机器学习常见的面试题：

1. percision怎么计算，recall怎么计算的？F1计算公式？

P = TP / (TP + FP)

R = TP / (TP + FN)

1. LR原理说一下，公式推导

逻辑回归是一种分类模型

1. 线性回归原理讲一下，公式推导

Y = wx + b

利用回归方程对一个或多个特征值和目标值建模

1. 过拟合和欠拟合怎么解决

解决过拟合：

增加训练数量、降低模型复杂度、使用L1、L2正则化

解决欠拟合：

增加模型复杂度、特征工程、泛化特征、减少正则化参数

1. K近邻的原理讲一下

在已有数据中找到离它最近的K个数据，如果这K个数据大多数属于某个类别，则该样本属于此类别。

1. 朴素贝叶斯的公式推导
2. 朴素贝叶斯的朴素是什么意思？
3. 决策树的原理讲一下？

是一种树形结构，每个内部节点代表一个属性判断

1. 什么是生成式什么是判别式概念？
2. 你学过的算法哪些是生成式哪些是判别式？
3. 特征处理常见方法？

过采样、欠采样

1. 分类时候数据不均衡怎么处理的？
2. 聚类算法有哪些？kmeans了解吗，怎么确定K值？

KMeans、DBSCAN

KMeans算法对每个样本点计算到各个中心点的距离，并将该样本点分配给距离最近的中心点代表的类别，一次迭代完成后，根据聚类结果更新每个类别的中心点，然后重复之前操作再次迭代，直到前后两次分类结果没有差别。

深度学习算法：

1. 讲一下什么是CNN（因为你们目前只学了cnn其实需要学很多）

CNN是卷积神经网络

1. 常见的激活函数都有哪些

sigmiod、tanh、sotfmax、swish、ReLu、Leaky\_ReLu

1. 如何选择激活函数

隐藏层：

优先选择RELU激活函数，然后是Leaky ReLu。

使用Relu时避免出现大的梯度从而导致过多的神经元死亡。

输出层：

⼆分类问题选择sigmoid激活函数

多分类问题选择softmax激活函数

回归问题选择identity激活函数

1. 常见的损失函数

分类任务：交叉熵损失

回归任务：MAE 平均绝对误差

MSE 欧氏距离

smoothL1 损失函数

1. 梯度下降算法的优化方法

动量算法

AdaGrad

RMSprop

Adam

1. 什么样的数据集不适合深度学习
2. 什么造成梯度消失问题
3. 过拟合怎么解决的

增加训练数量、降低模型复杂度、使用L1、L2正则化

1. L1和L2区别

L1正则化的权重值减少到0

L2正则化的权重趋向于0

1. BN（批归一化）的作用

计算神经元的均值和方差

1. 什么是梯度消失和爆炸，怎么解决
2. 说一下你对神经网络的理解

神经网络是一种模拟人类神经系统工作机制的计算模型，由大量神经元连接组成。

1. 你学了之后认为，神经网络能解决什么问题

分类问题、回归问题、图像分割。