**1 简述KNN分类算法的基本原理并说明该算法的主要缺点。**

算法的原理是：对于一个未知样本，找到其最近的 K 个已知样本，然后通过这 K 个样

本的类别和属性进行决策来预测该未知样本的类别或属性。

主要缺点：当训练样本数目过多时速度较慢

假设训练数据如下表所示：

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 标签 | A | A | A | A | A | A | B | B | B | B |
| 数值 | 1.2 | 1.4 | 1.6 | 1.7 | 2.4 | 2.6 | 2.8 | 3.0 | 3.1 | 3.2 |

设置分别设置K=5 和K=9时 测试样本2.7的类别分别是？（距离定义 d=|x-y|）

2.7与各个点的距离

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| A | A | A | A | A | A | B | B | B | B |

1.5, 1.3, 1.1, 1.0 , 0.3, 0.1, 0.1, 0.3, 0.4, 0.5

K =5 时 最近的5个点为[A,A,B,B,B]分类为B

K =9时 最近的9个点为[A,A,A,A,A,B,B,B,B]分类为 A

2、**假设有2个事件A和B，**

**情况一 : P(A)=0.5、P(B)=0.5**

**情况二：P(A)=0.25、P(B)=0.75**

**计算两种情况下的信息熵, 哪种情况下系统更混乱，在分类模型的训练中希望系统的信息熵沿着那种趋势变化？**

**3 简述利用 ID3 法进行决策树构建的过程**

(1) 找到可以令平均熵最小的特征维度对数据集进行分割

(2) 对分割后的数据集,再找寻可以使平均熵最小的特征维度，再对数据集进行分割

(3) 重复上面步骤直到 用完所有特征、或者子集中目标标签全部相同

(4) 如果所有特征都用完，最终的子集中，目标标签仍不一致，则使用最多标签作为最终输出

4、写出贝叶斯公式，并简述普通贝叶斯分类器的工作原理。



（1）训练阶段

在分类器的训练阶段，我们需要先准备一些已经标注好的训练数据集。对于每个类别，我们需要统计其中每个特征（或属性）在该类别中出现的概率，以及该类别中所有样本的总数。

（2）预测阶段

在预测阶段，当我们需要对一个新的样本进行分类时，我们需要根据之前统计得到的各项概率，来计算该样本属于每个类别的概率。

（3）最后，我们可以选择概率最大的类别作为该样本的分类结果。

5、简述kmeans算法的基本原理

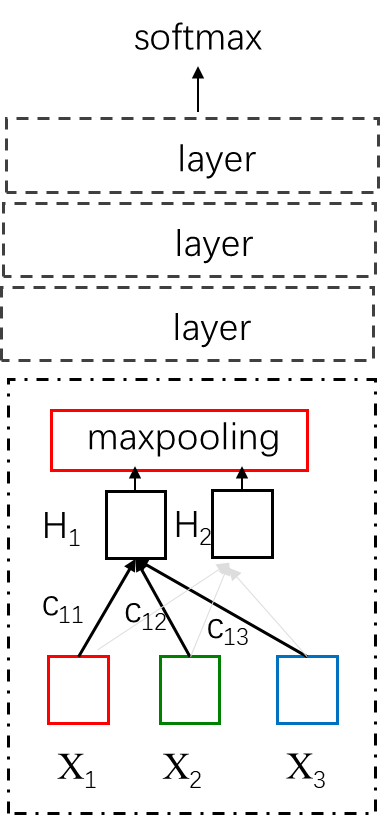
**（1）选择 k 个初始质心。在聚类开始之前，我们需要先决定要将数据点分成几个簇，然后随机选择 k 个数据点作为初始质心。**

**（2）按照每个数据点距离最近的质心，将其归入相应的簇中。对于每个数据点，我们计算它和每个质心之间的距离，将其归入距离最近的质心所代表的簇中。**

**（3）重新计算每个簇的质心。对于每个簇，我们将其中所有数据点的坐标取平均值，作为该簇的新质心。**

**（4）重复步骤 2 和 3，直到质心不再改变或者达到预定的迭代次数。**

**6**



一个由卷积神经网络（CNN）构成的图像分类器结构如图所示。第一层输入图像，大小为 256\*256，通道数为3。经过2维卷积后，通道数为2，如果卷积核的大小为5\*5，那么这一层神经网络用到了多少个参数（不考虑bias）？ 经过了maxpooling后输出特征图的大小为[2,128,128],那么maxpooling层的步进（stride size）是多少？经过训练后，底层的卷积层和顶层的卷积层分别捕获了图像的什么信息？

答：参数量 5\*5\*3\*2=150

步进2\*2

底层捕获细节信息

顶层捕获整体信息

7、简述AGENS聚类算法的原理与步骤？已知有两个簇分别是

A [ 2,5,7]

B [3,8]

那么采用single-linkage、complete-linkage以及average-linkage时两个簇之间的距离各自是多少？ （采用1范数距离 d=|x-y|）

8、 简述人脸识别系统构建的基本流程。欧氏距离打分和余弦距离打分的区别：

（1）模型训练：采集大量id已知的人脸数据训练一个人脸特征提取器。

（2）人脸注册：进行注册人脸收集，进行人脸检测，并截取人脸部分，进行人脸特征的提取与收集。

（3）人脸识别：利用人脸检测器，提取人脸区域后，提取人脸特征，将提取的特征和人脸注册阶段收集的人脸特征进行比较，找到相似度最大的一张人脸，并进行打分，如果分数满足一定的阈值则相应将人脸id输出，否则的话输出unknown。

欧式距离越小相似度越高，余弦距离越大相似度越高。

**9、一段语音采用16Khz采样率，16bit量化编码，时长为5秒，那么这个音频中含有多少个采样点？占据多少字节？对该信号进行短时傅里叶变换，设置FFT点数为512，帧长20ms 帧移动10ms，那么得到的特征的维度是多少？**

答： 样本点 5\*16000 = 80000个

字节 80000\*2 = 160000个

帧数 5\*1000/10 = 500

特征维度 512/2+1 = 257

最终特征 257\*500

**10、 梅尔滤波器组在频带上有什么特点？为什么会有这样的设置？**

答: 特点： 低频段分布比较稠密、高频段分布比较稀疏

目的： 人耳的听觉对低频段比较敏感因此需要较为密集的采样，对高频信息不敏感因此采样较为稀疏

**11、利用直接谱减法进行语音增强会出现那种缺陷？通常采用那种方式来对这种缺陷进行弥补？**

答： 缺陷是由于噪声估计误差导致出现的响铃噪声现象。

通常采用过减的方法加大减去噪声的比例来进行弥补.

**12 简述GMM-UBM说话人认证系统的构建方法及流程，并说明说话人认证过程中的打分公式。**

模型构建方法：

（1） 收集大量的语音训练通用背景模型UBM

（2）采集注册语音的特征利用MAP算法训练GMM作为表征说话人特征的模型

（3）测试过程中将测试语音特征分别送入GMM和UBM计算似然值的差值作为说话人认证的分数

打分公式：



（4）如果分数大于预设的阈值则让语音通过，反之则拒绝。