

Assignment 5
CS 4783/5783
Due: 11/28/2022 11:59 pm

Consider the following data about the possible taste of a particular fruit based on some visual characteristics.

Taste	Farm climate	Visual defects	Size
Meh	Warm	Some	Small
Meh	Cold	None	Large
Meh	Cold	None	Large
Yummy	Cold	Many	Small
Yummy	Warm	Many	Small
Meh	Warm	Some	Large
Yummy	Warm	Many	Large
Yummy	Cold	None	Small
Yummy	Cold	None	Small
Meh	Warm	Some	Large

Answer the following questions based on your understanding of decision trees and Naïve Bayes.

[Question 1] **[3 points]**

Support you want to build a decision tree. What is the initial entropy of the target variable *taste*?

[Question 2] **[3 points]**

Consider that the variable *Visual defects* is chosen as the root of the decision tree. What is the information gain of the decision tree?

[Question 3] **[4 points]**

What is entropy $H(\text{Taste} | \text{Visual Defect} == \text{Some})$ and the entropy $H(\text{Taste} | \text{Visual Defect} == \text{None})$?

[Question 4] **[10 points]**

In this question, you will be using k-means to perform image compression. Implement a naïve version of the k-means algorithm based on your understanding. Your code must take the number of clusters k as input and perform k-means clustering on the given image. Once the algorithm finishes running, the cluster centroids represent the top-k common colors in the image. Iterate through each pixel in the image and assign the closest color to each pixel. Save and visualize the resulting image. For reading and writing images, you can use [OpenCV](#), which is an open-source computer vision toolkit. The following code will load the image into a NumPy array. You can use this as input to your K-Means algorithm.

```
import cv2
img = cv2.imread('input.png')
height, width, channels = np.shape(img)
for i in width:
    for j in height:
        pixel = img[j][i] # Read the pixel at location (i,j)
        img[j][i] = newValue # Assign a new value to the pixel
```

Note: You cannot use any library for k-means. You can only use OpenCV or other equivalent library to load images.