

天翼云自然语言处理产品 用户操作手册

中国电信股份有限公司云计算分公司

目 录

1 产品介绍	1
1.1 产品简介	1
1.2 产品功能	1
1.3 使用场景	1
2 定价	2
自然语言处理产品定价	2
3 快速入门	4
3.1 创建自然语言处理应用	4
3.1.1 准备工作	4
3.1.2 创建自然语言处理应用订单	5
3.1.3 订单确认	5
3.2 自然语言处理用户控制台	8
4 API 文档	10
4.1 情感分析接口说明	10
4.2 文本纠错接口说明	14
4.3 智能工单推荐	18
5 常见问题	23

修订历史

版本号	日期	修订人	修订内容	修订原因
V1.0	2020.10		初版	

1 产品介绍

1.1 产品简介

天翼云自然语言处理产品可帮助用户进行情感分析、文本纠错与智能工单推荐。

1.2 产品功能

- **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。

- **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。

- **智能工单推荐：**

对投诉类、网络故障类等历史工单进行阅读，将与当前工单相似的历史工单推荐给用户，以达到故障处理借鉴的效果。例如，对于“上海用户拨打本地电话正常，而无法拨打江苏的电话”这样的故障类工单，推荐历史工单中类似的工单给用户，从而让用户得到该故障的参考解决方案

1.3 使用场景

- **网络舆情分析**

了解大众对热门事件的情感倾向，掌握舆论导向，从而更及时有效的进行舆情监控

- **写作文本纠错**

对内容编写平台增加文本纠错功能，降低用户创作的时间成本，提升编写内容质量

- **工单处理**

帮助相关机构客服或运维部门进行投诉类、故障类工单辅助处理。

2 定价

自然语言处理产品定价

产品公测期间免费试用不收费。

自然语言处理产品采用封顶资源包的方式订购：

实例类型	示例规格	价格
情感分析	10 万次/年	待定
情感分析	40 万次/年	待定
情感分析	80 万次/年	待定
情感分析	400 万次/年	待定
情感分析	800 万次/年	待定
文本纠错	10 万次/年	待定
文本纠错	40 万次/年	待定
文本纠错	80 万次/年	待定
文本纠错	400 万次/年	待定
文本纠错	800 万次/年	待定
智能工单推荐	10 万次/年	待定
智能工单推荐	40 万次/年	待定
智能工单推荐	80 万次/年	待定
智能工单推荐	400 万次/年	待定

智能工单推荐	800 万次/年	待定
--------	----------	----

备注：产品资源包价格以价格发文为准

订购说明：

1、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 10 万次/年的 API 调用次数包和 1 个 40 万次/年的 API 调用次数包。

2. 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。

3 快速入门

3.1 创建自然语言处理应用

3.1.1 准备工作

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【大数据与AI】，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“情感分析”产品详情页。



点击【立即使用】按钮，进入自然语言处理应用创建页面，点击“新的应用”

在应用订购页面，用户根据自己需求，创建人脸识别应用订单。



3.1.2 创建自然语言处理应用订单

用户根据自身需要输入“应用名称”与“应用概述”。

应用名称：长度为 4-15 个字符。



新建应用

* 应用名称: 情感分析应用场景 ①

应用概述: 情感分析应用场景 ①

取消 确定

用户创建完应用后，点击“确定”和“下一步”，进入应用订购页面

我的能力

自然语言处理	人脸识别	安全生产	图像识别	印刷文字识别	内容安全
情感分析					
API	状态	调用量限制	购买次数包	领取免费额度	
情感分析	待付费	—	购买	立即领取	
智能工单推荐					
API	状态	调用量限制	购买次数包	领取免费额度	
智能工单推荐	待付费	—	购买	立即领取	
命名实体识别					
API	状态	调用量限制	购买次数包	领取免费额度	
命名实体识别	限时免费	—	—	—	

3.1.3 订单确认

在“情感分析”列点击“购买”，进入产品规格及数量配置页面。

情感分析

选择配置

规格	公测截止时间	单价	购买数量
10万次	2020-11-30 23:59:59	¥0.00	- 0 +
40万次	2020-11-30 23:59:59	¥0.00	- 1 +
80万次	2020-11-30 23:59:59	¥0.00	- 0 +
400万次	2020-11-30 23:59:59	¥0.00	- 0 +
800万次	2020-11-30 23:59:59	¥0.00	- 0 +

配置清单

清除

规格	购买数量	价格
40万次	1	0

费用合计

¥0元

配置费用 ¥ 0

参考价格，具体配置请以配置为准。 [了解计费详情](#)

下一步：确认配置

用户根据需求选择要开通的资源包规格及数量，点击“下一步：确认配置”，进入配置确认界面。

配置

产品名称	计费方式	规格	公测截止时间	购买数量	单价
情感分析	次数包	40万次	2020-11-30 23:59:59	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

参考价格，具体配置请以配置为准。 [了解计费详情](#)

上一步

立即购买

用户勾选服务协议，点击“立即购买”，进入订单支付页面

订单管理 - 云订单 > 订单详情 > 订单支付

订单号: 20201106112141246340 产品类型: 包周期新购 创建时间: 2020-11-06 11:21:41 更新时间: 2020-11-06 11:21:41

产品1 待支付 0.00元					
产品	配置	订购数量	所属资源池	周期	金额 (元)
自然语言处理	-	x 1	—	1次	0.00

费用合计: 0.00元

郑重提示: 对云主机的使用请遵循国家相关法律法规之规定。对于违反相关法律法规的行为, 服务商将关闭服务器, 并视情况决定是否关闭用户帐号, 停止所有服务, 不退余款。弹性IP不支持降级操作, 购买时请慎重选择。

1、支付订单请务必确认所有者已进行实名认证, 如无, 请点击 [立即实名认证](#)

2、云市场订单不能使用优惠券和代金券支付

3、订单不支持同时使用代金券和优惠券

支付方式

● 预付费

账户余额支付 (元)

无可用账户余额

代金券支付 (元)

无可用代金券

优惠券

无可用优惠券

订单费用: + 0.00元

账户余额支付: - 0.00元

代金券支付: - 0.00元

优惠券优惠: - 0.00元

需支付 **0.00元**

立即支付

用户确认订单费用是否准确，费用无误点击“立即支付”。完成支付后进入“支付结果”界面。

订单管理 - 云订单 > 订单详情 > 支付结果

✔ 支付成功

返回订单列表

订单号: 20201106112141246340 产品类型: 包周期新购 创建时间: 2020-11-06 11:21:41 更新时间: 2020-11-06 11:25:16

产品1 已支付 0.00元					
产品	配置	订购数量	所属资源池	周期	金额 (元)
自然语言处理	-	x 1	—	1次	0.00

费用合计: 0.00元

订单费用: + 0.00元

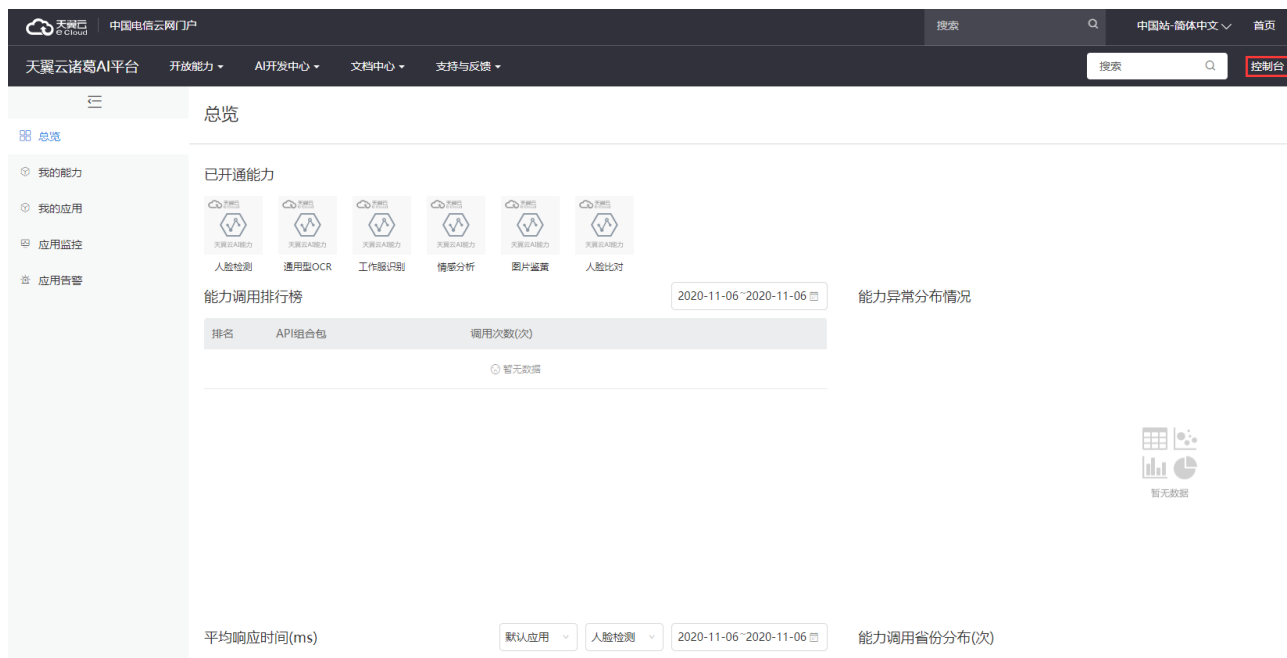
账户余额支付: - 0.00元

代金券支付: - 0.00元

实际支付: **0.00元**

3.2 自然语言处理用户控制台

点击页面右上角“控制台”，进入控制中心。



用户点击“我的应用”，可以查看已经开通的应用。



用户点击“应用监控”，可以查看所创建应用的请求次数、响应时间与请求流量。



用户点击“应用告警”，可以了解应用使用过程中系统所反馈的异常情况。



4 API 文档

4.1 情感分析接口说明

接口描述

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）

接口要求

文本格式：

- 1. 单段文本长度不超过 256，编码格式为 UTF-8 编码
- 2. 仅支持简体中文

请求说明

请求示例

HTTP 方法：POST

请求 URL：[https://ai.ctyun.cn/v1/aiop/api/\[系统生成的服务码\]/text/textanalyse/api/sentiment](https://ai.ctyun.cn/v1/aiop/api/[系统生成的服务码]/text/textanalyse/api/sentiment)

请求头说明

参数	值(说明)
Content-Type	application/json
Sign	MD5(AppKey + AppSecret + Timestamp + Path + Http Body), 32 位小写的 MD5 加密串
Timestamp	13 位时间戳

AppKey	平台分配给应用独立的 AppKey
--------	-------------------

请求参数

参数	是否必选	类型	可选值范围	说明
data	true	string	-	请求参数放置于 http body 中，为待分析的文本字符串

请求示例代码

Bash:

```
curl -X POST https://ai.ctyun.cn/v1/aiop/api/[系统生成的服务码]/text/textanalyse/api/sentiment -H "Content-Type: application/json" -H 'Sign: XXXXX' -H 'Timestamp: XXXXX' -H 'AppKey: XXXXX' --data '{"data": "string"}
```

Python:

```
# -*- coding: utf-8 -*-
import requests
import base64
import hashlib
import time
import json

'''
文本情感分析
'''

def md5(key):
    input_name = hashlib.md5()
    input_name.update(key.encode("utf-8"))
    return input_name.hexdigest()

request_url = "文本情感分析的 API URL"
params = {"data": "待分析的文本"}
app_key = '平台分配的 APP Key'
app_secret = '平台分配的 APP Secret '
time_stamp = str(int(round(time.time() * 1000)))
```

```

headers = {
    'content-type': 'application/json',
    'Sign': md5(app_key + app_secret + time_stamp + '/v1/aiop/api/[系统生成的服务码]/text/textanalyse/api/sentiment' +
    json.dumps(params)),
    'Timestamp': time_stamp,
    'AppKey': app_key
}
response = requests.post(request_url, data=json.dumps(params), headers=headers)
if response:
    print(response.json())

```

返回说明

API 调用成功返回值

返回 字段	是否 必选	类 型	字段说明
code	true	int	返回状态，返回 0 表示成功
message	true	string	返回对应的可读消息，值为 success
result	true	object	API 具体返回信息
pro_neg	true	float	文本中消极倾向概率
pro_pos	true	float	文本中积极倾向概率

API 调用失败返回值

返回 字段	是否 必选	类型	字段说明
code	true	int	错误码，参见本文档错误码说明部分
message	true	string	返回对应的可读消息，值为 error

details	true	string	错误消息
---------	------	--------	------

返回示例

HTTP/1.1 200 OK

Content-Type: application/json

```
{
  'code': 0,
  'message': 'success',
  'result': {
    "pro_neg" : 0.23586,
    "pro_pos" : 0.76414
  }
}
```

错误码说明

错误代码	错误消息	错误可能原因
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
4002	请求参数异常	请求的参数格式不正确或者参数名错误
4003	请求的 API URL 不存在	请求路径非指定路径
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 255
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%
4006	请求方法非指定方法	请求方法非指定的 POST 方法

5000	服务内部错误	服务内部发生错误，需要联系管理员处理
------	--------	--------------------

4.2 文本纠错接口说明

接口描述

用于自动识别并纠正一段文本中的字词语法错误

接口要求

文本格式：

1. 仅支持单段文本，长度不超过 1000
2. 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
3. 对于未超过 50%的非中文，自动略过，不予纠错

请求说明

请求示例

HTTP 方法：POST

请求 URL：[https://ai.ctyun.cn/v1/aiop/api/\[系统生成的服务码\]/nlp/text-correct](https://ai.ctyun.cn/v1/aiop/api/[系统生成的服务码]/nlp/text-correct)

请求头说明

参数	值(说明)
Content-Type	application/json
Sign	MD5(AppKey + AppSecret + Timestamp + Path + Http Body)，32 位小写的 MD5 加密串
Timestamp	13 位时间戳

AppKey	平台分配给应用独立的 AppKey
--------	-------------------

请求参数

参数	是否必选	类型	可选值范围	说明
data	true	string	-	请求参数放置于 http body 中，支持一段文本的处理

请求示例代码

Bash:

```
curl -X POST https://ai.ctyun.cn/v1/aiop/api/[系统生成的服务码]/nlp/text-correct -H "Content-Type: application/json" -H 'Sign: XXXXX' -H 'Timestamp: XXXXX' -H 'AppKey: XXXXX' --data '{"data": "string"}'
```

Python:

```
# -*- coding: utf-8 -*-
import requests
import base64
import hashlib
import time
import json

'''
文本纠错
'''
def md5(key):
    input_name = hashlib.md5()
    input_name.update(key.encode("utf-8"))
    return input_name.hexdigest()

request_url = "文本纠错的 API URL"
params = {"data": "string"}
app_key = '平台分配的 APP Key'
app_secret = '平台分配的 APP Secret '
time_stamp = str(int(round(time.time() * 1000)))

headers = {
    'content-type': 'application/json',
```

```

    'Sign': md5(app_key + app_secret + time_stamp + '/v1/aiop/api/[系统生成的服务码]/nlp/text-co
rrect' + json.dumps(params)),
    'Timestamp': time_stamp,
    'AppKey': app_key
}
response = requests.post(request_url, data=json.dumps(params), headers=headers)
if response:
    print(response.json())

```

返回说明

API 调用成功返回值

返回 字段	是否 必选	类 型	字段说明
code	true	int	返回状态，返回 0 表示成功
message	true	string	返回对应的可读消息，值为 success
result	true	object	API 具体返回信息
corrected	true	string	纠正后的文本字符串
details	true	list of list	错误的详细信息：[[错误字符，纠正字符，开始索引，结束索引]]

API 调用失败返回值

返回 字段	是否 必选	类型	字段说明
code	true	int	错误码，参见本文档错误码说明部分
message	true	string	返回对应的可读消息，值为 error
detail	true	string	错误消息

ls		g	
----	--	---	--

返回示例

```
HTTP/1.1 200 OK
Content-Type: application/json
{
  'code': 0,
  'message': 'success',
  'result': [
    {
      'corrected': '我们是中国电信的员工',
      'details': [
        ['我们', '我们', 0, 2],
        ['中过', '中国', 3, 5],
        ['点信', '电信', 5, 7],
        ['源工', '员工', 8, 10]
      ]
    }
  ]
}
```

错误码说明

错误代码	错误消息	错误可能原因
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
4002	请求参数异常	请求的参数格式不正确或者参数名错误
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

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
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%
4006	不支持的请求方法	仅支持 HTTP POST 请求方法
4007	请求的 API 不存在	请求的 API 路径错误
5000	服务内部错误	服务内部发生错误，需要联系管理员处理

4.3 智能工单推荐

接口描述

对投诉类、网络故障类等历史工单进行阅读，将与当前工单相似的历史工单推荐给用户，以达到故障处理借鉴的效果

接口要求

文本格式：

1. 每次请求文本长度不超过 500 个字符，且仅支持单条文本处理
2. 仅支持简体中文文本，如果文本内容超过 50%非中文，则返回错误，具体见错误码 4004 对应内容
3. 文本字符串不能为空，且不能全为空格或者换行等转义符

请求说明

请求示例

HTTP 方法：POST

请求 URL：[https://ai.ctyun.cn/v1/aiop/api/\[系统生成的服务码\]/shai/promt](https://ai.ctyun.cn/v1/aiop/api/[系统生成的服务码]/shai/promt)

请求头说明

参数	值(说明)
Content-Type	application/json
Sign	MD5(AppKey + AppSecret + Timestamp + Path + Http Body), 32 位小写的 MD5 加密串
Timestamp	13 位时间戳
AppKey	平台分配给应用独立的 AppKey

请求参数

参数	是否必选	类型	可选值范围	说明
data	true	string	-	请求参数放置于 http body 中，为一段文本字符串

请求示例代码

Bash:

```
curl -X POST https://ai.ctyun.cn/v1/aiop/api/[系统生成的服务码]/shai/promt -H "Content-Type: application/json" -H 'Sign: XXXXX' -H 'Timestamp: XXXXX' -H 'AppKey: XXXXX' --data '{"data": "string"}
```

Python:

```
# -*- coding: utf-8 -*-
import requests
import base64
import hashlib
import time
import json
```

```

...
智能工单推荐
...
def md5(key):
    input_name = hashlib.md5()
    input_name.update(key.encode("utf-8"))
    return input_name.hexdigest()

request_url = "智能工单推荐 API URL"
params = {"data": "string"}
app_key = '平台分配的 APP Key'
app_secret = '平台分配的 APP Secret '
time_stamp = str(int(round(time.time() * 1000)))

headers = {
    'content-type': 'application/json',
    'Sign': md5(app_key + app_secret + time_stamp + '/v1/aiop/api/[系统生成的服务码]/shai/promt' + json.dumps(p
arams)),
    'Timestamp': time_stamp,
    'AppKey': app_key
}
response = requests.post(request_url, data=json.dumps(params), headers=headers)
if response:
    print(response.json())

```

返回说明

API 调用成功返回值

返回 字段	是否 必选	类型	字段说明
code	true	int	返回状态，返回 0 表示成功
message	true	string	返回对应的可读消息，值为 success
result	true	object	API 具体返回信息
故障 描述	true	list of str	按相似度推荐的工单列表
相似 度	true	list of float	工单对应的相似度列表

		t	
--	--	---	--

API 调用失败返回值

返回 字段	是否 必选	类型	字段说明
code	true	int	错误码，参见本文档错误码说明部分
message	true	string	返回对应的可读消息，值为 error
details	true	string	错误消息

返回示例

```
HTTP/1.1 200 OK
Content-Type: application/json
{
```

```
  "code": 0,
```

```
  "message": "success"
```

```
  "result":{
```

```
    "故障描述":[
```

"用户反映无法拨打美国的号码 23:15 左右 161XXXXXXXX，提示你所拨打号码不存在，用移动手机拨打正常，请查证处理，谢谢。*****HLR 与 CRM 数据比对一致，查用户数据正常，国际长途功能开通，查看呼叫记录，用户拨打该号码的记录我方已送至上海 ISC3，记录已上传，请转国际局协查。";

"安徽 173XXXXXXXX 漫游上海，普通用户；1 日 14: 50 拨打合肥拨打合肥，提示拨打号码不正确；查信令上海 XXX 返回 XXX 错误。请上海集团 NOC 核查。记录见附件。";

"用户来电反映 199XXXXXXXX 漫游在山东济南无法接听电话，例如 152XXXXXXXX 是北京移动手机（号码）拨打 199XXXXXXXX 此号码提示空号，接续消息已经送到（详见附件），请集团上海帮忙核查。";

"0571 用户反映无法拨打山东济南的移动联通电信的电话都打不通,比如 2019-08-01 09:18:25 主叫号码 057 呼叫 1895 已经送至上海了，对方回的释放。请集团和山东协查";

"北京电信用户 177XXXXXXXX 投诉拨打 041XXXXXXXX 不通。我方用测试机 133XXXXXXXX 于 2018.1.1 10:38 拨打也无法拨通，请协查，谢谢。";

"95XXX 用户反映 XXXX 95XXX 热线打进忙音，我处 025-865XXXXXXXX 在 13:05 左右拨测 95XXX 提示忙音，拨打用户落地号码 031XXXXXXXX 正常回铃音，请平台核实"，

"四次拨打都失败，时间分别是 2017-12-31 19:02:31,2017-12-31 19:09:11,2017-12-31 19:26:40,2017-12-31 19:38:45。用主叫 85XXXXXXXX 拨打 867XXXXXXXX，听到 2-3 次铃声然后掉线，请集团协查。"，

"079XXXXXXXX 用户来电反映无法拨打长途，用户数据正常有国内长途权限，用户拨测了 153XXXXXXXX 和 156XXXXXXXX 两个手机,我方跟踪了消息,一个送给 XXXXXXXX,是 XXXXXXXX 回的无效号码格式;另一个送给了 XXXXXXXX,是 XXXXXXXX 回的无效号码格式(消息请见附件),请集团核查。"，

"请集团协助封堵主叫 869XXXXXXXX，我处查看有大量此主叫拨打 WESTERN SAMOA 造成 XXX,请查看附件"，

"境外申告漫游无法呼叫:目的地：CHINA-MOBILE CM 症状：CALL FAIL（MSRN） 目的地代码：0-9173 CDR 日志如下 日期和时间 GMT # DEST # B 号码#A 号码 # 详细呼叫记录见附件"

],

"相似度":[

0.7276202006427118,

0.7205485385931023,

0.6688286744010508,

0.6562693145083968,

0.6042671435338854,

0.5678871524051599,

0.5662423034792552,

0.5525977575781975,

0.5129372699940562,

0.5095703137781431

]

}

}

错误码说明

错误代码	错误消息	错误可能原因
------	------	--------

4001	请求中未包含内容	请求体中参数为空字符串或者全为空格等转义字符
4002	请求参数异常	请求的参数格式不正确或者参数名错误
4003	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 500
4004	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%
4005	不支持的请求方法	仅支持 HTTP POST 方法
4006	请求的 API 不存在	请求的 API 路径错误
5000	服务内部错误	服务内部发生错误，需要联系管理员处理

5 常见问题

1. Q: 自然语言处理产品有哪些 API 服务？

A: 有三个 API 服务：情感分析、文本纠错与智能工单推荐。

2. Q: 情感分析服务对于文本格式是否有要求？

A: （1）单段文本长度不超过 256，编码格式为 UTF-8 编码；（2）仅支持简体中文

3. Q: 文本纠错服务对于文本格式是否有要求？

A: （1）仅支持单段文本，长度不超过 100；（2）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；（3）对于未超过 50%的非中文，自动略过，不予纠错。

4. Q: 智能工单推荐服务对于文本格式是否有要求？

A: （1）每次请求文本长度不超过 500 个字符，且仅支持单条文本处理；（2）仅支持简体中文文本，如果文本内容超过 50%非中文，则返回错误；（3）对于未超过 50%的非中文，自动略过，不

予纠错。

5. Q: 自然语言处理产品是 HTTP GET 请求还是 HTTP POST 请求?

A: HTTP POST 请求。