研究方向

机器学习

人类一直试图让机器具有智能，也就是人工智能（Artificial Intelligence）。从上世纪50年代，人工智能的发展经历了“推理期”，通过赋予机器逻辑推理能力使机器获得智能，当时的AI程序能够证明一些著名的数学定理，但由于机器缺乏知识，远不能实现真正的智能。因此，70年代，人工智能的发展进入“知识期”，即将人类的知识总结出来教给机器，使机器获得智能。在这一时期，大量的专家系统问世，在很多领域取得大量成果，但由于人类知识量巨大，故出现“知识工程瓶颈”。

无论是“推理期”还是“知识期”，机器都是按照人类设定的规则和总结的知识运作，永远无法超越其创造者，其次人力成本太高。于是，一些学者就想到，如果机器能够自我学习问题不就迎刃而解了吗！机器学习（Machine Learning）方法应运而生，人工智能进入“机器学习时期”。“机器学习时期”也分为三个阶段，80年代，连接主义较为流行，代表工作有感知机（Perceptron）和神经网络（Neural Network）。90年代，统计学习方法开始占据主流舞台，代表性方法有支持向量机（Support Vector Machine），进入21世纪，深度神经网络被提出，连接主义卷土从来，随着数据量和计算能力的不断提升，以深度学习（Deep Learning）为基础的诸多AI应用逐渐成熟。



计算机视觉



目前，计算机视觉是深度学习领域最热门的研究领域之一。计算机视觉实际上是一个跨领域的交叉学科，包括计算机科学（图形、算法、理论、系统、体系结构），数学（信息检索、机器学习），工程学（机器人、语音、自然语言处理、图像处理），物理学（光学 ），生物学（神经科学）和心理学（认知科学）等等。许多科学家认为，计算机视觉为人工智能的发展开拓了道路。

计算机视觉本质上是研究视觉感知问题。视觉感知，根据维科百基（Wikipedia）的定义, 是指对“环境表达和理解中，对视觉信息的组织、识别和解释的过程”。根据这种定义，计算机视觉的目标是对环境的表达和理解，核心问题是研究如何对输入的图像信息进行组织，对物体和场景进行识别，进而对图像内容给予解释。

计算机视觉是以图像（视频）为输入，以对环境的表达（representation）和理解为目标，研究图像信息组织、物体和场景识别、进而对事件给予解释的学科。从目前的研究现状看，目前还主要聚焦在图像信息的组织和识别阶段，对事件解释还鲜有涉及，至少还处于非常初级的阶段。

尽管人们对计算机视觉这门学科的起始时间和发展历史有不同的看法，但应该说， 1982年马尔( David Marr )《视觉》（Marr, 1982）一书的问世，标志着计算机视觉成为了一门独立学科。计算机视觉的研究内容，大体可以分为物体视觉（object vision）和空间视觉（spatial vision）二大部分. 物体视觉在于对物体进行精细分类和鉴别，而空间视觉在于确定物体的位置和形状，为“动作（action）” 服务。正像著名的认知心理学家J.J. Gibson 所言，视觉的主要功能在于“适应外界环境，控制自身运动”。 适应外界环境和控制自身运动，是生物生存的需求，这些功能的实现需要靠物体视觉和空间视觉协调完成。

计算机视觉40多年的发展中，尽管人们提出了大量的理论和方法，但总体上说，计算机视觉经历了4个主要历程。即： 马尔计算视觉、主动和目的视觉、多视几何与分层三维重建和基于学习的视觉。

自然语言处理

简单地说，自然语言处理（Natural Language Processing，简称NLP）就是用计算机来处理、理解以及运用人类语言(如中文、英文等)，它属于人工智能的一个分支，是计算机科学与语言学的交叉学科，又常被称为计算语言学。由于自然语言是人类区别于其他动物的根本标志。没有语言，人类的思维也就无从谈起，所以自然语言处理体现了人工智能的最高任务与境界，也就是说，只有当计算机具备了处理自然语言的能力时，机器才算实现了真正的智能。

从研究内容来看，自然语言处理包括语法分析、语义分析、篇章理解等。从应用角度来看，自然语言处理具有广泛的应用前景。特别是在信息时代，自然语言处理的应用包罗万象，例如：机器翻译、手写体和印刷体字符识别、语音识别及文语转换、信息检索、信息抽取与过滤、文本分类与聚类、舆情分析和观点挖掘等，它涉及与语言处理相关的数据挖掘、机器学习、知识获取、知识工程、人工智能研究和与语言计算相关的语言学研究等。

值得一提的是，自然语言处理的兴起与机器翻译这一具体任务有着密切联系。机器翻译指的是利用计算机自动地将一种自然语言翻译为另外一种自然语言。例如自动将英文“I like Beijing Tiananmen Square”翻译为“我爱北京天安门”，或者反过来将“我爱北京天安门”翻译为“I like Beijing Tiananmen Square”。由于人工进行翻译需要训练有素的双语专家，翻译工作非常耗时耗力。更不用说需要翻译一些专业领域文献时，还需要翻译者了解该领域的基本知识。世界上有超过几千种语言，而仅联合国的工作语言就有六种之多。如果能够通过机器翻译准确地进行语言间的翻译，将大大提高人类沟通和了解的效率。

《圣经》里有一个故事说巴比伦人想建造一座塔直通天堂。建塔的人都说着同一种语言，心意相通、齐心协力。上帝看到人类竟然敢做这种事情，就让他们的语言变得不一样。因为人们听不懂对方在讲什么，于是大家整天吵吵闹闹，无法继续建塔。后来人们把这座塔叫作巴别塔，而“巴别”的意思就是“分歧”。虽然巴别塔停建了，但一个梦想却始终萦绕在人们心中：人类什么时候才能拥有相通的语言，重建巴别塔呢？机器翻译被视为“重建巴别塔”的伟大创举。假如能够实现不同语言之间的机器翻译，我们就可以理解世界上任何人说的话，与他们进行交流和沟通，再也不必为相互不能理解而困扰。

事实上，“人工智能”被作为一个研究问题正式提出来的时候，创始人把计算机国际象棋和机器翻译作为两个标志性的任务，认为只要国际象棋系统能够打败人类世界冠军，机器翻译系统达到人类翻译水平，就可以宣告人工智能的胜利。四十年后的1997年，IBM公司的深蓝超级计算机已经能够打败国际象棋世界冠军卡斯帕罗夫。而机器翻译到现在仍无法与人类翻译水平相比，从此可以看出自然语言处理有多么困难。

目前，人们主要通过两种思路来进行自然语言处理，一种是基于规则的理性主义，另外一种是基于统计的经验主义。理性主义方法认为，人类语言主要是由语言规则来产生和描述的，因此只要能够用适当的形式将人类语言规则表示出来，就能够理解人类语言，并实现语言之间的翻译等各种自然语言处理任务。而经验主义方法则认为，从语言数据中获取语言统计知识，有效建立语言的统计模型。因此只要能够有足够多的用于统计的语言数据，就能够理解人类语言。然而，当面对现实世界充满模糊与不确定性时，这两种方法都面临着各自无法解决的问题。例如，人类语言虽然有一定的规则，但是在真实使用中往往伴随大量的噪音和不规范性。理性主义方法的一大弱点就是鲁棒性差，只要与规则稍有偏离便无法处理。而对于经验主义方法而言，又不能无限地获取语言数据进行统计学习，因此也不能够完美地理解人类语言。二十世纪八十年代以来的趋势就是，基于语言规则的理性主义方法不断受到质疑，大规模语言数据处理成为目前和未来一段时期内自然语言处理的主要研究目标。统计学习方法越来越受到重视，自然语言处理中越来越多地使用机器自动学习的方法来获取语言知识。

自然语言处理作为一个年龄尚不足一个世纪的新兴学科，正在进行着突飞猛进的发展。回顾自然语言处理的发展历程，并不是一帆风顺，有过低谷，也有过高潮。而现在我们正面临着新的挑战和机遇。例如，目前网络搜索引擎基本上还停留在关键词匹配，缺乏深层次的自然语言处理和理解。语音识别、文字识别、问答系统、机器翻译等目前也只能达到很基本的水平。路漫漫其修远兮，自然语言处理作为一个高度交叉的新兴学科，不论是探究自然本质还是付诸实际应用，在将来必定会有令人期待的惊喜和异常快速的发展。

人工神经网络

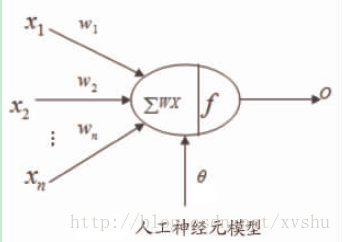
人工神经网络（Artificial Neural Network，即ANN ），是20世纪80 年代以来人工智能领域兴起的研究热点。它从信息处理角度对人脑神经元网络进行抽象， 建立某种简单模型，按不同的连接方式组成不同的网络。在工程与学术界也常直接简称为神经网络或类神经网络。神经网络是一种运算模型，由大量的节点（或称神经元）之间相互联接构成。每个节点代表一种特定的输出函数，称为激励函数（activation function）。每两个节点间的连接都代表一个对于通过该连接信号的加权值，称之为权重，这相当于人工神经网络的记忆。网络的输出则依网络的连接方式，权重值和激励函数的不同而不同。而网络自身通常都是对自然界某种算法或者函数的逼近，也可能是对一种逻辑策略的表达。

最近十多年来，人工神经网络的研究工作不断深入，已经取得了很大的进展，其在模式识别、智能机器人、自动控制、预测估计、生物、医学、经济等领域已成功地解决了许多现代计算机难以解决的实际问题，表现出了良好的智能特性。

数学模型

人工神经元结构：人工神经元的研究源于脑神经元学说，19世纪末，在生物、生理学领域，Waldeger等人创建了神经元学说。

人工神经网络是由大量处理单元经广泛互连而组成的人工网络，用来模拟脑神经系统的结构和功能。而这些处理单元我们把它称作人工神经元。人工神经网络可看成是以人工神经元为节点，用有向加权弧连接起来的有向图。在此有向图中，人工神经元就是对生物神经元的模拟，而有向弧则是轴突—-突触—-树突对的模拟。有向弧的权值表示相互连接的两个人工神经元间相互作用的强弱。人工神经元结构如下图所示：



神经网络从两个方面模拟大脑：

(1)、神经网络获取的知识是从外界环境中学习得来的。

(2)、内部神经元的连接强度，即突触权值，用于储存获取的知识。

神经网络系统由能够处理人类大脑不同部分之间信息传递的由大量神经元连接形成的拓扑结构组成，依赖于这些庞大的神经元数目和它们之间的联系，人类的大脑能够收到输入的信息的刺激由分布式并行处理的神经元相互连接进行非线性映射处理，从而实现复杂的信息处理和推理任务。

对于某个处理单元（神经元）来说，假设来自其他处理单元（神经元）i的信息为Xi，它们与本处理单元的互相作用强度即连接权值为Wi, i=0,1,…,n-1,处理单元的内部阈值为θ。

那么本处理单元（神经元）的输入为20171018094426537：，而处理单元的输出为：20171018094445009式中，xi为第i个元素的输入，wi为第i个处理单元与本处理单元的互联权重即神经元连接权值。f称为激活函数或作用函数，它决定节点（神经元）的输出。θ表示隐含层神经节点的阈值。

结合以上数学模型，神经网络的主要工作是建立模型和确定权值，一般有前向型和反馈型两种网络结构。通常神经网络的学习和训练需要一组输入数据和输出数据对，选择网络模型和传递、训练函数后，神经网络计算得到输出结果，根据实际输出和期望输出之间的误差进行权值的修正，在网络进行判断的时候就只有输入数据而没有预期的输出结果。神经网络一个相当重要的能力是其网络能通过它的神经元权值和阈值的不断调整从环境中进行学习，直到网络的输出误差达到预期的结果，就认为网络训练结束。

对于这样一种多输入、单输出的基本单元可以进一步从生物化学、电生物学、数学等方面给出描述其功能的模型。利用大量神经元相互连接组成的人工神经网络，将显示出人脑的若干特征，人工神经网络也具有初步的自适应与自组织能力。在学习或训练过程中改变突触权重wij值，以适应周围环境的要求。同一网络因学习方式及内容不同可具有不同的功能。人工神经网络是一个具有学习能力的系统，可以发展知识，以至超过设计者原有的知识水平。通常，它的学习(或训练)方式可分为两种，一种是有监督(supervised)或称有导师的学习，这时利用给定的样本标准进行分类或模仿；另一种是无监督(unsupervised)学习或称无导师学习，这时，只规定学习方式或某些规则，而具体的学习内容随系统所处环境(即输入信号情况)而异，系统可以自动发现环境特征和规律性，具有更近似于人脑的功能。