

人工智能初步

浙江大学计算机学院

吴飞

2020年1月

内容提纲

全书概略

基于搜索问题的求解

决策树

回归分析

贝叶斯分析

神经网络学习

神经网络学习：感知机 (perceptrons)

在图2.7.3所示的感知机模型中，有 x_1 ， x_2 ， x_3 三个输入项、一个神经元（用圆圈表示）和一个输出项，每个输入项均通过一定的权重与神经元相连（如 w_1 是 x_1 与神经元相连的权重）。当然，在实际中，输入项可以有很多。一旦给定了输入项和权重，神经元如何进行操作得到输出项呢？

神经元可采用如下的简单方法进行工作：先计算输入项传递给神经元的消息总和，即 $w_1 \times x_1 + w_2 \times x_2 + w_3 \times x_3$ ，如果这个总和大于某个预先设定阈值（比如0.5），那么神经元就输出1，否则就输出0。于是，我们可以写出神经元进行信息处理的数学模型：

$F((w_1 \times x_1 + w_2 \times x_2 + w_3 \times x_3) > 0.5) \rightarrow 1$ 。

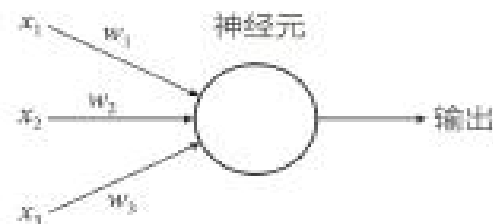


图2.7.3 感知机模型示意图

神经网络学习：感知机 (perceptrons)

感知机中神经元有两个操作：一是“汇总”与之相连的输入项传递而来的所有信息；二是对汇总后的信息做一个处理（图2.7.3 中神经元的处理规则是判断其汇总信息是否大于0.5，然后基于这个判断输出1 或0）。

如果有更多的输入，感知机的处理方法也是一样的，只不过需要更多权值 w 。这就是感知机的工作方法。

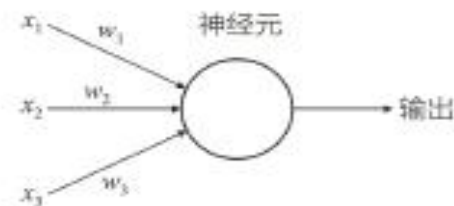
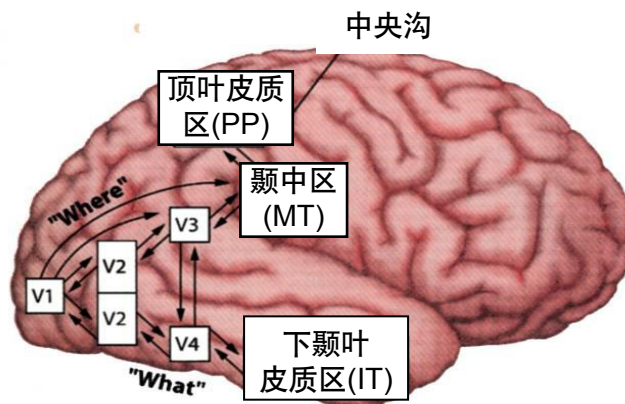
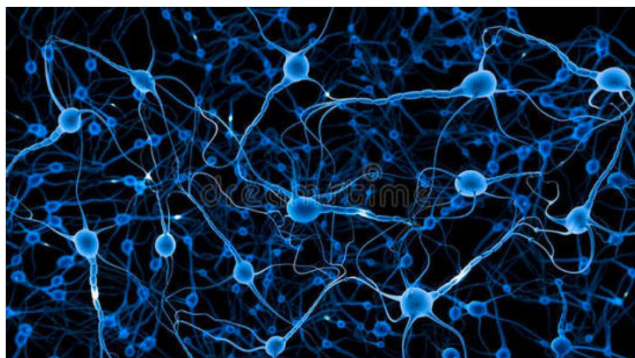


图2.7.3 感知机模型示意图



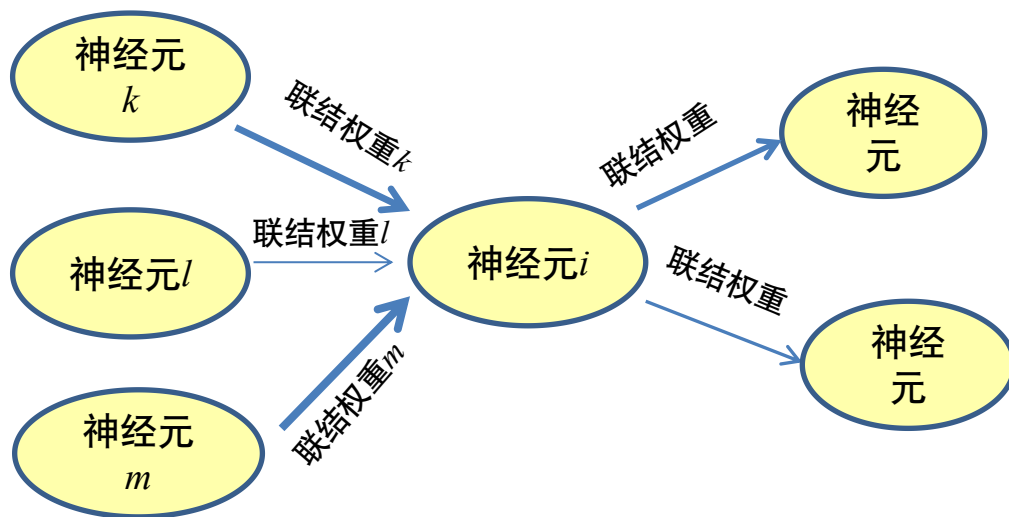
视觉感知存在层次回路、而非“开门见山”。因此要将网络结构“变深”

神经网络学习：神经学假设



- 物理反应：把前序神经元所传递过来的信息按照联结权重累加

神经元 i 获得信息 = 神经元 k 信息 \times 联结权重 k
+ 神经元 l 信息 \times 联结权重 l + 神经元 m 信息 \times 联结权重 m



- 化学反应：对神经元 i 获得信息施以一个非线性变换（通过激活函数），激活若干信息、而非“来者不拒”。
- 信息流通：将神经元非线性变换后所得信息继续向后传递

层层递进、逐层抽象

神经网络学习：模型结构

神经网络构成：输入层、输出层以及隐藏层，每个隐藏层中包含了若干神经元。每个神经元完成物理反应、化学反应和信息流通任务

从输入层开始，第一个隐藏层包含了3个神经元（用圆圈表示）、第二个隐藏层包含了4个神经元。给定某个隐藏层，该隐藏层中的神经元之间没有任何连接。每个神经元与其前后相邻层中的每个神经元均相连。

第二个隐藏层中的每个神经元接受与之相连第一个隐藏层中所有神经元通过不同权重传递过来的“信息总和”，然后对汇总所得的信息进行处理，再将处理后信息向后传递。

称为全连接神经网络，即每个神经元均与其前后相邻层中的神经元相连。可见，一旦给出了图2.7.4神经网络，输入数据从输入层开始，经过两个隐藏层，逐层抽象，最后在输出层得到输出结果。

层层递进、逐层抽象

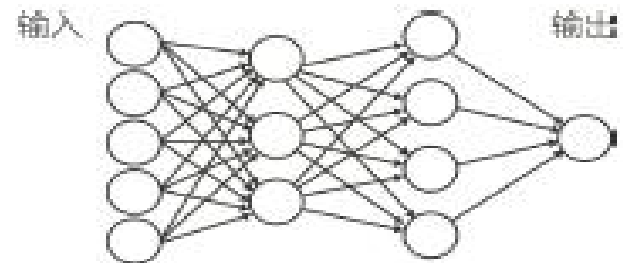


图2.7.4 神经网络架构示意图

有1个输入层，2个隐藏层（hidden layer）和1个输出层

神经网络学习：参数学习

神经网络的参数：联结权重

学习参数的算法超出了高中学生应掌握知识点！

太难了

梯度下降和误差后
向传播

层层递进、逐层抽象

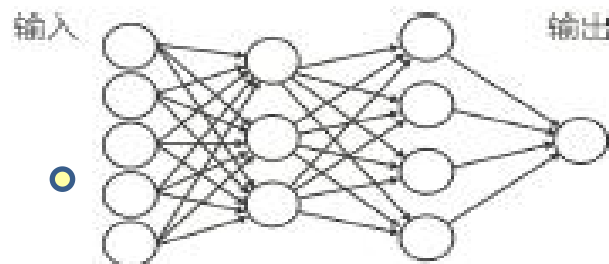


图2.7.4 神经网络架构示意图

有1 个输入层，2 个隐藏层 (hidden layer) 和1 个输出层

数据

+

模型

+

算力



燃料、引擎、加速器

神经网络学习：神经网络用于识别任务

神经网络的参数：联结权重

- ◆ 输入层与第一个隐藏层所有神经元之间的连接权重，第一个隐藏层中神经元和第二个隐藏层中神经元之间的连接权重，第二个隐藏层中神经元与输出层之间的连接权重。
- ◆ 针对某个特定任务（如人脸识别），需要收集与该任务相关的数据来训练神经网络，得到这个神经网络的最佳参数组合，使该神经网络能完成该特定任务。因此，我们可将神经元之间的连接视为“记忆”，即神经网络“记存”了该特定任务的模式。

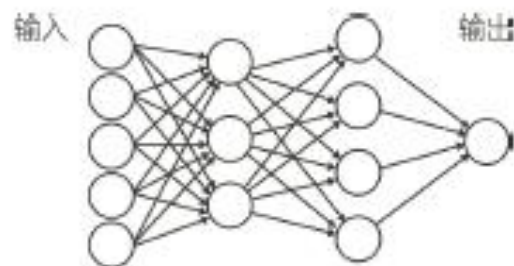


图2.7.4 神经网络架构示意图

拓展链接

赫布（Hebb）可塑性理论（记忆的可能由来）

在巴甫洛夫条件反射中，每次给狗喂食都摇铃，经过多次刺激以后，狗听到铃声就流口水。这一实验说明条件反射具有一定的神经机制。赫布理论所描述的突触可塑性基本原理，在一定程度上揭示了记忆的可能由来。在赫布理论中，突触前神经元向突触后神经元进行持续重复刺激，可导致突触传递效能的增加，即持续重复刺激使得神经元之间“记住”了这种刺激。这是记忆由来的一个猜想。

神经网络学习：神经网络

每个隐藏层中有512 个神经元， 化学反应函数为Relu

```
model = Sequential()  
model.add(Dense(512,input_shape=(784,), kernel_initializer='he_normal'))  
model.add(Activation('relu'))  
  
model.add(Dense(512, kernel_initializer='he_normal'))  
model.add(Activation('relu'))  
model.add(Dense(nb_classes))  
model.add(Activation('softmax'))
```

输入数据大小为784

通过softmax实现概率输出

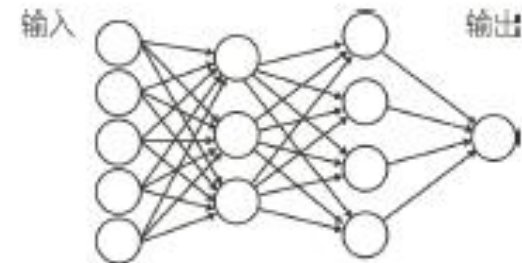
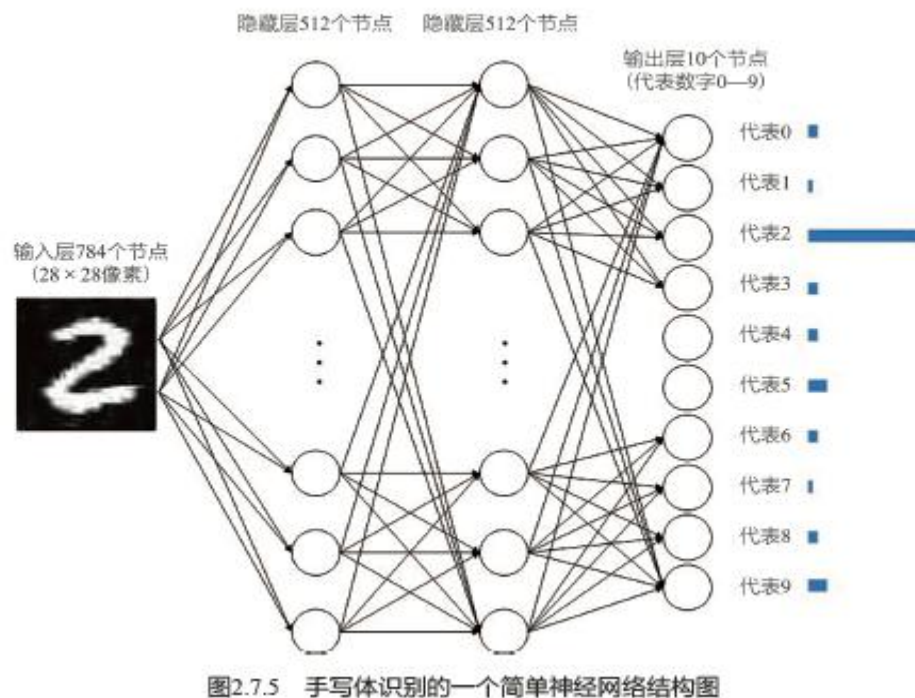


图2.7.4 神经网络架构示意图

一旦神经网络搭建完成后，就可用它来完成手写体识别等任务

神经网络学习：神经网络用于手写体识别

- 输入层接受一幅图像的所有像素点信息，输出层是一个10维的向量，每个维度对应0—9这十个数字。
- 对于给定一幅图像，经过图2.7.5神经网络处理后，输出10维向量，取10维向量中值最高的那个维度作为识别结果。
- 比如，图2.7.5中输入的图像是“2”，那么所输出10维向量中第三维（注意：第一维对应的是0）概率值应该最大，从而将输入图像识别为2。
- 其他识别任务？



图像 -----> 模式 -----> 类别

一种特殊且常用的深度学习模型：卷积神经网络

卷积滤波

卷积神经网络中最为重要的概念是“卷积”（convolution）。下面通过图3.2.4来解释卷积的基本概念。在图3.2.4中，滤波器（filter）用来对图像进行处理。这个 3×3 滤波器从上到下、从左到右的9个值为（1, 2, 1, 2, 4, 2, 1, 2, 1），这9个值称为滤波器的权重参数。

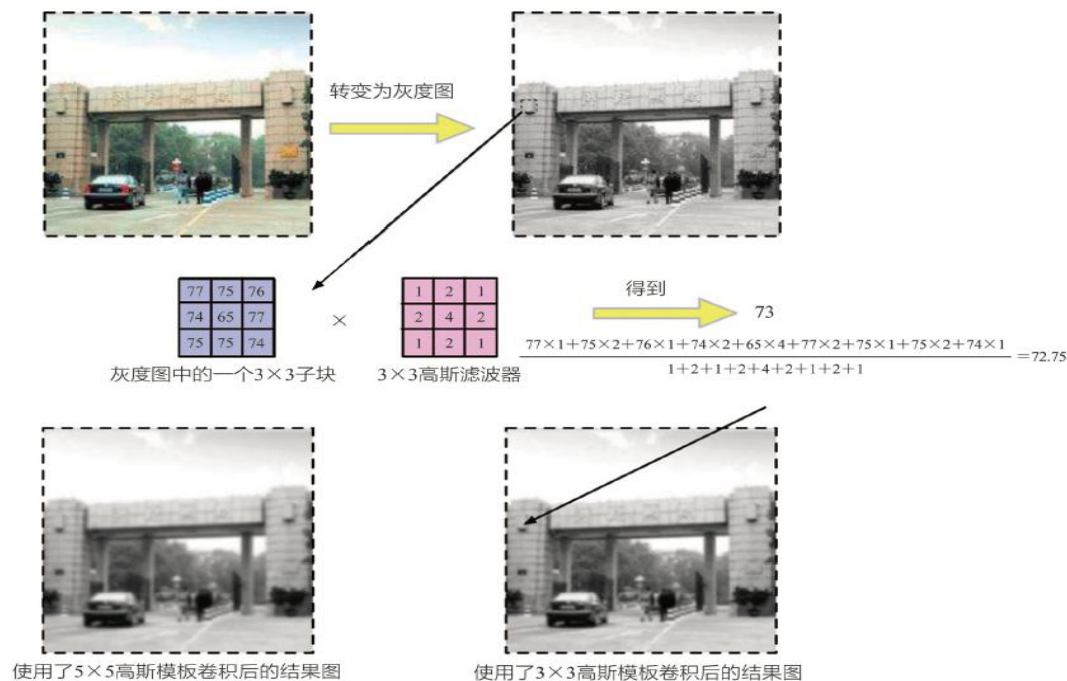


图3.2.4 图像卷积操作示意图

一种特殊且常用的深度学习模型：卷积神经网络

卷积滤波

在图像滤波中，将被滤波图像中的某个像素点作为中心像素点，形成一个与滤波器大小一样的 3×3 分辨率图像子块，然后将 3×3 滤波器中每个方格位置的值与 3×3 图像子块相应位置的值相乘，并将乘积累加后除以 3×3 滤波器所有权重之和，得到的结果作为该中心像素点的值。例如，图3.2.4中“65”这个中心像素点被 3×3 滤波器滤波后，其值变为73。图像中所有像素点都这样处理后，得到的结果称为图像滤波后结果，即卷积后结果。当然，以图像中的边界像素点为中心无法形成 3×3 的子块，因此，边界像素点不参与卷积计算。

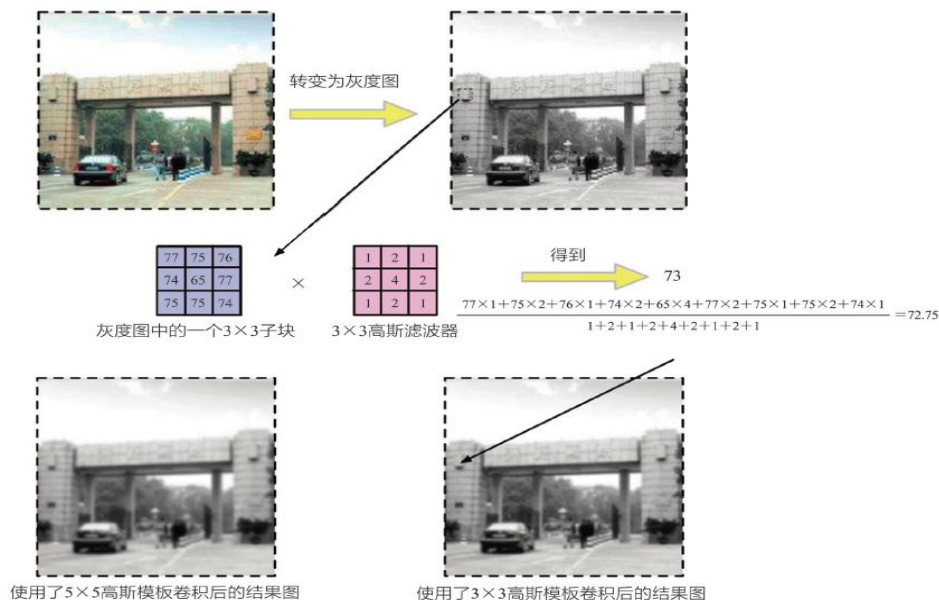


图3.2.4 图像卷积操作示意图

一种特殊且常用的深度学习模型：卷积神经网络

激活函数

- 视觉理解的目的是将图像像素点这一信息映射为语义内容，实现从像素点空间到语义空间的一次复杂变换。如图3.2.5所示，在视觉理解中，“非线性映射”将左图图像（像素点集合）映射到高层语义（如冬雪、行人与树木等）。
- 深度学习中也应该考虑非线性映射机制，用以增强机器学习的性能。在深度学习中，非线性映射这一机制由激活函数来完成。从神经网络角度来说，激活函数考虑的是什么信息可以被激活而传递给后续神经元、什么信息因其未能激活而无法传递给后续神经元。



图3.2.5 视觉理解中的非线性映射：从像素点到语义

一种特殊且常用的深度学习模型：卷积神经网络

池化操作

- 人脑对外界信息会不断地抽象，去繁就简，如在理解一幅100万像素点构成的人脸图像时，人脑会很快对这100万像素点进行压缩，从中提取出有意义的信息。
- 深度学习通过池化操作（pooling）将原始信息“由厚变薄”，保留应该保留的信息。如图3.2.6中给出了最大池化（max-pooling）和均值池化（average-pooling）操作的结果。在将原始图像划分为 2×2 子块后，前者从每个子块中选择一个值最大的像素点，后者从每个子块中选择该子块中像素点均值，来完成原始图像的池化操作。

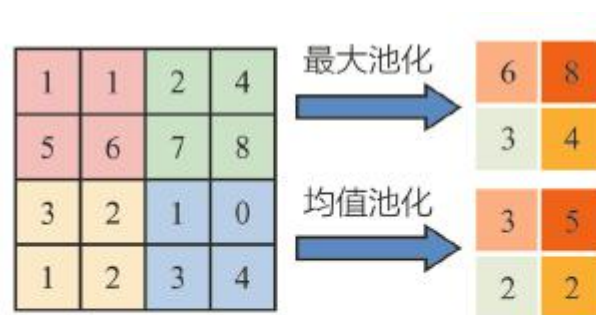


图3.2.6 最大池化和均值池化操作示意图

一种特殊且常用的深度学习模型：卷积神经网络

全连接层

全连接层就是前面讲过的感知机模型，相邻层之间神经元相互连接，将卷积和池化操作所得的结果转变成向量，以便更好地进行分类。全连接层中层与层之间的连接权重均是要训练学习得到。

一种特殊且常用的深度学习模型：卷积神经网络

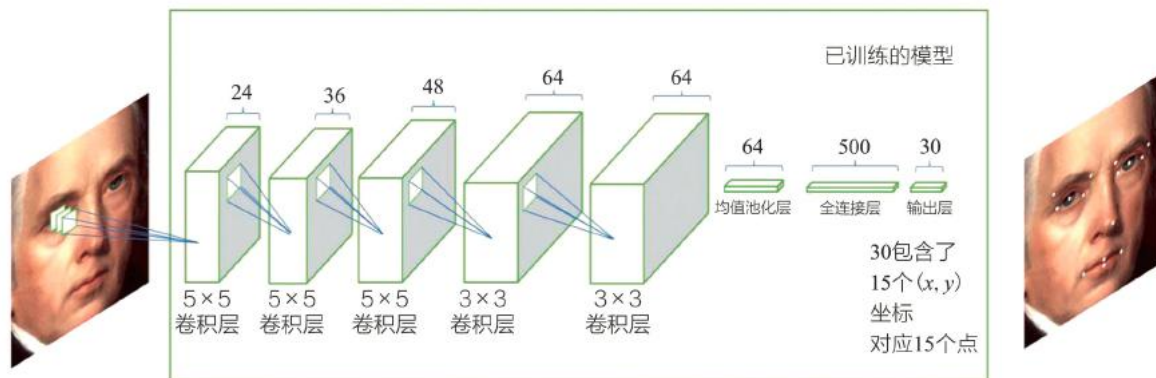


图3.2.7 通过卷积神经网络检测人脸关键点所输出的网络结构

图3.2.7给出了人脸关键点检测的流程图，其包含了五个卷积层。先是三个 5×5 的卷积层，卷积滤波器个数分别是24，36，48；接着是两个 3×3 的卷积层，卷积滤波器个数分别是64，64；然后是一个均值池化层；最后是一个全连接层和一个输出层。输出层输出15个关键点的横、纵坐标位置，因此输出层是30维向量。

表3.2.1 卷积神经网络学习中不同操作的解释

操作名称	作用	参数学习
卷积计算	挖掘中心像素点与周围像素点之间存在的模式。卷积（滤波器）个数越多，挖掘出的像素点空间模式越多	卷积的权重参数要通过深度学习训练得到
池化操作	保留图像中的重要信息，对图像进行约减抽象	没有需要训练的参数
激活函数	对神经网络进行非线性变换	没有需要训练的参数
全连接层	将特征映射为向量	连接权重主要通过深度学习训练得到

表3.2.1列出了深度学习模型中一种有代表性的模型——卷积神经网络不同操作的作用和参数学习等信息。在标注大数据驱动学习下，一旦训练得到一个良好的深度模型，则该模型可提取数据中最能表达语义内容的特征，然后将这个特征用于分类和识别。因此，深度学习也被称为特征学习（feature learning）。很显然，由于利用了类别标注信息，卷积神经网络这种学习模型是有监督学习。

算法与描述对应关系

