

1 Youtube 论文中 Watchtime 推导

Youtube 论文中, 基于预估 ctr 的逻辑回归, 对正样本 (click) 按照其对应的观看时长, 做样本增强, 然后使用不同的激活函数获得相应的时长。假定 N 为总的样本量, 其中 k 为 click 样本量, $N - k$ 则为 non-click show 的样本量。另假定每个 click 对应的时长为 T_i , 在训练时对第 i 个 click 样本进行 T_i 倍的增强, 在这种正负样本比例下, 用逻辑回归所学习出来的 odds(事件发生概率/事件不发生概率)

$$Odds_{watch} = \frac{\sum_i T_i}{N - k} \quad (1)$$

这里记作 $Odds_{watch}$ 用于区分这是 watchtime 训练, 区别于 ctr 训练中的 $Odds_{ctr}$ 。对以上式分子分母进行变换如下

$$Odds_{watch} = \frac{\sum_i T_i}{N - k} = \frac{\frac{\sum_i T_i}{N}}{\frac{N - k}{N}} = \frac{E(T)}{1 - \frac{k}{N}} = \frac{E(T)}{1 - p_{ctr}} \quad (2)$$

其中 $E(T)$ 为样本的期望观看时长 (针对所有样本), p_{ctr} 为学习出来的点击率。由于在 Youtube 系统中 p_{ctr} 小, 而 $f(x) = 1/(1 - x)$ 按泰勒展开后, $f(x) = 1 + x + x^2 + x^3 + \dots$, 原文中对其二次方以后的部分做了省略, 因此得到

$$Odds_{watch} = \frac{E(T)}{1 - p_{ctr}} \approx E(T)(1 + p_{ctr}) \approx E(T) \quad (3)$$

如果采用样本增强后, 使用逻辑回归预测出来的概率记作 p_w , 则有

$$\frac{p_w}{1 - p_w} = Odds_{watch} \approx E(T) \quad (4)$$

对于 sigmoid 激活下的输入 x 和其输出 p_w , $\frac{p_w}{1 - p_w} = e^x$ 成立, 因此, 采用 e^x 作为预测阶段的激活函数, 近似得到期望时长 $E(T)$ 。

2 针对我手使用场景的微调

2.1 微调原理

对于我们在短视频上的时长预估场景, p_{ctr} 已经不足以小到可以忽略, 关注页和发现页短视频的 emp ctr 达到 10% 以上, 同城页 8% 以上。因此

不能直接套用原文中省略 p_{ctr} 的逻辑，应该使用原始的推导如下：

$$\frac{p_w}{1 - p_w} = Odds_{watch} = \frac{E(T)}{1 - p_{ctr}} \quad (5)$$

因此有 $E(T) = \frac{p_w}{1 - p_w}(1 - p_{ctr})$ 成立，最终的 watchtime 预估应该结合 p_{ctr} ，因此在使用增强样本方式训练 watchtime 逻辑回归的同时，也应该训练一个 ctr 目标用于修正 watchtime 预估。

2.2 具体做法

由于直接扩充样本的存储和传输成本较高，在训练 watchtime 的时候，直接对 click 样本的梯度按照实际的观看时长做增强，对于样本 $click_i$ ，假设其观看时长为 T_i ，对其梯度放大 T_i 倍。对于负样本，梯度保持不变。

- 对于正样本，梯度为 $T_i(label - pred)$ ，对于负样本梯度为 $(0 - pred)$ ，其中 $pred$ 为 sigmoid 激活之后的输出，对应上文中的 p_w 。
- 同时在另一个网络同步训练 ctr，输出 p_{ctr} ，用于 watchtime 预测阶段。

3 扩展方法 (by 白老师)

以上的 watchtime 训练，是将所有样本都用来训练 watch time 的方法。一种扩展方式，是只对 click 样本进行 watch time 训练。在这种情况下，训练的 Odds 表达如下：

$$Odds_{watch} = \frac{\sum_i T_i}{N - k} = \frac{\frac{\sum_i T_i}{k}}{\frac{N - k}{k}} = \frac{E_c(T)}{\frac{1 - p_{ctr}}{p_{ctr}}} = \frac{E_c(T)p_{ctr}}{1 - p_{ctr}} \quad (6)$$

这里计算的时长期望定义与 Youtube 方法相比有所变化，考虑的范围不是所有的样本，而是 click 样本，用 $E_c(T)$ 来表示。如果用 p_w 表示 watchtime 训练所得到的预测值，则 $Odds_{watch} = \frac{p_w}{1 - p_w}$ 。因此，如下等式成立：

$$\frac{p_w}{1 - p_w} = Odds_{watch} = \frac{E_c(T)p_{ctr}}{1 - p_{ctr}} \quad (7)$$

进而得出：

$$E_c(T) = \frac{p_w}{1 - p_w} \frac{1 - p_{ctr}}{p_{ctr}} \quad (8)$$

$E_c(T)$ 为 click 发生条件下的预估时长。

在梯度更新的时候, 所采用的方法与 2.2 中所提方法类似, 但由于这种情况的 watchtime 训练中, 没有所谓的“负样本”, 所以梯度公式调整为

$$T_i(1 - pred) + (0 - pred) \quad (9)$$

即为每个样本都设置一个单位负样本梯度。目前滑滑板的时长预估采用的为此种训练方式。