隐私计算

常见问题

产品文档

腾讯云

常见问题

最近更新时间：2022-05-10 10:46:17

腾讯云隐私计算是什么?

腾讯云隐私计算（Privacy Computing）是腾讯云推出的以联邦学习（FL）、安全多方计算（MPC）、可信执行环境（TEE）等隐私数据保护技术为基础的隐私计算平台，产品针对机器学习算法进行定制化的隐私保护改造，保证原始数据不出本地即可完成联合建模，同时支持安全多方 PSI（隐私保护集合求交技术）、安全隐私查询、安全统计分析，提供基于硬件的 TEE 可信执行环境。通过腾讯云隐私计算，各合作机构既能保障数据安全，又能发挥数据最大价值，很好地解决了业界数据孤岛的难题。

什么是联邦学习?

联邦学习又名联邦机器学习、联合学习、联盟学习。联邦机器学习是一个分布式机器学习框架，它将传统的机器学习进行定制化的隐私保护改造，能够在合规的基础上帮助多个机构，在进行数据合作及联合建模时，有效地保护用户隐私及数据安全。

联邦学习有哪些使用场景?

在跨机构数据合作中，所有涉及到数据联合建模的场景，都有使用联邦学习的潜在诉求，联邦学习保障数据不出本地即可完成模型训练，在模型效果无损的情况下最大化保障数据安全。

联邦学习目前主要应用在风控和营销场景。风控主要是金融信贷风控和保险业务风控；营销场景一般不限行业，所有需要降低获客成本、提升获客质量的场景都可以满足，包括保险、教育、游戏、电商等行业。

什么是安全多方计算（MPC）?

- 安全多方计算（MPC）是指在无可信第三方的情况下，多个参与方协同计算一个约定的函数，并且保证每一方仅获取自己的计算结果，无法通过计算过程中的交互数据推测出其他任意一方的输入和输出数据。
- 安全多方计算（MPC）包括同态加密、秘密共享、混淆电路、零知识证明、不经意传输（OT）等技术。

安全多方计算有哪些使用场景?

安全多方计算的应用场景不同于联邦学习，一般不涉及到机器学习联合建模，多用在匿踪查询、统计分析、多方协同运算等场景，具体有如下示例可供参考：

- 多头借贷场景
  每当有用户申请贷款时，信贷业务机构都想知道借款人是否在其它机构也有同样的借款行为，但又不想暴露借款人信息。
- 数据安全查询场景
  商业机构需要查询信息用作商业用途，但数据方不希望数据被泄露，同时商业机构也不希望数据方知道其查询细节。利用安全多方计算技术，可以实现数据的安全查询。
- 联合数据分析场景
  跨机构的合作经常需要做统计、相关性等联合数据分析，应用安全多方计算技术可以使原始数据在无需归集与共

享的情况下实现联合数据分析，保护各方数据隐私。

使用腾讯云隐私计算进行联合建模，合作双方需要对齐数据吗?

使用腾讯云隐私计算平台进行数据合作的双方或多方，首先需要在数据规范上对齐。平台的数据类型分为 ID 数据、样本/标签、用户特征，所有数据类型前两个字段都必须是日期字段和 ID 字段：

- 日期字段只支持 yyyy、yyyymm、yyyymmdd、yyyymmddhh 这样的数据格式
- ID 字段目前支持 IMEI、IDFA、手机号、QQ 号等类型，联合建模之前需要先进行 ID 对齐

样本/标签数据包含标签字段，样本数据还包含特征字段，用户特征数据一般不包含标签字段：

- 标签字段为 y，取值“1”为正样本、取“0”为负样本

数据参考如下：

ID数据

	A	B
1	date	id
2	202002	cfcd208495d565ef66e7dff9f98764da
3	202002	eccbc87e4b5ce2fe28308fd9f2a7baf3

标签数据

	A	B	C
1	date	id	y
2	202002	cfcd208495d565ef66e7dff9f98764da	1
3	202002	eccbc87e4b5ce2fe28308fd9f2a7baf3	0

样本数据

	A	B	C	D	E
1	date	id	y	x0	x1
2	202002	cfcd208495d565ef66e7dff9f98764da	1	0.170308	0.122642
3	202002	eccbc87e4b5ce2fe28308fd9f2a7baf3	0	-0.9018	0

用户特征

	A	B	C	D	E
1	date	id	x0	x1	x2
2	202002	cfcd208495d565ef66e7dff9f98764da	0.170308	0.122642	0
3	202002	eccbc87e4b5ce2fe28308fd9f2a7baf3	-0.9018	0	0.004491

什么是可信执行环境（TEE）?

可信执行环境（TEE）是 CPU 内的一个安全区域，它是与操作系统并行运行且隔离执行的一个独立的环境，可以保证　部加载的代码和数据的机密性和完整性都得到保护。

可信执行环境（TEE）适合哪些场景?

可信执行环境（TEE）内的计算与明文计算逻辑基本相同，可以适用于任何有数据合作和数据隐私保护需求的场景，但其特性相比安全多方计算和联邦学习更适合：

- 实时性较高的机器学习应用场景
  例如部分推荐场景（在线训练、推荐模型日更等）对实时性要求高，这种需要联合数据实时推荐的场景和可信执行环境更契合。
- 海量数据联合分析
  跨机构的数据合作时，如果其中一方有大量数据需要处理，在联合分析时往往会受限于单方的性能，基于 TEE

可以提供更优异的计算性能。

腾讯云隐私计算是怎么保障用户的数据安全的?

- 在联邦学习应用场景中，区别于市面主流的联邦学习协议，腾讯云隐私计算平台首先改造剥离了可信中间方，模型训练过程中的任何加密中间参数都在合作双方或多方之间直接传递，不会经由任何第三方包括我们的隐私计算平台，这样最大化降低了任何数据方受到合谋攻击的风险。
- 在安全多方计算应用场景中，基于密码学设计的安全多方计算协议，所有计算过程中只会涉及到密态中间参数或者同态密文下的计算，防止恶意攻击下的隐私信息泄露。
- 在可信执行环境（TEE）应用场景中，数据只能在一个特定的安全区域进行解密使用，其业务逻辑代码也在该安全区域内也受到安全性及完整性的保护。

综上所述，3种方式均可以保障数据可用不可见，最大化保护数据隐私安全。

使用腾讯云隐私计算，需要什么样的机器资源?

在联邦学习应用场景下，所需要的机器资源根据模型训练的数据量级有所不同。Host 侧比较耗费机器资源，按20W样本量级、500维特征的数据量级进行评估，一般要准备16C/64G/2T磁盘/5M宽带的资源，Guest 侧推荐准备16C/32G/2T磁盘/5M宽带。

除了两方联邦，腾讯云隐私计算平台可以解决多方数据合作的问题吗?

在稳定高效的两方联邦学习基础上，目前腾讯云隐私计算已经实现了3方、4方联邦，已经能够满足主流的多方联邦应用场景，更多方联邦也会在后续陆续支持。

腾讯云隐私计算在金融风控场景有什么落地案例?

目前在信贷风控场景已经有多个落地案例，头部银行、消金公司跟数据提供方进行合作，通过腾讯云隐私计算进行联合建模，通过不停地优化迭代算法，效果稳定提升。

腾讯云隐私计算在营销场景有什么落地案例?

目前已有包括金融信贷、保险、教育等多个行业的营销落地案例，其中一个信贷营销案例中服务商 A 为了扩大服务投放业务并稳定效果，联合数据方 B 展开了联邦合作，联邦建模效果比服务方 A 独立建模效果有稳定提升，营销投放业务持续正常开展。

- 点击模型：AUC从0.62提升到0.72，KS从0.16提升到0.32
- 进件模型：AUC从0.63提升到0.66，KS从0.22提升到0.27

如何使用腾讯云隐私计算?

单击 立即申请，填写申请单，填写完成后，单击**提交申请**，完成线上申请。

腾讯云隐私计算平台性能表现如何?

腾讯云隐私计算平台基于分布式可扩展的集群动态提供计算资源后，可以轻松支持千万级数据量的模型训练，分钟级完成千万数据的推理。