1、Alexnet模型有几层隐藏层，具体怎么实现的，为什么要使用Alexnet模型

2、sigmod函数，如何求导的

函数：f(z) = 1 / (1 + exp( ??z))

导数：f(z)' =?f(z)(1 ??f(z))

3、欠拟合，过拟合的解决方法

欠拟合解决方法：

1）添加其他特征项，有时候我们模型出现欠拟合的时候是因为特征项不够导致的，可以添加其他特征项来很好地解决。例如，“组合”、“泛化”、“相关性”三类特征是特征添加的重要手段，无论在什么场景，都可以照葫芦画瓢，总会得到意想不到的效果。除上面的特征之外，“上下文特征”、“平台特征”等等，都可以作为特征添加的首选项。

2）添加多项式特征，这个在机器学习算法里面用的很普遍，例如将线性模型通过添加二次项或者三次项使模型泛化能力更强。例如上面的图片的例子。

3）减少正则化参数，正则化的目的是用来防止过拟合的，但是现在模型出现了欠拟合，则需要减少正则化参数

过拟合解决方法：

1）重新清洗数据，导致过拟合的一个原因也有可能是数据不纯导致的，如果出现了过拟合就需要我们重新清洗数据。

2）增大数据的训练量，还有一个原因就是我们用于训练的数据量太小导致的，训练数据占总数据的比例过小。

3）采用正则化方法。正则化方法包括L0正则、L1正则和L2正则，而正则一般是在目标函数之后加上对于的范数。但是在机器学习中一般使用L2正则，下面看具体的原因。

L0范数是指向量中非0的元素的个数。L1范数是指向量中各个元素绝对值之和，也叫“稀疏规则算子”（Lasso regularization）。两者都可以实现稀疏性，既然L0可以实现稀疏，为什么不用L0，而要用L1呢？个人理解一是因为L0范数很难优化求解（NP难问题），二是L1范数是L0范数的最优凸近似，而且它比L0范数要容易优化求解。所以大家才把目光和万千宠爱转于L1范数。

L2范数是指向量各元素的平方和然后求平方根。可以使得W的每个元素都很小，都接近于0，但与L1范数不同，它不会让它等于0，而是接近于0。L2正则项起到使得参数w变小加剧的效果，但是为什么可以防止过拟合呢？一个通俗的理解便是：更小的参数值w意味着模型的复杂度更低，对训练数据的拟合刚刚好（奥卡姆剃刀），不会过分拟合训练数据，从而使得不会过拟合，以提高模型的泛化能力。还有就是看到有人说L2范数有助于处理 condition number不好的情况下矩阵求逆很困难的问题（具体这儿我也不是太理解）。

4）采用dropout方法。这个方法在神经网络里面很常用。

4、你知道的统计软件有哪些，你擅长哪一个，为什么

SPSS SAS

5、人脸识别的根据什么特征进行分类

提取特征一般都是先求特征向量，通过训练样本得到，得到特征向量之后，

用特征向量提取训练样的特征（也就是把样本矩阵降维），

同样用特征向量提取测试样本的特征，比较测试样本与训练样本就可以进行识别，

分类器可以选择最近距离或者最近临域都可以