Grant Davis
CS 458
HW3
10/27/2022

1.a. The Rule accuracy
    R1: 9/9+0 = 1
    R2: 65/(65+40) = .619
    R1 is best rule
  b. The Laplace measure
    R1: 19+1/19+2 = 20/21 = 0.9523
    R2: 65+1/105+2 = 66/107 = 0.6168
    R1 is best rule
  c. The m-estimate measure
    70/200 = 7/20 = 0.35
    R1=19+0.7/19+2 = 19.7/21 = 0.9381
    R2=65+0.7/19.42=65.7/107=0.6140
    R2 is best rule

2.a. P(A=1|+)
    =P(+|A=1).P(A=1) / (P(+|A=1).P(A=1) + P(+|A=0).P(A=0))
    =(5/6 * 6/10) / ((5/6 * 6/10) + (2/4 * 4/10))
    = 5/7
  P(B=1|+)
    =P(+|B=1).P(B=1) / (P(+|B=1).P(B=1) + P(+|B=0).P(B=0))
    =(4/5* 5/10) / ((4/5 * 5/10) + (3/5 * 5/10))
    = 4/7
  P(C=1|+)
    =P(+|C=1).P(C=1) / (P(+|C=1).P(C=1) + P(+|C=0).P(C=0))
    =(1/4* 4/10) / ((1/4 * 4/10) + (6/6 * 6/10))
    = 1/7
  P(A=1|-)
    =P(-|A=1).P(A=1) / (P(-|A=1).P(A=1) + P(-|A=0).P(A=0))
    =(1/6 * 6/10) / ((1/6 * 6/10) + (2/4 * 4/10))
    = 1/3
  P(B=1|-)
    =P(-|B=1).P(B=1) / (P(-|B=1).P(B=1) + P(-|B=0).P(B=0))
    =(1/5* 5/10) / ((1/5 * 5/10) + (2/5 * 5/10))
    = 1/3
  P(C=1|-)
    =P(-|C=1).P(C=1) / (P(-|C=1).P(C=1) + P(-|C=0).P(C=0))
    =(3/4* 4/10) / ((3/4 * 4/10) + (0/6 * 6/10))
    = 1

b. P(+| A=0, B=1, C=1)

    = P(A=0|+).P(B=1|+).P(C=1|+).P(+) / (P(A=0,B=1,C=1)

    $\propto$ P(A=0|+).P(B=1|+).P(C=1|+).P(+)   [As denominator is same for both the classes]

    $\propto$ (2/7)*(4/7)*(1/7)*(7/10)

    $\propto$ 0.016

  P(-| A=0, B=1, C=1)

    = P(A=0|-).P(B=1|-).P(C=1|-).P(-) / (P(A=0,B=1,C=1)

    $\propto$ P(A=0|-).P(B=1|-).P(C=1|-).P(-)   [As denominator is same for both the classes]

    $\propto$ (2/3)*(1/3)*(1/1)*(3/10)

    $\propto$ 0.066

  0.066 is greater than 0.016, so, the class label is negative

c. Suppose the number of data points = n

  Number of features = d

  Number of classes = c

  Brute force approach time complexity at training phase is O(n*d*c)

  Optimization approach time complexity at training phase is O(n*d)

  If d is small, time complexity will be O(n)

d. Suppose the number of data points = n

  Number of features = d

  Number of classes = c

  Time complexity at test phase is O(d*c)

3.a. P(Value=High|Engine=Good, Air Cond=Working) = 0.750

  P(Value=High|Engine=Good, Air Cond=Broken) = 0.667

  P(Value=High|Engine=Bad, Air Cond=Working) = 0.222

  P(Value=High|Engine=Bad, Air Cond=Broken) = 0

b. P(Engine = Bad, Air Conditioner = Broken) = 0.1453