

Bronco ID: | 0 | 1 | 4 | 5 | 5 | 6 | 3 | 7 | 3 |

Last Name: Leos

First Name: Hugo

Assignment 2

1. S

a. $Accuracy = \frac{a+d}{a+b+c+d} = \frac{TP+TN}{TP+TN+FP+FN}$

Age	Spectacle	Astigmatism	Tear	Lenses (Ground Truth)	Prediction	Result
Young	Hypermetrope	Yes	Normal	Yes	No	FN
Young	Hypermetrope	No	Normal	Yes	Yes	TP
Young	Myope	No	Reduced	No	No	TN
Presbyopic	Hypermetrope	No	Reduced	No	No	TN
Presbyopic	Myope	No	Normal	No	Yes	FP
Presbyopic	Myope	Yes	Reduced	No	No	TN
Prepresbyopic	Myope	Yes	Normal	Yes	Yes	TP
Prepresbyopic	Myope	No	Reduced	No	No	TN

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} = \frac{2 + 4}{2 + 4 + 1 + 1} = \frac{6}{8} = \frac{3}{4} = 0.75\%$$

b. $precision\ p = \frac{TP}{TP+FP}$ $recall\ r = \frac{TP}{TP+FN}$ $F_1 = \frac{2rp}{r+p}$

$$p = \frac{2}{2+1} = \frac{2}{3} = 0.66$$
$$r = \frac{2}{2+2} = \frac{2}{4} = 0.5$$
$$F_1 = \frac{2(0.5)(0.66)}{0.5+0.66} = \frac{0.66}{1.16} = 0.569$$

Bronco ID: | 0 | 1 | 4 | 5 | 5 | 6 | 3 | 7 | 3 |

Last Name: Leos

First Name: Hugo

2. [link](#)

Bronco ID: | 0 | 1 | 4 | 5 | 5 | 6 | 3 | 7 | 3 |

Last Name: Leos

First Name: Hugo

3.

a. $error\ rate = \frac{number\ of\ wrong\ predictions}{total\ number\ of\ predictions}$

1NN			
Location	Truth	Prediction	Result
(0,5)	negative	negative	TN
(0,3)	negative	negative	TN
(1,4)	negative	positive	FP
(2,4)	positive	negative	FN
(2,1)	negative	positive	FP
(3,3)	positive	positive	TP
(3,2)	positive	positive	TP
(4,4)	positive	positive	TP
(4,3)	positive	positive	TP
(4,1)	negative	positive	FP

$$error\ rate = \frac{4}{10} = 40\%$$

b.

3NN			
Location	Truth	Prediction	Result
(0,5)	negative	negative	TN
(0,3)	negative	negative	TN
(1,4)	negative	negative	TN
(2,4)	positive	positive	TP
(2,1)	negative	positive	FP
(3,3)	positive	positive	TP
(3,2)	positive	positive	TP
(4,4)	positive	positive	TP
(4,3)	positive	positive	TP
(4,1)	negative	positive	FP

$$error\ rate = \frac{2}{10} = 20\%$$

c.

9NN			
Location	Truth	Prediction	Result
(0,5)	negative	positive	FP
(0,3)	negative	positive	FP
(1,4)	negative	positive	FP
(2,4)	positive	negative	FN
(2,1)	negative	positive	FP
(3,3)	positive	negative	FN

Bronco ID: | 0 | 1 | 4 | 5 | 5 | 6 | 3 | 7 | 3 |

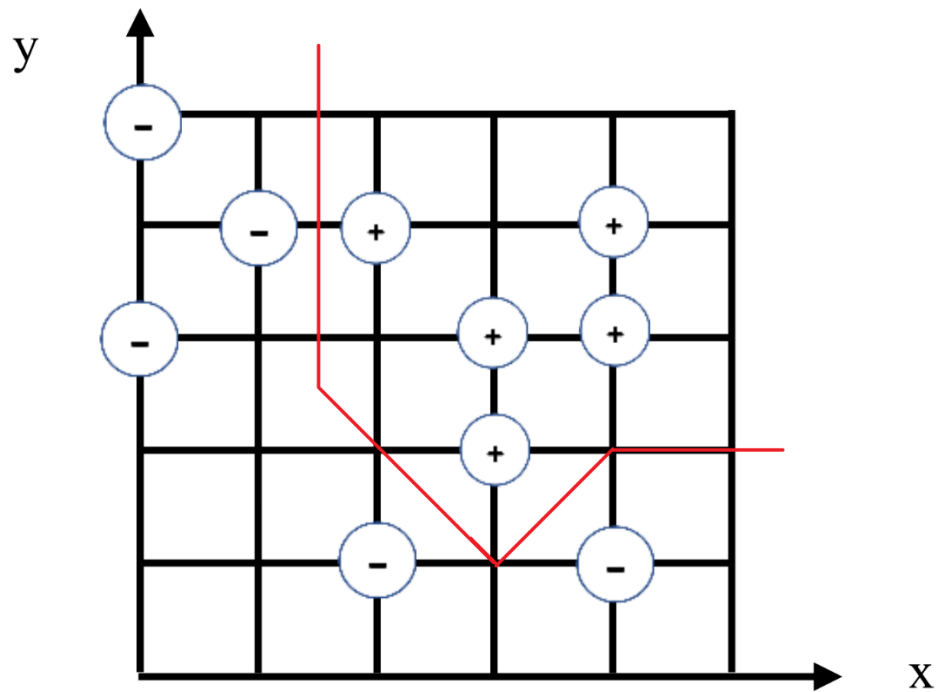
Last Name: Leos

First Name: Hugo

(3,2)	positive	negative	FN
(4,4)	positive	negative	FN
(4,3)	positive	negative	FN
(4,1)	negative	positive	FP

$error\ rate = \frac{10}{10} = 100\%$ since there is an even amount of positive and negative, when using a positive for test then the majority of training data is negative and vice-versa for the negative test.

d.



e. [Link](#)

Bronco ID: | 0 | 1 | 4 | 5 | 5 | 6 | 3 | 7 | 3 |

Last Name: Leos

First Name: Hugo

$$4. d(x, y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2}$$

$$\begin{aligned} d(1, 10) &= \sqrt{(154 - 220)^2 + (205 - 20)^2 + (50 - 60)^2} \\ &= \sqrt{(-66)^2 + (185)^2 + (-10)^2} = \sqrt{38681} = 196.675 \end{aligned}$$

$$\begin{aligned} d(2, 10) &= \sqrt{(154 - 255)^2 + (205 - 99)^2 + (50 - 21)^2} \\ &= \sqrt{(-101)^2 + (106)^2 + (29)^2} = \sqrt{22278} = 149.258 \end{aligned}$$

$$\begin{aligned} d(3, 10) &= \sqrt{(154 - 250)^2 + (205 - 128)^2 + (50 - 14)^2} \\ &= \sqrt{(-96)^2 + (77)^2 + (36)^2} = \sqrt{16441} = 128.222 \end{aligned}$$

$$\begin{aligned} d(4, 10) &= \sqrt{(154 - 144)^2 + (205 - 238)^2 + (50 - 144)^2} \\ &= \sqrt{(10)^2 + (-33)^2 + (-94)^2} = \sqrt{10025} = 100.125 \end{aligned}$$

$$\begin{aligned} d(5, 10) &= \sqrt{(154 - 107)^2 + (205 - 142)^2 + (50 - 35)^2} = \sqrt{(47)^2 + (63)^2 + (15)^2} \\ &= \sqrt{6403} = 80.019 \end{aligned}$$

$$\begin{aligned} d(6, 10) &= \sqrt{(154 - 46)^2 + (205 - 139)^2 + (50 - 87)^2} \\ &= \sqrt{(108)^2 + (66)^2 + (-37)^2} = \sqrt{17389} = 131.867 \end{aligned}$$

$$\begin{aligned} d(7, 10) &= \sqrt{(154 - 64)^2 + (205 - 224)^2 + (50 - 208)^2} \\ &= \sqrt{(90)^2 + (-19)^2 + (-158)^2} = \sqrt{33425} = 182.825 \end{aligned}$$

$$\begin{aligned} d(8, 10) &= \sqrt{(154 - 176)^2 + (205 - 224)^2 + (50 - 23)^2} \\ &= \sqrt{(-22)^2 + (-19)^2 + (27)^2} = \sqrt{1574} = 39.674 \end{aligned}$$

$$\begin{aligned} d(9, 10) &= \sqrt{(154 - 100)^2 + (205 - 149)^2 + (50 - 237)^2} \\ &= \sqrt{(54)^2 + (56)^2 + (-187)^2} = \sqrt{41021} = 202.536 \end{aligned}$$

ID	Red	Green	Blue	Class	Distance to 10
8	176	224	23	3	39.674
5	107	142	35	2	80.019
4	144	238	144	2	100.125
3	250	128	14	1	128.222
6	46	139	87	2	131.867
2	255	99	21	1	149.258
7	64	224	208	3	182.825
1	220	20	60	1	196.675
9	100	149	237	3	202.536
10	154	205	50	?????	

The predicted class will be 2 based on 3NN.

5.

- a. $P(\text{class} = \text{no} \mid \text{Outlook} = \text{Sunny}, \text{Temp} = \text{Mild}, \text{Humidity} = \text{Normal}, \text{Wind} = \text{Weak})$

$$\begin{aligned}
 &= P(\text{Outlook} = \text{Sunny} \mid \text{class} = \text{no}) * P(\text{Temp} = \text{Mild} \mid \text{class} = \text{no}) * \\
 &P(\text{Humidity} = \text{Normal} \mid \text{class} = \text{no}) * \\
 &P(\text{Wind} = \text{Weak} \mid \text{class} = \text{no}) * P(\text{class} = \text{no}) \\
 &= \frac{3}{5} * \frac{2}{5} * \frac{1}{5} * \frac{2}{5} * \frac{5}{14} = 0.00686
 \end{aligned}$$

$$P(\text{class} = \text{yes} \mid \text{Outlook} = \text{Sunny}, \text{Temp} = \text{Mild}, \text{Humidity} = \text{Normal}, \text{Wind} = \text{Weak})$$

$$\begin{aligned}
 &= P(\text{Outlook} = \text{Sunny} \mid \text{class} = \text{yes}) * \\
 &P(\text{Temp} = \text{Mild} \mid \text{class} = \text{yes}) * \\
 &P(\text{Humidity} = \text{Normal} \mid \text{class} = \text{yes}) * \\
 &P(\text{Wind} = \text{Weak} \mid \text{class} = \text{yes}) * P(\text{class} = \text{yes}) \\
 &= \frac{2}{9} * \frac{4}{9} * \frac{6}{9} * \frac{6}{9} * \frac{9}{14} = 0.0282
 \end{aligned}$$

Normalized:

$$\begin{aligned}
 &P(\text{class} = \text{no} \mid \text{Outlook} = \text{Sunny}, \text{Temp} = \text{Mild}, \text{Humidity} = \text{Normal}, \text{Wind} \\
 &= \text{Weak}) = \frac{0.00686}{0.00686 + 0.0282} = 0.196
 \end{aligned}$$

$$\begin{aligned}
 &P(\text{class} = \text{yes} \mid \text{Outlook} = \text{Sunny}, \text{Temp} = \text{Mild}, \text{Humidity} = \\
 &\text{Normal}, \text{Wind} = \text{Weak}) = \frac{0.0282}{0.00686 + 0.0282} = 0.804
 \end{aligned}$$

- b. [LINK](#)