第1章 人工智能概述

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| 课程名称 | 人工智能基础与应用 | | 章名 | | | 人工智能概述 | |
| 教学内容 | 人工智能概述 | | | | 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 | 了解人工智能和深度学习  熟悉人工智能发展现状与人工智能机器学习框架  了解怎样学习人工智能 | | | | | | |
| 教学内容 | 1. 人工智能 2. 深度学习 3. 人工智能发展现状 4. 人工智能机器学习框架 5. 怎样学习人工智能 | | | | | | |
| 教学重点 | 人工智能发展现状、人工智能机器学习框架 | | | | | | |
| 教学难点 | 人工智能机器学习框架 | | | | | | |
| 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】  从人工智能的应用引入本节课学习内容。  【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** | 1.1 了解人工智能  想了解人工智能，需要先了解它的几个常见应用。  1．人脸识别  目前，手机屏幕解锁、支付宝的脸部ID支付、门禁闸机、企业考勤、金融开户等都在使用人脸识别技术。人脸识别还可以对人的眉毛、眼睛、鼻子、嘴巴以及轮廓等关键点进行检测，在解锁或支付过程中，找到眼睛的关键点后可以识别眼睛是否为睁开的状态，如果闭着眼睛就不能解锁手机屏幕。  再复杂一些，人脸识别还可以识别出被检测者的大概年龄、性别、种族、表情情绪、是否戴眼镜以及当前的头部姿态等特征。通过识别这些特征，可以对相册中的人物进行自动分类等操作。  2．智能音箱  现在几百元钱就可以买到一个能够对话、播放音乐的智能音箱，智能音箱在联网之后可以对用户说的话进行语音识别，并对其内容进行语义理解，最后将要回应的答案一方面通过语音合成反馈给用户，另一方面针对用户的需求或者要求的动作进行执行，如听歌等。  当然，手机的语音助手也可以有相同的功能，如Siri、小爱同学等。  3．机器翻译  在日常学习或者工作中，人们经常会有一些单词不认识或者要把某句话翻译成英语，目前借助软件或者网页就可以实现中英文互译。这里举一个例子，当翻译“我在周日看了一本书”的时候，人工翻译可以很好地翻译出“I read a book on Sunday”，但是对于机器而言，它有着不同的翻译结果，首先就是一词多义的问题，如“看”这个词，可以翻译为“look”“watch”“read”等，其次就是语序问题，“在周日”这样的时间状语一般习惯放在句子后面，在翻译时需要做到“信、达、雅”，现在基于深度学习的翻译基本上可以做到“信、达”。基于深度学习的百度翻译结果如图所示。    基于深度学习的百度翻译结果  4．无人驾驶  百度百科对无人驾驶汽车给出的定义是“无人驾驶汽车是智能汽车的一种，也称为轮式移动机器人，主要依靠车内以计算机系统为主的智能驾驶仪来实现无人驾驶的目的”。  无人驾驶汽车是一项集合了自动控制、人工智能、传感器技术等多项技术的高度发展的产物。  目前，我国的百度公司、清华大学都在致力于无人驾驶汽车的研发。  人工智能已经在多个方面融入人们的日常生活中，并给人们带来了诸多便利。  简言之，人工智能就是通过一些科学的计算方法，让机器做一些人类能够做的事情，例如，人类的视觉——目标认知、图像识别，人类的听觉——语音识别，人类的思考——对图像的分析、对语言中语义的分析、理解以及回答等。  1.2 了解深度学习  在了解深度学习之前，需要对它和人工智能之间的关系进行了解与分析，人工智能与深度学习的关系如图所示。  1-2  人工智能与深度学习的关系  由上图可知，深度学习是隶属于机器学习范畴的，机器学习又是隶属于人工智能范畴的，先解释一下什么是机器学习。  卡内基梅隆大学（Carnegie Mellon University）的汤姆·迈克尔·米切尔（Tom Michael Mitchell）教授在*Machine learning*一书中对机器学习给出了比较专业的定义，即“如果一个程序在任务T上，随着经验E的增加，效果P也可以随之增加，则称这个程序可以从经验中学习”。通俗地讲，就是如果机器学习算法（一个程序）要实现一个预测、分类问题（任务T），那么需要对数据（经验E）进行分析，如果数据越多，最后实现预测、分类时的准确率（效果P）越高，那么就称这个机器学习算法可以从数据中学习。  机器学习是人工智能的一个分支，是实现人工智能的方法，是计算机通过学习来提高性能的一种方式。机器学习过程并不是告诉机器该怎么做，而是告诉它该怎么学习，在这个学习的过程中机器从数据里提取特征，当然这需要大量的数据。对于一些复杂的问题，机器学习并不能解决。如需要在一张照片上找到所有人的面部，利用机器学习解决这个问题是非常困难的，因为有的人留着长发，有的人戴着眼镜，还有拍照时的不同表情等，所以并不能完全保证利用这些特征能够找到人脸。  而深度学习可以解决这个问题，这个在早期试图模仿人类大脑神经元之间的学习机理、将各个特征进行联系从而组成更为复杂特征的方法，在图像识别和语音识别等领域具有突破性的进展。  深度学习是机器学习的一个分支，深度学习和机器学习的实现过程如图所示。  1-3  深度学习和机器学习的实现过程  深度学习这个词，除了从传统机器学习那里继承了“学习”之外，“深度”是其区别于其他方法的特征之一。在人工智能中，深度学习是人工神经网络或者深层神经网络的代名词，“深度”指的是神经网络的网络层次，最基本的神经网络结构如图所示。  1-4  最基本的神经网络结构  在上图中，最左侧的一层称为输入层，最右侧的一层叫作输出层，中间所有节点组成的层称为隐藏层。最基本的神经网络结构中只有一层隐藏层，在实际的神经网络中，输入层*x*需输入训练数据，中间有一个或多个隐藏层*a*，输出层*y*输出结果。  在21世纪前，深度神经网络由于受到计算力、数据量以及算法的限制，一直没有很好的表现，当时传统机器学习的表现要好于神经网络，如支持向量机（Support Vector Machine，SVM）。步入21世纪，随着计算机性能的不断提高，大数据、云计算以及各种专用于计算的芯片的出现与普及，计算力以及数据量不再像之前那样限制神经网络的发展。在2012年的大规模视觉识别挑战赛（ImageNet Large Scale Visual Recognition Challenge，ILSVRC）中，辛顿（Hinton）和他的学生克里泽夫斯基（Alex Krizhevsky）设计的深度学习网络AlexNet获得了冠军；到2015年，深度学习在图像分类方面的错误率已经低于人工标注的错误率；现在，深度学习被广泛应用于各个方向并取得了非常好的效果。  深度学习最早兴起于计算机视觉领域，在短短数年内，深度学习的方法推广到了越来越多的领域中，如语音识别、自然语言处理等领域。现如今，图像分类、图像识别已经运用于包括无人驾驶、按图识物、人脸识别、美颜、文字识别等在内的多个方向。语音识别也被运用于智能音箱、手机语音助手、语音输入法、同声传译等方向。自然语言处理被运用到聊天机器人、机器翻译等方向。  1.3 人工智能发展现状  在竞赛领域，在ILSVRC的图像分类比赛中，2012年，AlexNet网络（一个7层的卷积神经网络）将图像分类的错误率从2011年的26%下降到16%。2014年，VGG和GoogleNet分别将错误率下降到了7.3%和6.7%，VGG模型在多个迁移学习任务中的表现要优于GoogleNet。到了2015年，微软亚洲研究院提出的深度残差网络（ResNet）将错误率下降到了3.57%，低于人类5.1%的分类错误率，ResNet是一个152层的卷积神经网络。  在实际应用中，人工智能在语音识别、语音合成上取得了非常瞩目的结果。2016年10月，由微软美国研究院发布的语音识别最新应用实现了错误率为5.9%的新突破，这是第一次用人工智能技术得到跟人类近似的语音识别错误率。  人工智能已经被运用于农业上，2017年，蓝河公司（BlueRiver）的喷药机器人开始使用计算机视觉来识别需要肥料的植物，如棉花、生菜，以及其他特色植物，为它们喷洒农药或除草剂等。  在医学上，谷歌大脑与Alphabet旗下子公司Verily联合开发了一款能用来诊断乳腺癌的人工智能产品。为了确定这个功能的可用性以及准确率，谷歌专门安排了一场人类与人工智能进行病理分析的大比拼。这场比拼中，一位资深病理学家花了整整30个小时，仔仔细细分析了130张切片，但以73.3%的准确率败给准确率达88.5%的人工智能。  在电商领域，阿里的人工智能系统“鲁班”在2017年的“双十一”网络促销日期间，根据用户行为和偏好，智能地为手机淘宝自动生成了4亿张不重复的海报。  2017年，百度的小度机器人在《最强大脑》中战胜人类“脑王”，搜狗的问答机器人汪仔在《一站到底》中战胜哈佛女学霸。同年，百度CEO李彦宏将无人驾驶平台Apollo汽车开上五环。  2018年，Google的Duplex代替人类自动接打电话、预订餐厅。  2019年的央视网络春晚，人工智能主播“小小撒”携手撒贝宁，一同亮相舞台。  现在，人工智能在动作识别、人脸识别、人体姿态估计、图像分类、图像生成、图像分割等计算机视觉领域，以及在问答、常识推理、机器翻译等自然语言处理领域都取得了很大的进展。  1.4 人工智能机器学习框架  在实现机器学习或深度学习的过程中，往往需要写大量的代码来实现某个功能或算法，这时就需要一些框架把这些大量且重复的代码进行整合。在使用过程中，只需要调用这个框架下的某个方法或者某几个方法，就能实现原来大量代码实现的功能。 1.4.1 机器学习框架简介 常用的人工智能机器学习的基本框架包括TensorFlow、PaddlePaddle、Caffe、PyTorch、MXNet等，其说明如表所示。  TensorFlow、PaddlePaddle、Caffe、PyTorch、MXNet框架的说明   |  |  |  | | --- | --- | --- | | 框架名称 | 接口语言 | 是否开源 | | TensorFlow | C++、Java、Python、Go、C#等 | 开源 | | PaddlePaddle | C++、Python | 开源 | | Caffe | C++、MATLAB、Python | 开源 | | PyTorch | C++、Python等 | 开源 | | MXNet | C++、Python、R、MATLAB等 | 开源 |  1.4.2 TensorFlow TensorFlow的关注数在GitHub上表示为star的数量，截止到2019年6月，已经超过了12万，不同框架的star数如图所示。    不同框架的star数  TensorFlow中文社区对自己的框架给出的定义是：TensorFlow是一个采用数据流图（Data Flow Graphs）的，用于数值计算的开源软件库。它灵活的架构允许用户在多种平台上展开计算，如台式计算机中的一个或多个CPU（或GPU）、服务器、移动设备等。  该社区还对TensorFlow的六大优势做了重点阐述。  （1）高度的灵活性。TensorFlow不是一个严格的“神经网络”库。只要用户可以将计算表示为一个数据流图，就可以使用TensorFlow。用户通过构建图描写驱动计算的内部循环。TensorFlow提供了有用的工具来帮助用户组装“子图”（常用于神经网络），用户可以在TensorFlow的基础上完成“上层库”的搭建，并且不会产生性能损耗。若找不到想要的底层数据操作，可以通过C++代码来丰富其底层操作。  （2）可移植性（Portability）。TensorFlow可以在CPU和GPU上运行，如可以运行在台式机、服务器、手机移动设备上等。TensorFlow可以在用户的笔记本电脑上实现机器学习算法，或者不改变代码就可以实现将训练的模型在多个CPU上规模化运算。TensorFlow还可以将训练好的模型放入手机App里，并可以将模型作为云端服务运行在服务器上，或者运行在Docker容器里。  （3）将科研和产品联系在一起。过去如果要将科研中的机器学习算法用到产品中，需要大量的代码重写工作。开发者可使用TensorFlow尝试新的算法，产品团队使用TensorFlow来训练和使用计算模型，并直接提供给在线用户。使用TensorFlow可以让应用型研究者将想法迅速运用到产品中，也可以让学术型研究者更直接地分享代码，从而提高科研产出率。  （4）自动求微分。基于梯度的机器学习算法会受益于TensorFlow自动求微分的功能。用户只需要定义预测模型的结构，将这个结构和目标函数（Objective Function）结合在一起，并添加数据，TensorFlow就将自动为用户计算相关的微分导数。计算某个变量相对于其他变量的导数仅仅是通过扩展图来完成的，所以用户能一直了解计算过程。  （5）多语言支持。TensorFlow具有合理的C++使用界面，也有一个易用的Python使用界面来构建和执行Graphs。用户可以直接写Python、C++程序，也可以通过交互式的IPython界面用TensorFlow尝试某些想法，它可以帮助用户将笔记、代码、可视化内容等有条理地归置好。  支持TensorFlow的语言包括Python、C++、Java、C#等，如图所示，但运用最多的还是Python。    支持TensorFlow的语言  （6）性能最优化。由于TensorFlow提供了对线程、队列、异步操作等的支持，因此可以将用户硬件的计算潜能全部发挥出来。用户可以自由地将TensorFlow图中的计算元素分配到不同设备上，TensorFlow可以管理好这些不同的计算副本。  现在，很多公司都在使用TensorFlow做产品。谷歌在2017年就有6000多个产品在使用TensorFlow。当然，国内的京东、小米等公司也在使用TensorFlow做人工智能的开发。 1.4.3  PaddlePaddle PaddlePaddle是百度推出的开源深度学习平台，最初由百度科学家和工程师开发，有着全面、准确的中文使用文档，为国内的开发者建立了友好的生态环境。  在百度的GitHub代码仓库上，其对PaddlePaddle的四大优势做了重点阐述。  （1）灵活性。PaddlePaddle支持丰富的神经网络架构和优化算法，易于配置复杂模型，例如，带有注意力机制或复杂记忆连接的神经网络机器翻译模型。  （2）高效性。为了高效地使用异步计算资源，PaddlePaddle对框架的不同层进行优化，包括计算、存储、架构和通信。下面是一些样例：通过SSE/AVX内置函数、BLAS库（如MKL、OpenBLAS、cuBLAS）或定制的CPU/GPU内核优化数学操作；通过MKL-DNN库优化CNN网络，高度优化循环网络，无须执行padding操作即可处理变长序列，针对高维稀疏数据模型，优化了局部和分布式训练。  （3）稳定性。PaddlePaddle使利用各种CPU、GPU和机器来加速的训练变得简单。PaddlePaddle通过优化通信可以实现巨大的吞吐量，并可以快速执行。  （4）与产品相连。PaddlePaddle的设计也使其易于部署。在百度，PaddlePaddle已经部署到具有巨大用户量的产品和服务上，包括广告点击率（CTR）预测、大规模图像分类、光学字符识别（OCR）、搜索排序、计算机病毒检测、推荐系统等。PaddlePaddle广泛应用于百度产品中，产生了非常重要的影响。  1.5 怎样学习人工智能  学习人工智能，准确地说应该是学习机器学习和深度学习的方法，前提是要会使用一些工具，包括语言工具及编程工具。本书使用Python语言对TensorFlow的框架进行介绍。  1．掌握一门编程语言  Python是一种高层次的结合了解释性、编译性、互动性的面向对象的脚本语言，它的使用方法非常简单。最重要的是，早期的时候，要构建TensorFlow图，Python是唯一的选择。对于初学者而言，Python语言是一个非常不错的选择，而且学习Python对学习TensorFlow也有帮助。  由于TensorFlow是开源的，长期以来，在社区的支持下，更多的语言开始支持使用TensorFlow学习深度学习，所以也可以用Java、C++、C#等构建网络。  2．学会使用一种工具  拥有一种方便易用的编程工具会事半功倍，Jupyter Notebook、Sublime、Notepad++、Spyder、PyCharm等都可以编写Python和深度学习的网络，本书选择PyCharm作为工具。  PyCharm是一个专门为Python打造的集成开发环境（Integrated Development Environment，IDE），具有一整套可以帮助用户在使用Python语言开发时提高效率的工具，如调试、语法高亮、Project管理、代码跳转、智能提示、自动完成、单元测试、版本控制。  3．多学习一些论文  每年都会有很多关于深度学习的论文发表。在学习初期，需要研读一些入门的网络论文。如果想了解图像领域的LeNet、AlexNet框架，可以先阅读相关的论文，这些经典的网络论文往往会让读者对这个领域的知识有更加深入的理解。  4．尝试修改并训练网络  对一个初学者而言，修改网络参数（即调参）是一件很困难的事，但是很多开源的网络都可以用自己的数据集去训练。在训练过程中，初学者会对这个网络有更加深刻的认识。 |
| **总结评价** | 本节课先列出了生活中常见的人工智能应用；之后对人工智能、机器学习、深度学习之间的关系做了分析，并对深度学习和神经网络进行了分析；接着对人工智能的发展现状用举例的形式做了描述；然后对实现深度学习的多个框架做了简单介绍，并详细介绍了TensorFlow以及PaddlePaddle；最后对如何学习人工智能做了简单介绍。 |