人工智能个人报告

2015202008

水治禹

关于人工智能，我们首先需要明确它的含义。通常我们的理解是让计算机去做只有人才能做的富有智能的工作，从另一个角度，也有一种说法是“人工智能是一个知识信息处理系统“。那么机器如何才算完成了这样一个从知识信息处理呢？

举个例子，想让机器买一张回家的车票，它需要了解哪些知识？它需要知道车票是什么东西，起始站和目的地，全国的（车票）交通网络，购票的方法，规划路线等，归纳一下，也就是：（1）现状等背景知识；（2）对掌握的知识的分析、选择、归纳和总结的知识；（3）解决问题的策略；（4）问题本身的专业知识。如果计算机掌握了上述知识，具备解决问题的能力，那么就可以认为该系统具有智能了。

总之，人工智能是一门综合性的边缘学科，它借助智能系统完成诸如模式识别、自然语言理解、程序自动设计、定理自动证明、机器人、专家系统等智能活动。下面我会选几个我比较感兴趣的人工智能的领域来继续报告。

1. 模式识别

每当春运抢票时，往往需要填验证码，现在春运的验证码比过去复杂了很多。不是简单的对数字或者汉字识别，举个例子，系统会给出许多图标，让你选出哪些是汽车的商标。这是一个高级模式识别的好例子，计算机如果想要成功通过测试，需要具备非常广泛的生活常识和归纳能力。

模式识别主要有两种方法：句法方法和统计决策法。也就是基于知识/结构或是基于数据。举个例子，我们在进行语言翻译的时候，最开始，我们是希望按照主、谓、宾、补语、状语等将一句话分成不同结构来进行翻译的，这就是基于知识的方法，典型代表就是专家系统。但是，自然语言并不像机器语言那么精准，它充满了随意性和不确定性，很多时候一个句子根本没有正确的句式结构，这就产生了另外一种方法，基于数据的方法。这个方法看起来更加简单粗暴，我们直接给出机器大量的数据，告诉他无数句话的意思，在翻译时，从数据中分析出它最可能的意思是什么。这个过程是把大量数据给机器进行学习，并建立判别函数。显然，句法方法适合于结构清晰，噪声很少的模式识别，比如指纹、染色体等识别。

而基于数据的方法是目前机器学习中研究较多的方法。获取大量样本数据后，根据统计学原理来构造分类器，来对未来样本进行预测。在分类过程中，又涉及到有监督和无监督的区别。当本身有足够的先验知识时，分类是比较容易的，根据特征向量确定判别函数，在对未知模式进行分类，这就是监督分类。非监督分类，也就是聚类分析的用途则更为广泛。显示中很多情况下我们都没有一个准确的分类，聚类分析方法用数学的方法分析特征向量之间的距离及分散程度。有些特征向量可能聚集成若干个群，按各个群之间的距离远近分类。方法一般有分级聚类方法和迭代最优化法。关于聚类，课堂上老师也做了很多讲解，重要的是选取合适的聚类准则和类间相似性的测度，使类内相似度尽量大、类间相似度尽量小。

一个模式识别系统通常包括典型的四个部分：对原始数据的获取和预处理，特征提取与特征选择，分来或聚类，后处理。这是整个系统的核心所在，也是模式识别学科的主要研究内容。

1. 专家系统

专家系统是目前人工智能应用中最成功的一个领域。它的目标是：把专门领域中若干个人类专家的知识和思考、解决问题的方法以适当方式存储在计算机中，使计算机能在推理机的控制下模仿人类专家去解决问题。这种设想显然非常具有实际意义。专家系统目前也确实取得了一些不错的成就。比如说，美国在伊拉克发动的沙漠风暴之战中，利用专家系统完成了对1500磅物资的规划运输。同时，专家系统在数据解释、诊断、监测、控制、银行业务决策中也都各有建树。

专家系统的基础是专家经验性的知识，中心是知识库和推理中心，它的智能软件系统结构可以表示为：知识+推理=系统。

按照知识获取机制，专家系统将大量的专家知识存储在知识库中，同时，具备了推理机制和解释机制，推理过程记在动态存储器中。比较出乎我预料的，专家系统的瓶颈竟然是在于知识获取。知识获取对于人力财力消耗比较多，无论是人工获取还是机器学习，或者文字识别。专家系统是比较有针对性的，它针对它所要作用的领域，专门设定好工作方式。

总之，专家系统是面向一个专业领域定制的一个知识获取到应用的系统。是人工智能在应用方面的优先的示范。

1. 机器学习

桑克说：”一台计算机若不会学习，就不能说它有智能。“机器学习应该是什么样的：（1）有学习环境，包括信息来源；（2）有学习能力，能在与环境多次相互作用下逐步学到有关知识，在加以实践、评价（就像我们通常的演绎、反馈、加以调整的过程）；（3）能用所学的知识解决问题，也就是要具备可靠的分析、决策、控制能力；（4）能提高系统性能，通过学习，随着学习知识的增长，解决问题的能力要得到提高。

典型的机器学习的模型应该是：环境 -> 学习环节 -> 知识库 ->执行环节，执行环节又反作用于学习环节。按照字面意思并不是很难理解，中心思想是这样的，在具体不同的系统中又有细节构造上的不同。

如果把机器当作人的话，它的学习也可以类比，当老师讲解的很细时，学生就学的很轻松；而老师如果只是启发式的教育，那么学生就需要很多思考很多学习。按照机器负担由轻到重，有下面一些机器学习的分类：机械学习、讲授学习、演绎学习、解释学习、类比学习、归纳学习。这些方法都更倾向于接收知识，复杂一些的对知识进行分类、处理、发现，但都不太能满足当下的需求。当前又有了新一层级的基于系统性的分类：归纳学习、分析学习、联结学习、遗传算法。从分析学习开始，系统已经不仅仅是接受了，分析学习的推理过程是基于演绎的。遗传算法则成为了一个新的重要分支。

AlphaGo掀起了人工智能热的一波高潮，它学习了人类千年的棋谱，深度学习，通过走棋网络、快速走子、估值网络、蒙特卡罗树搜索等方法打败了世界冠军李世石。而后又有了AlphaGo Zero，在没有任何棋谱和人类先验知识的情况下，左右互博490万棋局学会下棋。它在掌握规则后自我演绎、反馈后得出了知识库。不经意间克服了人类难以跳出的局部最优性。AlphaGo的DNN模型的缺点——训练过程需要消耗大量人类标注样本也得到了解决，看起来很完美。

当然，在现实生活中，并不像围棋这样有着精准的规则和为我们画好的棋盘，它有很多不确定性。相对而言，围棋的演绎是可以实现的，而对现实生活中的很多领域，并不是那么容易实现的。

有AlphaGo我们也可以来推断一个过程，首先我们需要有棋谱，这是实例空间；接下来，我们要告诉机器，棋谱的含义，什么是棋子、什么是棋盘、大小什么的，这是解释例子；然后我们需要告诉机器下棋的规则，这个规则正是从前面的实例也就是棋谱中归纳出来的，这是规则空间；最后有了规则空间后，搜索规则空间就是下棋了，这是实验规则。在具体系统的设计中，当然会有不断的归纳分析多次测试，但核心思想是相似的。

事实上，我们目前的机器学习更多的是从海量的人类智能行为中归纳出规则空间和从规则比较清晰的领域演绎出策略，更多随机性较强的领域还没有很好的发展，距离人们的想象还有一段距离。但随着人工智能的发展，机器的智能化是必然的。

1. 机器人

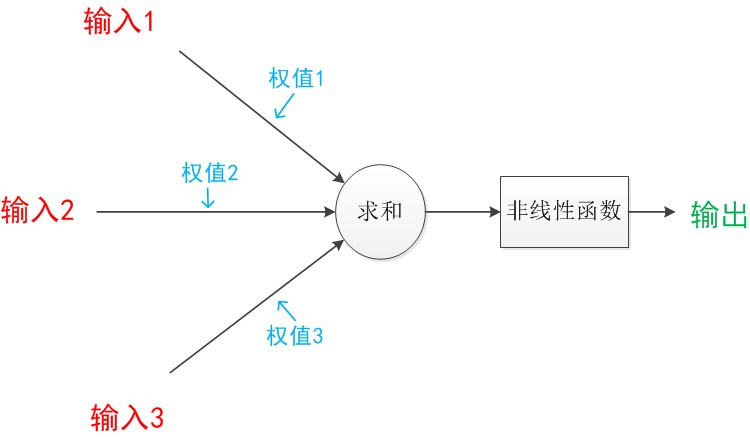
机器人是我们最早知道ai的事物，从电影、小说中我们产生了很多对机器人的幻想。图灵测试是说我们无法区分是在和机器还是人交流，加以延伸一下，当人工智能从交流延伸到声音、行为、思维上时，我们或许就再也无法区分人与机器人。事实上，现在模拟的电子音与人声是非常接近的，siri对问题的回答也具有一定的逻辑性，在加上图像识别、语音识别技术，仿真机器人的出现似乎很有希望。

话归原处，从技术实现上来说，机器人需要实现的功能主要有下面两个：（1）模式识别，相当于让它有了视觉、触觉；（2）运动协调推理，也就是接受外界刺激之后做出相应的反应。模式识别我们在第一部分中已经介绍过了，而运动协调推理主要包含了一个决策的行为规划。

目前我们对于机器人的开放应该是有各个方向的，有很大一部分在它说话和一些既定的动作，不具备太多的智能。如果我们能够加以对人类行为的归纳和演绎，模拟出人类更多的应激（甚至没有刺激）行为，赋予机器，那么机器人很有可能成为电影中的那样，和人类几乎没有差别。

1. 神经网络

在研究人工智能时，我们不免想到，人类的智能来源——大脑。那么智能与大脑的结构是否存在关系呢？人类对于大脑中神经元的研究由来已久，一个神经元通常具有多个树突，主要用来接受传入信息；而轴突只有一条，轴突尾端有许多轴突末梢可以给其他多个神经元传递信息。轴突末梢跟其他神经元的树突产生连接，从而传递信号。这个连接的位置在生物学上叫做“突触”。1943年，心理学家McCulloch和数学家Pitts参考了生物神经元的结构，发表了抽象的神经元模型MP。神经元模型是一个包含输入，输出与计算功能的模型。输入可以类比为神经元的树突，而输出可以类比为神经元的轴突，计算则可以类比为细胞核。



每一个连接上都有一个权重。一个神经网络的训练算法就是让权重的值调整到最佳，以使得整个网络的预测效果最好。

接下来，有了单层神经网络——感知器。它是首个可以学习的人工神经网络，可以学习识别简单图形。在“感知器”中，有两个层次。分别是输入层和输出层。输入层里的“输入单元”只负责传输数据，不做计算。输出层里的“输出单元”则需要对前面一层的输入进行计算。我们把需要计算的层次称之为“计算层”，并把拥有一个计算层的网络称之为“单层神经网络”。

感知器类似一个逻辑回归模型，可以做线性分类任务。我们可以用决策分界来形象的表达分类的效果。决策分界就是在二维的数据平面中划出一条直线，当数据的维度是3维的时候，就是划出一个平面，当数据的维度是n维时，就是划出一个n-1维的超平面。

神经网络的大范围推广与使用是两层神经网络（多层感知器）的产生时。两层神经网络除了包含一个输入层，一个输出层以外，还增加了一个中间层。与单层神经网络不同。理论证明，两层神经网络可以无限逼近任意连续函数。这是什么意思呢？也就是说，面对复杂的非线性分类任务，两层（带一个隐藏层）神经网络可以分类的很好。从输入层到隐藏层时，数据发生了空间变换。也就是说，两层神经网络中，隐藏层对原始的数据进行了一个空间变换，使其可以被线性分类，然后输出层的决策分界划出了一个线性分类分界线，对其进行分类。这样就导出了两层神经网络可以做非线性分类的关键--隐藏层。联想到我们一开始推导出的矩阵公式，我们知道，矩阵和向量相乘，本质上就是对向量的坐标空间进行一个变换。因此，隐藏层的参数矩阵的作用就是使得数据的原始坐标空间从线性不可分，转换成了线性可分。

两层神经网络通过两层的线性模型模拟了数据内真实的非线性函数。因此，多层的神经网络的本质就是复杂函数拟合。

BP网络是1986年由Rumelhart和McCelland为首的科学家小组提出，是一种按误差逆传播算法训练的多层前馈网络，是目前应用最广泛的神经网络模型之一。BP网络能学习和存贮大量的输入-输出模式映射关系，而无需事前揭示描述这种映射关系的数学方程。它的学习规则是使用最速下降法，通过反向传播来不断调整网络的权值和阈值，使网络的误差平方和最小。BP神经网络模型拓扑结构包括输入层、隐层和输出层。

周而复始的信息正向传播和误差反向传播过程，是各层权值不断调整的过程，也是神经网络学习训练的过程，此过程一直进行到网络输出的误差减少到可以接受的程度，或者预先设定的学习次数为止。具体步骤如下；

1、从训练集中取出某一样本，把信息输入网络中。

2、通过各节点间的连接情况正向逐层处理后，得到神经网络的实际输出。 3、计算网络实际输出与期望输出的误差。

4、将误差逐层反向回传至之前各层，并按一定原则将误差信号加载到连接权值上，使整个神经网络的连接权值向误差减小的方向转化。

5、対训练集中每一个输入—输出样本对重复以上步骤，直到整个训练样本集的误差减小到符合要求为止。

但是神经网络仍然存在若干的问题：尽管使用了BP算法，一次神经网络的训练仍然耗时太久，而且困扰训练优化的一个问题就是局部最优解问题，这使得神经网络的优化较为困难。同时，隐藏层的节点数需要调参，这使得使用不太方便。

到现在，神经网络发展到了多层神经网络，也就是我们熟悉的深度学习。我们延续两层神经网络的方式来设计一个多层神经网络。　　在两层神经网络的输出层后面，继续添加层次。原来的输出层变成中间层，新加的层次成为新的输出层。增加更多的层次有什么好处？更深入的表示特征，以及更强的函数模拟能力。

　　更深入的表示特征可以这样理解，随着网络的层数增加，每一层对于前一层次的抽象表示更深入。在神经网络中，每一层神经元学习到的是前一层神经元值的更抽象的表示。例如第一个隐藏层学习到的是“边缘”的特征，第二个隐藏层学习到的是由“边缘”组成的“形状”的特征，第三个隐藏层学习到的是由“形状”组成的“图案”的特征，最后的隐藏层学习到的是由“图案”组成的“目标”的特征。通过抽取更抽象的特征来对事物进行区分，从而获得更好的区分与分类能力。

多层神经网络比较典型的有CNN(卷积神经网络)、RNN(循环神经网络)、DNN(深度神经网络)。神经网络的层数直接决定了它对现实的刻画能力——利用每层更少的神经元拟合更加复杂的函数。CNN对参数处理的需求已经下降了，简单来说，训练CNN在相当意义上是在训练每一个卷积层的滤波器。让这些滤波器组对特定的模式有高的激活，以达到CNN网络的分类/检测等目的。DNN是一个大类，层次CNN是一个典型空间深度上的神经网络，RNN是在时间上深度的神经网络。在RNN中，神经元的输出可以在下一个时间戳直接作用在自身，当然，，如果担心在时间上的梯度消失，可以使用长短时间记忆单元LSTM，通过门的开关实现时间上的记忆功能，防止梯度消失。

更高层的神经网络由于计算性能的问题，以及一些计算方法的问题，更多的优势无法得到体现。当然，我们可以寄期待于量子计算机强大的计算能力。

最后，关于此门课程大作业，我们首先是对机械方面小车的搭建制造。小车uno板的代码并不算晦涩，阅读老师给的样例之后举一反三并不算困难。之后对图像识别minist的代码进行了修改调试，把输入由设定的测试集改为了输入自己拍摄到的数字图片。刚开始的KNN模型识别正确率较低，后来改用了Softmax之后提高了正确率，测试结果比较好。小车有一点显著的不足是，由于材料购买的比较随意，所以可能质量不太好，导致小车的运行不太稳定，经常出现硬件问题，对实验进行比较有影响。

总的来说，这门课程给了我人工智能的入门引导，对人工智能了解过、接触过、使用过，自学人工智能无疑是非常困难的，现在接下来进一步的学习给具备了一定的基础，有助于进一步的学习和发展。