



天翼云自然语言处理

用户操作手册

天翼云科技有限公司

1 产品简介	1
1.1 产品定义	1
1.2 功能特性	1
1.3 产品优势	2
1.4 应用场景	2
2 计费说明	4
2.1 计费模式	4
2.2 产品续订	5
2.3 产品退订	5
3 用户指南	6
3.1 购买服务	6
3.2 创建应用及开通服务	8
3.3. 自然语言控制台	12
4 API 参考	15
4.1 API 概览	15
4.1.1 概述	15
4.1.2 API 概览	15
4.1.3 状态码	15
4.2 如何调用 API	17
4.1.1 终端节点	17

4.1.2 构造请求.....	17
4.1.3 认证鉴权.....	18
4.1.4 Python3调用示例.....	22
4.3API.....	24
4.3.1 情感分析.....	24
4.3.2 文本纠错.....	29
4.3.4 商品评价解析.....	33
4.3.4 语音合成.....	38
4.4 更新历史.....	43
5 常见问题.....	44
5.1 计费类.....	44
5.2 购买类.....	45
5.3 操作类.....	47
5.4 使用限制.....	47

1 产品简介

1.1 产品定义

天翼云自然语言处理（Natural Language Processing，简称 NLP）产品可广泛应用于内容推荐、舆论分析、投诉分析等各类场景，满足各行各业文本智能需求。当前天翼云自然语言处理提供语音合成，情感分析、文本纠错、商品评价解析能力。

1.2 功能特性

情感分析

通过分析带有主观描述的评论观点，自动判断评论观点的情感极性类别，情感极性分为积极/消极。

语音合成

语音合成（Text To Speech，TTS）将文本转成拟人化的语音。支持中文语音合成，提供男、女两种音色的选择，支持自定义语调、语速等参数。

文本纠错

针对一段文本，自动识别纠正其中的语法或字词错误，包括形似字、音似字，漏字等文法错误的纠正和探测。

商品评价解析

智能解析评论文本内容，进行评论观点的抽取与分析，自动输出评论观点标签及评论观点极性，将文本转化为结构化的属性字段。

1.3 产品优势

广泛的领域覆盖：覆盖自然语言处理的多个应用领域，如情感分析、文本纠错等，满足用户个性化需求。

优良的处理效果：采用大量、多主题的语料数据进行模型训练，使得能力具备良好的使用效果。

快速响应：配备高配处理硬件，采用高并发服务框架，可对用户请求快速响应，提高用户使用体验。

接口易用：标准化接口封装，通过云计算调用可快速使用工具，大大降低开发人力成本。

1.4 应用场景

文本情感分析

在数字时代，总结文档和信息含义的能力变得越来越重要。情感分析作为一种常见的自然语言处理方法的应用，可以让我们能够从大量数据中识别和吸收相关信息，而且还可以理解更深层次的含义。比如，企业分析消费者对产品的反馈信息，或者检测在线评论中的差评信息等。

个性化推荐

自然语言处理可以依据大数据和历史行为记录，学习出用户的兴趣爱好，预测出用户对给定物品的评分或偏好，实现对用户意图的精准理解，同时对语言进行匹配计算，实现精准匹配。例如，电商领域，通过对用户评论的智能解析，可以进行后续商品个性化推荐、新闻定制服务等应用方向，提升用户粘性。

对话机器人

通过自然语言处理技术，快速解析用户输入的自然语言，进行语音合成，并根据分析结果应用于对话机器人中槽位提取、句式模板分析等应用方向，为对话机器人的底层对话平台搭建提供能力支撑。

2 计费说明

2.1 计费模式

产品公测期间免费试用不收费。以下是价格说明：

自然语言处理产品采用封顶资源包的计费方式，一些在售卖的 api 服务会免费赠送调用额度供测试使用，若有已购买且在有效期内的商品资源包，则可进行抵扣。

目前自然语言处理的语音合成、文本纠错、情感分析这三个能力处于公测阶段，公测期间0元购买不收费。

除以上三个能力以外的其他自然语言处理的能力目前是处于限时免费的状态，无需购买，只需开通即可使用。

为满足客户不同业务使用量需求，每类 API 设置五档套餐，如：用户评估认为应用每年情感分析的 API 调用量大概为 10 万次，则可以购买一个 10 万次/年的 API 调用次数包，若大概为 50 万次，则可以购买 1 个 40 万次/年的 API 调用次数包与 1 个 10 万次/年的 API 调用次数包。

用户预先购买 N 个固定额度的资源包，后续使用过程中产生的接口调用次数从资源包中抵扣，有效期内未使用的资源不会流转至下一年。有效期内资源包中的次数用完之后，用户无法再调用 AI 产品 API。对于有效期内未使用的资源包剩余调用次数，本产品不支持退订。

自然语言处理 API 服务免费额度如下：

API 名称	免费额度	具体说明
语音合成	10000 次/年	以单个 API 为统计维

文本纠错	10000 次/年	度, 当年有效, 过期作废
情感分析	10000 次/年	

自然语言处理 API 资源包价格如下:

API 名称	10 万次 (元)	40 万次 (元)	80 万次 (元)	400 万次 (元)	800 万次 (元)
文本纠错	0	0	0	0	0
情感分析	0	0	0	0	0

API 名称	10 万次 (元)	100 万次 (元)	5000 万次 (元)	1000 万次 (元)
语音合成	0	0	0	0

备注：表格价格仅供参考，具体购买价格以价格发文为准。

2.2 产品续订

续订说明：当已订购的资源包订单即将到期或即将用完时，可通过订购新的资源包进行续订。

2.3 产品退订

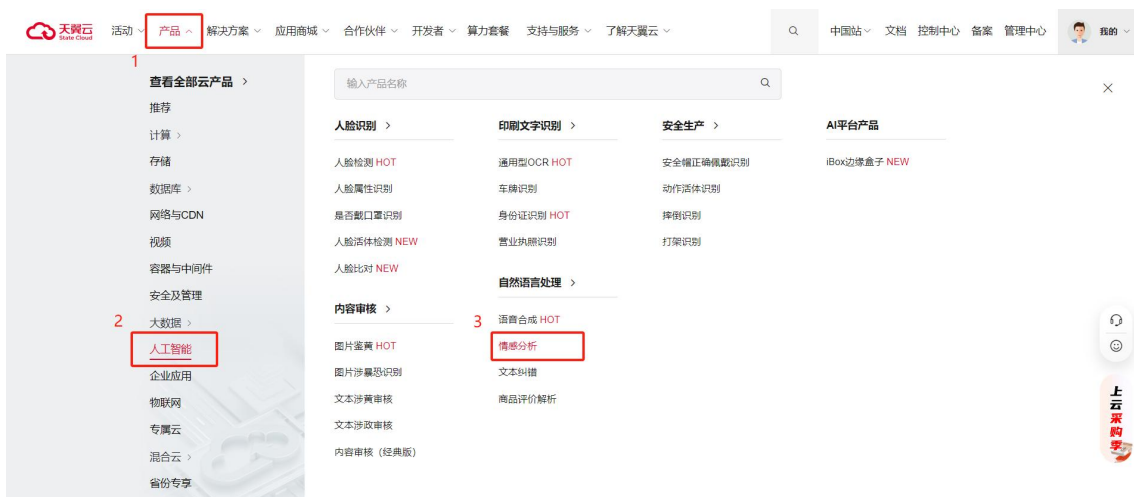
退订说明：已购订单内的资源包不支持退订。

3 用户指南

3.1 购买服务

1、进入产品详情页

首先，用户使用天翼云网门户账号登录天翼云网门户后，点击菜单栏“产品”项，进入【人工智能】，即可看到自然语言处理相应服务。以“情感分析”为例，点击“情感分析”名称即可进入产品详情页。



2.购买资源包

点击【立即购买】选择相应规格，确认订单费用是否准确，费用无误后，点击“立即支付”。会跳转到第三方支付页面进行支付，目前支持支付宝、微信、以及银行卡等支付方式。

公测 情感分析

对带有情感色彩的主观性文本进行推理与分析，自动进行文本情感倾向性判断，输出该文本的情感极性类别，情感极性分为积极、消极。

[立即购买](#) [公测开通](#) [管理控制台](#) > [产品文档](#)

算力套餐
让算力触手可及

云上钜惠
领取5万元大礼包更实惠

云主机特惠
新老同享云主机3折起

- 动态 查看情感分析的API
- 动态 支持用户查看AK, SK
- 动态 查看自然语言处理的认证鉴权

天翼云 控制中心

费用 工单 备案 支持 合作

< AI能力开放平台

情感分析

选择配置 单次限购1个规格

规格	公测截止时间	调用有效期	单价	购买数量
10万次	另行通知，届时将提前10个自然日在网站内发布公告	1年	¥0.00	- 1 +
40万次	另行通知，届时将提前10个自然日在网站内发布公告	1年	¥0.00	- 0 +
80万次	另行通知，届时将提前10个自然日在网站内发布公告	1年	¥0.00	- 0 +
400万次	另行通知，届时将提前10个自然日在网站内发布公告	1年	¥0.00	- 0 +
800万次	另行通知，届时将提前10个自然日在网站内发布公告	1年	¥0.00	- 0 +

配置费用 **¥0.00**
参考价格，具体扣费请以账单为准。了解计费详情

[取消](#) [下一步：确认配置](#)

天翼云 控制中心

费用 工单 备案 支持 合作

< AI能力开放平台

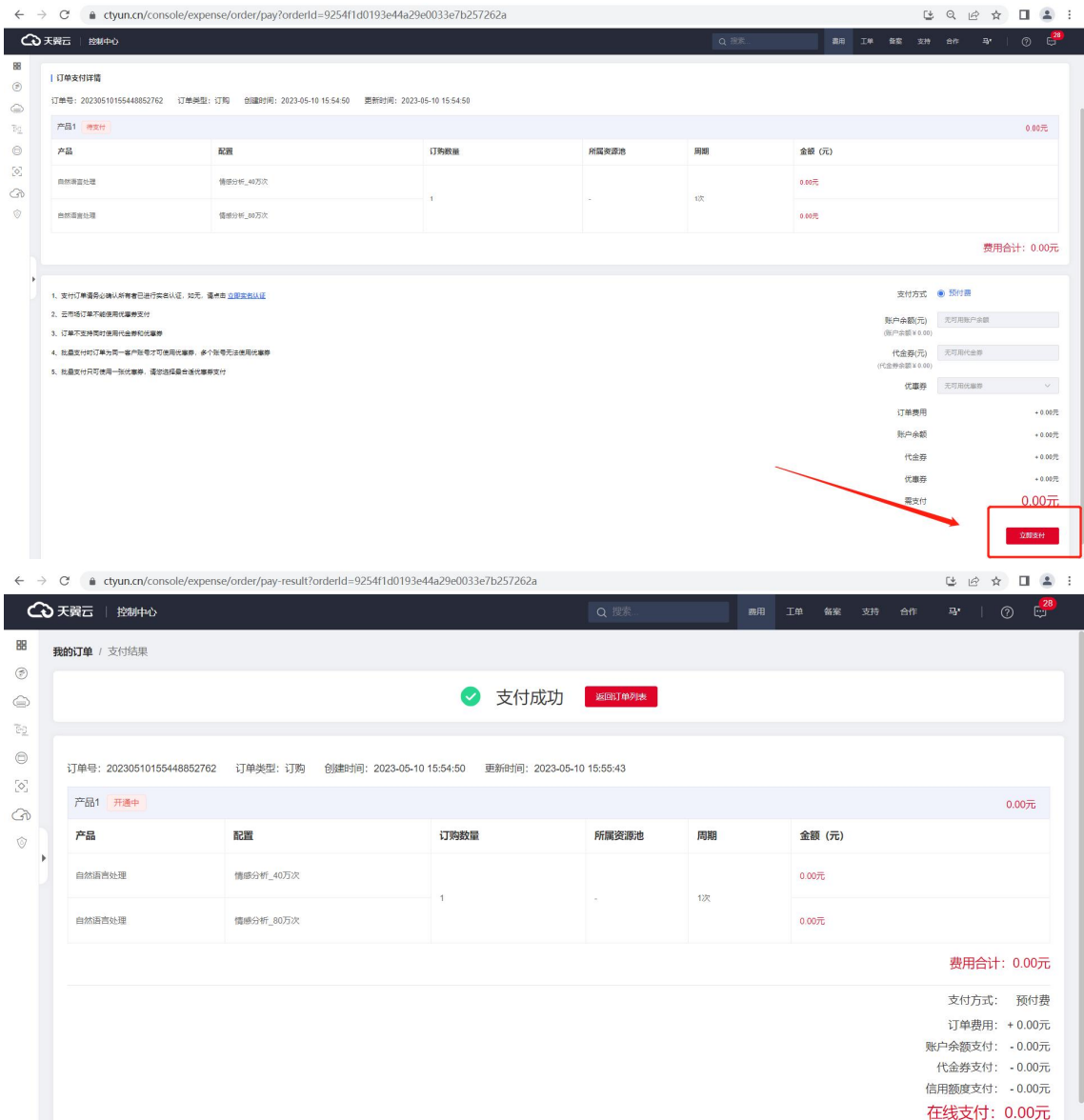
产品名称	计费方式	规格	调用有效期	公测截止时间	购买数量	单价
情感分析	包周期	10万次	1年	另行通知，届时将提前10个自然日在网站内发布公告	1	¥0.00

协议

我已阅读并同意相关协议 [《天翼云公测产品服务协议》](#)

配置费用 **¥0.00**
参考价格，具体扣费请以账单为准。了解计费详情

[上一步](#) [立即购买](#)



3.2 创建应用及开通服务

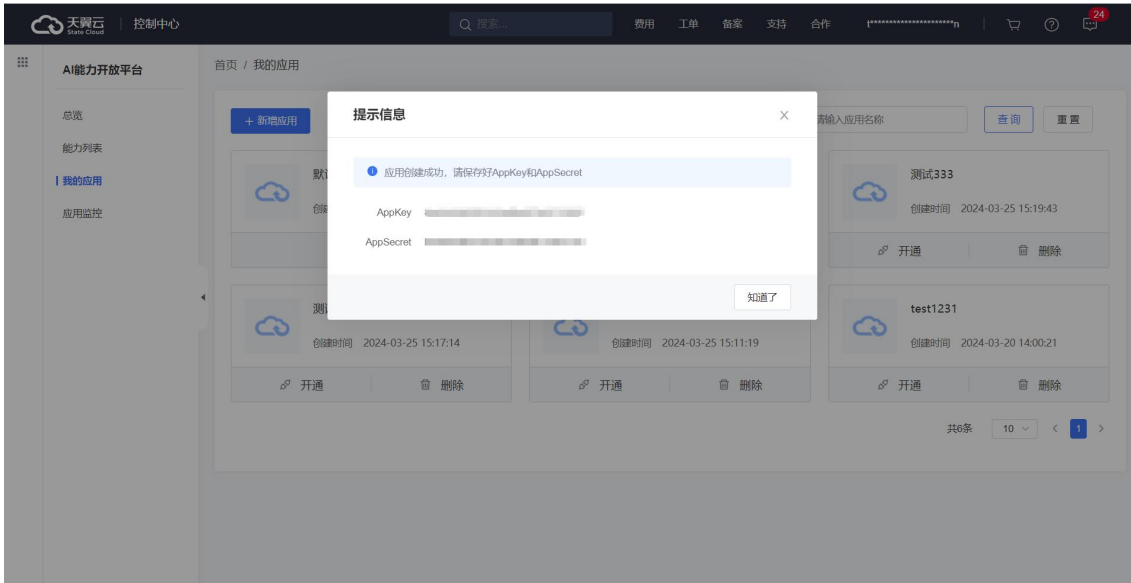
1.首先，用户使用天翼云网门户账号登录天翼云网门户后，点击菜单栏“产品”项，进入【人工智能】，即可看到自然语言处理产品相应服务。以“情感分析”为例，点击“情感分析”名称即可进入产品详情页。



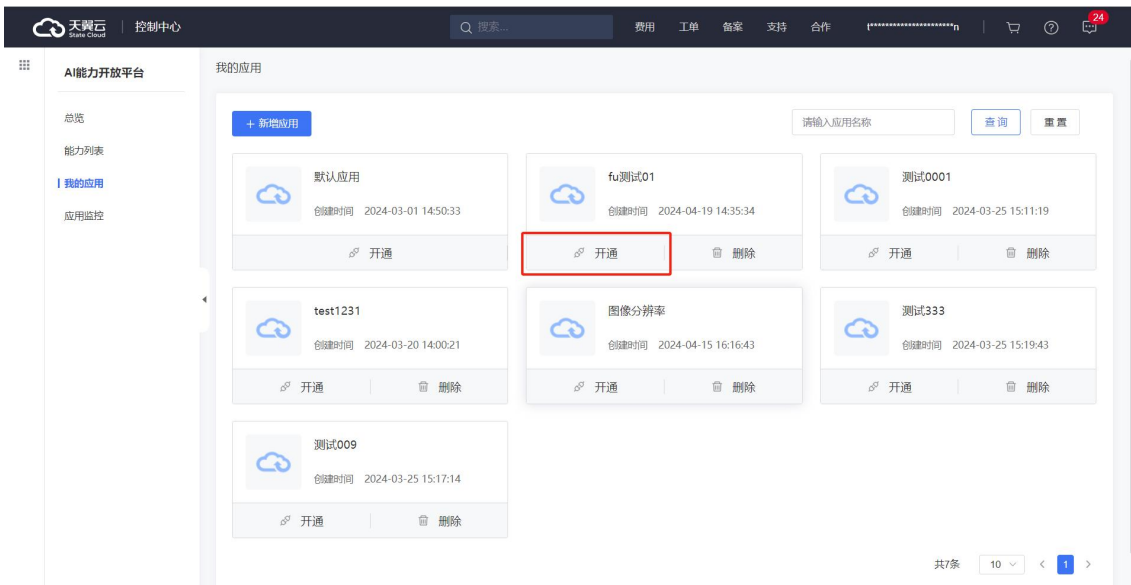
2. 点击【公测开通】按钮，进入控制台【我的应用】，点击【新增应用】，页面出现新建应用的弹窗，填写应用名称及应用概述，点击【确定】按钮。



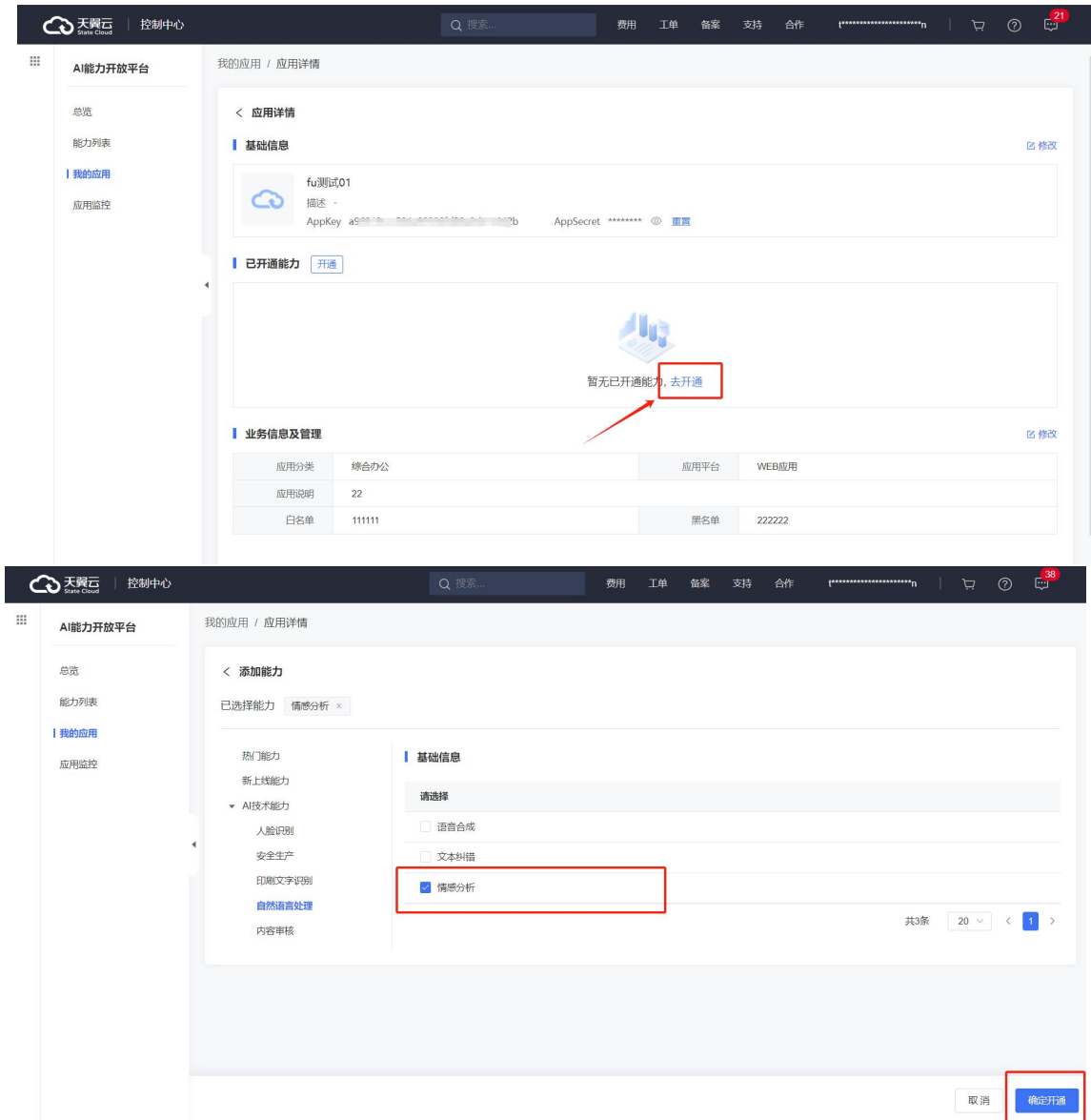
3. 再次弹出小弹窗，提示应用创建成功，请保存好AppKey和AppSecret，点击【知道了】即新应用创建成功。



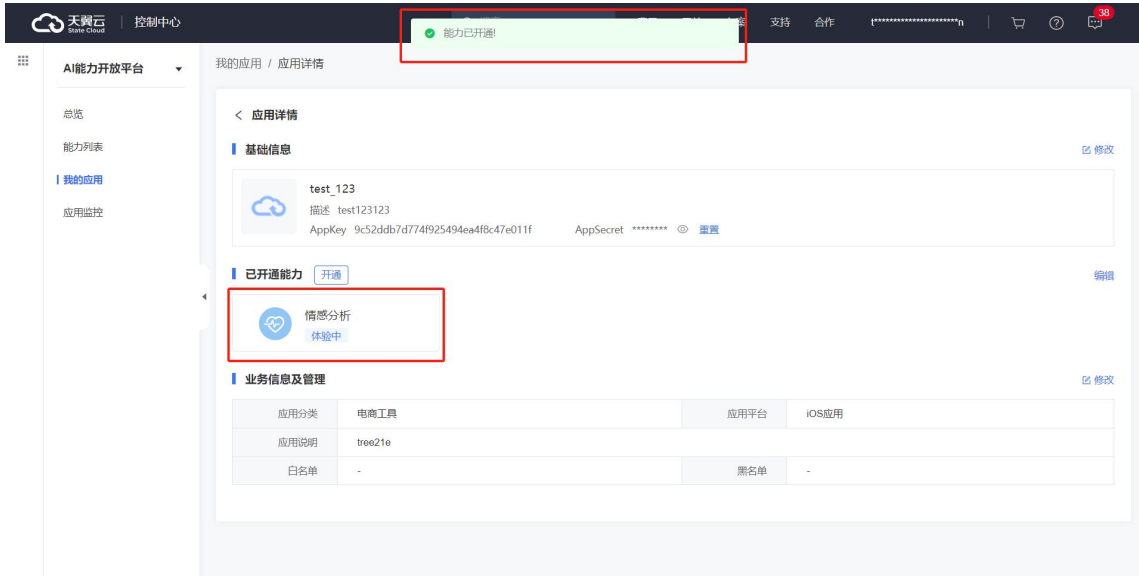
4. 根据需要选择应用, 点击【开通】, 跳转到应用详情页。



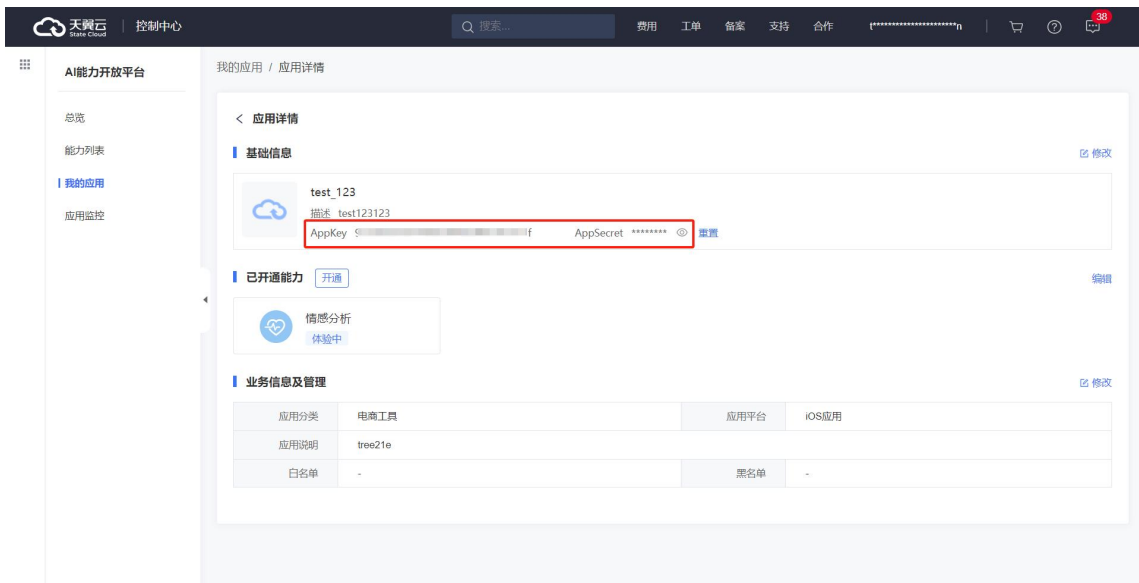
5. 点击【去开通】, 添加能力。



6. 点击确认开通，即反馈【能力已开通】，并返回应用详情，在能力管理中展示了已开通的能力。



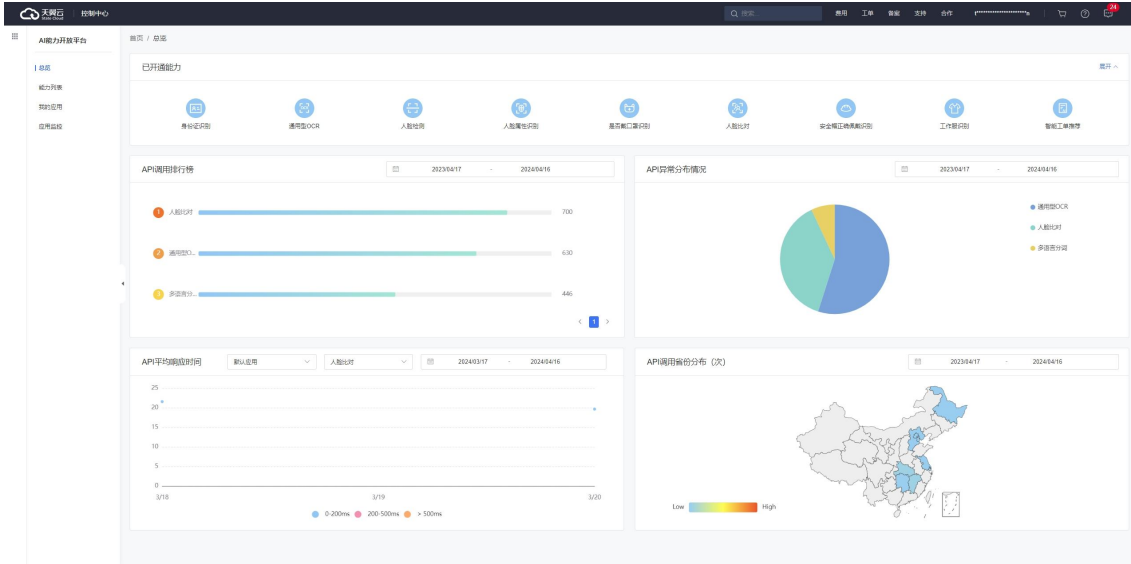
7. 点击应用，在应用详情页可以查看对应的AppKey、AppSecret。



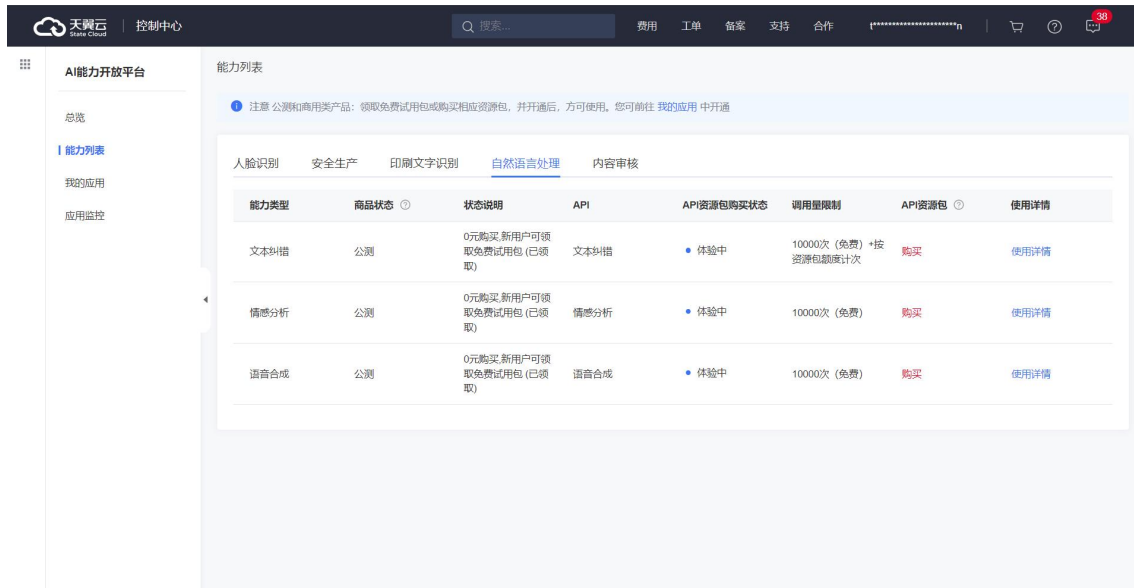
3.3. 自然语言控制台

1、入口：点击产品详情页左上角的【管理控制台】，页面跳转到控制台总览页面。

2、总览：点击左侧菜单栏【总览】，可以查看已开通能力、能力调用排行榜以及平均响应时间等内容。

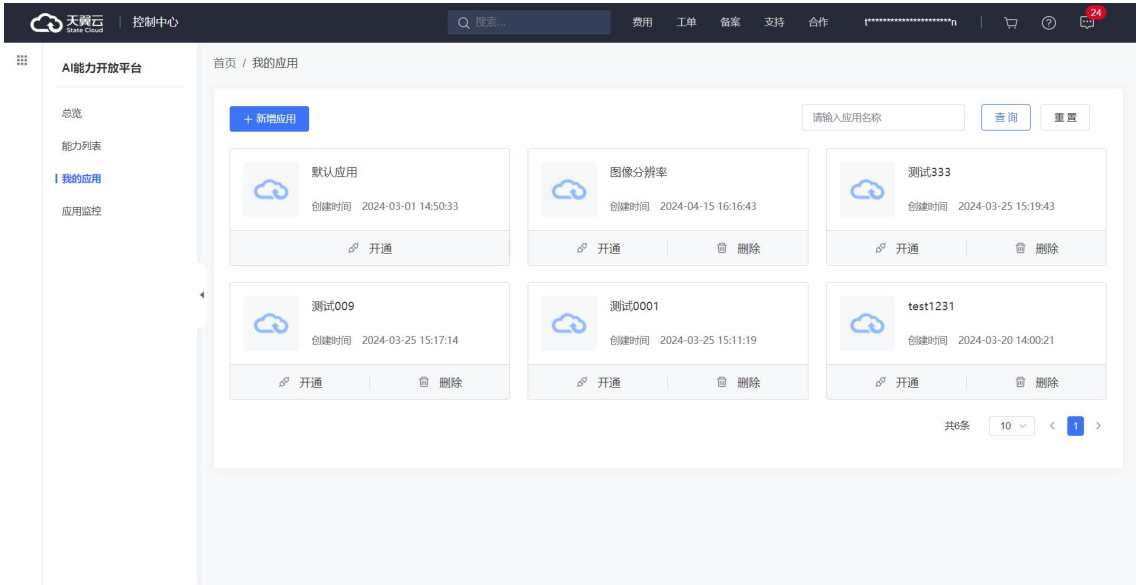


3、能力列表：点击左侧菜单栏【能力列表】，可以查看已有能力。

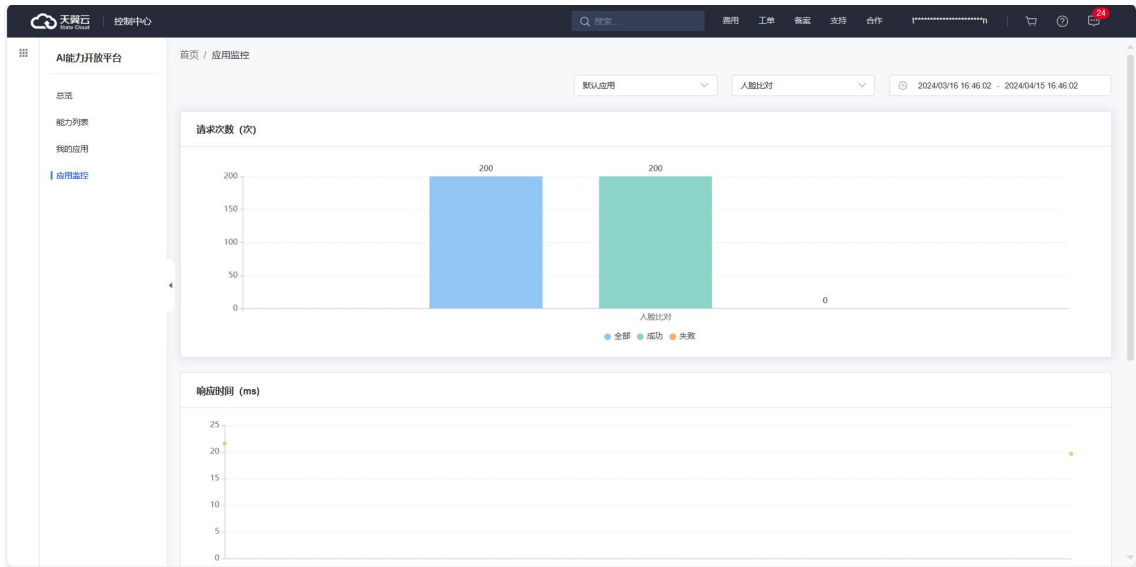


能力类型	商品状态	状态说明	API	API资源包购买状态	调用量限制	API资源包	使用详情
文本纠错	公测	0元购买,新用户可领取免费试用包(已领取)	文本纠错	体验中	10000次(免费)+按资源包超额计费	购买	使用详情
情感分析	公测	0元购买,新用户可领取免费试用包(已领取)	情感分析	体验中	10000次(免费)	购买	使用详情
语音合成	公测	0元购买,新用户可领取免费试用包(已领取)	语音合成	体验中	10000次(免费)	购买	使用详情

4、我的应用：点击左侧菜单栏【我的应用】，可以查看已经创建的应用。



5、应用监控：点击左侧菜单栏【应用监控】，可以查看所创建应用请求次数、响应时间与请求流量。



4 API 参考

4.1 API 概览

4.1.1 概述

本说明提供了自然语言处理产品 API 的描述、语法、参数说明及示例等内容。

4.1.2 API 概览

类型	描述
情感分析	用于自动分析用户评论的情感倾向性，并给出评论观点的情感极性类别（积极或消极）
文本纠错	用于自动识别并纠正一段文本中的字词语法错误
商品评价解析	智能解析评论文本内容，将文本转化为结构化的属性字段，自动甄别正负面评论观点标签
语音合成	语音合成（Text To Speech, TTS）将文本转成拟人化的语音。支持中文语音合成，提供男、女两种音色的选择，支持自定义语调、语速等参数

4.1.3 状态码

1、请求状态码

正常状态码	描述
-------	----

200	请求成功
3**	请求转移
4**	客户端错误
5**	服务端错误

2、全局请求返回错误码

错误码	描述
10002	生成签名时官网 ak 信息错误
10020	签名错误
40002	缺少 appkey 头
40006	无效的 appkey
40008	不支持的请求类型
40009	IP 未被 APP 授权
40010	IP 未被 API 授权
50000	服务内部错误
50001	服务未注册
50002	应用未开通

50003	API 中无效的 URL 请求
51001	购买服务已过期
51002	收费 API 未购买
51003	API 可用次数已不足

4.2 如何调用 API

4.1.1 终端节点

- 主要作用：用户信息的发送和接收，信令信息的控制处理、安全保护等作用。
- 终端节点地址：<https://ai-global.ctapi.ctyun.cn>。

4.1.2 构造请求

请求地址：{终端节点地址}+{对应接口 URI}

请求头：

Key	Value(说明)
Content-Type	application/json
ctyun-eop-request-id	用户请求 id，通过 uuid 生成，形如 33dfa732-b27b-464f-b15a-21ed6845afd5
Eop-Authorization	由天翼云官网 accessKey 和 securityKey 经签名后生成，签名逻辑详见后续说明
eop-date	请求时间，形如 yyyymmddTHHMMSSZ，例如20211221T163014Z
host	终端节点域名

appkey

控制台-我的应用中每个应用具有的 AppKey 信息，鉴权时需要加入header

4.1.3 认证鉴权

信息的获取

云网平台获取

登录云网门户，在“控制台” -> “个人中心” -> “安全设置” -> “用户 AccessKey” 点击“查看”获取。

如之前没有获取，请点击“新建”。



用户AccessKey

支持新建、查询用户AccessKey，用于第三方控制管理



基本签名流程

ctyun-eop-ak/ctyun-eop-sk 基本签名流程

- 1、待签字符串：使用规范请求和其他信息创建待签字符串；
- 2、计算密钥：使用 HEADER、ctyun-eop-sk、ctyun-eop-ak 来创建 Hmac 算法的密钥；
- 3、计算签名：使用第三步的密钥和待签字符串在通过 hmacsha256 来计算签名。
- 4、签名应用：将生成的签名信息作为请求消息头添加到 HTTP 请求中。

创建待签名字符串

待签名字符串的构造规则如下：

待签名字符串 = 需要进行签名的 Header 排序后的组合列表 + "\n" + 排序的 query + "\n" + toHex(sha256(原封的 body))

需要进行签名的 Header 排序后的组合列表 (排序的 header)	header 以 header_name:header_value 来一个一个通过\n 拼接起来, EOP 是强制要求 ctyun-eop-request-id 和 eop-date 这个头作为 Header 中的一部分, 并且必须是待签名 Header 里的一个。需要进行签名算法的 Header 需要进行排序 (将它们的 header_name 以 26 个英文字母的顺序来排序), 将排序后得到的列表进行遍历组装成待签名的 header
排序的 query	query 以 & 作为拼接, key 和值以 = 连接, 排序规则使用 26 个英文字母的顺序来排序, Query 参数全部都需要进行签名
toHex(sha256(原封的 body))	传进来的 body 参数进行 sha256 摘要, 对摘要出来的结果转十六进制

排序的 header 例子:

- 假设你需要将 ctyun-eop-request-id、eop-date、host 都要签名, 则待签名的 header 构造出来是:

```
ctyun-eop-request-id:123456789\n
eop-date:20210531T100101Z\n
host:1.1.1.1:9080\n
```

ctyun-eop-request-id、eop-date 和 host 的排序就是这个顺序。

- 如果你加入一个 ccad 的 header; 同时这个 header 也要是进行签名, 则待签名的 header 组合:

```
ccda:123\n
ctyun-eop-request-id:123456789\n
eop-date:20210531T100101Z\n
host:1.1.1.1:9080\n
```

构造动态密钥

发起请求时, 需要构造一个 eop-date 的时间, 这个时间的格式是

yyyymmddTHHMMSSZ; 言简意赅一些, 就是年月日T时分秒Z。

- 1、先是拿你申请来的 ctyun-eop-sk 作为密钥, eop-date 作为数据, 算出 ktime。
- 2、拿 ktime 作为密钥, 你申请来的 ctyun-eop-ak 数据, 算出 kAk。
- 3、拿 kAk 作为密钥, eop-date 的年月日值作为数据, 算出 kdate。

eop-date	yyyymmddTHHMMSSZ (20211221T163614Z) (年月日 T 时 分秒 Z)
Ktime	使用 ctyun-eop-sk 作为密钥, eop-date 作为数据, 算出 ktime; Ktime = hmacSha256(ctyun-eop-sk, eop-date)
kAk	使用 ktime 作为密钥, 你申请来的 ctyun-eop-ak 数据, 算出 kAk; kAk = hmacsha256(ktime,ctyun-eop-ak)
kdate	使用 kAk 作为密钥, eop-date 的年月日值作为数据; 算出 kdate; kdate = hmacsha256(kAk, eop-date)

签名应用及示例

由“构造动态密钥”和“创建待签名字符串”分别的出来的待签名字符串 string_sigture、kdate 生成出 Signature。

Signature	待签名字符串 string_sigture、kdate; 再根据 hmacsha256(kdate,string_sigture)得出的结果, 再将结果进行 base64 编码得出 Signature
Eop-Authorization	ctyun-eop-ak Header=你构造待签名字符串时的 header 排序 Signature (ctyun-eop-ak 后及 Signature 都有一个空格) header 排序以分号“;” 拼接例子所述: 你待签名的字符串 header 顺序是 eop-date 和 host, 那么你加到 header 里的值就是 Eop-Authorization: ctyun-eop-ak Headers =eop-date; host Signature=xad01/ada

由上得到 Eop-Authorization , 然后将数据整合成 HEADER 放在 http_client 内 , 发出即可。

http_client 所需请求头部如下:

Eop-Authorization: ctyun-eop-ak Headers = ctyun-eop-request-id;eop-date
Signature =xad01/ada

eop-date:20211221T163614Z

ctyun-eop-request-id: 123456789

(注：若需要进行签名的 Header 不止默认的 ctyun-eop-request-id 和 eop-date，需要在 http_client 的请求头部中加上，并且 Eop-Authorization 中也需要增加。)

4.1.4 Python3调用示例

```
import hmac
import base64
import hashlib
import json
import time
import uuid
import requests

from urllib.parse import urlparse

def sha256(content):
    x = hashlib.sha256()
    x.update(content.encode())
    return x.hexdigest().upper()

def hmac_sha256(key, content):
    sign = hmac.new(key, content, digestmod="sha256").digest()
    ret = base64.b64encode(sign)
    return ret

# 计算签名

def get_signature(ak, sk, app_key, params):
    # 创建待签名字符串

    # 一、header部分

    # 主要包括3个header需要作为签名内容：appkey、ctyun-eop-request-id、eop-date

    # 1. 首先通过uuid生成ctyun-eop-request-id
    request_id = str(uuid.uuid1())

    # 2. 获取当前时间戳并对时间进行格式化
    now_time = time.localtime()
    eop_date = time.strftime("%Y%m%dT%H%M%S", now_time)
    eop_date_simple = time.strftime("%Y%m%d", now_time)

    # 3. 对header部分按照字母顺序进行排序并格式化
    camp_header = "appkey:{0}\nctyun-eop-request-id:{1}\neop-date:{2}\n".format(app_key, request_id, eop_date)
```

```
# 二、query部分
# 对url的query部分进行排序
parsed_url = urlparse(request_url)
query = parsed_url.query
query_params = sorted(query.split("&"))
after_query = ""
for query_param in query_params:
    if len(after_query) < 1:
        after_query += query_param
    else:
        after_query += "&" + query_param

# 三、body参数进行sha256摘要
# sha256 body
content_hash = sha256(json.dumps(params)).lower()
# 完成创建待签名字符串
pre_signature = camp_header + "\n" + after_query + "\n" + content_hash
# 构造动态密钥
k_time = hmac_sha256(sk.encode("utf-8"), eop_date.encode("utf-8"))
k_ak = hmac_sha256(base64.b64decode(k_time), ak.encode("utf-8"))
k_date = hmac_sha256(base64.b64decode(k_ak), eop_date_simple.encode("utf-8"))
# 签名的使用
signature = hmac_sha256(base64.b64decode(k_date), pre_signature.encode("utf-8"))
# 将数据整合得到真正的header中的内容
sign_header = "{0} Headers=appkey;ctyun-eop-request-id;eop-date Signature={1}".format(ak, signature.decode())
# 返回request-id eop-date和sign_header
return request_id, eop_date, sign_header

# 向服务发送请求
def do_post(url, headers, params):
    response = requests.post(url, data=json.dumps(params), headers=headers)
    try:
        print(response.status_code)
        print(response.json())
    except AttributeError:
        print("请求失败")

if __name__ == '__main__':
    # 请求地址
    request_url = "https://ai-global.ctapi.ctyun.cn/v1/aiop/api/2f6hqix09mv4/face/PERSON/person/detectFaceFromBase64"
    # 官网accessKey
    ctyun_ak = accessKey
```

```
# 官网securityKey
ctyun_sk = 'securityKey'
# 控制台-我的应用中获取的appKey
ai_app_key = 'appKey'
# body内容从本地文件中获取
# 打开图片文件
f = open(r'test.jpeg', 'rb')
img_base64 = base64.b64encode(f.read()).decode()
# body内容
params = {"imageContent": img_base64}
# 调用get_signature方法获取签名
request_id, eop_date, sign_header = get_signature(ctyun_ak, ctyun_sk, ai_app_key, params)
# 生成请求header
# 请求header
headers = {
    'Content-Type': 'application/json;charset=UTF-8',
    'ctyun-eop-request-id': request_id,
    'appkey': ai_app_key,
    'Eop-Authorization': sign_header,
    'Eop-date': eop_date,
    'Host': 'ai-global.ctapi.ctyun.cn'
}
print("请求头部:")
print(headers)
# 执行post请求
do_post(request_url, headers, params)
```

4.3 API

4.3.1 情感分析

1、接口描述

用于分析用户评论的情感倾向性，并给出评论观点的情感极性类别（积极或消极），目前仅支持中文识别。

2、请求方法

POST

3、接口要求

- 文本长度不超过256。
- 中文字符占比需高于50%，这里的中文字符仅包括汉字，不包括中英文标点、数字、特殊符号、英文等。

4、请求 URL

/v1/aiop/api/2gyv07uqvroc/text/textanalyse/api/sentiment

5、请求参数

请求头 header 参数

参数	是否必填	参数类型	说明	示例	下级对象
Content-Type	是	string	json格式	application/json	
appkey	是	string	应用appkey	562b89493b1a40e1b97ea05e50dd8170	
ctyun-eop-request-id	是	string	用户请求 id, 由用户构造, 用户可以通过 uuid 等方法自行生成唯一字符串, 用于日志请求追踪。	3dfa732-b27b-464f-b15a-21ed6845afd5	
eop-date	是	string	请求时间, 由用户构造, 形如 yyyymmddTHHMMSSZ	20211221T163014Z	
host	是	string	终端节点域名, 固定字段	ai-global.ctapi.ctyun.cn	
Eop-Authorization	是	string	由天翼云官网 accessKey 和 securityKey 经签名后生成, 参与签名生成的字段包括天翼云官网 accessKey、securityKey、平台的 appkey (非必须), 用户请求 id (非必须), 请求时间, 终端节点域名 (非必须) 以及请求体内容		

请求体 body 参数

参数	是否必填	参数类型	说明	示例	下级对象
data	是	string	待分析的文本字符串	今天天气不错	

6、请求代码示例

```
Curl -X POST
"https://ai-global.ctapi.ctyun.cn/v1/aiop/api/2gyv07uqvroc/text/textanalyse/api/sentiment " -H "Content-Type: application/json"
-H "ctyun-eop-request-id:33dfa732-b27b-464f-b15a-21ed6845afd5" -H "AppKey:XXX"
-H "Eop-Authorization:XXX"
-H "Eop-date:20211109T104641Z"
-H "Host:ai-global.ctapi.ctyun.cn"
--data '{"data":"今天天气不错"}
```

7、返回值说明

请求成功返回响应参数

参数	是否必填	参数类型	说明	示例	下级对象
code	是	int	返回状态, 返回 0 表示成功, 返回错误代码 参考下面的错误代码列表。	0	
message	是	string	如果 code 为 0, 返回 success; 如果 code 非 0, 则返回对应的可读错误信息。	success	
returnObj	是	list	返回的情感分析结果对象		returnObj

表表 returnObj

参数	是否必填	参数类型	说明	示例	下级对象
----	------	------	----	----	------

pro_neg	是	float	请求文本所表达的情绪是负面的概率	0.5	
pro_pos	是	float	请求文本所表达的情绪是正面的概率	0.5	

请求失败返回响应参数

参数	是否必填	参数类型	说明	示例	下级对象
statusCode	是	int	接口状态码	500001	
error	是	string	返回错误码，代码参考下面的错误代码列表	'AI_OP_500001'	
message	是	string	返回对应的错误信息	"服务接口异常，请联系管理员"	
details	是	string	返回对应的错误描述	"需要联系管理员处理"	

8、返回值示例

请求成功返回值示例

```
{
  "statusCode": 0,
  "message": "success",
  "returnObj": {
    "pro_neg": 0.0033719746861606,
    "pro_pos": 0.9966280460357666,
  }
}
```

请求失败返回值示例

```
{
  "statusCode": 500001,
  "error": 'AI_OP_500001',
  "message": "服务接口异常，请联系管理员",
  "details": "需要联系管理员处理"
}
```

9、状态码

状态码	描述
200	表示请求成功。

10、错误码说明

6位错误码。4 开头为业务错误码，5 开头为服务错误码。

错误码	错误信息	错误描述
AI_OP_400001	请求路径错误	请求路径错误
AI_OP_400002	请求方法错误	请求方法错误，请使用 POST 请求
AI_OP_400003	请求体内容为空	请求体请求数据为空，没有包含内容
AI_OP_400004	请求体非 json 格式	请求体内容需要符合 json 要求
AI_OP_400005	请求体类型错误	请求体需为字典，不能为其他类型
AI_OP_400006	必传的参数未传	必须的参数 (data) 未传
AI_OP_400008	请求体的参数字段类型错误	data 字段应该是 string 类型
AI_OP_400009	请求体的参数字段值为空	data 字段值为空
AI_OP_420001	文本长度超过限制	文本输入过长，请参考接口文档说明
AI_OP_420003	中文占比过低	文本中的中文（仅包含汉字，不包含任何数字、符号）占比不低于 50%
AI_OP_500001	服务接口异常，请联系管理员	需要联系管理员处理

11、base64 编码规则：使用常规的 safe base64 编码方式

- python 中推荐使用 `base64.urlsafe_b64encode()` 函数进行编码。
- java 中推荐使用 `BASE64.getUrlEncoder().encodeToString()` 函数进行编码。

4.3.2 文本纠错

1、接口描述

用于自动识别并纠正一段文本中的音似字、形似字错误

2、请求方法

POST

3、接口要求

- 文本长度不超过1000个字符。
- 仅支持简体中文和繁体中文文本，如果文本内容超过50%非中文，则返回错误，具体见错误码4005对应内容。
- 对于未超过50%的非中文，自动略过，不予纠错。
- 不支持空字符串或者全为空格以及转义字符等。

4、请求 URL

`/v1/aiop/api/2grogewxrdwk/nlp/text-correct`

5、请求参数

请求头 header 参数

参数	是否必填	参数类型	说明	示例	下级对象
Content-Type	是	string	json格式	application/json	
appkey	是	string	应用appkey	562b89493b1a40e1b97ea05e50dd8170	
ctyun-eop-request-id	是	string	用户请求 id, 由用户构造, 用户可以通过 uuid 等方法自行生成唯一字符串, 用于日志请求追踪。	3dfa732-b27b-464f-b15a-21ed6845afd5	

eop-date	是	string	请求时间, 由用户构造, 形如 yyyymmddTHHMMSSZ	20211221T163014Z	
host	是	string	终端节点域名, 固定字段	ai- global.ctapi.ctyun.c n	
Eop- Authorization	是	string	由天翼云官网 accessKey 和 securityKey 经签名后生成, 参与签 名生成的字段包括天翼云官网 accessKey、securityKey、 appkey (非必须), 用户请求 id (非必须), 请求时间, 终端节点 域名 (非必须) 以及请求体内容		

请求体 body 参数

参数	是否必填	参数类型	说明	示例	下级对象
data	是	string	待处理文本	测试文本	

6、请求代码示例

```

Curl -X POST
"https://ai-global.ctapi.ctyun.cn/v1/aiop/api/2grogewxrdwk/nlp/text-correct"
-H "Content-Type: application/json"
-H "ctyun-eop-request-id:33dfa732-b27b-464f-b15a-21ed6845afd5"
-H "appkey:XXX"
-H "Eop-Authorization:XXX"
-H "eop-date:20211109T104641Z"
-H "host:ai-global.ctapi.ctyun.cn"
--data '{"data": "我们是中国点信的源工"}'

```

7、返回值说明

请求成功返回响应参数

参数	是否必填	参数类型	说明	示例	下级对象
code	是	string	返回状态, 返回 0 表示成功, 返回错误代码 参考下面的错误代码列表。	0	
message	是	string	如果 code 为 0, 返回 success; 如果 code 非 0, 则返回对应的可读错误信息。	success	
returnObj	是	object	返回对象		returnObj

表 returnObj

参数	是否必填	参数类型	说明	示例	下级对象
corrected	是	String	纠正后的文本	我们是中国电信的员工	
details	是	List	具体错误词以及对应索引区间和正确词	[['我们', '我们', 0, 2], ['点信', '电信', 5, 7], ['源工', '员工', 8, 10]]	

请求失败返回响应参数

参数	是否必填	参数类型	说明	示例	下级对象
statusCode	是	Int	错误码, 放置API对应的错误码	5000	
message	是	String	返回对应的错误信息	error	
details	是	String	返回对应的错误描述	服务内部错误	
error	是	String	返回对应的错误码	AI_OP_5000	

8、返回值示例

请求成功返回值示例

```
{
  "code": 0,
  "message": "成功",
  "success": true, "result": [
    {
      "corrected": "我们是中国电信的员工",
      "details": [
        [
          "我们",
          "我们",
          0,
          2
        ],
        [
          "点信",
          "电信",
          5,
          7
        ]
      ]
    },
    [
      "源工",
      "员工",
      8,
      10
    ]
  ]
}
```

请求失败返回值示例

```
{
  "statusCode": 4001,
  "message": "error",
  "details": "请求未包含内容",
  "error": "AI_OP_4001"
}
```

9、状态码

状态码	描述
200	表示请求成功

10、错误码说明

4 位错误码。4 开头为业务错误码，5 开头为服务错误码。

错误码	错误信息	错误描述
AI_OP_4001	请求未包含内容	请求体中存在空字符串或者全为空格等转义字符
AI_OP_4002	请求参数异常	请求的参数格式不正确或者参数名错误
AI_OP_4003	单次批量请求数量超过指定数目	传入了多段文本，目前仅支持单段文本
AI_OP_4004	单条文本长度超过限度	传递的文本中存在超过规定最大长度的文本，最大长度为1000
AI_OP_4005	请求文本为非法字符串	传递的文本中存在大量的非中文字符，非中文字符占比不能超过50%
AI_OP_5000	服务内部错误	服务内部发生错误，需要联系管理员处理

11、base64 编码规则：使用常规的 safe base64 编码方式

- python 中推荐使用 `base64.urlsafe_b64encode()` 函数进行编码。
- java 中推荐使用 `BASE64.getUrlEncoder().encodeToString()` 函数进行编码。

4.3.4 商品评价解析

1、接口描述

智能解析评论文本内容，将文本转化为结构化的属性字段，自动甄别正负面评论观点标签

2、请求方法

POST

3、接口要求

- 文本长度不超过512。

- 中文字符占比需高于50%，这里的中文字符仅包括汉字，不包括中英文标点、数字、特殊符号、英文等。

4、请求 URL

/v1/aiop/api/2gyuqiiopt6o/text/textanalyse/api/textopinion

5、请求参数

请求头 header 参数

参数	是否必填	参数类型	说明	示例	下级对象
Content-Type	是	string	json格式	application/json	
appkey	是	string	应用appkey	562b89493b1a40e1b97ea05e50dd8170	
ctyun-eop-request-id	是	string	用户请求 id，由用户构造，用户可以通过 uuid 等方法自行生成唯一字符串，用于日志请求追踪。	3dfa732-b27b-464f-b15a-21ed6845afd5	
eop-date	是	string	请求时间，由用户构造，形如 yyyymmddTHHMMSSZ	20211221T163014Z	
host	是	string	终端节点域名，固定字段	ai-global.ctapi.ctyun.cn	
Eop- Authorization	是	string	由天翼云官网 accessKey 和 securityKey 经签名后生成，参与签名生成的字段包括天翼云官网 accessKey、securityKey、平台的 appkey（非必须），用户请求 id（非必须），请求时间，终端节点域名（非必须）以及请求体内容		

请求体 body 参数

参数	是否必填	参数类型	说明	示例	下级对象

data	是	Object	由评价文本内容和评价文本所属类别构成的dict		data
------	---	--------	-------------------------	--	------

表data

参数	是否必填	参数类型	说明	示例	下级对象
text	是	String	评价文本内容	环境很不错...	data
cate	是	String	评价文本所属类别，目前适用的类别有hotel、catering、shopping三类	catering	

7、请求代码示例

```
Curl -X POST
"https://ai-global.ctapi.ctyun.cn/v1/aiop/api/2gyuqiiopt6o/text/textanalyse/api/textopinion"
-H "Content-Type: application/json"
-H "ctyun-eop-request-id:33dfa732-b27b-464f-b15a-21ed6845afd5"
-H "AppKey:XXX"
-H "Eop-Authorization:XXX"
-H "Eop-date:20211109T104641Z"
-H "Host:ai-global.ctapi.ctyun.cn"
--data '{"data":{"text":"环境很不错，不愧是老店~服务热情，倒车停车还有工作人员主动帮忙，三文鱼尤其好吃，五星好评","cate":"catering"}}'
```

7、返回值说明

请求成功返回响应参数

参数	是否必填	参数类型	说明	示例	下级对象
----	------	------	----	----	------

code	是	Int	返回状态, 返回 0 表示成功, 返回错误代码 参考下面的错误代码列表	0	
message	是	String	如果 code 为 0, 返回 success; 如果 code 非 0, 则返回对应的可读错误信息。	success	
returnObj	是	Object	返回的商品评价文本解析的结果对象		returnObj

表 returnObj

参数	是否必填	参数类型	说明	示例	下级对象
Aspect	否	String	文本中提到的商品属性		
Opinion	否	String	文本中与商品属性相对应的评论观点		
Category	否	String	商品属性所属类别		
Polarity	否	String	与文本中提到的商品属性相对应的情感极性		

请求失败返回响应参数

参数	是否必填	参数类型	说明	示例	下级对象
statusCode	是	Int	错误码, 放置API对应的错误码	4001	
message	是	String	返回对应的错误信息	error	
details	是	String	返回对应的错误描述	请求中未包含文本	

error	是	String	返回对应的错误码	AI_OP_4001	
-------	---	--------	----------	------------	--

8、返回值示例

请求成功返回值示例

```

{
  "statusCode": 0,
  "message": "success",
  "returnObj": [
    {
      "Aspect": "服务",
      "Opinion": "好",
      "Category": "服务",
      "Polarity": "正面"
    },
    {
      "Aspect": "房费",
      "Opinion": "贵",
      "Category": "价格",
      "Polarity": "负面"
    }
  ]
}

```

请求失败返回值示例

```

{
  "statusCode": 4001,
  "message": "error",
  "details": "请求中未包含文本",
  "error": "AI_OP_4001"
}

```

9、状态码

状态码	描述
200	表示请求成功。

10、错误码说明

4 位错误码。4 开头为业务错误码，5 开头为服务错误码。

错误码	错误信息	错误描述
-----	------	------

错误码	错误信息	错误描述
AI_OP_4001	请求中未包含文本	请求体中未传入待分析的有效文本字符串
AI_OP_4002	请求参数异常	请求的参数格式不正确或者参数名错误或参数缺失，如仅传入text未传入cate
AI_OP_4004	单条文本长度超过限度	传递的文本超过规定的最大文本长度，最大长度为512
AI_OP_4005	请求文本为非法字符串	传递的文本中存在大量的非中文字符，中文字符占比需高于50%
AI_OP_5000	服务内部错误	服务内部发生错误，需要联系管理员处理

11、base64 编码规则：使用常规的 safe base64 编码方式

- python 中推荐使用 `base64.urlsafe_b64encode()` 函数进行编码。
- java 中推荐使用 `BASE64.getUrlEncoder().encodeToString()` 函数进行编码。

4.3.4 语音合成

1、接口描述

语音合成 (Text To Speech, TTS) 将文本转成拟人化的语音。目前仅支持中文语音合成，提供男、女两种音色的选择，支持自定义语调、语速等参数，更符合业务场景的需要

2、请求方法

POST

3、接口要求

- 目前仅支持 base64 编码方式请求，暂不支持 url 方式请求
- 一次请求仅支持一条数据

4、请求 URL

/v1/aiop/api/2z0yhhrzgv0g/tts/predict

5、请求参数

请求头 header 参数

参数	是否必填	参数类型	说明	示例	下级对象
Content-Type	是	string	json格式	application/json	
appkey	是	string	应用appkey	562b89493b1a40e1b97ea05e50dd8170	
ctyun-eop-request-id	是	string	用户请求 id, 由用户构造, 用户可以通过 uuid 等方法自行生成唯一字符串, 用于日志请求追踪。	3dfa732-b27b-464f-b15a-21ed6845afd5	
eop-date	是	string	请求时间, 由用户构造, 形如 yyyymmddTHHMMSSZ	20211221T163014Z	
host	是	string	终端节点域名, 固定字段	ai-global.ctapi.ctyun.cn	
Eop-Authorization	是	string	由天翼云官网 accessKey 和 securityKey 经签名后生成, 参与签名生成的字段包括天翼云官网 accessKey、securityKey、平台的 appkey (非必须), 用户请求 id (非必须), 请求时间, 终端节点域名 (非必须) 以及请求体内容		

请求体 body 参数

参数	是否必填	参数类型	说明	示例	下级对象
Action	是	String	算法名称	TTS	

TextData	是	String	合成语音的 文本 ，按 UTF-8 编码。最小支持 3 个汉字，最多支持 150 个汉字。英文、数字、标点均等同于一个字符。强烈建议输入文本加上标点符号，便于合成更正确的语音数据。仅支持 中文 语音合成。	今晚去吃火锅吗	
VoiceType	否	Int	声音类型 。目前仅支持两种声音。0 表示男生声音，1 表示女生声音。	0	
Pitch	否	Float	语调 。返回范围 [0.8, 2]，默认值为 1.0，原调。	1.0	
Speed	否	Float	语速 。返回范围 [0.5, 2]，默认值为 1.0，原速。"1.0"		

8、请求代码示例

```

Curl -X POST
  "https://ai-global.ctapi.ctyun.cn/v1/aiop/api/2z0yhhrzgv0g/tts/predict"
-H "Content-Type: application/json"
-H "ctyun-eop-request-id:33dfa732-b27b-464f-b15a-21ed6845afd5"
-H "appkey:XXX"
-H "Eop-Authorization:XXX"
-H "eop-date:20211109T104641Z"
-H "host:ai-global.ctapi.ctyun.cn"
--data '{"Action": "TTS", "TextData": "今晚去吃火锅吗"}'

```

7、返回值说明

请求成功返回响应参数

参数	是否必填	参数类型	说明	示例	下级对象
code	是	Int	返回状态，返回 0 表示成功，返回错误代码 参考下面的错误代码列表	0	
message	是	String	如果 code 为 0，返回 success；如果 code 非 0，则返回对应的可读错误信息。	success	
returnObj	是	Object	返回结果对象		returnObj

表returnObj

参数	是否必填	参数类型	说明	示例	下级对象
Audio	是	String	**语音编码数据。 **Base64 编码"_9j_4AAQSkZJRgA.....Z 的 wav 格式音频数据。	uv_9k="	

请求失败返回响应参数

参数	是否必填	参数类型	说明	示例	下级对象
statusCode	是	Int	错误码，放置API对应的错误码	500001	
message	是	String	返回对应的错误信息	服务接口异常，请联系管理员	
details	是	String	返回对应的错误描述	需要联系管理员处理	
error	是	String	返回对应的错误码	AI_OP_500001	

8、返回值示例

请求成功返回值示例

```

{
  "statusCode": 0,
  "message": "success",
  "returnObj": {
    "Audio": "_9j_4AAQSkZJRgA.....Zuv_9k="
  }
}

```

请求失败返回值示例

```

{
  "statusCode": 400003,
  "message": "Body 内容为空",
  "details": "Body 请求数据为空，没有包含内容",
  "error": "AI_OP_400003"
}

```

9、状态码

状态码	描述
200	表示请求成功。

10、错误码说明

6 位错误码。4 开头为业务错误码，5 开头为服务错误码。

错误码	错误信息	错误描述
AI_OP_400003	Body 内容为空	Body 请求数据为空，没有包含内容
AI_OP_400004	Body 请求体非 json 格式	Body内容需要符合 json 要求
AI_OP_400005	请求体类型错误	请求体需为字典，不能为其他类型
AI_OP_410001	必传的参数未传	必须的参数（Action、TextData）未传
AI_OP_410002	请求体的字段类型错误	请求体的字段（Action、TextData）类型错误，类型只能为字符串，不能为其他类型
AI_OP_400009	Action 值设置错误	Action 值设置错误
AI_OP_410003	TextData 字段值为空字符	TextData 字段值为空字符
AI_OP_410006	文本输入过短	文本输入过短，请参考接口文档说明
AI_OP_410007	文本输入过长	文本输入过长，请参考接口文档说明
AI_OP_450001	VoiceType 字段类型错误	VoiceType 字段应该是 int 类型
AI_OP_450002	VoiceType 字段不符合规范	VoiceType 字段不符合规范，请参考接口文档说明
AI_OP_450003	Pitch 字段类型错误	Pitch 字段应该是 float 类型

错误码	错误信息	错误描述
AI_OP_450004	Pitch 字段不符合规范	Pitch 字段不符合规范, 请参考接口文档说明
AI_OP_450005	Speed 字段类型错误	Speed 字段应该是 float 类型
AI_OP_450006	Speed 字段不符合规范	Speed 字段不符合规范, 请参考接口文档说明
AI_OP_500001	服务接口异常, 请联系管理员	需要联系管理员处理

11、base64 编码规则：使用常规的 safe base64 编码方式

- python 中推荐使用 `base64.urlsafe_b64encode()` 函数进行编码。
- java 中推荐使用 `BASE64.getUrlEncoder().encodeToString()` 函数进行编码。

4.4 更新历史

更新日期	更新内容
020-12-11	第一次正式发布
2022-10-25	第二次正式发布。本次更新说明如下：修改API文档格式
2022-11-03	更新了情感分析, 商品评价解析接口文档
2022-11-29	第三次正式发布。本次更新说明如下：修改API文档格式
2023-12-21	第四次正式发布。本次更新说明如下：修改API文档内容

5 常见问题

5.1 计费类

自然语言处理支持哪些计费方式？

目前自然语言处理的文本纠错、情感分析、语音合成处于公测阶段，公测期间0元购买不收费。

除以上能力外的其他自然语言处理的能力目前是处于限时免费的状态，无需购买，只需开通即可使用。

开通应用可以参考自然语言处理产品文档中的创建应用及开通应用。

调用量在哪里查看？

调用明细您可在控制台内查看，进入控制台选择【能力列表】里面对应能力的条目，点击【使用详情】按钮即可查看。

查看路径：控制台—能力列表—使用详情。

查看步骤：

- 在产品详情页，点击【管理控制台】进入控制台。
- 点击左侧【能力列表】，选择希望查看的原子能力。
- 点击【使用详情】按钮即可查看已调用量、剩余可调用量、有效期。

调用量的扣费顺序是？

调用量的扣减顺序为：免费试用包-付费资源包，购买资源包后，将按照资源包下单顺序抵扣额度。

如果当前扣减额度为付费资源包，而后又领取了免费试用包，则当前付费资源包使用完毕后，优先抵扣免费试用包，然后按照资源包下单顺序抵扣。

在购买之前，可先领取免费试用包，待体验过后再自行购买。

什么情况下会扣资源包次数，识别报错会扣次数吗？

资源包抵扣规则按照：成功调用算入计费次数。

若因网络错误未成功调用则不计算次数。

自然语言处理服务剩余次数如何查询？

如您已购买自然语言处理资源包，资源包剩余次数可至控制台内查看，进入控制台选择对应能力的条目，点击“使用详情”按钮查询；

注意：如您购买的资源包已过期，会显示已过期；如果资源包已使用完毕，则显示已用完。

自然语言处理资源包可以转移到别人的账号么？

目前不支持转移，建议购买前评估真实使用量，资源包用完可以叠加购买。

若后期支持转移功能，我们会发公告通知，请您留心注意。

自然语言处理服务开通后能否关闭？不使用是否会产生费用？

目前自然语言处理服务均处于公测或者限免阶段，不会产生费用。

服务开通后无法关闭，如您后续无使用需求，不调用已开通的接口即可。

同时，请您保管好AppKey和AppSecret信息，保证后期无人调用。

5.2 购买类

自然语言处理服务是否支持试用？

当前自然语言处理的文本纠错、情感分析、语音合成处于公测阶段，其他产品属于限时免费状态。

限时免费产品可以直接开通应用进行试用。

公测产品可以在控制台领取免费试用包进行试用。

自然语言处理是否支持续订？

资源包是一次性产品，不支持标准续订操作。

当已订购的资源包订单即将到期或即将用完时，可通过订购新的资源包进行续订，按照次数包下单顺序抵扣额度。

资源包买错了可以退款吗？

若因产品介绍说明有误导导致的购买错误且资源包尚未使用，您可通过提交天翼云官网工单或者拨打客服电话【400-810-9889】联系我们，申请手动退款款项会原路退回，可能需要数个工作日审批，请耐心等待。在审批期间，如您急需使用其他产品，可以先购买相应资源包。

资源包使用到一半可以退吗？

目前自然语言处理产品处于公测或者限免状态，不会计费。如若不想再使用，停止调用服务即可。

若自然语言产品商用后，订购的资源包一经售卖且产生正常调用抵扣，不允许退款。

资源包属于预付费产品，建议您在购买资源包的时候按实际调用情况预估调用量以免造成资源浪费，谢谢配合。

5.3 操作类

自然语言处理服务在什么时候生效？

自然语言处理产品目前处于公测或者限免状态。

限免状态产品开通应用后即可进行服务测试。

公测产品在领取或者订购资源包，并且开通应用后即可使用服务。

自然语言处理服务的并发是多少？

默认支持5个QPS，建议您在程序中进行一定的请求限制，避免收到大量限流报错。

如果因业务需要QPS超过5个，请提前线下咨询沟通再购买下单。

注意：如果您的程序在失败时有重试机制，当您扩大并发量后接口返回错误码时，请不要重试，否则可能加重限流报错情况。

5.4 使用限制

情感分析服务对于文本格式是否有要求？

具体情感分析服务文本格式要求也可以参考情感分析API。

- 文本长度不超过256。
- 中文字符占比需高于50%，这里的中文字符仅包括汉字，不包括中英文标点、数字、特殊符号、英文等。

文本纠错服务对于文本格式是否有要求？

- 具体文本纠错服务文本格式要求也可以参考文本纠错API。
- 文本长度不超过1000个字符。

-
- 仅支持简体中文和繁体中文文本，如果文本内容超过50%非中文，则返回错误，具体见错误码4005对应内容。
 - 对于未超过50%的非中文，自动略过，不予纠错。
 - 不支持空字符串或者全为空格以及转义字符等。

商品评价解析服务对于文本格式是否有要求？

具体商品评价解析服务文本格式要求也可以参考商品评价解析API。

- 文本长度不超过512。
- 中文字符占比需高于50%，这里的中文字符仅包括汉字，不包括中英文标点、数字、特殊符号、英文等。

语音合成服务对于文本格式是否有要求？

具体语音合成服务文本格式要求也可以参考语音合成API。

- 目前仅支持 base64 编码方式请求，暂不支持 url 方式请求。
- 一次请求仅支持一条数据。

自然语言处理产品是HTTP GET请求还是HTTP POST请求？

自然语言处理产品以API的方式提供服务，支持HTTP POST请求。

HTTP POST提供了加密传输、身份验证和授权以及请求参数验证和过滤等机制，从而可以更好的保障用户的数据安全。

自然语言处理能否提供100%识别准确率？

自然语言处理准确率与上传的文本内容质量相关，同时也存在一定概率的误差，无法做到100%识别准确率。

如您对当前使用的自然语言处理产品服务有识别准确率相关问题，您可通过提交天翼云官网工单或者拨打客服电话【400-810-9889转1】联系我们。

QPS不够用怎么办？

默认支持5个QPS，建议您在程序中进行一定的请求限制，避免收到大量限流报错。

如果因业务需要QPS超过5个，请提前线下咨询沟通再购买下单。

注意：如果您的程序在失败时有重试机制，当您扩大并发量后接口返回错误码时，请不要重试，否则可能加重限流报错情况。

自然语言处理服务是否支持私有化部署？

自然语言处理支持公有云、私有化部署。

您可通过提交天翼云官网工单或者拨打客服电话【400-810-9889转1】沟通私有化部署相关合作。

注意

如用户操作手册说明与官网帮助中心说明不一致，以官网帮助中心说明为准。