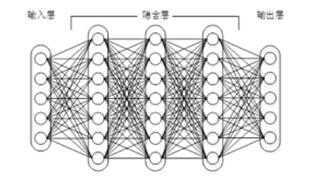
上世纪八十年代Rumelhart、Williams、Hinton、LeCun等人发明了多层感知机(multilayer perceptron)。层感知机的结构：



多层感知机可以摆脱早期离散传输函数的束缚，使用sigmoid（https://gss3.bdstatic.com/-Po3dSag_xI4khGkpoWK1HF6hhy/baike/s%3D99/sign=a46bd6f1dd33c895a27e9472d01340df/0df3d7ca7bcb0a4659502a5f6f63f6246b60af62.jpgSigmoid函数常被用作神经网络的阈值函数，将变量映射到0,1之间。）或tanh（https://gss2.bdstatic.com/-fo3dSag_xI4khGkpoWK1HF6hhy/baike/s%3D177/sign=adf59bb1f51fbe09185ec7135c610c30/96dda144ad345982648941550bf431adcaef84f2.jpg）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。这就是我们现在所说的神经网络NN。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

（BP算法：它建立在梯度下降法的基础上。BP网络的输入输出关系实质上是一种映射关系：一个n输入m输出的BP神经网络所完成的功能是从n维欧氏空间向m维欧氏空间中一有限域的连续映射，这一映射具有高度非线性。BP算法的学习过程由正向传播过程和反向传播过程组成。在正向传播过程中，输入信息通过输入层经隐含层，逐层处理并传向输出层。如果在输出层得不到期望的输出值，则取输出与期望的误差的平方和作为目标函数，转入反向传播，逐层求出目标函数对各神经元权值的偏导数，构成目标函数对权值向量的梯量，作为修改权值的依据，网络的学习在权值修改过程中完成。误差达到所期望值时，网络学习结束。

**激励传播**

每次[迭代](https://baike.baidu.com/item/%E8%BF%AD%E4%BB%A3/8415523)中的传播环节包含两步：

1. (前向传播阶段)将训练输入送入网络以获得激励响应；
2. (反向传播阶段)将激励响应同训练输入对应的目标输出求差，从而获得隐层和[输出层](https://baike.baidu.com/item/%E8%BE%93%E5%87%BA%E5%B1%82/7202179)的响应误差。

**权重更新**

对于每个突触上的权重，按照以下步骤进行更新：

1. 将输入激励和响应误差相乘，从而获得权重的梯度；
2. 将这个梯度乘上一个比例并取反后加到权重上。
3. 这个比例将会影响到训练过程的速度和效果，因此称为“训练因子”。梯度的方向指明了误差扩大的方向，因此在更新权重的时候需要对其取反，从而减小权重引起的误差。

BP网络可以较好地实现宽频带、小噪比、信号模式较少情况下的信号识别和信噪分离。

）

感知机缺陷：

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](http://ad.doubleclick.net/ddm/trackclk/N7442.5006CHINABYTE/B10313247.138166523;dc_trk_aid=310538354;dc_trk_cid=74205219)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，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。

解决方案：

2006年hinton利用与训练方法缓解了局部最优解问题,隐含层——>7层（深度无固定定义，语音识别中4层网络就被认为较深，图像识别20层以上网络常见）

克服梯度消失：ReLUctant、maxout等传输函数代替了sigmoid——》》》形成DNN，结构上全连接的DNN与多层感知机无区别。2016年出现了Highway network高速公路网络和deep residual learning 深度残差学习进一步避免了梯度消失，网络层数达到一百多层。

全连接的DNN结构里上下层神经元和所有上层神经元都能够形成连接

潜在问题：

参数数量膨胀。如：像素１０００×１０００的图像隐含层１Ｍ个节点，一层有１０^１２个权重要训练，容易过拟合，极易陷入局部最优，且图像中的固有局部模式可以利用，如轮廓，边界，人眼，嘴　鼻子（适应需求：ＣＮＮ：并不是所有的上下层神经元都直接相连，而是通过卷积和作为中介图像通过卷积后仍保留原先位置）

无法对时间序列上的变化进行建模（适应需求ＲＮＮ）

深度学习（也称为深层结构学习或层次学习）是[人工神经网络](https://en.wikipedia.org/wiki/Artificial_neural_networks)（ANN）在包含多个[隐藏层的](https://en.wikipedia.org/wiki/Multilayer_perceptron#Layers)学习任务中的应用。深度学习是基于[学习数据表示](https://en.wikipedia.org/wiki/Learning_representation)的更广泛的[机器学习](https://en.wikipedia.org/wiki/Machine_learning)方法系列的一部分，而不是任务特定的算法。学习可以[监督](https://en.wikipedia.org/wiki/Supervised_learning)，部分监督或[无监督](https://en.wikipedia.org/wiki/Unsupervised_learning)。

深度学习是一类[机器学习](https://en.wikipedia.org/wiki/Machine_learning) [算法](https://en.wikipedia.org/wiki/Algorithm)：

* 使用多层[非线性处理](https://en.wikipedia.org/wiki/Nonlinear_filter)单元的级联来进行[特征提取](https://en.wikipedia.org/wiki/Feature_extraction)和转换。每个连续的层使用前一层的输出作为输入。算法可以被[监督](https://en.wikipedia.org/wiki/Supervised_learning)或[无监督](https://en.wikipedia.org/wiki/Unsupervised_learning)，应用包括模式分析（无监督）和分类（监督）。
* 是基于（无监督）学习多个级别的特征或数据表示。较高级别的特征源于较低级别的特征以形成分层表示。
* 是更广泛的机器学习领域的学习表示数据的一部分。
* 学习对应于不同抽象层次的多层次的表示; 这些层次形成了一个概念层次。

**困难和挑战**：（以图像为例与声场类比）

（声音方位变化）（声音响度变化）（声音音调，音色变化）（噪声干扰）（同一人/物发出不同声音）

**数据驱动方法**：创造一个声场分类的算法——给计算机很多数据，然后实现学习算法，让计算机学习到每个类的外形。这种方法，就是数据驱动方法。

* 对于声场来说，第一步就是收集已经做好分类标注的声音片段来作为训练集
* 模型：

对数据集使用的特征和分类器的综合考量，其中包括已知的梅尔频谱倒谱系数或更多的专门特征，例如声音事件的直方图或从时间 - 频率表示中学习的渐变直方图，和隐马尔可夫模型（HMM），高斯混合模型（GMM）或支持向量机（SVM）的声学模型。

* （**图像分类**就是输入一个元素为像素值的数组，然后给它分配一个分类标签。完整流程如下：

**输入**：输入是包含N个图像的集合，每个图像的标签是K种分类标签中的一种。这个集合称为*训练集。*

**学习**：这一步的任务是使用训练集来学习每个类到底长什么样。一般该步骤叫做*训练分类器*或者*学习一个模型*。

**评价**：让分类器来预测它未曾见过的图像的分类标签，并以此来评价分类器的质量。我们会把分类器预测的标签和图像真正的分类标签对比。毫无疑问，分类器预测的分类标签和图像真正的分类标签如果一致，那就是好事，这样的情况越多越好。）

在**图像分类**问题中，给出一个由被标注了分类标签的图像组成的集合，要求算法能预测没有标签的图像的分类标签，并根据算法预测准确率进行评价。

* 了解了一个简单的图像分类器：**最近邻分类器(Nearest Neighbor classifier)**。分类器中存在不同的超参数(比如k值或距离类型的选取)，选取超参数的正确方法是：将原始训练集分为训练集和**验证集**，我们在验证集上尝试不同的超参数，最后保留表现最好那个。
* 如果训练数据量不够，使用**交叉验证**方法，它能帮助我们在选取最优超参数的时候减少噪音。
* 一旦找到最优的超参数，就让算法以该参数在测试集跑且只跑一次，并根据测试结果评价算法。
* 最近邻分类器能够在CIFAR-10上得到将近40%的准确率。该算法简单易实现，但需要存储所有训练数据，并且在测试的时候过于耗费计算能力。
* 最后，我们知道了仅仅使用L1和L2范数来进行像素比较是不够的，图像更多的是按照背景和颜色被分类，而不是语义主体分身。