**人工智能学习总结**

2014201870 冯晋嘉

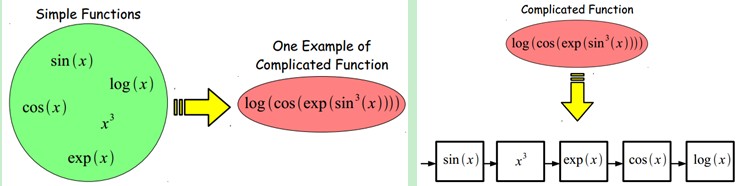
人工智能知识架构机器学习是一种很特别的方法，它与传统的编程不一样，传统的编程需要通过人为给定明确的指令去工作，而机器学习却只需要人投入大量数据，它自己会完成直接编程无法做到的工作，因此有人说机器学习赋予了机器学习的能力。从实践的意义来说，机器学习就是利用数据训练出模型，再通过模型进行预测的方法。机器学习适用于解决或是优化这些问题：已有的解决方法需要大量手动调整工作的问题，传统思路下找不到解决方法的复杂问题，不稳定环境中的问题，数据集过于庞大的问题。通过机器学习，一方面是可以提高原来解决方案的工作效率，另一方案是针对原来找不到解决办法的问题给出解答。

人工智能领域的知识框架中包含了大量这些知识体系结构，也逐渐建立起自己对人工智能的理解。

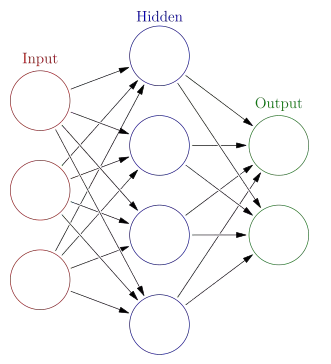
人工智能首先涉及到的是数学和计算机专业的基础知识，进一步的是机器学习的几大算法和模型，其中包括回归模型、决策树模型、支持向量机模型、贝叶斯分类模型、神经网络、聚类算法等。图中还清晰的展示了机器学习的分类，监督学习是已有训练样本，并有其对的输出，通过这样的训练样本来训练模型，无监督学习是没有样本的，直接对数据进行建模，比如聚类算法，还有一种是半监督学习，是指综合利用有类标的和没有类标的数据来生成类函数。其他的分类还有在线学习，指的是每输入一个样本都计算一次误差，进一步调整一下参数，而批量学习是指的一次性就输入所有的样本进行训练。除此之外还有基于实例的学习基于模型的学习，迁移学习和强化学习。

深度学习的概念由Hinton等人于2006年提出。如果把学习结构看作一个网络，则深度学习的核心思路是：无监督学习用于每一层网络的pre-train；每次用无监督学习只训练一层，将其训练结果作为其高一层的输入；用自顶而下的监督算法去调整所有层。

深度学习的概念源于人工神经网络的研究。简而言之，深度学习网络是一个能够根据输入训练集的参数进行自适应和自调整的一组节点，其模型的原型参考了人类大脑内的脑细胞神经元之间的神经连接的处理方式。神经-中枢-大脑的工作过程，是一个不断迭代、不断抽象的过程。这里的关键词有两个，一个是抽象，一个是迭代。从原始信号，做低级抽象，逐渐向高级抽象迭代。人类的逻辑思维，经常使用高度抽象的概念。



例如，从原始信号摄入开始（瞳孔摄入像素 Pixels），接着做初步处理（大脑皮层某些细胞发现边缘和方向），然后抽象（大脑判定，眼前的物体的形状，是圆形的），然后进一步抽象（大脑进一步判定该物体是只气球）。脑细胞通过突触之间的连接接受神经元的电刺激，然后对电信号进行处理，之后元再将处理之后的电信号发送到下一个神经元，这便是深度学习的基本模式：由节点构成的层，以及层与层之间节点的传输。神经网络技术起源于上世纪五、六十年代，当时叫感知机（perceptron），拥有输入层、输出层和一个隐含层。输入的特征向量通过隐含层变换达到输出层，在输出层得到分类结果。虽然感知机在功能上十分弱小（甚至不能解决异或问题），但所有的现代神经网络都是在感知机的基础框架上的变种，无论是卷积神经网络，循环神经网络还是深度神经网络，无论起怎样变化，都离不开“节点->输出->节点”的模式。

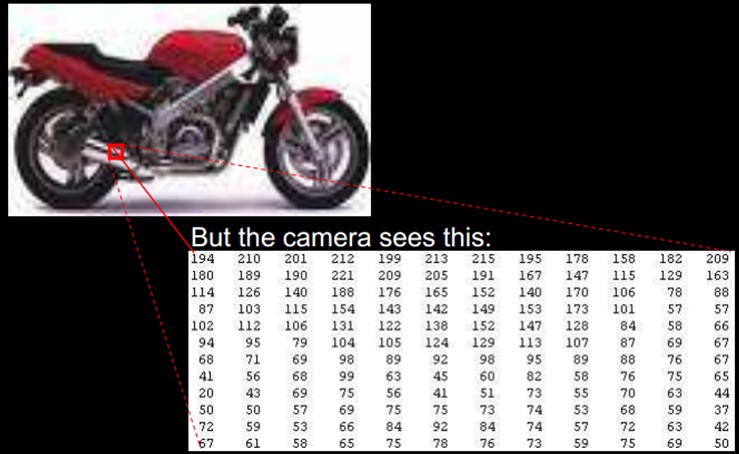


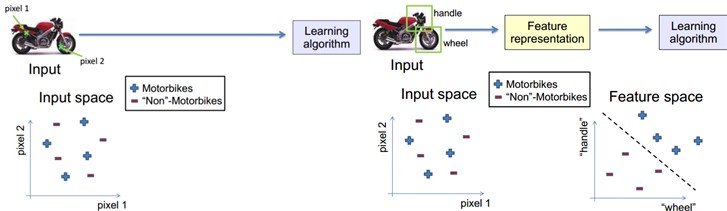
简单来说，含多隐含层的多层感知器就是一种深度学习结构。通过多层隐含层的特征处理和传输，不仅能够克服单层感知机的不能解决异或问题，而且通过层与层之间和同一层中节点之间的权值调整，使得模型能够处理非线性问题，这对神经网络的结果分析而言是十分重要的，如果只能处理线性问题，那我们对就必须将问题的输入转化为线性，虽然这并不是特别困难，但是绝大多数问题都是非线性的，这样就会给我们的输入带来很大的限制。而现在深度学习通过组合低层特征形成更加抽象的高层表示属性类别或特征，通过层之间的权值的调整，使得输入的特征能够更加明显或者更加抽象的表示出来，以发现数据的分布式特征。多层感知机可以摆脱早期离散传输函数的束缚，使用sigmoid或tanh等连续函数模拟神经元对激励的响应，在训练算法上则使用Werbos发明的反向传播BP算法。多层感知机解决了之前无法模拟异或逻辑的缺陷，同时更多的层数也让网络更能够刻画现实世界中的复杂情形。

多层感知机给我们带来的启示是，神经网络的层数直接决定了它对现实的刻画能力，因为层数越多，在神经元数量相同的情况下，越多的层数意味着越少的每层神经元个数，就越能够提取更多的特征和拟合更复杂的函数。

但是层数变多随之带来了两个重大的问题：梯度消失和参数数量膨胀，梯度消失随着神经网络层数的加深，优化函数越来越容易陷入局部最优解，并且这个“陷阱”越来越偏离真正的全局最优。利用有限数据训练的深层网络，性能还不如较浅层网络。同时，另一个不可忽略的问题是随着网络层数增加，“梯度消失”现象更加严重。具体来说，我们常常使用sigmoid作为神经元的输入输出函数。对于幅度为1的信号，在BP反向传播梯度时，每传递一层，梯度衰减为原来的0.25。层数一多，梯度指数衰减后低层基本上接受不到有效的训练信号。

学习[算法](http://lib.csdn.net/base/datastructure" \o "算法与数据结构知识库" \t "http://blog.csdn.net/zouxy09/article/details/_blank)在一个什么粒度上的特征表示，才有能发挥作用？就一个图片来说，像素级的特征根本没有价值。例如下面的摩托车，从像素级别，根本得不到任何信息，其无法进行摩托车和非摩托车的区分。而如果特征是一个具有结构性（或者说有含义）的时候，比如是否具有车把手（handle），是否具有车轮（wheel），就很容易把摩托车和非摩托车区分，学习算法才能发挥作用。





基于深度置信网络(DBN)提出非监督贪心逐层训练算法，为解决深层结构相关的优化难题带来希望，随后提出多层自动编码器深层结构。此外Lecun等人提出的卷积神经网络（CNN）是第一个真正多层结构学习算法，它利用空间相对关系减少参数数目以提高训练性能。

深度学习（deep learning）的基本训练过程如下：如果对所有层同时训练，时间复杂度会太高；如果每次训练一层，偏差就会逐层传递。这会面临跟上面监督学习中相反的问题，会严重欠拟合（因为深度网络的神经元和参数太多了）。hinton提出了在非监督数据上建立多层神经网络的一个有效方法，简单的说，分为两步，一是每次训练一层网络，二是调优，使原始表示x向上生成的高级表示r和该高级表示r向下生成的x'尽可能一致。方法是：

1）首先逐层构建单层神经元，这样每次都是训练一个单层网络。

2）当所有层训练完后，Hinton使用wake-sleep算法进行调优。

将除最顶层的其它层间的权重变为双向的，这样最顶层仍然是一个单层神经网络，而其它层则变为了图模型。向上的权重用于“认知”，向下的权重用于“生成”。然后使用Wake-Sleep算法调整所有的权重。让认知和生成达成一致，也就是保证生成的最顶层表示能够尽可能正确的复原底层的结点。比如顶层的一个结点表示人脸，那么所有人脸的图像应该激活这个结点，并且这个结果向下生成的图像应该能够表现为一个大概的人脸图像。Wake-Sleep算法分为醒（wake）和睡（sleep）两个部分。

1）wake阶段：认知过程，通过外界的特征和向上的权重（认知权重）产生每一层的抽象表示（结点状态），并且使用梯度下降修改层间的下行权重（生成权重）。也就是“如果现实跟我想象的不一样，改变我的权重使得我想象的东西就是这样的”。

2）sl+eep阶段：生成过程，通过顶层表示（醒时学得的概念）和向下权重，生成底层的状态，同时修改层间向上的权重。也就是“如果梦中的景象不是我脑中的相应概念，改变我的认知权重使得这种景象在我看来就是这个概念”。

deep learning训练过程具体如下：

1）使用自下上升非监督学习（就是从底层开始，一层一层的往顶层训练）：采用无标定数据（有标定数据也可）分层训练各层参数。在传统的神经网络的中，每一层在学习之后都会将经过本层训练的特征传向下一层，而深度学习网络直接使用原始数据对每一层（除了最后一层）都进行训练。这一步可以看作是一个无监督训练过程，是和传统神经网络区别最大的部分（这个过程可以看作是feature learning过程）。具体的，先用无标定数据训练第一层，训练时先学习第一层的参数（这一层可以看作是得到一个使得输出和输入差别最小的三层神经网络的隐层），由于模型capacity的限制以及稀疏性约束，使得得到的模型能够学习到数据本身的结构（也是就能够更为抽象的提取出数据的特征），从而得到比输入更具有表示能力的特征；在学习到第n-1层后，将n-1层的输出作为第n层的输入，训练第n层，由此分别得到各层的参数；

2）自顶向下的监督学习（就是通过带标签的数据去训练，误差自顶向下传输，对网络进行微调）：基于第一步得到的各层参数进一步fine-tune整个多层模型的参数，这一步是一个有监督训练过程；第一步类似神经网络的随机初始化初值过程，由于DL的第一步不是随机初始化，而是通过学习输入数据的结构得到的，因而这个初值更接近全局最优，从而能够取得更好的效果；所以deep learning效果好很大程度上归功于第一步的feature learning过程。

深度学习是机器学习研究中的一个新的领域，其动机在于建立、模拟人脑进行分析学习的神经网络，它模仿人脑的机制来解释数据，例如图像，声音和文本。深度学习的实质，是通过构建具有很多隐层的机器学习模型和海量的训练数据，来学习更有用的特征，从而最终提升分类或预测的准确性。因此，“深度模型”是手段，“特征学习”是目的。区别于传统的浅层学习，深度学习的不同在于：1）强调了模型结构的深度，通常有5层、6层，甚至10多层的隐层节点；2）明确突出了特征学习的重要性，也就是说，通过逐层特征变换，将样本在原空间的特征表示变换到一个新特征空间，从而使分类或预测更加容易。与人工规则构造特征的方法相比，利用[大数据](http://lib.csdn.net/base/hadoop" \o "Hadoop知识库" \t "http://blog.csdn.net/zouxy09/article/details/_blank)来学习特征，更能够刻画数据的丰富内在信息。

**智能小车实验**

在这门课的动手实践部分，我们以小组为单位，进行了人工智能小车的开发，前期主要

按照老师的要求去完成小车结构的搭建，以及实现一些基本功能，后期我们逐渐加入了自己

的想法，利用课堂理论知识所学，将机器学习的一些算法和模型融合进小车，丰富它的功能，

期间我们遇到各种各样的问题，思考了许多的解决方案，我们小组每个星期都约定时间讨论，

不得不说这样的实践经验极大地锻炼了我们的工程能力、动手能力、发现问题、解决问题的

能力，更让人欣喜的是，在实现了这些原本以为非常高深莫测的功能以后，获得了深深的成

就感，并且在亲自动手的过程，在脑海里留下了深刻的印象，这对于我们日后再进行这方面

的开发或是研究，有着极大的作用。接下来我将阐述我们小组在实践部分做的一些工作，我

们的想法，我们遇到的问题以及我们的解决方案，还有我作为小组成员在其中担任的角色，

以及我自己的收获。

对于我们小组的智能小车，我的想法是实现一个智能机器人配送货物的流程。作为送货

机器人，基础功能就是小车可以自行前往指定的目的地，并且在途中做到自动避开障碍物，

以及识别红绿灯的功能，为了实时监控小车的位置，我们利用 GPS 模块获取其地理位置信

息，以此为基础对其发出转向的指令，让它始终保持在前往目的地的正确方向上，为了简化

问题，我们暂时还没有考虑实际的道路规划和实时路况信息，今后如果要完善我们的功能，

可以加入地图中的实际道路数据和道路交通情况等信息，同时还需要做出最短路径规划的决

策。在小车行进途中，我们一方面是需要沿着道路行进，这就需要实现一个寻迹的功能，具

体方法就是通过图像识别，另一方面是要检测红绿灯，通过图像识别，判断当前是否是绿灯，

是否可以继续前进，还是需要等待。我们原本的想法是让小车可以对交通信号灯的倒计时进

行判别，但是因为实际情况很复杂，小车获取到的图像非常复杂，要想从中提取出我们想要

的区域比较困难，目前这一部分还在考虑中。另一个功能是跟踪指定物体，我们同样是利用

图像识别来完成，应用场景就是希望小车可以跟着人行动，例如帮助人搬运重物。当送货机

器人到达目的地以后，需要找到具体的一户人家，因而就涉及到了门牌号的识别，在这一块

我们是从网络上找到了门牌号的训练数据集，利用神经网络的模型对其进行了训练。除此之

外，还需要完成的是人脸识别，我们目前可以实现的是从一张图像中找到人脸的区域，这部

分也是通过图像识别来完成，在未来我们希望将人脸区域提取出来以后，可以真正实现人脸

识别，需要进一步做的工作就是对训练数据的搜集。

在进行智能小车的开发中，其中很重要的一部分是对硬件方面的学习，因为以前未曾接

触过这方面的内容，因此进行实践的第一步就是了解 Arduino 开发板的工作机制，以及一些 感应器和促动器的工作方式，经过学习，我们知道开发版支持的编程语言正是最基础的 C

语言，再加上许多功能模块都有最基本的调用示例代码，因此在硬件方面我们很快就上手了。

首先，需要用到的功能模块有超声波回声测距模块，用来识别距前方障碍物的距离，以

便小车及时转弯，避开障碍物，不过它的工作模式有一定缺陷，即当小车以较大角度驶向障

碍物时，由于波的反射原理，使得模块没有检测到侧前方的障碍物，这个问题我们目前还没

有找到可行的解决方案，未来计划通过图像识别的方法来智能避障。

另一个模块是 GPS 模块，用于获取小车实时的地理位置（经纬度），由此我们就可以

得到小车的行动轨迹，也便于我们对它进行远程的监视和控制。

蓝牙模块是用于通信的模块，它负责将小车搜集到的各种数据上传到我们的服务器（个

人电脑/手机），也负责接收我们向它输送的指令，我们小组耗费了非常多的时间在蓝牙通

讯这一部分。在前期我们遇到过蓝牙连不上、只能单向通讯、接收不到指令等等问题，通过

不断使用测试代码对其进行测试，建立断点以定位出错的区域，最后解决了这一系列的问题。

在软件部分，我们编写代码、烧录程序是用的 Arduino 配套软件，对于数据的分析和处 理我们用 python 实现，期间使用 python 完成通讯的部分我们遇到一些困难，图像数据传输 的通讯通道很不稳定，有时会被迫中止，却找不到合理的原因，后来我们改蓝牙通讯为 wifi

通讯，目前的执行情况稍有提高。