Computer Vision Exam 2

Name: Yasuo Ignacio Maidana Pérez   **Student ID**:A01328427

From the following confusion matrix:

		Actual		
		Truck	Car	Motorcycle
Predicted	Truck	4	8	15
	Car	16	23	42
	Motorcycle	6	8	15

1. Calculate the sensitivity for the car class. (Note: **Only** calculate the sensitivity for the car class)

		Actual		
		Truck	Car	Motorcycle
Predicted	Truck	4	8 F_n	15
	Car	16	23 T_p	42
	Motorcycle	6	8 F_n	15

$$Sensitivity = \frac{T.P.}{T.P.+F.N.}$$

$$S_{car} = \frac{23}{8+23+8} = 0.5897 \cong 58.97\% \checkmark$$

2. Calculate the specificity for the truck class. (Note: **Only** calculate the specificity for the truck class)

		Actual		
		Truck	Car	Motorcycle
Predicted	Truck	4	8 F_p	15 F_p
	Car	16	23 T_n	42 T_n
	Motorcycle	6	8 T_n	15 T_n

$$Specificity = \frac{T.N.}{T.N.+F.P.}$$

$$Sp_{Truck} = \frac{23+42+8+15}{23+42+8+15+8+15} = 0.79$$

$$\cong 79.279\% \checkmark$$

3. Your local zoo requests you to create an image classifier for the following animals: Lions, Zebras, Monkeys and Elephants. After using your image web scraper to collect data, you noticed that the number of images per class is not evenly distributed:
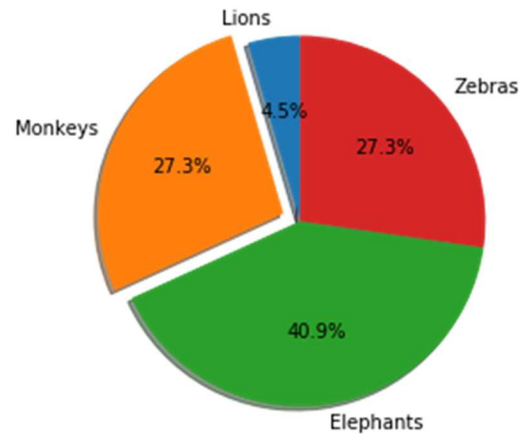
Describe at least 3 ways you would handle this imbalanced class issue. Note: You do not have access to a camera for taking additional pictures and you cannot rerun you web scraper in a different webpage.

1. Transform our data, modifying some aspect of them. For example, for lions' pictures (that's a minority class) we could apply some size or rotational transformations in order to increase the amount of data of this class, we can make this operation with each class until we got a reasonable quantity of data (data augmentation)
2. Another but inefficient option could be only taking the same quantity of data of our smallest class from other classes, however this implies that we will be wasting data, since we won't be using all the data
3. Another option could be creating subgroups of data, these subgroups will be created taking random images of each class, this selection will be done under the idea that we can replace the elements from the original source. Then we will have some batches of data that we could use to train our model

4. Using the following data, where X, Y and Z are input variables which value is either 1 or 0, and R is the output variable which value is also either 1 or 0.
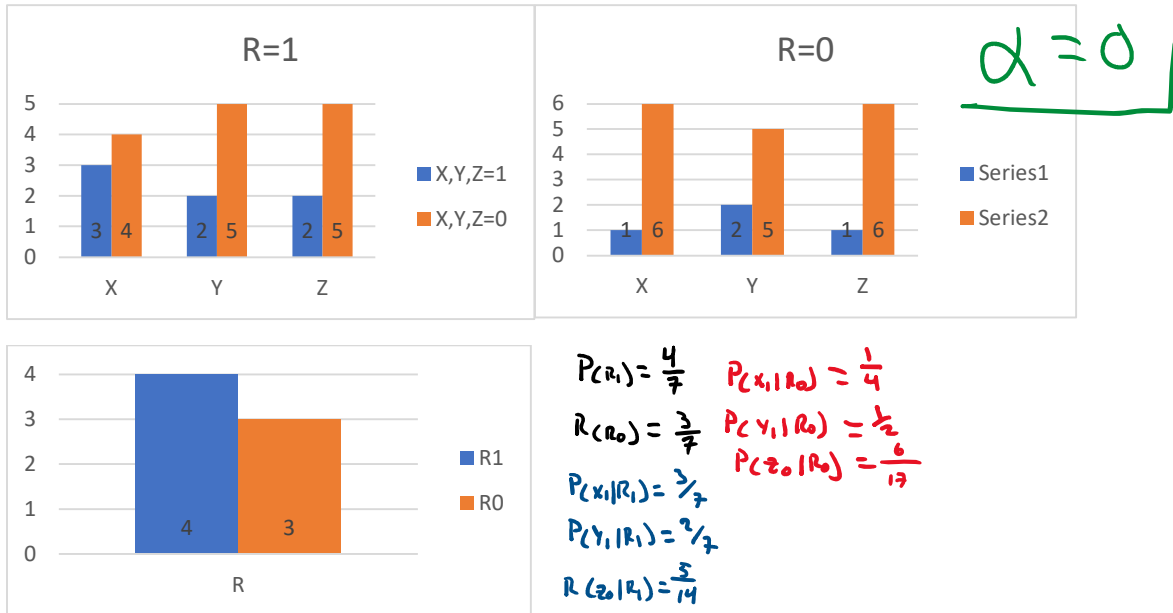
X	Y	Z	R
1	1	0	1
0	0	1	0
1	0	0	1
0	1	0	0
1	1	1	1
1	1	0	0
0	0	1	1

Show how a **Naïve Bayes** classifier predict **R** given the following input:

$$X = 1, Y = 1, Z = 0$$

If there is a tie in the output between predicting 0 or 1, the classifier always prefers to predict 1 for R.

IMPORTANT: SHOW YOUR PROCEDURE, INDICATE HOW DID YOU ARRIVE TO YOUR ANSWER.

$\alpha = 0$

$$P(R_1) = \frac{4}{7} \quad P(x_1 | R_0) = \frac{1}{4}$$
$$P(R_0) = \frac{3}{7} \quad P(y_1 | R_0) = \frac{1}{2}$$
$$P(z_0 | R_0) = \frac{6}{17}$$
$$P(x_1 | R_1) = \frac{3}{7}$$
$$P(y_1 | R_1) = \frac{2}{7}$$
$$P(z_0 | R_1) = \frac{5}{14}$$

Test R=1

$$\left(\frac{4}{7}\right)\left(\frac{3}{7}\right)\left(\frac{2}{7}\right)\left(\frac{5}{14}\right) = 0.02498 \cong 2.498\%$$

Test R=0

$$\left(\frac{3}{7}\right)\left(\frac{1}{4}\right)\left(\frac{1}{2}\right)\left(\frac{6}{17}\right) = 0.18907 \cong 1.8907\%$$

$$T_{R_1} > T_{R_0} \Rightarrow R = 1$$
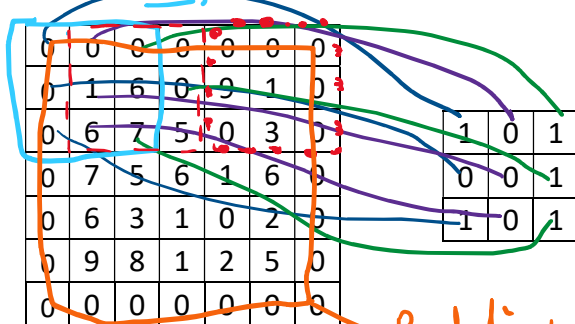
5. Given the following input matrix A:

1	6	0	9	1
6	7	5	0	3
7	5	6	1	6
6	3	1	0	2
9	8	1	2	5

And the following kernel B:

1	0	1
0	0	1
1	0	1

a) Calculate the convolution of A*B **with** zero padding and a stride of 1:

Note: Do **NOT** flip the kernel before doing the convolution.

Padded matrix:

0	0	0	0	0	0	0
0	1	6	0	9	1	0
0	6	7	5	0	3	0
0	7	5	6	1	6	0
0	6	3	1	0	2	0
0	9	8	1	2	5	0
0	0	0	0	0	0	0

Padding

Kernel:

1	0	1
0	0	1
1	0	1

$$= 0(1) + 0(6) + 0(1) + 0(6) + 1(0) + 6(1)$$
$$+ 0(4) + 6(6) + 7(1) =$$

Result:

13	11	16	9	0
18	19	21	16	10
15	24	11	17	0
16	24	16	20	3
11	8	5	8	0

← Result

b) Calculate the convolution A*B **without** zero padding and a stride of 1:

Note: Do **NOT** flip the kernel before doing the convolution.

1	6	0	9
6	7	5	0
7	5	6	1
6	3	1	0
9	8	1	2

$*$

1	0	1
0	0	1
1	0	1

$\Rightarrow$

19	21	16
24	11	17
24	16	20

✓

6. Let H[u,v] be a 3x3 matrix. Recall the definitions of
   Erosion:
   G[x,y] = 1 if F[x+u-1,y+v-1] is 1 everywhere that H[u,v] is 1
   G[x,y] = 0 otherwise

   Dilation:
   G[x,y] = 1 if F[x+u-1,y+v-1] is 1 somewhere that H[u,v] is 1
   G[x,y] = 0 otherwise

Input image F

Desired image D

The image D above can be generated from F using a dilation or erosion operation followed by a second dilation or erosion. Assume H is a 3x3 matrix containing all ones. Ignore the border pixels. Below, specify the two operations (circle one), and for each one draw the entries of the image after that operation has been performed.

0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	1	0	1	0	0	0	0	0	0
0	0	0	1	1	1	0	1	1	1	0	0
0	0	0	1	1	1	1	1	1	1	0	0
0	0	1	1	1	0	1	1	1	0	0	0
0	0	0	1	0	1	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0

Adding padding to the matrix and coloring for easy understanding

If we apply the erosion function this will happen

0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	1	0	1	0	0	0	0	0	0	0
0	0	0	1	1	1	0	1	1	1	0	0	0
0	0	0	1	1	1	1	1	1	1	0	0	0
0	0	1	1	1	1	0	1	1	1	0	0	0
0	0	0	1	0	1	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0

0

$3 < 9 \therefore R = 0$

$9 = 9 \therefore R = 1$

0	1	1	2	1	1	0	0	0	0	0
0	2	3	5	3	3	2	3	2	1	0
0	3	5	8	6	6	5	6	4	2	0
1	4	7	9	7	7	7	9	6	3	0
1	4	6	8	6	6	5	6	4	2	0
1	3	4	5	3	3	2	3	2	1	0
0	1	1	2	1	1	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0

Only the cell with 9 value satisfies the convolution threshold so

0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0
0	0	0	1	0	0	0	1	0	0	0
0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0

After erosion

Then we apply again padding

0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	1	0	0	0	1	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0

✓

After applying dilatation

0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0
0	0	1	1	1	0	1	1	1	0	0
0	0	1	1	1	0	1	1	1	0	0
0	0	1	1	1	0	1	1	1	0	0
0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0

✓

I = H applied to F
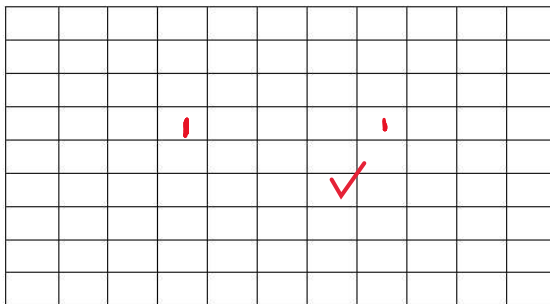
D = H applied to I

Circle one: dilation or (erosion) ✓
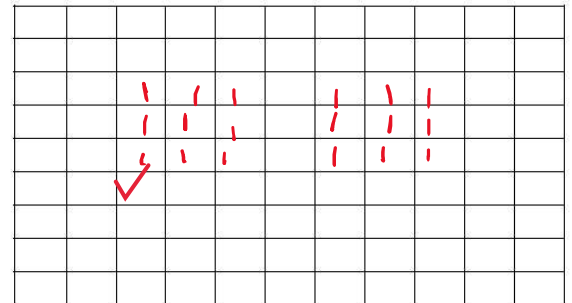
Circle one: (dilation) or erosion ✓

NOTE: Image I is the Intermediate image generated after applying the first erosion or dilation operation to F. Image D is the Desired final image generated by applying a dilation or erosion operation to I.

7. For each of the following, <u>briefly</u> explain how it may be used to help solve a computer vision problem.

a) Bag of Words. It is a technique that use histogram of "small" detections, so in accordance with the characteristic more common the algorithm assigns a class or property. It is useful when want to classify objects ✓

b) Median filter. It is a filter that through a kernel obtains the "mean" of a region, some kernels are 3x3 with each element equal to 1/9. It is useful to prepare a picture before edge detection algorithm because the mean help to get sharp limits ✓

c) Edge detection. It is technique that through a kernel allows us to get the region where a change of features occurs, it uses a kernel like this [[1,0,-1], [1,0,-1] ,[1,0,-1]]

d) Gaussian filter. It is a filter with a kernel based on the gaussian function. It is very used to blur, or to filter noise, it could help to get soft limits. ✓

8. Given the following input matrix:

a) Perform a max pooling convolution on it using a 2x2 kernel with a stride of 2. The name of this output matrix is **Result8A**
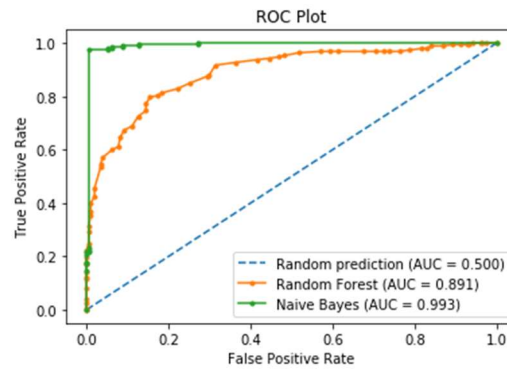
9 3 7 9
9 7 5 9
4 9 6 7
6 9 1 4 ✓

b) Perform a max pooling convolution on **Result8A** using a 2x2 kernel with a stride of 2.
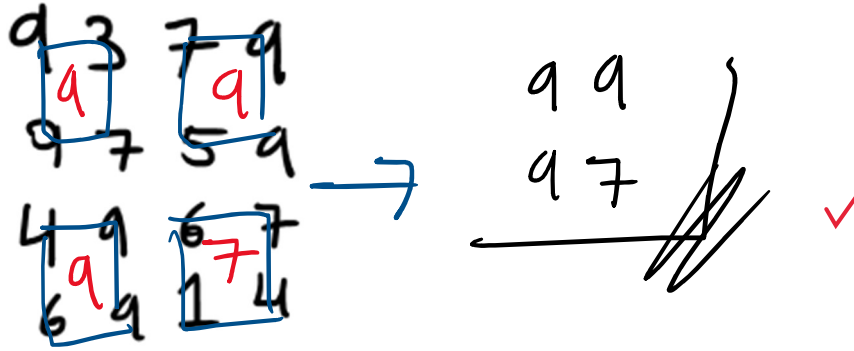
9. For an experiment you designed three classifiers, naïve Bayes, random forest and a model that predicts classes at random. The ROC of the three models is the following:


ROC Plot

Random prediction (AUC = 0.500)
Random Forest (AUC = 0.891)
Naive Bayes (AUC = 0.993)

a) Which of the 3 models made a better job at classifying? Justify your answer. **NOTE:** No partial credit will be assigned without an explanation of your answer.
The green one, because the area under its curve is greater.

b) The following ROC graph has 5 points (A, B, C, D and E) each one associated to a threshold value.

$$TPR = \frac{TP}{TP + FP}$$

$$FPr = \frac{FP}{FP + TF}$$
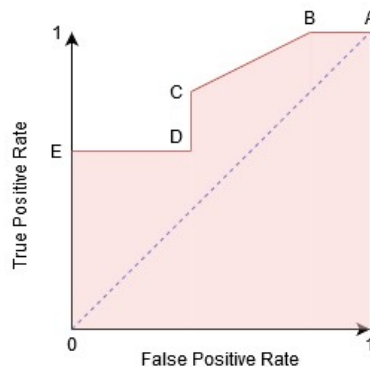
$$Sensitivity = \frac{T.P.}{T.P. + F.N.}$$

Considering that **<u>sensitivity</u>** is the most important metric for this model, which of the 5 threshold values is the optimal? Justify your answer.

The point B is the point with the highest TPR(sensitivity) with the smallest false positive rate, so this should be the point to choose ✓

10. Explain a concept of image processing that was not included in this exam that you learned in this second week.
    - Morphological transformation. A set of operation that allows us to manipulate through kernels to obtain shapes desired, as skeleton, edges or areas.
    - Semantic branch. The difference in the way that humans process images compared to computers
    - Image warping. An image processing where we modify the dimensions or orientation of an image, compared to filtering, where we modify the values of the pixels. The difference relies on while the modification of image warping is focused on the "position" of the pixels the filtering processing is focused on the values of the pixels

✓