

## 一. 第一章简答题

1. 简述人工智能、机器学习、深度学习联系与区别。

① 人工智能是计算机科学的一个分支，是研究、开发用于模拟、延伸和拓展人所具有的智能的理论、方法、技术及应用系统的技术科学。

② 机器学习是人工智能的一个分支；机器学习又包括监督学习、无监督学习、强化学习。

③ 深度学习是机器学习的一个分支（学习范式），它可以用于监督学习、无监督学习、强化学习

无监督学习、强化学习

考到

2. “分类”与“聚类”区别？解释“回归”“泛化”术语

① 分类是有监督的算法，聚类是无监督的算法。（有监督↓）

② 分类：根据输入数据，把样本分配到事先定义好的离散类别中

③ 聚类：在没有标签的情况下，根据数据本身的相似性，把样本自动

分成若干组。（无监督）

④ 回归：根据已知的数据进行规律的总结，并基于该规律对未来发展趋势进行预测。

⑤ 泛化：是指机器学习算法对新样本的适应能力。

### 3. 什么是“监督学习”

利用一组已知类别的样本训练分类器参数，使其达到所要求性能的过程。

### 4. 机器学习中，“超参数”“参数”区别是什么？CNN/RNN参数/超参数举例

① 参数是模型可以根据数据训练得到的变量

② 超参数是学习过程之前设置好的参数，而不是训练而得。 **考到**

③

CNN

参数: 卷积核权重、偏置、全连接权重

超参数: 卷积核大小、卷积核个数、训练轮数、步长

④

RNN

参数: 输入到隐藏层的权重、隐藏层到隐藏层权重、隐藏层到输出权重

超参数: 隐藏层维数、RNN层数、训练轮数

### 5. 机器学习数据集中，“<sup>测试</sup>数据集”、“验证集”、“训练集”区别是什么？

① 测试集: 用于评估训练出的模型效果，但不会改变模型参数及效果，一般验证模型是否欠拟合，决定训练重训/早法验证。

② 验证集: 因为训练集和测试集均源自同一分布，随着时间流逝，近期样本的分布与训练样本分布会有出入，需要校验好的模型在近期样本（即验证集）上效果如何。（围绕模型稳定性、鲁棒性、泛化能力）

③ 训练集: 用于让模型学习数据规律并更新模型参数的数据集。

### 6. 机器学习的过程

获取数据 → 数据预处理（数据、异常值处理） → 特征工程（仅用知识技巧处理数据，使得特征在比之上发挥更好） → 模型训练 → 模型评估

## 7. "支持向量机"的分类器工作原理

根据支持向量数据点以及间隔最大化原则找到最佳分类线。

8. 先解释“欠拟合”“过拟合”概念，再画示意图。

① 欠拟合：由于模型太简单，特征太少，训练不充分

模型未能很好学习训练中的规律，→ 拟合能力不足

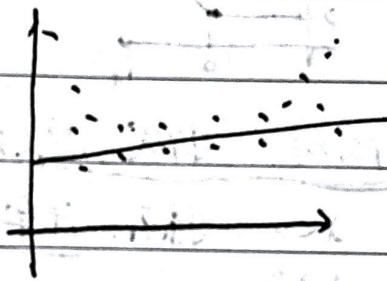
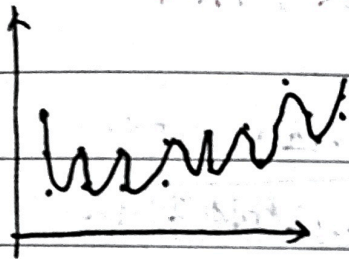
② 过拟合：由于模型过于复杂，数据过少，参数过多

泛化能力不足

模型在训练集上表现很好，但在测试集/新数据表现差。

③ 过拟合

欠拟合



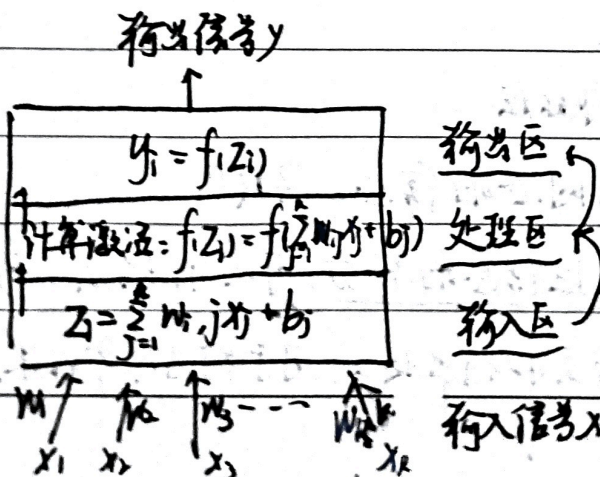
## 2. 第二章简答题

1. 数学模型表示人工神经元 (画图)。

人工神经元是模拟生物神经元的数学模型，同时也是多输入单输出的非线性元件。神经元有  $R$  个元素输入，每个输入权重  $w_j$ ，并偏置  $b$ ，经过函数  $f$  激活后，

得到结果  $f$ 。

神经元  $i$



简述与应用

均考到

2. 神经网络设计的激活函数作用:

对于一个没有激活函数的神经元, 其输出/输入都是线性组合;

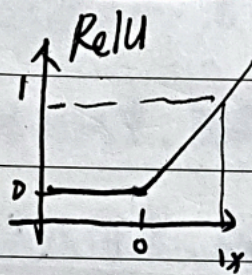
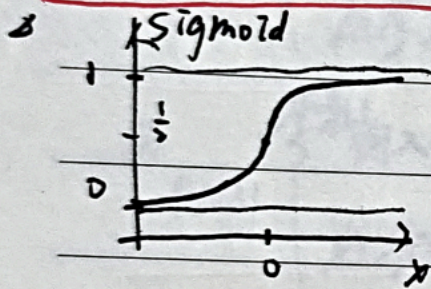
如果引入非线性激活, 就可以使得神经网络任意逼近任何非线性函数 (引入非线性从而应用到各类非线性任务).

3. 简述 Sigmoid、ReLU 函数图像

① Sigmoid:  $\sigma(x) = \frac{1}{1+e^{-x}}$  定义域  $(-\infty, \infty)$  值域  $(0, 1)$

函数连续可导, 易出现梯度消失

$x \rightarrow -\infty, \sigma(x) \rightarrow 0; x \rightarrow +\infty, \sigma(x) \rightarrow 1$



考到, 考两个  
函数特点

②  $\text{ReLU}(x) = \max\{0, x\}$  折线, 定义域  $(-\infty, +\infty)$ , 值域  $(0, +\infty)$

$= \begin{cases} 0 & x < 0 \\ x & x \geq 0 \end{cases}$  特点: 计算简单收敛快, 使梯度消失, 可能出现“神经元死亡”

③ Sigmoid { 用途: 把输入映射到  $[0, 1]$  之间, 常用二分类的输层 (概率)  
特点: 值域  $(0, 1)$ , 可以表示概率.

ReLU { 用途: 引入非线性, 神经网络能学习复杂特征,  
常用于隐层.

4. 简述人工神经网络的工作原理

① 输入层: 读取输入到神经网络的信息。

② 隐层: 神经网络实际处理信息的部分, 可以有任意层, 形成深层神经网络。

③ 输出层: 对接最后的隐层和输出接口, 用于输出计算结果。

④ 神经元之间互相连接, 权重不同。通过学习, 神经网络可以自动为其中神经元找到对于当前数据最优的权重/偏置。

5. 解释下列神经网络相关术语。

① 前馈神经网络: 数据正向流动, 输出由输入/网络权重决定, 无反馈。

② 误差反向传播网络 (BP网络):

权重问题上学习了反向传播学习算法, 处理数据时, 数据从输入层经隐藏层逐层向后传播; 训练网络时, 从输出层开始逐层向前修正网络权重。

6. 训练网络过程中, “损失函数”之作用。

“损失函数”用于衡量网络的输出数据与正确数据之间总体误差, 并求得权重参数。

7. 如此理解下列神经网络的“学习”? 如何理解“学习率”?

① “学习”即网络训练, 可以看作一个调整神经元参数以使误差接近正确值的过程。

② “学习率”是控制学习过程的一个参数, 合适的学习率能使误差在一定时间内收敛到局部最小值。

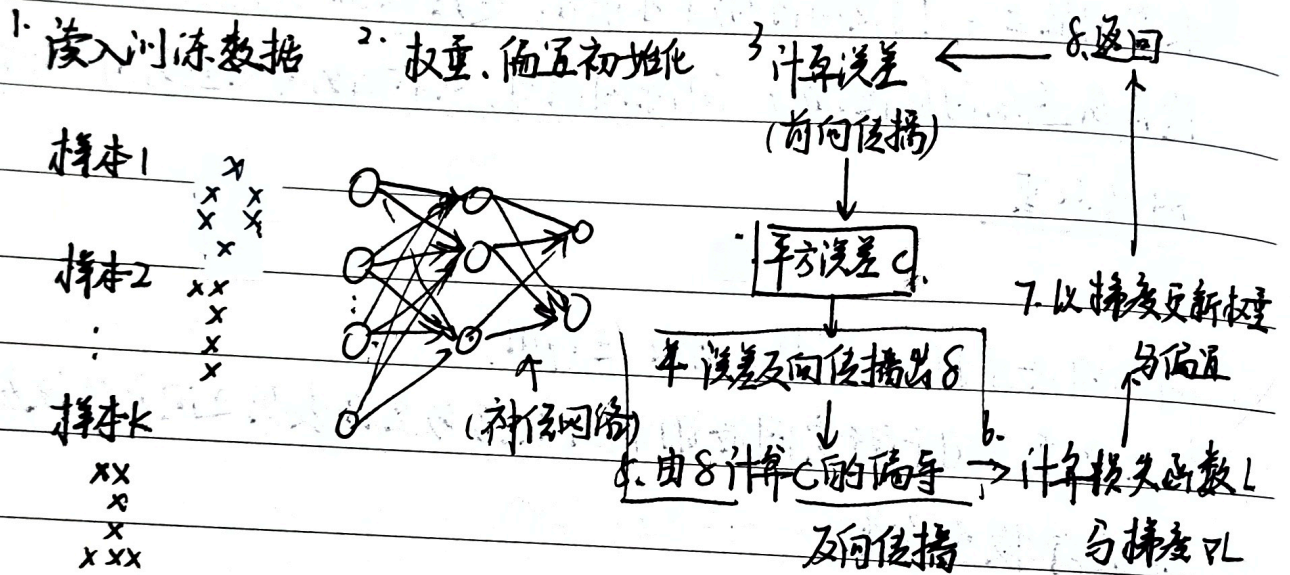
8. “梯度下降法”是常用的寻找损失函数最小值的方法, 简述原理。

即“学习”过程为在山脉中找到最低的那个低谷。应用梯度下降法, 每一步梯度下降方向找最低点指引, 以此接近最低点。

9. 解释介绍“误差反向传播算法”，并画图。

① 解释：误差反向传播算法是一种用于训练神经网络算法，通过前向传播计算网络输出，利用损失函数计算误差，再通过链式法则将误差从输出层反向传播到各隐藏层，计算各参数的梯度，并采用梯度下降方法更新参数。

② 示意图：



10. 如何采用神经网络进行多分类问题

首先进行数据预处理，然后应用神经网络提取特征，最后应用 softmax 网络进行多分类。Softmax 输出等于多分类中类别的个数。

## 三. 第三章简答题

1. CNN/RNN由哪些部分组成?

①. CNN = <sup>输入层</sup>卷积层, 池化层, 全连接层, 输出层

②. RNN: 输入层, 循环隐藏层, 输出层

卷积

2. CNN中, 卷积层的作用: 通过卷积核对输入信息进行卷积从而提取特征

3. CNN中, 池化层的作用: ① 减少特征图尺寸 ② 引入不变性

③ 方法: 最大池化为平均池化.

4. CNN的“参数共享”是什么? 意义是什么?

① 参数共享作用: 卷积核参数共享 ② 减少参数数量, 使运算简洁高效, 能够适应大规模样本运算.

5. CNN中, 哪些参数需要训练确定.

卷积核权重  $W$  与偏置  $b$  (及全连接权重).

6. CNN的训练目的: 优化更新参数, 降低网络输出值与理想目标值误差.

7. 在Lenet的第一个卷积层  $C1$  中, 输入图像  $32 \times 32$ , 输入图片通道 1 个, 卷积核  $5 \times 5$ , 卷积步长 1. 请问本层输出的特征图个数为多少? 本层有多少个参数需训练?

① Lenet的  $C1$  层 6 个卷积核  $\rightarrow$  6 个特征图.

① 一个卷积核的权重参数  $5 \times 5 = 25$ , 外加偏置 1 参数

每个卷积核 26 个参数  $\rightarrow$  总参数  $26 \times 6 = 156$  个

### 深度学习

#### 卷积输出尺寸公式

其中  $N$  输入(图像尺寸),  $F$  卷积核尺寸,

$$\Delta \text{ 输出尺寸} = \frac{N - F + 2P}{S} + 1$$

$P$  (padding, 四周补图宽度),  $S$  步长

此题: 输出尺寸 =  $\frac{32 - 5 + 0}{1} + 1 = 28$

9. LeNet 的 S4 池化层输出特征图 16 个  $\Rightarrow$  C5 卷积层即有 16 个子卷积核.

C5 卷积核 120 个  $\Rightarrow$  输出 120 张特征图

10. Alexnet 相比于 LeNet 的抑制过拟合方法.

#### Dropout 和数据扩充

### 四. 第四章习题(与扩充)

1. ① 在机器学习领域, 什么是“迁移学习”

迁移学习是把领域(即源领域)的知识, 迁移到另一个领域, 使得目标领域有更好的学习效果.

② 步骤:

程序也是此步骤

1. 加载图像数据, 划分为训练集/验证集
2. 加载预先训练好的网络
3. 对网络结构改造(迁移)
4. 训练数据集 (预处理 + 数据增强 + 尺寸匹配)
5. 训练网络
6. 进行验证并显示效果 (若验证不过关, 返回 5)

③ 迁移学习应用场景: 训练数据规模小, 计算资源/硬件资源有限(嵌入式)  
模型需快速收敛, 目标/源任务相似