

---

# 2018 CCF TCARCH 计算机体系结构挑战赛

## -人工智能加速器应用与设计

### 1.比赛题目背景:

广告是互联网行业的重要商业模式之一。因此，如何准确的预测用户的点击行为是互联网广告行业需要解决的重要问题。一方面，通过准确预测，展示用户最可能点击的广告，能方便用户找到所需，提升用户体验；另一方面，通过准确预测，增加用户点击次数，可以有效提高互联网企业的收入。同时，通过有效预测用户点击后产生的购买行为，也能提升广告主通过广告产生的收入，并进一步增加互联网平台的吸引力，形成正向循环。

用户点击行为预测包含很多指标，点击率（CTR, Click-Through-Rate）是其中一个重要指标。点击率指用户点击某条广告或商品的概率，是搜索排序的一个重要参考指标。点击率预测已广泛应用在各种互联网产品中，例如百度网页搜索、百度 App 等。

点击率预测模型和算法需要考虑很多特征，包括通用的用户画像、用户历史点击行为、广告特征、用户使用产品当前上下文等。点击率预测模型的评价包含两部分，一是预测结果的准确性，也即预测结果和最后的用户实际点击行为是否匹配。二是预测算法的性能，也即对于给定用户数和资源，预测所有候选广告所需的时间。如何根据给定数据，综合考虑预测结果的准确性和预测算法的性能，设计预测模型和算法是点击率预测需要解决的主要问题，也是互联网企业和学术界关注的重点。

### 比赛规则:

本次比赛以一个互联网应用（如 CTR）为切入点，比赛协办方将提供资源和数据集，初赛将考察参赛队伍搭建、训练网络的能力，复赛和决赛将进一步考察参赛队伍修改深度学习框架、模型压缩（主要是量化）能力，最终以计算模型质量（AUC）和预测评价速度（以平均用时计）评比。同时，为了鼓励参赛队伍对 AI 加速器设计进行探索，增设 AI 加速器设计特别奖，自主设计 AI 加速器的队伍有资格参与 FPGA AI 加速器设计特别奖评选。

---

## 2. 数据组成、提交方式和评测标准

### 2.1 数据组成

初赛的数据集包括 3 个部分，复赛的数据集另外追加了一个部分。

#### 2.1.1 训练集

即 `training set`。该数据集提供给选手，带有标注，同时会附带一份格式说明，说明有哪些字段，每个字段的类型（例如离散型、连续型等等）和意义（有些是脱敏或者混淆之后当作 `anonymous categories` 特征的）。

#### 2.1.2 开发集

即 `develop set`。该数据集提供给选手，带有标注，同时会附带一份格式说明。该数据集的原始日志时间保证在训练集之后。

#### 2.1.3 测试集 A

即 `public test set`。该数据集提供给选手，不带有标注，同时会附带一份格式说明。该数据集的原始日志时间保证在开发集之后。在选手提交自己的结果（注意，初赛和复赛的提交的东西有所不同，评价指标也有所不同）之后，平台会自动进行打分，并且将所有选手提交结果的打分进行排名并且进行展示。

#### 2.1.4 测试集 B

即 `private test set`。该数据集不提供给选手。该数据集和测试集 A 同源，并且原始日志时间保证在开发集之后。复赛最终的评分和评奖是用测试集 B 进行的。因此请选手注意，在测试集 A 上获得好效果的模型未必在测试集 B 上表现良好。

## 2.2 提交方式及评测标准

### 2.2.1 初赛

选手可以自行使用任何方法（包括但不限于，任何的开源软件或者框架，任何自研的软件或者框架），对测试集 A 的每一条样本进行预测，然后将预测结果提交至平台。请注意测试集 A 中附带的提交格式说明。

**评测标准：**和组委会保留的真实数据的真实 `label` 相比较，计算选手提交的测试集 A 的 `AUC` 指标，并且按照这个指标进行排序，排序之后最高的若干名（或若干个团队）进入复赛。

---

### 2.2.2 复赛与决赛

相比初赛，参赛队伍需要进一步完成的工作包括，参赛队伍可以选择使用百度 Paddlepaddle 开源框架或用其他框架。

a)对百度 Paddlepaddle 开源框架源代码进行修改使其能够调用百度云 FPGA 服务器（基于 Xilinx FPGA）上的 AI 加速器进行加速，注意 FPGA-AI 加速器目前支持 int16 以及 int8。参赛选手要在修改后的平台上完成训练以及预测的工作。

复赛时提交修改后的 Paddle 源代码，源代码在平台上编译后使用测试集 B 在选手提交的 int16/int8 模型上进行在线预测，计算模型质量（AUC）和预测评价速度（以平均用时计）

b)对于不使用百度 Paddlepaddle 框架自主设计 AI 加速器的队伍，复赛时需在必百度云 FPGA 环境（基于 Xilinx FPGA）上演示自行搭建的 AI 加速器并完成预测任务，评价指标亦为计算模型质量（AUC）和预测评价速度（以平均用时计）

**评测标准：**如果模型质量低于组委会预先训练的模型质量超过某个阈值，则提交无效。如果提交有效，则按照预测的平均速度进行排序，以确定名次。