1、深度学习介绍

1、深度学习相比传统机器学习的主要不同在是什么？

传统机器学习需要人工设计特征，而深度学习不需要人工设计特征，可以自动从原始数据中学习到特征。

2、深度学习相比传统机器学习的优越性体现在什么地方？

传统机器学习人工设计特征是非常困难，需要专家经验，费时费力，效果还不一定好，但是深度学习呢可以免去人工设计特征的麻烦，自动学习特征，效果还比人工设计的特征好。

3、深度学习的三大要素是什么

模型、数据、运算资源

4、神经网络在几十年前就出现了，为什么深度学习最近几年才火起来？

因为，最近几年互联网发展了，所以在网络上出现了大量的数据可以用于神经网络的训练，同时也有了GPU这种运算资源，共同造就了现在深度学习的繁荣，如果没有大数据和计算资源，只有神经网络，也是无法像现在这样繁荣的。

5、对输入数据进行归一化/标准化的目的和好处是什么？

在进行归一化或者标准化后：
可以确保输入数据的数值稳定性，不会有过大过小的数据出现，防止出现训练不收敛
归一化或标准化后，可以提升神经网络训练的收敛速度，对提高最终的训练精度也有帮助，
如果是有多种输入数据，归一化或标准化后，可以去掉不同数据的量纲，即让不同数据的重要性相同，防止出现一种数据把别的数据抹掉的问题。

6、为什么使用 one-hot 作为标签？优势是什么？

使用 one-hot 编码可以让不同类别之间的特征的距离计算比较合理，可以使 10 个类别任意两两的距离都是相同的（比如汉明距离都是 2），这样比较合理。但是如果直接用数字作为标签，那么 6 和 7 之间的距离就会比 1 和 6 之间小，这意味着数字 7 和 6 会更像，但是 1 和 6 更加不像，这是不合理的，因为图像识别的假设是不同的类别直接都是不像的，要尽量分开。

7、什么叫模型的泛化能力？

机器学习的主要挑战是希望算法能够在先前未观测到新数据上表现良好，而不只是在训练集上表现好。在先前未观测到的新数据上表现良好的能力被称为泛化能力。**这个意思是**，在训练集上训练好模型以后，比如训练了一个手写数字识别的模型，然后来了一张新的手写数字的图片，这个新图片在训练集里没有，所以是模型之前没有见过的，然后输入到模型里，模型也能输出正确的图片分类结果，那么这种能力就叫做泛化能力。

8、数据集中样本分布的基本假设是什么？

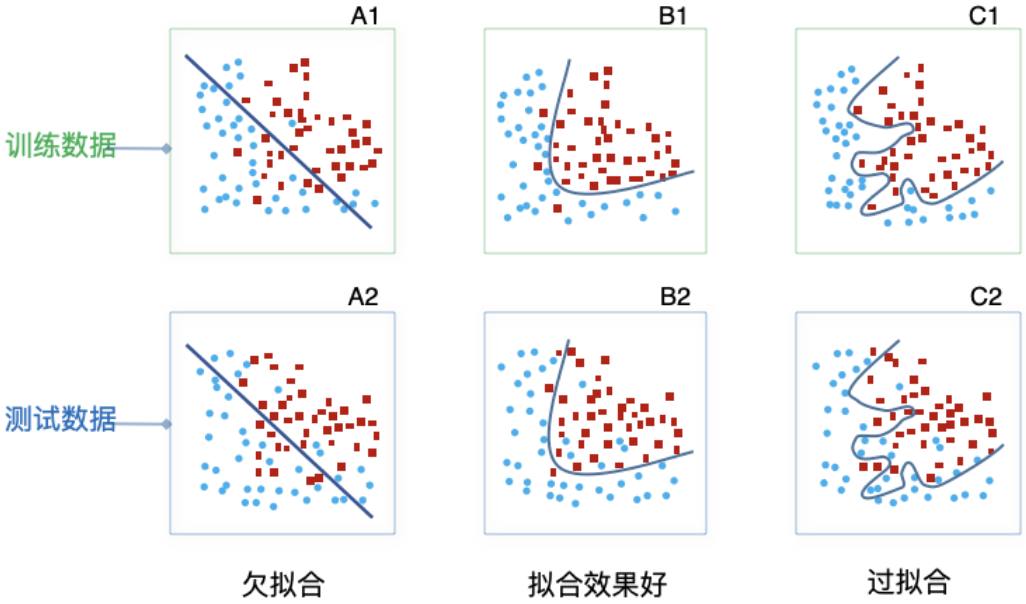
为什么当模型可以具有泛化能力，在没有见过的样本上也可以效果很好呢？这就涉及到一个数据集构建时候的基本假设，假设数据集中的样本都是独立同分布的，训练集和测试集都是独立同分布的。那么神经网络模型其实就是通过训练数据在学习对这个分布进行拟合。当模型训练完以后，那么模型对这个分布就拟合的比较好了，然后来一张没有见过的测试图片，虽然这个测试图片本身没有见过，但是它也是服从这个分布的，这个分布模型已经拟合的比较好了，所以对于没有见过的测试图片也可以获得很好的识别结果。

9、为什么数据集需要独立同分布？

因为独立同分布的数据集训练和测试出来的模型，泛化能力比较好。

10、什么是训练数据和测试数据独立同分布？什么是过拟合？什么是欠拟合？

如图所示：



训练数据和测试数据独立同分布是一个二分类问题，红色是一类，蓝色是另一类。可以看到虽然训练和测试数据的点的具体位置不一样，但整体来讲红色和蓝色点的位置分布是相似的，都是红色在右上角区域，各个点之间也是相互独立的，这就叫做独立同分布。

过拟合就是在训练集上效果好，但在测试集上效果差，泛化误差大，模型泛化能力差的。**欠拟合**，是在训练集和测试集上效果都比较差，说明模型的拟合能力不足。

11、在实验中怎么判断是否发生了过拟合？

就需要分别在训练集和测试集上进行测试，如果训练集上效果好，测试集上效果差，就说明发生了过拟合。

12、欠拟合和过拟合产生的原因分别是什么？

当模型复杂度不够，就会出现**欠拟合**。如果模型的复杂度过高，或者数据集太小，就会产生**过拟合**的问题。

13、如何解决欠拟合和过拟合的问题？

解决欠拟合的问题，需要增加模型复杂度；**解决过拟合**的问题，就需要降低模型复杂度，或者增大训练数据量。