第10章 人工智能开放平台应用

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| 课程名称 | 人工智能基础与应用 | | 章名 | | | 人工智能开放平台应用 | |
| 教学内容 | 人工智能开放平台应用 | | | | 课时 | 2 | |
| 项目性质 | □演示性 □验证性 □设计性 √综合性 | | | | | | |
| 授课班级 |  | 授课日期 | |  | | 授课地点 |  |
| 教学目标 | 了解AI开放平台介绍  掌握百度AI开放平台应用  了解更多AI开放平台实践 | | | | | | |
| 教学内容 | 1. AI开放平台介绍 2. 百度AI开放平台应用 3. 更多AI开放平台实践 | | | | | | |
| 教学重点 | 基于百度AI开放平台的图像、语音、人脸识别 | | | | | | |
| 教学难点 | 基于百度AI开放平台的图像、语音、人脸识别 | | | | | | |
| 教学准备 | 装有Python的计算机  教学课件PPT  教材：《人工智能基础与应用（微课版）》 | | | | | | |
| 作业设计 |  | | | | | | |

教学过程

|  |  |
| --- | --- |
| **教学环节** | **教学内容与过程**  **（教学内容、教学方法、组织形式、教学手段）** |
| **课前组织** | 做好上课前的各项准备工作（打开计算机、打开课件、打开软件、打开授课计划、教案等），吸引学生注意力。 |
| **课程说明** | 【课前说明】  从AI开放平台的基本概念开始进行初步的了解。  【目的】  使学生从了解本节课的学习目标、学习重点、考评方式等方面明确课程学习的要求和目标。 |
| **课程内容描述** | 10.1 AI开放平台介绍  AI开放平台是一些企业推出的人工智能接口，开发者无须了解人工智能的算法、网络以及训练过程，只需按照特定的方式接入，就可以使用接口提供的产品。基于该平台，开发者能够很快完成人工智能方向的应用开发。  AI开放平台大多都可以实现Web API，即在线使用这些平台接口，完成功能实现，不过该过程需要调用接口的设备能够上网。有些比较小的模型接口可以实现Mobile SDK，集成Mobile SDK到终端设备上，在移动设备上离线调用接口。  这些开放平台的应用遍及人工智能多个领域，如语音技术（识别、合成等）、图像技术（食品识别、动物识别、车牌识别等）、人脸识别（关键点检测、识别、对比等）、人体识别（关键点识别、手势识别等）、文字识别（场景下文字识别、证件文字识别等），还包括知识图谱、智能问答等领域。  不同的企业由于业务偏向不同，所以会针对性地开放不同类型的接口，基于常年的技术积累以及海量的数据积累，其常兼具稳定性以及泛化能力。下表所示为部分AI开放平台（排名不分先后）。  部分AI开放平台   |  | | --- | | AI开放平台 | | 百度AI开放平台 | | 腾讯AI开放平台 | | 阿里AI开放平台 | | 京东AI开放平台 | | 小爱AI开放平台 | | 讯飞AI开放平台 |   10.2 百度AI开放平台应用  百度AI开放平台是一个开放的AI使用平台，其提供了图像技术、语音技术、人脸与人体识别技术、视频技术、自然语言处理技术、数据智能技术、知识图谱技术等多项智能平台的接口。 10.2.1 百度AI开放平台介绍 百度AI开放平台的开放能力如图所示。    百度AI开放平台的开放能力  百度AI开放平台支持的接口语言有Java、PHP、Python、C++、C#、Node.js，支持的平台有PC、Android、iOS以及嵌入式Linux平台。 10.2.2 基于百度AI开放平台的图像识别 本案例基于百度AI开放平台实现图像识别项目。本项目使用Web API，需要在有网络的情况下完成。  1．密钥申请  （1）登录并使用  首先登录百度AI开放平台，进入通用图像分析的主界面，完成百度账号登录，通用图像分析界面如图所示。    通用图像分析界面  （2）创建应用  在创建新应用界面输入应用名称，选择应用类型，接口默认选择“图像识别”，添加应用描述，单击“立即创建”按钮，如图所示。  10-3  “创建新应用”界面  （3）记录密钥  在“应用详情”界面记录AppID、API Key、Secret Key这3个密钥，应用详情界面如图所示。    应用详情界面  （4）下载SDK  在SDK下载界面下载Python版本的SDK，SDK下载界面如图所示。  10-5  SDK下载界面    2．调用API实现识别  SDK目录下的aip目录下还有其他文件，用于其他识别，图像识别用到的文件如图所示。    图像识别用到的文件  在交互界面输入（联网状态）“pip install baidu-aip”命令，安装baidu-aip。  在Python目录下新建baidu\_api目录，解压下载好的SDK，在baidu\_api下新建baidu\_api\_image目录，AipImageClassify是图像识别的Python SDK客户端，为使用图像识别的开发人员提供了一系列的交互方法，在baidu\_api\_image目录下新建一个image\_predict.py文件，baidu\_api\_image目录如图所示。    baidu\_api\_image目录  【例10-1】 在baidu\_api\_image目录下放入一张需要识别的照片，并在image\_predict.py文件中编写代码，实现图像识别。   from aip import AipImageClassify  """ 你的 APPID AK SK """  APP\_ID = '你的 App ID'  API\_KEY = '你的 Api Key'  SECRET\_KEY = '你的 Secret Key'  client = AipImageClassify(APP\_ID, API\_KEY, SECRET\_KEY)  """ 读取图片 """  def get\_file\_content(filePath):  with open(filePath, 'rb') as fp:  return fp.read()  image = get\_file\_content('corn.png')  """ 调用图像识别 """  print (client.advancedGeneral(image))  """ 如果有可选参数 """  options = {}  options["baike\_num"] = 1  """ 带参数调用图像识别 """  print (client.advancedGeneral(image, options))  在image\_predict.py文件的代码中需要将APP\_ID、API\_KEY、SECRET\_KEY这3个密钥填入。  client.advancedGeneral()函数用于通用物体以及通用场景的识别，可以直接调用识别函数或者添加可选参数options，本代码中添加了可选参数options["baike\_num"]=1，作用是打印出识别到第一个结果的百度百科，识别的是一张玉米的照片。  识别结果：  {'log\_id': 4154135890364883769, 'result\_num': 5, 'result': [{'score': 0.882586,  'root': '商品-食品', 'keyword': '玉米'}, {'score': 0.693215, 'root': '植物-其他',  'keyword': '玉米棒'}, {'score': 0.512256, 'root': '植物-其他', 'keyword': '甜玉米'},  {'score': 0.224656, 'root': '植物-其他', 'keyword': '黏玉米'}, {'score': 0.031608, 'root':  '植物-葫芦科', 'keyword': '丝瓜'}]}  {'log\_id': 53680928017029305, 'result\_num': 5, 'result': [{'score': 0.882586, 'root': '商品-食品', 'baike\_info': {'baike\_url': 'http://baike.baidu.com/item/ %E7%8E%89%E7%B1%B3/18401', 'image\_url': 'http://imgsrc.baidu.com/baike/pic/item/ 5243fbf2b2119313d5a23cff6f380cd790238d98.jpg', 'description': '玉米(Zea mays L.)是禾本科的一年生草本植物。又名苞谷、苞米棒子、玉蜀黍、珍珠米等。原产于中美洲和南美洲，它是世界重要的粮食作物，广泛分布于美国、中国、巴西和其他国家。玉米与传统的水稻、小麦等粮食作物相比，具有很强的耐旱性、耐寒性、耐贫瘠性以及极好的环境适应性。玉米的营养价值较高，是优良的粮食作物。作为中国的高产粮食作物，是畜牧业、养殖业、水产养殖业等的重要饲料来源，也是食品、医疗卫生、轻工业、化工业等的不可或缺的原料之一。玉米资源极为丰富、廉价且易于获得，它们还具有许多生物活性，如抗氧化、抗肿瘤、降血糖、提高免疫力和抑菌杀菌等，具有广阔的开发及应用前景。2018年8月，财政部、农业农村部、银保监会印发通知，将玉米作物制种纳入中央财政农业保险保险费补贴目录。2018年，中国玉米产量25733万吨。'}, 'keyword':  '玉米'}, {'score': 0.693215, 'root': '植物-其他', 'keyword': '玉米棒'}, {'score': 0.512256,  'root': '植物-其他', 'keyword': '甜玉米'}, {'score': 0.224656, 'root': '植物-其他',  'keyword': '黏玉米'}, {'score': 0.031608, 'root': '植物-葫芦科', 'keyword': '丝瓜'}]}  除了通用物体识别外，图像识别还有菜品识别、车辆识别、Logo商标识别、动物识别、植物识别、地标识别等功能。 10.2.3 基于百度AI开放平台的语音识别 本案例基于百度AI开放平台实现语音识别项目。  1．密钥申请  语音识别密钥申请方式与图像识别类似，下载语音识别对应的SDK。  2．录音  百度AI开放平台的语音识别是将录音转换为文字，语音格式为原始PCM的录音，参数必须符合16kHz采样率、16位位深、单声道，支持的格式有PCM（不压缩）、WAV（不压缩，pcm编码）、AMR（压缩格式）。在识别之前需要先进行录音。  在baidu\_api目录下新建baidu\_api\_speech目录，将下载的语音SDK解压到该目录下，在交互界面输入（联网状态）“pip install pyaudio”命令安装pyaudio。  【例10-2】 在baidu\_api\_speech目录下新建recoder.py并尝试输入代码，完成录音功能，注意：运行此代码的设备必须有录音功能。  import pyaudio  import wave,time  CHUNK = 1024  FORMAT = pyaudio.paInt16  CHANNELS = 1  RATE = 16000  TIME = 3  def rec(file\_name, RECORD\_SECONDS):  p = pyaudio.PyAudio()  stream = p.open(format=FORMAT,  channels=CHANNELS,  rate=RATE,  input=True,  frames\_per\_buffer=CHUNK)  print("start......")  frames = []  for i in range(0, int(RATE / CHUNK \* RECORD\_SECONDS)):  data = stream.read(CHUNK)  frames.append(data)  print("end\n")  time.sleep(1)  stream.stop\_stream()  stream.close()  p.terminate()  wf = wave.open(file\_name, 'wb')  wf.setnchannels(CHANNELS)  wf.setsampwidth(p.get\_sample\_size(FORMAT))  wf.setframerate(RATE)  wf.writeframes(b''.join(frames))  wf.close()  if \_\_name\_\_ == "\_\_main\_\_":  wavfile = "./test.wav"  rec(wavfile, TIME)  代码中，TIME指的是录音时长，单位是秒。  3．调用API实现识别  录音完成后，实现语音识别功能。  【例10-3】 在baidu\_api\_speech目录下新建speech\_predict.py，在文件中编写代码。  from aip import AipSpeech  """ 你的 APPID AK SK """  APP\_ID = '你的 App ID'  API\_KEY = '你的 Api Key'  SECRET\_KEY = '你的 Secret Key'  client = AipSpeech(APP\_ID, API\_KEY, SECRET\_KEY)  # 读取文件  def get\_file\_content(filePath):  with open(filePath,'rb') as fp:  return fp.read()  # 识别本地文件  print (client.asr(get\_file\_content("test.wav"),'wav',16000,{'dev\_pid':1537,}))  运行代码可以得到识别结果。 10.2.4 基于百度AI开放平台的人脸识别 本案例基于百度AI开放平台实现人脸识别项目。  1．密钥申请  人脸识别密钥申请方式与图像识别类似，下载人脸识别对应的SDK。  2．人脸数据上传方式  在百度人脸识别API上，有3种方式上传人脸图片：BASE64字符串、URL字符串以及FACE\_TOKEN字符串。  （1）BASE64：BASE64编码的图片数据，编码后的图片大小不超过2MB。  （2）URL：图片的URL地址（可能由于网络等原因导致下载图片的时间较长）。  （3）FACE\_TOKEN：人脸图片的唯一标识。调用人脸检测接口时，会为每个人脸图片赋予一个唯一的FACE\_TOKEN，同一张图片多次检测得到的FACE\_TOKEN是同一个。  3．人脸检测  人脸检测区别于人脸识别。在一张照片中需要先找到人脸，才可以进行识别，但是一般而言，人脸检测之后也可以实现其他功能，如性别识别、年龄识别以及数字化妆等。  百度AI开放平台的人脸检测可以添加参数以返回当前人脸更为详细的信息，具体请求参数如表所示。  人脸检测具体请求参数   |  |  |  |  | | --- | --- | --- | --- | | 参数 | 必选 | 类型 | 说明 | | image | 是 | string | base64编码后的图片数据，需使用urlencode编码，编码后的图片大小不超过2MB | | max\_face\_num | 否 | uint32 | 最多处理人脸的数目，默认值为1，仅检测图片中面积最大的那个人脸 | | face\_fields | 否 | string | 包括age、beauty、expression、faceshape、gender、glasses、landmark、race、qualities信息，用逗号分隔，默认只返回人脸框、概率和旋转角度。如果要返回age等更多属性，请在此参数中添加 |   人脸检测部分返回参数如表所示。  人脸检测部分返回参数   |  |  |  |  | | --- | --- | --- | --- | | 参数 | 类型 | 必选 | 说明 | | log\_id | uint64 | 是 | 日志ID | | result\_num | uint32 | 是 | 人脸数目 | | result | object[] | 是 | 人脸属性对象的集合 | | +age | double | 否 | 年龄。face\_fields包含age时返回 | | +beauty | double | 否 | 美丑打分，范围为[0,100]，越大表示越美。face\_fields包含beauty时返回 | | +location | object | 是 | 人脸在图片中的位置 | | ++left | uint32 | 是 | 人脸区域离左边界的距离 | | ++top | uint32 | 是 | 人脸区域离上边界的距离 | | ++width | uint32 | 是 | 人脸区域的宽度 | | ++height | uint32 | 是 | 人脸区域的高度 | | +face\_probability | double | 是 | 人脸置信度，范围为[0,1] | | +rotation\_angle | int32 | 是 | 人脸框相对于竖直方向的顺时针旋转角，[-180,180] | | +yaw | double | 是 | 三维旋转之左右旋转角[-90(左), 90(右)] | | +pitch | double | 是 | 三维旋转之俯仰角度[-90(上), 90(下)] | | +roll | double | 是 | 平面内旋转角[-180(逆时针), 180(顺时针)] | | +expression | uint32 | 否 | 表情：0，不笑；1，微笑；2，大笑。face\_fields包含expression时返回 | | +expression\_probability | double | 否 | 表情置信度，范围为[0,1]。face\_fields包含expression时返回 | | +faceshape | object[] | 否 | 脸形置信度。face\_fields包含faceshape时返回 | | ++type | string | 是 | 脸形：square、triangle、oval、heart、round | | ++probability | double | 是 | 置信度：0～1 | | +gender | string | 否 | male、female。face\_fields包含gender时返回 | | +gender\_probability | double | 否 | 性别置信度，范围为[0,1]，face\_fields包含gender时返回 | | +glasses | uint32 | 否 | 是否戴眼镜，0-无眼镜，1-普通眼镜，2-墨镜。face\_fields包含glasses时返回 | | +glasses\_probability | double | 否 | 眼镜置信度，范围为[0,1]，face\_fields包含glasses时返回 |   （1）使用URL方式进行人脸检测  使用URL方式上传网络照片进行人脸检测时，可能会由于网络不通畅造成速度慢等问题，使用URL方式进行人脸检测的关键代码如下。  image = "网址"  imageType = "URL"  """ 调用人脸检测 """  # client.detect(image, imageType)  """ 如果有可选参数 """  options = {}  options["face\_field"] = "age,beauty"  options["max\_face\_num"] = 1  options["face\_type"] = "LIVE"  """ 带参数调用人脸检测 """  print (client.detect(image, imageType, options))  （2）使用BASE64方式进行人脸检测  使用BASE64方式上传本地照片进行人脸检测，关键代码如下。  """ 选择BASE64 """  filePath ="pic.png"  with open(filePath,"rb") as f:  # b64encode是编码  base64\_data = base64.b64encode(f.read())  image = str(base64\_data,'utf-8')  imageType = "BASE64"  """ 调用人脸检测 """  # client.detect(image, imageType)  """ 如果有可选参数 """  options = {}  options["face\_field"] = "age,beauty"  options["max\_face\_num"] = 1  options["face\_type"] = "LIVE"  """ 带参数调用人脸检测 """  print (client.detect(image, imageType, options))  4．添加人脸库  人脸识别需要构建人脸库来存放需要识别的人脸照片，人脸库在百度AI开放平台上。  需要新建组以及组下的用户，一个组中可以有多个用户，每个用户对应着一个人的照片。  5．调用API实现识别  【例10-4】 在baidu\_api目录下新建baidu\_api\_face目录，将下载的人脸识别的SDK解压到该目录下，在baidu\_api\_face目录下新建face\_predict.py文件。  from aip import AipFace  import base64  """ 你的 APPID AK SK """  APP\_ID = '你的 App ID'  API\_KEY = '你的 Api Key'  SECRET\_KEY = '你的 Secret Key'  client = AipFace(APP\_ID, API\_KEY, SECRET\_KEY)  """ 选择BASE64 """  filePath ="pic.png"  with open(filePath,"rb") as f:  # b64encode是编码  base64\_data = base64.b64encode(f.read())  image = str(base64\_data,'utf-8')  imageType = "BASE64"  groupIdList = "group"  """ 调用人脸搜索 """  print (client.search(image, imageType, groupIdList))  在groupIdList中可输入人脸库的用户组ID，client.search()函数可调用人脸识别。  10.3 更多AI开放平台实践  除了百度AI开放平台外，国内外还有很多开放平台，它们针对的方向不同，开放能力也有所不同。 10.3.1 腾讯AI开放平台 腾讯AI开放平台主页如图所示。    腾讯AI开放平台主页  腾讯AI开放平台的主要开放能力包括文字识别OCR、人脸与人体识别、图片特效、图片识别、敏感信息甄别、智能闲聊、机器翻译、基础文本分析、语义解析、语音识别与语音合成等。 10.3.2 阿里AI开放平台 阿里AI开放平台主页如图所示。    阿里AI开放平台主页  阿里AI开放平台的主要开放能力包括语音识别与自然语言处理等。 10.3.3 京东AI开放平台 京东AI开放平台主页如图所示。    京东AI开放平台主页  京东AI开放平台的主要开放能力包括文字识别、语音技术、人脸与人体识别、图像及视频理解与自然语言处理等。 10.3.4 小爱AI开放平台 小爱AI开放平台主页如图所示。    小爱AI开放平台主页  小爱AI开放平台的主要开放能力包括语音识别与自然语言处理等。 10.3.5 讯飞AI开放平台 讯飞AI开放平台主页如图所示。    讯飞AI开放平台主页  讯飞AI开放平台的主要开放能力包括语音识别、语音合成、图像识别与自然语言处 理等。 |
| **总结评价** | 本节课主要以百度AI开放平台为例介绍了AI开放平台的应用。 |