以下是对以上作业问题的解答：

**作业一：神经网络模型中为什么要引入激活函数？常见的激活函数有哪些，各自有什么优缺点？**

**为什么引入激活函数？**

激活函数的作用是引入非线性变换，使神经网络能够拟合复杂的非线性问题。若没有激活函数，神经网络只会是一个线性变换，无法处理复杂的决策边界。

**常见激活函数及优缺点：**

1. **Sigmoid函数**

• 优点：输出范围在 (0, 1)，适合处理概率问题。

• 缺点：容易引发梯度消失问题，计算开销较大。

2. **Tanh函数**

• 优点：输出范围在 (-1, 1)，收敛更快，适合对称数据。

• 缺点：也会导致梯度消失问题。

3. **ReLU (Rectified Linear Unit)**

• 优点：计算简单，能够缓解梯度消失问题。

• 缺点：在负值区域会导致神经元“死亡”，即输出恒为零。

4. **Leaky ReLU**

（当 ）；（当 , 通常 ）

• 优点：解决了ReLU的“死亡神经元”问题。

• 缺点：需要手动调整超参数 。

5. **Softmax**

用于多分类问题，输出一个概率分布。

• 优点：适合分类任务，结果可解释性强。

• 缺点：只用于输出层。

**作业二：感知机（单层神经网络）是线性模型还是非线性模型？多层神经网络呢？谈谈你的理由。**

• **感知机（单层神经网络）：**

是线性模型。原因是感知机仅由输入层和输出层构成，输出是输入的线性加权求和，无法表示非线性决策边界。

• **多层神经网络：**

是非线性模型。原因在于加入了隐藏层和非线性激活函数，能够拟合复杂的非线性关系，形成多样化的决策边界。

**作业三：神经网络的训练与预测相比缺少（或增加）了哪个过程，并说出理由。**

• **缺少/增加的过程：**

神经网络的训练过程相比预测，多了“反向传播”步骤，即通过计算误差的梯度来调整权重。

• **原因：**

预测过程仅进行前向传播，使用已有的模型参数生成输出。而训练需要通过反向传播和梯度下降方法，不断优化模型参数，使其更好地拟合训练数据。

**作业四：如何理解神经网络的过拟合现象？为了解决过拟合现象有哪些可行的方法？**

**过拟合的理解：**

过拟合是指神经网络在训练数据上表现很好，但在测试数据上的表现较差。这表明模型学习了训练数据中的噪声和细节，导致泛化能力下降。

**解决过拟合的可行方法：**

1. **增加训练数据**：通过增加多样化的数据，帮助模型学到更广泛的特征。

2. **正则化**：使用 或 正则化限制权重大小，防止模型过于复杂。

3. **Dropout**：在训练过程中随机“丢弃”部分神经元，减少神经元间的依赖性。

4. **提前停止**：在验证集误差开始增大时终止训练。

5. **数据增强**：对训练数据进行变换（如旋转、裁剪），增加数据多样性。

6. **降低模型复杂度**：减少网络层数或神经元数量，防止模型过度拟合。

若有其他问题需要详细说明，请随时提问！