

# Data Mining 2

CAT 2 - 2019/2020

Name \_\_\_\_\_ Surname \_\_\_\_\_ ID: \_\_\_\_\_

Test id. AUTO

Q1. Write down the constraints for a linear SVM nonseparable case.

A1. \_\_\_\_\_

Q2. In SVM in what consist the kernel trick?

A2. \_\_\_\_\_

Q3. For what the acronym CRISP-DM stands for?

A3. \_\_\_\_\_

Q4. In CRISP-DM what is done in the Modeling phase?

- 1) Evaluation of model
- 2) Generate test design
- 3) Verify data quality
- 4) Build model
- 5) Interpretation of model

A4. \_\_\_\_\_

N.B.: this question can have more than one correct answer

Q5. We need hidden layers in NN to ...

- 1) solve non linear problems
- 2) solve the XOR problem
- 3) speed up the calculus
- 4) use *tanh* as activation function
- 5) None of the others

A5. \_\_\_\_\_

N.B.: this question can have more than one correct answer

Q6. Train the linear perceptron on the training set in the Figure using the sign as activation function and classify the record in the test set.

train													
id	X <sub>1</sub>	X <sub>2</sub>	Y	it	W <sub>0</sub>	W <sub>1</sub>	W <sub>2</sub>	X.W	f(X.W)	error	delta <sub>0</sub>	delta <sub>1</sub>	delta <sub>2</sub>
a	1	1	1	1	-1	0	0						
b	0	2	1	2									
c	1	0	-1	3									
				4									
				5									
				6									
				7									
				8									
				9									
				10									
				11									
				12									

Lambda = 0.4  
f = sign

test			
id	X <sub>1</sub>	X <sub>2</sub>	Y
1	0	0	
2	-1	1	
3	2	0	

A6. \_\_\_\_\_

Q7. Which type of ensemble is specifically designed for decision trees?

- 1) Bagging
- 2) Naive Bayes
- 3) Random Forest
- 4) Boosting
- 5) None of the others

A7. \_\_\_\_\_

Q8. Given the dataset in the Figure run the first iteration of AdaBoost, find the best split, and fill the column 'norm weight' with the normalized new weights.

plan	sex	minutes	churn	weight	new weight	norm weight
travel	F	90	N			
travel	F	130	Y			
travel	M	70	N			
travel	M	80	N			
normal	M	90	Y			
normal	M	120	Y			
normal	F	100	Y			
normal	F	110	N			
travel	F	100	N			

Gain Function = Misclassification Error       $Z =$

Split by

Error =

Alpha =

$w^{j+1} =$

misclassified

$w^{j+1} =$   
correctly

classified

A8. \_\_\_\_\_