Bronco ID: 013001924
Last Name: Nalbandian
First Name: Andre

Assignment 2

Thursday, September 23, 2021 12:47 PM

1.

Age	Spectacle	Astigmatism	Tear	Lenses (ground truth)
Young	Hypermetrope	Yes	Normal	Yes
Young	Hypermetrope	No	Normal	Yes
Young	Myope	No	Reduced	No
Presbyopic	Hypermetrope	No	Reduced	No
Presbyopic	Myope	No	Normal	No
Presbyopic	Myope	Yes	Reduced	No
Prepresbyopic	Myope	Yes	Normal	Yes
Prepresbyopic	Myope	No	Reduced	No

A.

a. False Negative
b. True Positive
c. True Negative
d. True Negative
e. False Positive
f. True Negative
g. True Positive
h. True Negative

B. Precision, Recall and F-1 Measure

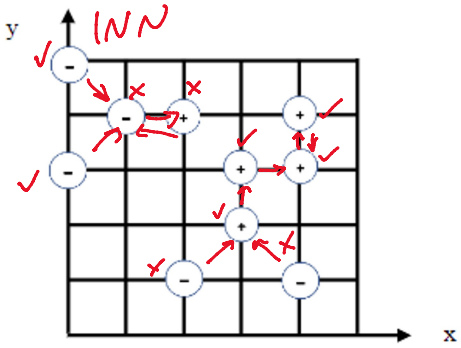
$P = \frac{TP}{TP+FP}$ $r = \frac{TP}{TP+FN}$ $F_1 = \frac{2rp}{r+p}$

$P = \frac{2}{2+1} = \boxed{\frac{2}{3}}$ $r = \frac{2}{2+1} = \boxed{\frac{2}{3}}$ $F_1 = \frac{2\left(\frac{2}{3}\right)\left(\frac{2}{3}\right)}{\frac{4}{3}} = \frac{\frac{4}{3}\left(\frac{2}{3}\right)}{\frac{4}{3}} = \boxed{\frac{2}{3}}$

2. https://github.com/NootCode/CS-4210-Asignment-2/blob/master/Question%202/decision_tree%20.py

3.

A.

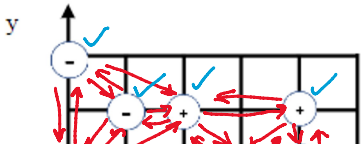


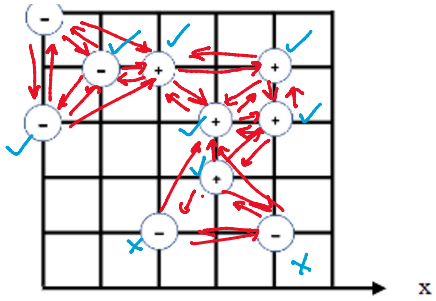
X = incorrect prediction
✓ = correct prediction

#ofX's = 4
#of✓'s = 6

Error rate = #ofX's / #oftotal = 4/10 = 0.4

B. 3NN

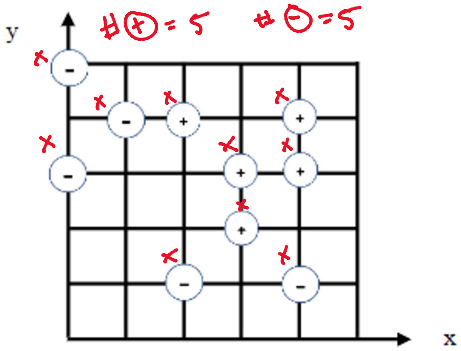




correct = 8
incorrect = 2

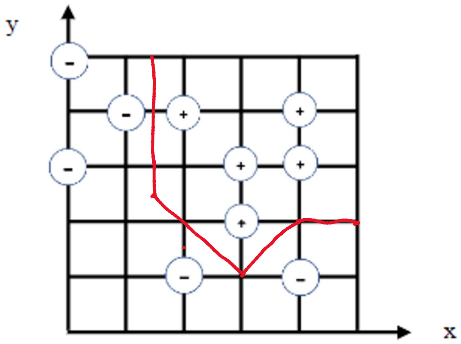
Error rate = 2/10 = 0.2

C. 9NN



When 9NN is called we take the majority value of the whole data set.
Since the data set is evenly distributed when a + value is left out the negative values become more dominant and vice versa.
incorrect = 10
total = 10
Error rate = 10/10
Therefore the error rate is 1.

D.



E. https://github.com/NootCode/CS-4210-Asignment-2/blob/master/Question%203/knn.py

4.

ID	Red	Green	Blue	Class
#1	220	20	60	1
#2	255	99	21	1
#3	250	128	14	1
#4	144	238	144	2
#5	107	142	35	2
#6	46	139	87	2
#7	64	224	208	3
#8	176	224	23	3
#9	100	149	237	3
#10	154	205	50	?

$d(1,10) = \sqrt{(154-220)^2 + (205-20)^2 + (50-60)^2}$

#9	100	149	237	3
#10	154	205	50	?

$d(1,10) = \sqrt{(154-220)^2 + (205-20)^2 + (50-60)^2}$

$= 196.67$

$d(2,10) = \sqrt{(154-255)^2 + (205-99)^2 + (50-21)^2}$

$= 149.25$

$d(3,10) = \sqrt{(154-240)^2 + (205-128)^2 + (50-14)^2}$

$= 128.22$

$d(4,10) = \sqrt{(154-144)^2 + (205-238)^2 + (50-144)^2}$

$= 100.12$

$d(5,10) = \sqrt{(154-107)^2 + (205-142)^2 + (50-35)^2}$

$= 80.01$

$d(6,10) = \sqrt{(154-46)^2 + (205-139)^2 + (50-87)^2}$

$= 131.87$

$d(7,10) = \sqrt{(154-64)^2 + (205-224)^2 + (50-208)^2}$

$= 182.82$

$d(8,10) = \sqrt{(154-176)^2 + (205-224)^2 + (50-23)^2}$

$= 39.67$

$d(9,10) = \sqrt{(154-100)^2 + (205-149)^2 + (50-232)^2}$

$= 202.53$

3NN = 8, 4, 5

Class(4) = 2

Class(5) = 2

Class(8) = 3

Class(10) = 2

5.

A.

Day	*Outlook*	*Temperature*	*Humidity*	*Wind*	*PlayTennis*
D1	Sunny	Hot	High	Weak	No
D2	Sunny	Hot	High	Strong	No
D3	Overcast	Hot	High	Weak	Yes
D4	Rain	Mild	High	Weak	Yes
D5	Rain	Cool	Normal	Weak	Yes
D6	Rain	Cool	Normal	Strong	No
D7	Overcast	Cool	Normal	Strong	Yes
D8	Sunny	Mild	High	Weak	No
D9	Sunny	Cool	Normal	Weak	Yes
D10	Rain	Mild	Normal	Weak	Yes
D11	Sunny	Mild	Normal	Strong	Yes
D12	Overcast	Mild	High	Strong	Yes
D13	Overcast	Hot	Normal	Weak	Yes
D14	Rain	Mild	High	Strong	No

<D15, Sunny, Mild, Normal, Weak>

$= \text{argmax}\ P(v_j) \prod_i P(a_i|v_j)$

$P(\text{class} = \text{yes} \mid A_1 = O\ \ A_2 = T\ \ A_3 = H\ \ A_4 = W)$

$P(\text{class} = \text{Yes} \mid \text{Sunny}, \text{Mild}, \text{Normal}, \text{Weak}) = \left(\prod_i P(A_i = x_i \mid (\text{class} = O)\right) * P(\text{class} = \text{Yes})$

$= \frac{2}{9} \cdot \frac{4}{9} \cdot \frac{6}{9} \cdot \frac{6}{9} \cdot \frac{9}{14} = .0282$

$P(\text{class} = \text{NO} \mid \text{Sunny}, \text{Mild}, \text{Normal}, \text{Weak})$

$= \frac{3}{5} * \frac{2}{5} * \frac{1}{5} * \frac{2}{5} * \frac{5}{14} = .0068$

Normalization

$P(\text{class} = \text{Yes} \mid \ldots) = .0282 / (.0282 + .0068) = .8057$

$P(\text{class} = \text{NO} \mid \ldots) = .0068 / (.0068 + .0282) = .194$

∴ Class of instance 15 is

Yes

B. https://github.com/NootCode/CS-4210-Asignment-2/blob/master/Question%205/naive_bayes.py