

深度学习调研报告

1611736 钟腾

什么是深度学习

深度学习的概念源于人工神经网络的研究。含多隐层的多层感知器就是一种深度学习结构。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，以发现数据的分布式特征表示。

深度学习的概念由Hinton等人于2006年提出。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。

深度学习是机器学习中一种基于对数据进行表征学习的方法。观测值（例如一幅图像）可以使用多种方式来表示，如每个像素强度值的向量，或者更抽象地表示成一系列边、特定形状的区域等。而使用某些特定的表示方法更容易从实例中学习任务（例如，人脸识别或面部表情识别）。深度学习的好处是使用非监督式或半监督式的特征学习和分层特征提取高效算法来替代手工获取特征。

深度学习是机器学习研究中的一个新领域，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。

同机器学习方法一样，深度机器学习方法也有监督学习与无监督学习之分。不同的学习框架下建立的学习模型很不同。例如，卷积神经网络（Convolutional neural networks，简称CNNs）就是一种深度的监督学习下的机器学习模型，而深度置信网（Deep Belief Nets，简称DBNs）就是一种无监督学习下的机器学习模型。

主流框架

在深度学习初始阶段，每个深度学习研究者都需要写大量的重复代码。为了提高工作效率，这些研究者就将这些代码写成了一个框架放到网上让所有研究者一起使用。接着，网上就出现了不同的框架。随着时间的推移，最为好用的几个框架被大量的人使用从而流行了起来。全世界最为流行的深度学习框架有PaddlePaddle、Tensorflow、Caffe、Theano、MXNet、Torch和PyTorch。

1)PaddlePaddle

PaddlePaddle是百度研发的开源开放的深度学习平台，是国内最早开源、也是当前唯一一个功能完备的深度学习平台。依托百度业务场景的长期锤炼，PaddlePaddle有最全面的官方支持的工业级应用模型，涵盖自然语言处理、计算机视觉、推荐引擎等多个领域，并开放多个领先的预训练中文模型，以及多个在国际范围内取得竞赛冠军的算法模型。

PaddlePaddle同时支持稠密参数和稀疏参数场景的超大规模深度学习并行训练，支持千亿规模参数、数百个节点的高效并行训练，也是最早提供如此强大的深度学习并行技术的深度学习框架。PaddlePaddle拥有强大的多端部署能力，支持服务器端、移动端等多种异构硬件设备的高速推理，预测性能有显著优势。目前PaddlePaddle已经实现了API的稳定和向后兼容，具有完善的中英双语使用文档，形成了易学易用、简洁高效的技术特色。

PaddlePaddle 3.0版本升级为全面的深度学习开发套件，除了核心框架，还开放了VisualDL、PARL、AutoDL、EasyDL、AI Studio等一整套的深度学习工具组件和服务平台，更好地满足不同层次的深度学习开发者的开发需求，具备了强大支持工业级应用的能力，已经被中国企业广泛使用，也拥有了活跃的开发社区生态。

2)Tensorflow

Google 开源的Tensorflow是一款使用C++语言开发的开源数学计算软件，使用数据流图(Data Flow Graph)的形式进行计算。图中的节点代表数学运算，而图中的线条表示多维数据数组(tensor)之间的交互。Tensorflow灵活的架构可以部署在一个或多个CPU、GPU的台式及服务器中，或者使用单一的API应用在移动设备中。Tensorflow最初是由研究人员和Google Brain 团队针对机器学习和深度神经网络进行研究而开发，开源之后几乎可以在各个领域适用。

Tensorflow是全世界使用人数最多、社区最为庞大的一个框架，因为Google公司出品，所以维护与更新比较频繁，并且有着Python和C++的接口，教程也非常完善，同时很多论文复现的第一个版本都是基于Tensorflow写的，所以是深度学习界框架默认的老大。

3)Caffe

和Tensorflow名气一样大的是深度学习框架Caffe，由加州大学伯克利的Phd贾扬清开发，全称是Convolutional Architecture for Fast Feature Embedding，是一个清晰而高效的开源深度学习框架，由伯克利视觉中心(Berkeley Vision and Learning Center, BVLC)进行维护。

从它的名字就可以看出其对于卷积网络的支持特别好，同时也是用C++写的，提供的C++接口，也提供了matlab接口和python接口。

Caffe之所以流行，是因为之前很多ImageNet比赛里面使用的网络都是用Caffe写的，所以如果你想使用这些比赛的网络模型就只能使用Caffe，这也就导致了很多人直接转到Caffe这个框架下面。

Caffe的缺点是不够灵活，同时内存占用高，只提供了C++的接口，Caffe的升级版本Caffe2已经开源了，修复了一些问题，同时工程水平得到了进一步提高。

4)MXNet

MXNet的主要作者是李沐，最早就是几个人抱着纯粹对技术和开发的热情做起来的，如今成了亚马逊的官方框架，有着非常好的分布式支持，而且性能特别好，占用显存低，同时其开发的语言接口不仅仅有Python和C++，还有R，Matlab，Scala，JavaScript，等等，可以说能够满足使用任何语言的人。

但是MXNet的缺点也很明显，教程不够完善，使用的人不多导致社区不大，同时每年很少有比赛和论文是基于MXNet实现的，这就使得MXNet的推广力度和知名度不高。

5)Torch

Torch是一个有大量机器学习算法支持的科学计算框架，其诞生已有十年之久，但是真正起势得益于Facebook开源了大量Torch的深度学习模块和扩展。Torch的特点在于特别灵活，但是另一个特殊之处是采用了编程语言Lua，在深度学习大部分以Python为编程语言的大环境之下，一个以Lua为编程语言的框架有着更多的劣势，这一项小众的语言增加了学习使用Torch这个框架的成本。

6)PyTorch

PyTorch的前身便是Torch，其底层和Torch框架一样，但是使用Python重新写了很多内容，不仅更加灵活，支持动态图，而且提供了Python接口。它是由Torch7团队开发，是一个以Python优先的深度学习框架，不仅能够实现强大的GPU加速，同时还支持动态神经网络，这是很多主流深度学习框架比如Tensorflow等都不支持的。

PyTorch既可以看作加入了GPU支持的numpy，同时也可以看成一个拥有自动求导功能的强大的深度神经网络。除了Facebook外，它已经被Twitter、CMU和Salesforce等机构采用。

个人认识

深度学习作为计算机行业新兴热门技术，作为软件工程的学生自然听说过该技术名词，然而对深度学习的首次接触还是在本学期的科研项目中。科研中为检测大量新兴的KPI流快速部署异常检测模型，用到了深度学习中的半监督学习。由于传统方法的低精度和慢部署，在项目中采用了能够自动部署的深度学习模式，并鉴于监督学习的人工标注量大以及参数调整困难和无监督学习的低精度采取了综合两者长处并消除两者短处的半监督学习模式，算是对深度学习的第一次实际应用。对"深度学习"的深度学习也会在机器学习课程上继续进行。