange fraud**. In Proceedings of the 19th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pages 1240–1248. ACM, 2013.

3. Zhang, W., Pan, Y., Zhou, T., and Wang, J. **An empirical study on display ad impression viewability measurements**. arXiv preprint arXiv:1505.05788, 2015.

上一篇　120 | 广告投放如何选择受众？如何扩展受众群？

下一篇　122 | 数据科学家必备套路之一：搜索套路

精选留言

写留言

> 由作者筛选后的优质留言将会公开显示，欢迎踊跃留言。