Data Mining 2

CAT 2 - 2019/2020

Name	Surname	ID:	Test id. AUTO
Q1. Write down the const	raints for a linear SVM nonsep	arable case.	
A1			
Q2. In SVM in what consis	st the kernel trick?		
A2			
Q3. For what the acronym	CRISP-DM stands for?		
АЗ			
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

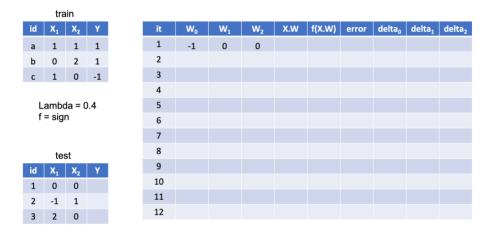
Q5. We need hidden layers in NN to ...

- 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.



A6. ___

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

1) Bagging	2) Naive Bayes	3) Random Forest	4) Boosting	5) None of the others
, 55 5	, ,	/	, 0	/

A7. _____

Q8. Given the dataset in the Figure run the first iteration of AdaBoost, fin 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	Ν			
travel	F	130	Y			
travel	М	70	Ν			
travel	М	80	N			
normal	М	90	Y			
normal	М	120	Y			
normal	F	100	Y			
normal	F	110	N			
travel	F	100	Ν			
Gain Function = Misclassification Error Z =						

Split by

Error =

Alpha =

w^{j+1} =

misclassified

w^{j+1} = correctly

classified

A8. _____