

智能推荐 开发指南 产品文档



腾讯云

【版权声明】

©2013-2020 腾讯云版权所有

本文档（含所有文字、数据、图片等内容）完整的著作权归腾讯云计算（北京）有限责任公司单独所有，未经腾讯云事先明确书面许可，任何主体不得以任何形式复制、修改、使用、抄袭、传播本文档全部或部分内容。前述行为构成对腾讯云著作权的侵犯，腾讯云将依法采取措施追究法律责任。

【商标声明】



及其它腾讯云服务相关的商标均为腾讯云计算（北京）有限责任公司及其关联公司所有。本文档涉及的第三方主体的商标，依法由权利人所有。未经腾讯云及有关权利人书面许可，任何主体不得以任何方式对前述商标进行使用、复制、修改、传播、抄录等行为，否则将构成对腾讯云及有关权利人商标权的侵犯，腾讯云将依法采取措施追究法律责任。

【服务声明】

本文档意在向您介绍腾讯云全部或部分产品、服务的当时的相关概况，部分产品、服务的内容可能不时有所调整。您所购买的腾讯云产品、服务的种类、服务标准等应由您与腾讯云之间的商业合同约定，除非双方另有约定，否则，腾讯云对本文档内容不做任何明示或默示的承诺或保证。

【联系我们】

我们致力于为您提供个性化的售前购买咨询服务，及相应的技术售后服务，任何问题请联系95716。

文档目录

开发指南

API 使用说明

统计分析

数据质量监控

开发指南

API 使用说明

最近更新时间：2020-04-28 16:02:32

简介

腾讯云智能推荐服务基于实时上报的 Item（物料）和 Action（行为），结合15亿+腾讯用户画像进行机器学习并实时返回推荐结果。为简化开发者上报 Item 和 Action 两类数据，同时方便获取智能推荐服务结果，腾讯云提供了相关协议标准。

术语表

- Item：物料，即被推荐物品，推荐系统使用物料属性信息、用户标签及行为，生成推荐结果。物料分 pool 管理，一个物料可以属于一个或多个 pool。获取推荐结果时，可以在指定 pool 中计算推荐结果。
- Pool：物料池，被推荐物品的集合，用于管理只需要特定场景下出现的物料集合。如果物料只希望在某个场景下出现，可以在物料上报时指定 pool，拉取推荐结果时，指定对应的 pool 即可。一个物料可以属于一个或多个 pool。
- Traceld：用户行为 ID，串联用户行为的会话 ID，用户行为包括曝光、点击、效果转换、点赞等，用户行为按时序依次产生，需要业务传递 traceld，标识点击、效果转换的动作是哪一次的曝光、点击产生的。

API 快速入门

通过智能推荐 API 使用腾讯云推荐服务，您需要依次完成如下几个步骤：

1. 提交业务接入申请表。
2. 开通业务白名单后，在腾讯云控制台创建一个业务。开通白名单和业务的审批请联系相关的同事。
3. 使用 Item 物料上报 API 上报物料，上报物料后，联系联调的同学查看上报物料数据质量。
4. 使用 Action 行为上报 API 上报用户行为，上报行为后，联系联调的同学查看上报行为数据质量。
5. 通过推荐结果 API 获取推荐结果。
6. 联调完成后，通知腾讯云联调同学配置上线算法。
7. 上线观察业务指标。

API 概览

- Item 上报：物料管理，包括物料上报、物料池划分、物料上下架等功能。
- Action 上报：用户行为上报，上报特定场景下，某个用户对某个物料产生的特定动作。
- 推荐结果拉取：从指定的物料池中获取特定场景下某个用户的推荐结果。

调用方式

智能推荐提供 HTTP+JSON 的服务，请求协议以 JSON 格式序列化，作为 HTTP 请求的 Body 部分，客户端以 POST 方式访问服务。服务地址：

- 推荐结果拉取：`http://service.dm.qcloud.com:8088`
- 物料与行为上报：`http://data.dm.qcloud.com:8088`

鉴权

协议使用 MD5 字段加密签名，防止竞争对手伪造数据包进行流量攻击。MD5 的计算方法为：BID、request_id、TOKEN 三个字符串用&字符连接，然后整个字符串的 MD5 值。

MD5 (BID&request_id&TOKEN)

注意：

目前 TOKEN 验证正在完善，可以先用一个字符串代替，后续再找腾讯云分配。目前接收请求的 CGI 还未针对此进行验证，因此前期数据接入时，MD5 可以随便填写。

Item 上报

接口描述

上报物料信息，包括物料 ID、物料有效期、物料标签和物料池等信息。相同的 item_id 可以重复上报，字段信息以最后一次上报为准。

开发者发送 http-post 获取服务结果，服务 URL：`http://data.dm.qcloud.com:8088`。

输入参数

POST 报文 Body 部分为 JSON 数据格式，如下所示：

Item 上报 JSON 数据格式：

```
{  
  
  "MD5": "40379db889f9124819228947faaeb1f7", //md5(bid&request_id&TOKEN)  
  "request_id": "request_id", //request_id 为 “毫秒级时间戳随机数”  
  "data_type": 1, //1 : item , 2 : action  
  "bid": "BID", //腾讯云为该业务分配的业务标识  
  "item_id": "item_id", //物料标识  
  "publish": 1, //1 : 上架 (默认) , 0 : 下架
```

```
"describe":"最新款黑色苹果7", //物料描述
"pool_id":"pool_id1;pool_id2;pool_id3", //物料池，多个池子用；号隔开
"tags":"电器;科技;电子;手机", //物料标签，多个 tag 用；号隔开
"item_time":"1386817569", //物料生成时间（默认为当前时间）
"expire_time":"1386917763", //物料过期时间（默认 item_time + 一个月）
"free":0, //0：免费（默认），1：付费
"score":9.99, //物料打分（默认为0.0）
"price":88.88, //物料价格（默认为0.0）
"platform":1, //平台，1：android（默认），2：iphone，3：PC
"big_type":"大类",
"middle_type":"中类",
"small_type":"小类",
"url":"URL",
"vender":"店铺，广告主",
"geo":{
  "latitude":-90.0~90.0,
  "longitude":-180.0~180.0,
  "country":"country code using ISO-3166-1-alpha-3",
  "city":"city name"
},
"extend":{
  "key1":"value1",
  "key2":"value2",
  "key3":"value3",
  ...
}
```

字段含义

字段类型及含义如下：

参数名称	类型	是否必填	含义
MD5	string	是	MD5(bid&request_id&TOKEN)
request_id	string	是	请求标识，格式：毫秒级时间戳_随机数，随机数建议三位以上
data_type	integer	是	协议类型，1：item，2：action，物料上报取值为1
bid	string	是	业务 ID，腾讯云分配的业务 ID，可以在控制台查询
item_id	string	是	物料 ID，物料唯一标识，可为数值和字符串
publish	integer	否	1：上架（默认），0：下架

参数名称	类型	是否必填	含义
describe	string	否	商品的描述信息，推荐系统会对商品的描述信息使用 NLP 技术，进行分词、提取关键词，作为 item 的特征使用
pool_id	string	否	物料池，多个物料用分号分割， 不要使用0作为 pool_id，0表示全局 pool_id ，任意物料我们都会加一个为0的 pool_id
tags	string	否	作为模型的特征使用，是对物品的一个重要的区分特征。tags 的丰富、区分程度，很大程度上决定了模型训练的好坏，默认 tags 字段需要上报三级类目、品牌。另外，如果有标签描述信息，也可以上报。tags 有区分性，主要用于对新 item 的冷启动。因此 tag 不能太粗，否则没有区分性，特征作用则较小；也不能太细，否则没有共性，曝光不充分，起不到冷启动的目的
item_time	string	否	item 生成时间，10位的秒级时间戳。如果不填，默认是系统当前时间
expire_time	integer	否	item 的过期时间，推荐结果会进行商品的过期过滤。如果不填，默认是 item_time + 1个月
free	string	否	0：免费（默认），1：付费
score	float	否	物料打分，默认0.0，没有可以不填写。推荐在离线时可以把 item 的 cvr 信息填写到 score，然后在 rerank 时，对排序进行加权
price	float	否	物料价格，默认0.0
platform	integer	否	平台，0：全平台，1：Android（默认），2：iphone，3：pc
big_type	string	否	大类，需要类目的中文，不能使用 ID
middle_type	string	否	中类，需要类目的中文，不能使用 ID
small_type	string	否	小类，需要类目的中文，不能使用 ID
url	string	否	物料相关的 url。一个作用是使用 DNN 提取图片的特征，放入到模型使用；一个作用是 URL 的域名，本身就可以作为特征使用，有一定的区分性
vender	string	否	可以传品牌、商家、店铺和广告主
geo.latitude	float	否	LBS 数据，部分 O2O 的商家，需要知道 item 可以在哪些商圈投放，因此需要填写 item 所属的商圈的经纬度数据。如果没有这个需求的，可以不填写
geo.longitude	float	否	物料相关地域维度

参数名称	类型	是否必填	含义
geo.country	string	否	物料相关国家
geo.city	string	否	物料相关城市
extend.key	string	否	扩展字段

item 上报返回 JSON 数据格式：

```
{  
  
  "request_id": "request_id", //request_id 原样返回  
  "code": 0, //-1：格式错误，-2：系统错误，-3：算法错误  
  "msg": "true"  
}
```

Action 上报

接口描述

上报某一用户在特定场景下的行为，用户的行为包括曝光、点击、转换、点赞等。上报用户行为时，必须指定用户行为的会话 ID。用户行为可以在客户端和服务端上报，建议在客户端上报，可控性更强些，遇到协议变更或者问题排查时，更容易处理。

开发者发送 http-post 获取服务结果，服务 URL：<http://data.dm.qqcloud.com:8088>。

输入参数

POST 报文 Body 部分为 JSON 数据格式，如下所示：

Action 上报 JSON 数据格式：

```
{  
  
  "MD5": "40379db889f9124819228947faaeb1f7", //md5(bid&request_id&TOKEN)  
  "request_id": "request_id", //request_id 为 "毫秒级时间戳随机数"  
  "data_type": 2, //1：item，2：action  
  "bid": "BID", //腾讯云为该业务分配的业务标识  
  "uid_type": 3, //0：QQ，1：微信号，3：设备号 imei/idfa 或其 MD5 值（默认），4：手机号  
  "uid": "userId", //QQ，微信号，imei/idfa 或其 MD5 值，手机号  
  "item_id": "item_id1;item_id2;item_id3", //物料列表,多个物料用 ; 号隔开  
  "scene_id": "1001", //广告展示场景。
```

```
"action_type":1, //行为类型。 1：曝光（浏览），2：点击（播放），3：转化（购买），4：点赞
"source":"tx", //流量标记。 tx：腾讯流量, bus：业务流量, def：默认流量
"trace_id":"trace_id", //跟踪点击和曝光的自定义跟踪 ID
"action_time":"1386817569", //行为发生秒级时间戳（默认当前时间）
"geo":{
  "latitude":-90.0~90.0,
  "longitude":-180.0~180.0,
  "country":"country code using ISO-3166-1-alpha-3",
  "city":"city name"
},
"extend":{
  "key1":"value1",
  "key2":"value2",
  "key3":"value3",
  ...
}
```

字段说明

输入参数各字段的含义如下：

参数名称	类型	是否必填	含义
MD5	string	是	MD5(bid&request_id&TOKEN)
request_id	string	是	请求标识，格式：毫秒级时间戳_随机数，随机数建议三位以上
data_type	integer	是	协议类型，1：item，2：action，行为上报传值为2
bid	string	是	业务 ID，腾讯云分配的业务 ID，可以在控制台查询
uid_type	integer	否	0：qq，1：微信号，3：imei/idfa 或其他 md5（默认），4：手机号
uid	string	是	qq、微信号、imei/idfa、手机号。如果是 imei 号，则是15位数字字符串；如果是 IFA 号，则是8-4-4-4-12，32个字符串；如果是 MD5（uid），则是32位的0-f的字符串，且需要是大写的 IDFA 进行的 MD5 加密
item_id	string	是	物料 ID，物料唯一标识，可为数值和字符串
scene_id	string	是	推荐场景 ID，例如有“猜你喜欢”、“热门商品”等推荐模块，每一个模块都有一个 scene_id 来表示。可在控制台创建场景和查看场景 ID

参数名称	类型	是否必填	含义
action_type	integer	是	用户行为，1：曝光，2：点击（播放），3：转换（购买），4：点赞
source	string	是	用于分流的标识字段。区分用户的行为是哪种算法，取值为 bus、tx、def 三者之一。其中 bus 表示业务自己的算法（business）；tx 为腾讯算法；def 为默认算法（default），可以理解为随机算法，或者对照组
trace_id	string	是	跟踪点击和曝光的自定义会话 ID。为了保证点击和曝光是同一个用户对同一个 item 的操作行为， 建议每次曝光分配一个 trace_id
action_time	string	是	行为发生的时间戳，秒级时间戳（默认为当前时间）。不能延迟太久，尽量实时上报，否则会影响推荐结果的准确性
geo.latitude	float	否	用户发生行为的纬度地理位置
geo.longitude	float	否	用户发生行为的经度地理位置
geo.country	string	否	用户发生行为的国家
geo.city	string	否	用户发生行为的城市
extend.key	string	否	扩展字段

Action 上报返回 JSON 数据格式：

```
{
  "request_id": "request_id",
  "code": 0, // -1：格式错误，-2：系统错误，-3：算法错误
  "msg": "true"
}
```

请求服务

接口描述

获取特定场景下，给特定用户的推荐结果。获取推荐时，需要指定场景、用户信息及需要使用的物料池，推荐系统返回用户在当前场景下，对各个物品喜好的权重。

开发者发送 http-post 获取服务结果，服务 URL：`http://service.dm.qcloud.com:8088`。

输入参数

POST 报文 Body 部分为 JSON 数据格式，如下所示：

智能推荐服务请求 JSON 数据格式：

```
{
  "MD5":"40379db889f9124819228947faaeb1f7", //md5(bid&request_id&TOKEN)
  "service_type":3, //服务类型。0：pCTR，1：流量优选，2：pCVR，3：个性化推荐（默认），4：物料优选
  "request_id":"request_id", //request_id 为“毫秒级时间戳随机数”
  "bid":"BID", //腾讯云为该业务分配的业务标识
  "uid_type":3, //0：QQ，1：微信号，3：设备号 imei/idfa 或其 MD5 值（默认），4：手机号
  "uid":"userId", //QQ，微信号，imei/idfa 或其 MD5 值，手机号
  "scene_id":"1001", //广告展示场景，请求服务时 scene_id 与上报数据时保持一致
  "request_num":50, //物料优选个数。默认每次请求50个
  "pool_id":"pool_id1;pool_id2", //物料池，多个池子用；号隔开
  "cid":"item_id", //当前物料（用于详情页推荐中指定当前物料）
  "geo":{"//当前 lbs 信息（用于考虑 lbs 场景的推荐）
    "latitude":-90.0~90.0,
    "longitude":-180.0~180.0,
    "country":"country code using ISO-3166-1-alpha-3",
    "city":"city name"
  },
}
```

字段含义

各个字段的含义如下：

参数名称	类型	是否必填	含义
MD5	string	是	MD5(bid&request_id&TOKEN)
request_id	string	是	请求标识，格式：毫秒级时间戳_随机数，随机数建议三位以上
service_type	integer	是	服务场景，0：pCtr，1：流量优选，2：pCvr，3：智能推荐（默认），4：物料优选
bid	string	是	业务 ID，腾讯云分配的业务 ID，可以在控制台查询
uid_type	integer	否	0：qq，1：微信号，3：imei/idfa 或其他 md5（默认），4：手机号

参数名称	类型	是否必填	含义
uid	string	是	qq、微信号、imei/idfa、手机号。如果是 imei 号，则是15位数字字符串；如果是 IFA 号，则是8-4-4-4-12，32个字符串；如果是 MD5 (uid)，则是32位的0 - f的字符串，且需要是大写的 IDFA 进行的 MD5 加密
scene_id	string	是	推荐场景 ID，请求服务的场景 ID，需要和行为上报时的 ID 一致
request_num	integer	否	物料优选的结果，默认50个。目前最多支持200个的 item 返回，如果返回个数更多，会影响性能，容易超时
pool_id	string	是	物料池，需要展示哪个 pool_id 下的商品，要和上报的 item 的 pool_id 一致
cid	string	是	当前页面的物料 ID，用于详情页面
geo.latitude	float	否	用户发生行为的纬度地理位置
geo.longitude	float	否	用户发生行为的经度地理位置
geo.country	string	否	用户行为相关的国家信息
geo.city	string	否	用户行为相关的城市信息

智能推荐服务返回 JSON 数据格式：

```
{  
  
  "request_id": "request_id",  
  "algid": 17,  
  "code": 0, //-1：格式错误，-2：系统错误，-3：算法错误  
  "scene_id": "1001",  
  "data": [  
    {  
      "item": "ItemId10",  
      "weight": 974  
    },  
    {  
      "item": "ItemId41",  
      "weight": 864  
    },  
    {  
      "item": "ItemId7",  
      "weight": 97  
    },  
  ],  
}
```

```
...
{
  "item": "ItemId86",
  "weight": 89
}
]
```

返回值

返回值名称	类型	含义
algid	Integer	算法 ID，可以在后续的行为上报中回传，跟踪算法效果
scene_id	string	场景 ID
data.item	lstring	推荐的物品 ID，即 item 上报中的 item_id
data.weigth	Integer	推荐物品的权重，取值范围[0,1000000]

错误返回如下：

```
{
  "request_id": "request_id",
  "code": -3,
  "msg": "empty result"
}
```

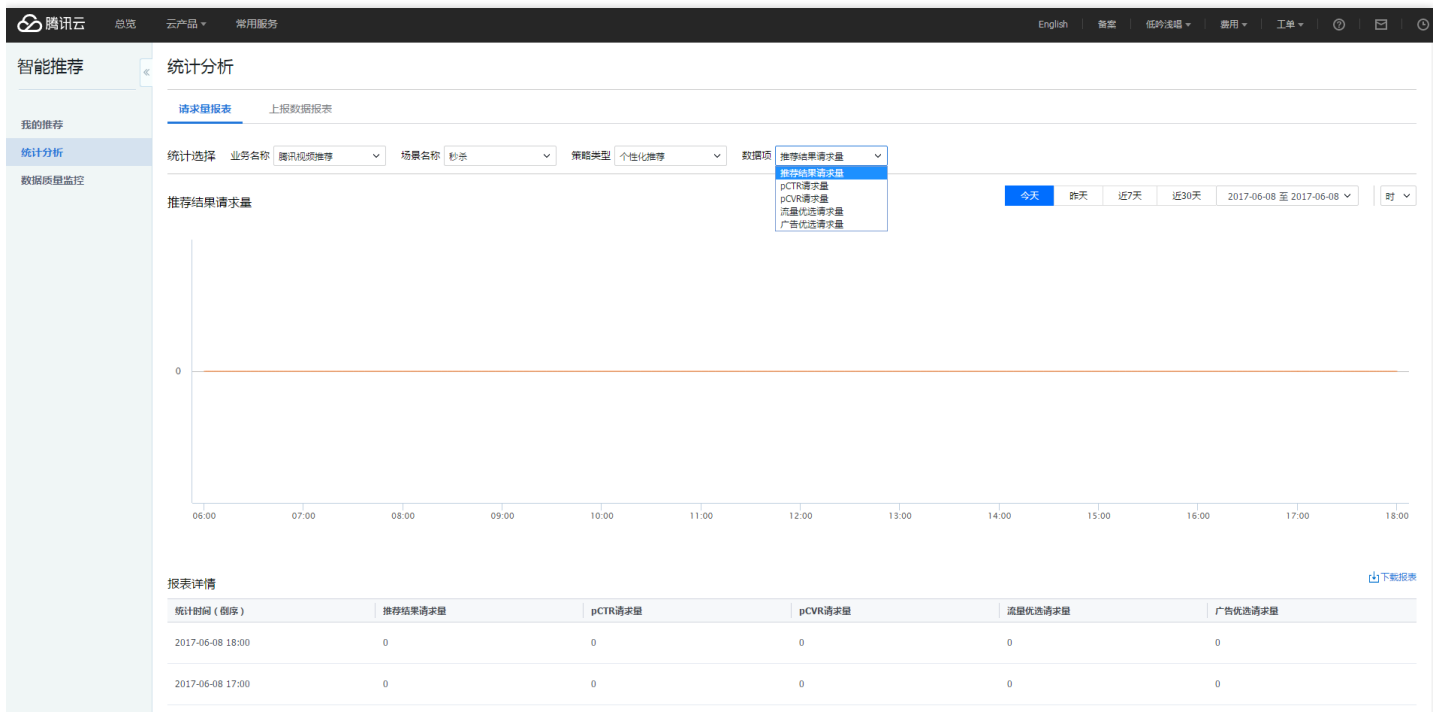
统计分析

最近更新时间：2020-04-28 16:02:13

上报 API 数据并经过业务接口联调、业务上线后，进入智能推荐控制台，单击左侧菜单栏【统计分析】，即可看到业务的统计分析。

请求量报表

选择【请求量报表】，在选定业务名称、场景名称、策略类型和数据项后，您可以查看今天、昨天、最近7天，最近30天或指定时间段的对应数据项的统计分析结果和报表详情。



上报数据报表

单击【上报数据报表】，在选定业务名称、场景名称、策略类型和数据项后，您可以查看今天、昨天、最近7天，最近30天或指定时间段的对应数据项的统计分析结果和报表详情。

版权所有：腾讯云计算（北京）有限责任公司

数据质量监控

最近更新时间：2020-04-28 16:01:56

上报 API 数据并经过业务接口联调、业务上线后，进入智能推荐控制台，单击左侧的【数据质量监控】后，可以看到您业务的数据质量监控情况。

在选定业务名称、场景名称、策略类型和数据项后，您可以查看今天、昨天、最近7天，最近30天或指定时间段的对应数据项的数据质量监控结果和报表详情。

