

一、

1. 奈奎斯特定理是指：要想不产生低频失真，则采样频率至少是录制的最高频率的 2 倍。其在数字媒体领域中主要应用于音频数字化、图像、视频压缩。例如：采样定理适用于相机系统，其中场景和镜头构成模拟空间信号源，图像传感器是空间采样设备。这些组件中的每一个都以调制传递函数（MTF）为特征，表示该组件中可用的精确分辨率。当镜头 MTF 和传感器 MTF 不匹配时，可能会出现混叠或模糊的影响。当传感器设备采样的光学图像包含比传感器更高的空间频率时，欠采样充当低通滤波器以减少或消除混叠。采样定理也适用于数字图像的后处理，例如上采样或下采样。混叠、模糊和锐化的效果可以通过在软件中实现的数字滤波进行调整。

2. JPEG 压缩的基本流程为：颜色模式转换及采样、DCT 变换、量化、编码。

3. 首先保证手机和显示器的两个屏幕平行。然后，照片的一条水平线和垂直线对齐手机屏幕，调整距离的远近，使得另一组再对齐。

4. 可以在数据中进行数据增强，先对数据进行一定的扰动，使其多出许多具有不用特征的同类数据，然后再进行分类器的训练，则其鲁棒性会更好。

5. 两个矩阵相乘的意义是将右边矩阵中的每一列向量 变换到左边矩阵中以每一行行向量为基所表示的空间中去。选择不同的基可以对同样一组数据给出不同的表示，如果基的数量少于向量本身的维数，则可以达到降维的效果。主成分分析通过正交变换将一组可能存在相关性的变量投影为一组线性不相关的变量，由于其只保留了信息最多的主成分，且投影到更低的维度，因此其一些无用信息被抛弃，即降噪了。

6. 假设在某一次反向传播中，有一个十分巨大的梯度通过反向传播传递到了 R。因为 R 是激活的，所以他会借助这个巨大的梯度通过反向传播算法到 R 得输出。这个将会导致 R 的前一层的给 R 提供输入的函数发生巨大的变化。这就意味着 R 的输入的分布会发生变化—我们又假设 R 的输入现在服从一个地方插均值为 -0.1 的高斯分布。之后大多数 R 的输入是负值且会导致 ReLU 关闭，而且会导致梯度无法通过 R 反向往回传播。因此 R 的输出的分布发生一个相对较小的改变（-0.2 的一个均值改变）导致了 R 的行为生的巨大改变。我们跨过了零点，现在 R 已经几乎总是关闭的了。现在的问题是这个关闭的 ReLU 无法再更新他的输入参数 w，所以死亡（总是关闭的）的 ReLU 会保持死亡。

解决方法有：采用更小的学习率；用 Leaky ReLU 代替；在训练时动态调整学习率。

7. 由于视频是有序数据，其帧间相似，因此可以使用帧间预测进行大幅度的有损压缩，其基本流程为：划分块、帧分组、帧间预测、残差 DCT 变换、CABAC 编码。

8.

p=1 时，为哈密顿距离， $d_{12}=4$ ， $d_{13}=6$ ， x_2 为最近点；

p=2 时，为欧式距离， $d_{12}=4$ ， $d_{13}=\sqrt{(4-1)^2+(4-1)^2}=3\sqrt{2}$ ， x_2 为最近点。

p=k 时， $d_{12}=4$ ， $d_{13}=\sqrt[k]{2*3^p}=3*\sqrt[k]{2}$ ，则当 $k>$

$\log_{4/3} 2$ 时, x_3 为最近点, 当 $k < \log_{4/3} 2$ 时, x_2 为最近点。

9.

(a)

$$H(\text{可食用} | \text{气味}=1) = -\log 1 = 0$$

$$H(\text{可食用} | \text{气味}=3) = -\log 1 = 0$$

$$H(\text{可食用} | \text{气味}=1 \text{ 或 } \text{气味}=3)$$

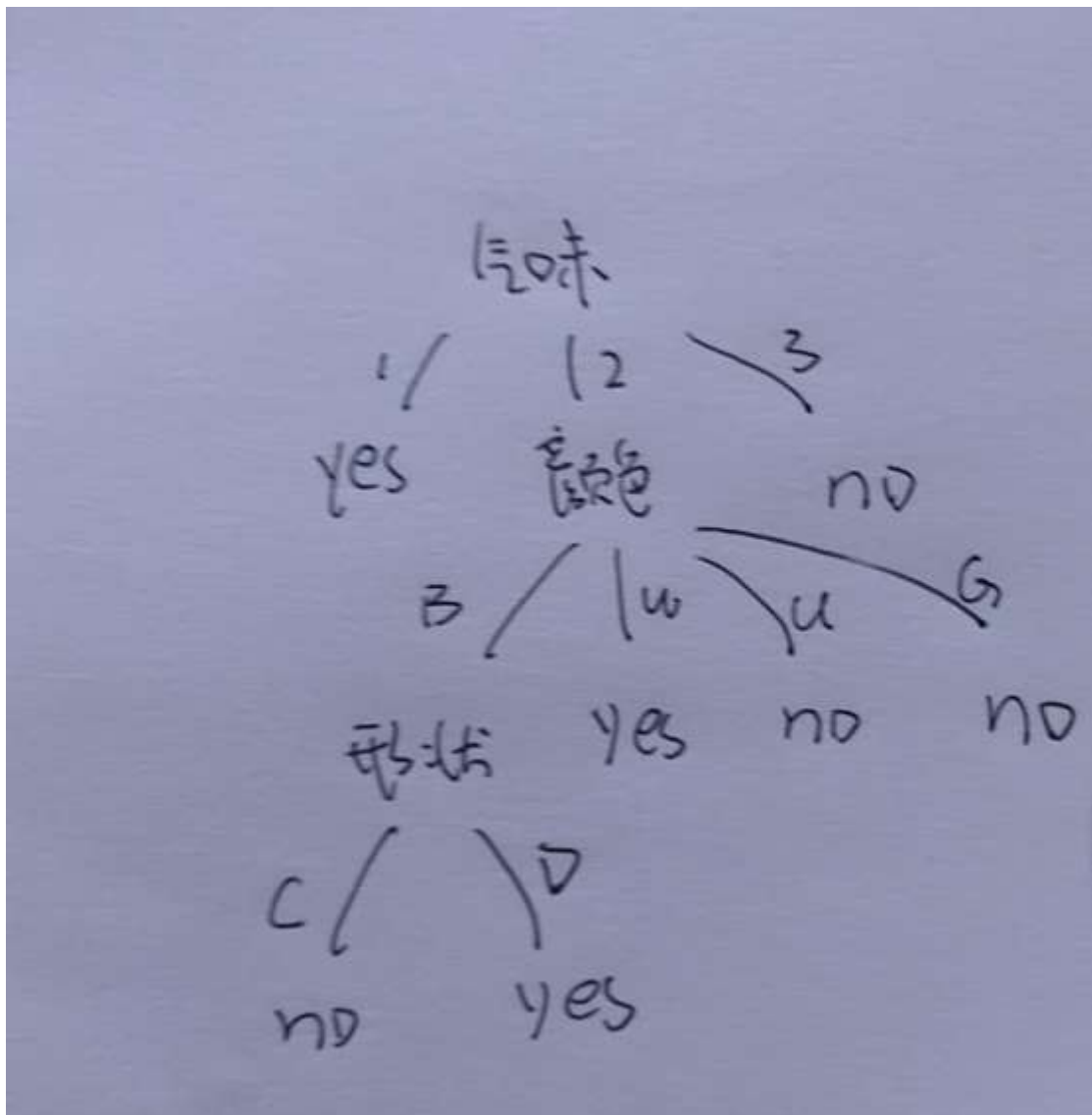
$$= p(\text{气味}=1) * H(\text{可食用} | \text{气味}=1) + p(\text{气味}=3) * H(\text{可食用} | \text{气味}=3)$$

$$= 0$$

(b)

形状、颜色、气味的条件熵中, 气味的条件熵最小, 因此根节点采用气味属性。。

(c)



二、

LZ77 是利用数据的重复结构信息来进行数据压缩。

LZ78 是在压缩过程创建字典，用字典里词语的序号（index）来表示重复出现的词语（新出现的重复字符串），从而减少需要压缩的文件长度。

LZW 是先建立动态字典，然后再像 LZ78 一样读入字符，没有的再进行编入字典。

三、

机器学习在多媒体信息处理中可能存在的挑战有：

1. 数据量庞大，
2. 信息的多元化导致信息维度过高

其解决方法有：

1. 对数据进行采样，减少数据规模
2. 采用分布式的计算平台，如 spark
3. 对数据进行降维，如 SVD、PCA 等等

四、

(a) 自拍照相亭的需求主要有两个：1. 拍证件照 2. 打印照片，其子需求有：拍照、更换底色、上传下载照片、编辑照片。

其系统模块分别为：通讯模块、拍照模块、图片编辑模块。

(b) 质量分析：核心问题是是否存在 bug，拍摄质量是否有问题，可通过对代码和硬件进行单元测试解决。

存储：核心问题是存储为何种格式，采用何种数据库。可将信息存储到数据库，然后照片加密后存储为文件。

传输：核心问题是如何让用户上传或下载照片。可使用微信小程序、公众号等等。

语言提示：核心问题是能否正确提示用户操作流程。可在测试中迭代优化。

(c) 拍照的主要应用场景为证件照，由于省去了一般照相馆的人工和场地成本，可将价格调整为照相馆的平均价格以下。

五、

(a) 监控系统的需求主要有：人员申报/登记、人群/车辆监控。

其系统模块可分为：管理模块、用户模块、人脸识别模块、车辆识别模块

(b) 隐私保护：核心问题是能否有效防止信息泄露。可通过网络隔离、权限分级等方式解决。

人脸检测识别：核心问题是能否识别出人员身份。可在所有出入口设置人脸识别打卡，在监控中加入人脸识别分析个人轨迹。

车辆检测：核心问题是能否识别出车辆号牌。可在车辆出入口设置车牌识别。

轨迹生成：核心问题是能否跟踪个人轨迹，分析异常人员。对能够不能正确识别的可疑人员需要记录并等待查验。

(c) 对于监控摄像头，人脸识别和车辆号牌识别都可以可使用 1080P 的红外摄像头，价格大概为 200-500 元/个。