Shivansh Suhane

CS383 Homework 1 Resubmit

# Theory Questions:

1. Using the formula,  
   A picture containing object

   Description automatically generated

(1+1)/10 \* H(1/2,1/2) + 4(1+0/10) \* H (1/1,0/1)) + 4((0+1)/10) \* H( 0/1,1/1))  
  
The first term: (1+1)/10 H(1/2,1/2) = (2/10)(0.6 log2 0.5) + (-0.5 log2(1/2) = 0.2  
Similarly, the second and third terms of equation equals 0.  
  
E(A) = 0.2 This is the entropy using class 1  
Inverse: 1-0.2 = 0.8 = IG(A)

1. Second feature only has one duplicate element (‘1’) so using the average entropy formula:  
     
   (2+1)/10 \* H (2/3,1/3) + 3 ((1+0)/10 \* H (1/1,0/1)) + 4 ((0+1)/10 \* H (0/1,1/1))   
     
   The first term: ((2+1)/10 ) \* H H(2/3,1/3) = ((-0.66log2(0.666)) + (-1/3 log2(1/3)) = 0.275  
   The second and third terms of the equation equals 0.  
     
   E(B) = 0.275  
   IG(B) = 1-0.275 = 0.725
2. Since IG(A) > IG(B), the most discriminating feature is feature 1.
3. X= [2 1

5 4

3 1

0 3

8 11

2 5

1 0

5 1

1 3

6 1];

Principal components are [0.70 0,70

0.70 -0.70]

Question 2:   
I Tried this question in both python and Matlab.   
  
Python Scatter Plot:

A close up of a white background

Description automatically generated

Matlab scatter plot

A screenshot of a cell phone

Description automatically generated