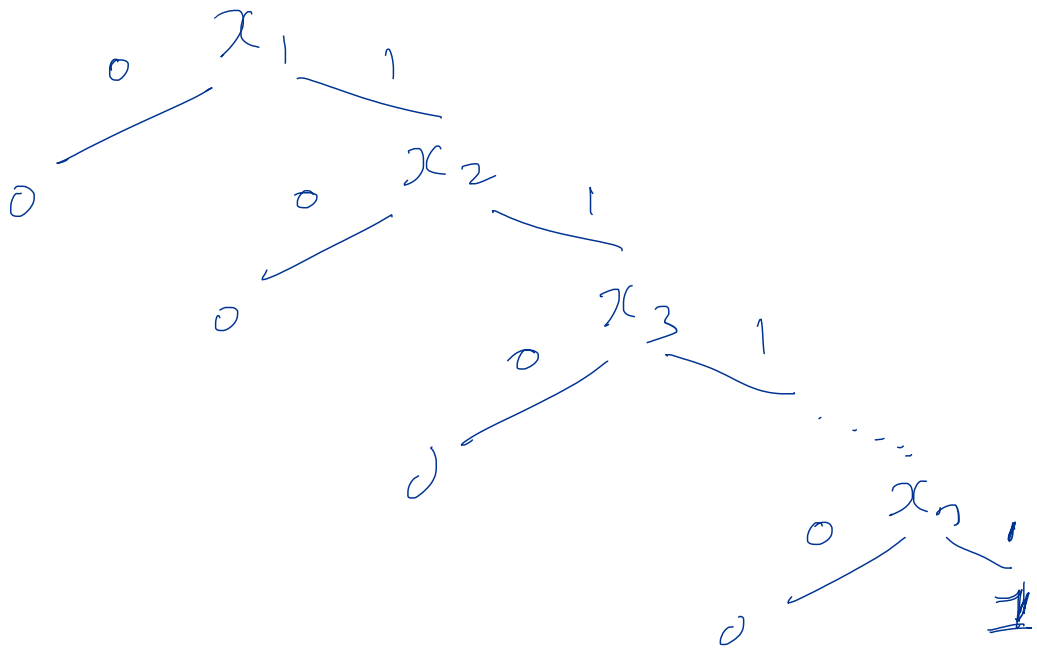
ML Sample Questions

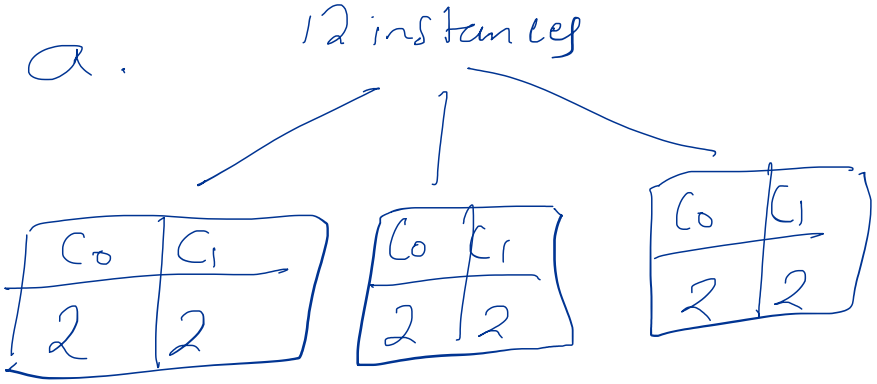
(1) Draw a decision tree for

$x_1 \wedge x_2 \wedge \ldots\ldots x_n$

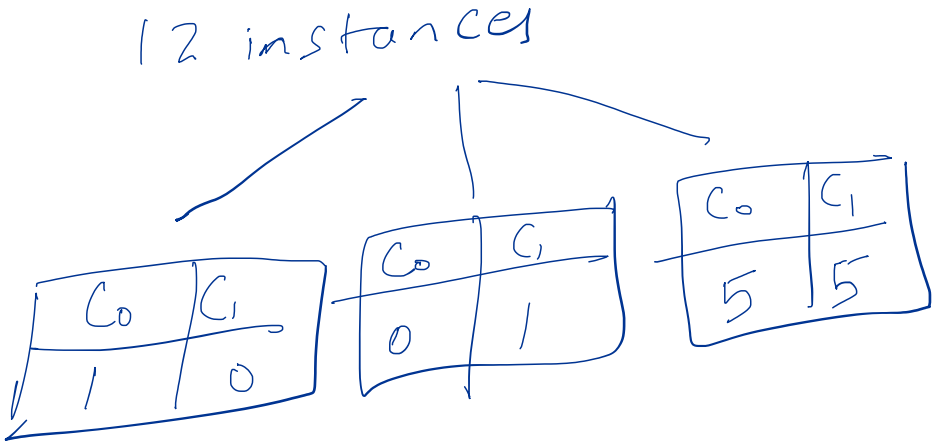


② Given 12 instances having Class labels either C_0 or C_1, which of the following is a better split

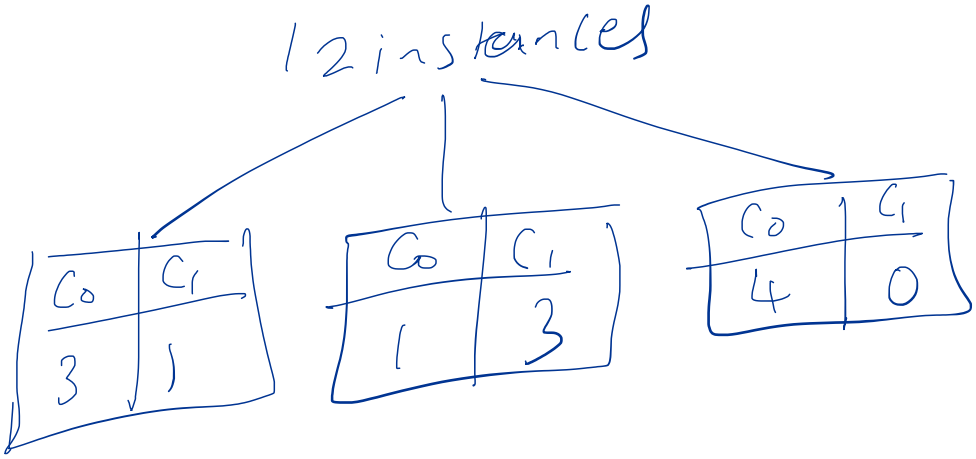
a.



b.



c.



C. is a better split since the entropy is nearly 0 after conditioning.

(2) Given the following dataset from a classification algorithm Compute precision & recall w.r.t the "+" class

True label	Output label
+	−
−	+
+	+
−	−
+	+

True positives = 2

False positives = 1

False negatives = 1

Precision: $2/3$

Recall: $2/3$

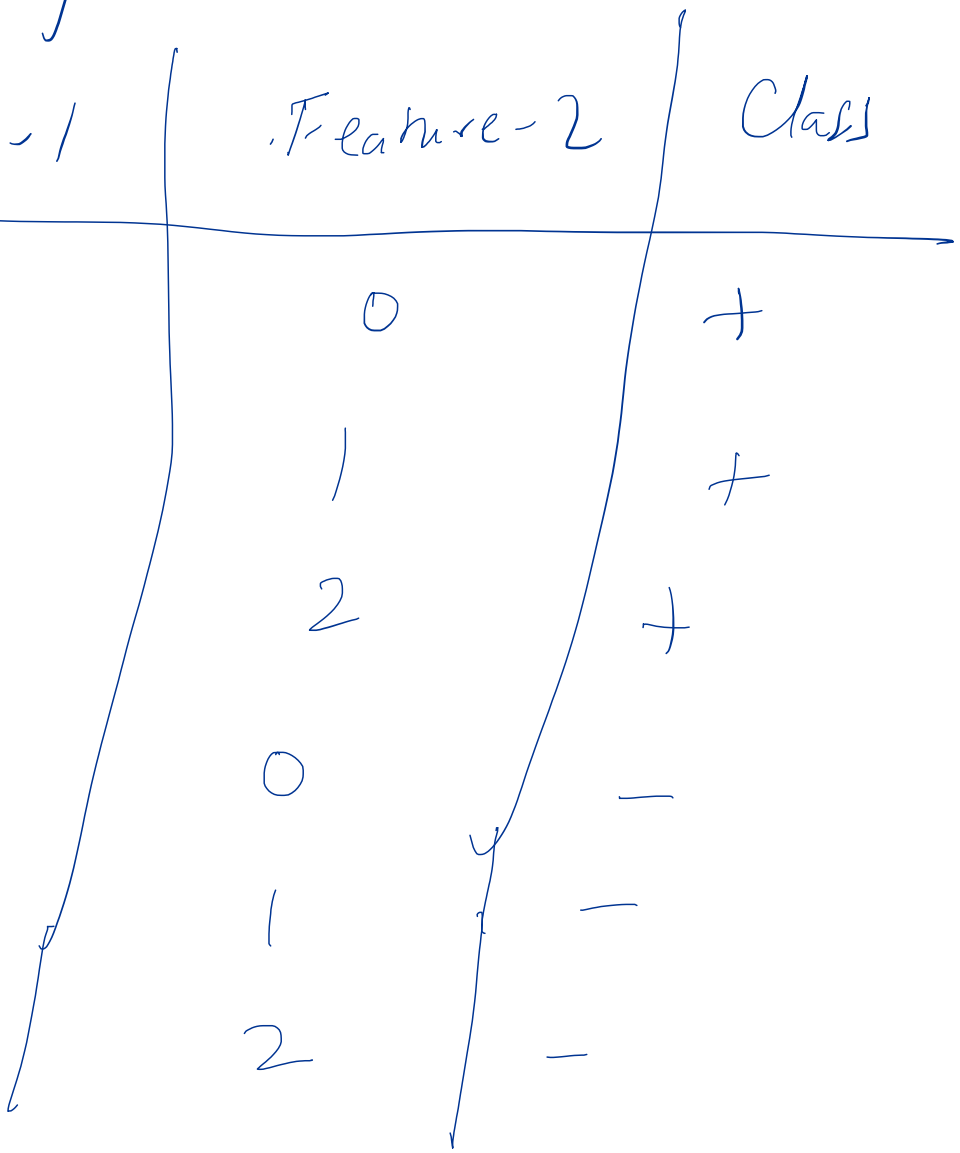
(3) - Suppose a classification algorithm works as follows. If all inputs in my training data is "+" output "+" else output "-" for any training instance. What is its Leave-one-out Cross Validation error if a dataset has 50% +ve & 50% -ve examples

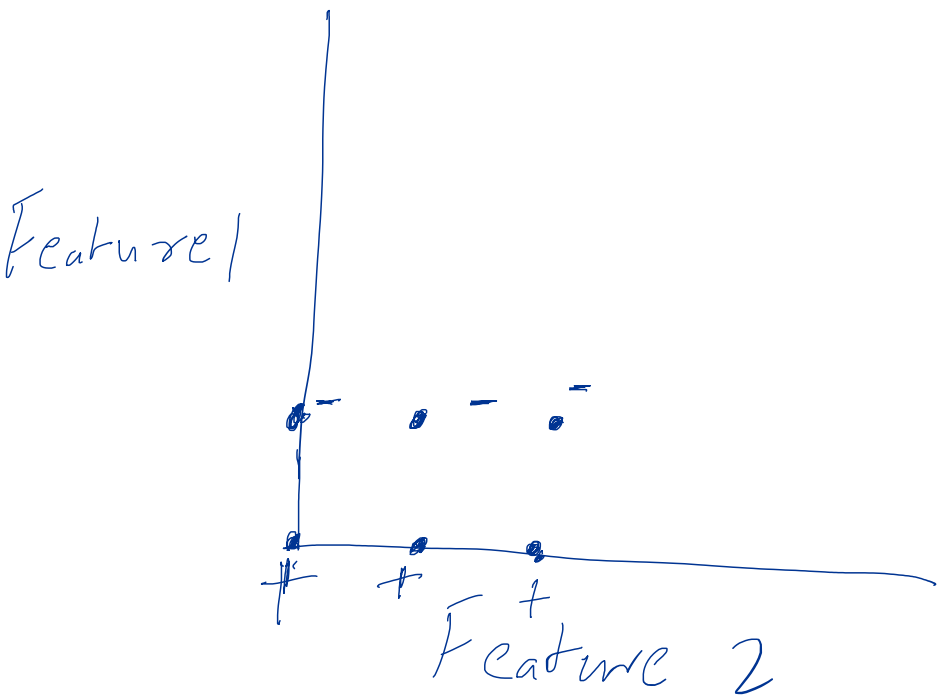
Soln

The algorithm gets the output correctly if the label of the instance "left-out" is -ve, else it gets the output wrong. ∴ LOOCV error = 50%

4. Can a perceptron with 2 weights classify the following dataset correctly.

Feature-1	Feature-2	Class
0	0	+
0	1	+
0	2	+
1	0	−
1	1	−
1	2	−





We cannot find a line passing through the origin that separates "+" from "-".

∴ Perceptron with 2 weights will not get all the labels correctly.

5. Given the following dataset

Car-ahead	Ice	Speed	Brake
No	No	Low	Yes
Yes	Yes	high	Yes
No	Yes	high	No
No	Yes	Low	No
Yes	No	low	No

what will the Naive-Bayes algorithm predict for brake given Car-ahead: Yes, Ice: Yes, Speed: Low

$$P(\text{Car-ahead: Yes} \mid \text{Brake: Yes}) = 1/2$$

$$P(\text{Car-ahead: Yes} \mid \text{Brake: No}) = 1/3$$

$$P(\text{Ice: Yes} \mid \text{Brake: Yes}) = 1/2$$

$$P(\text{Ice: Yes} \mid \text{Brake: No}) = 2/3$$

$$P(\text{Speed: low} \mid \text{Brake: Yes}) = 1/2$$

$$P(\text{Speed: low} \mid \text{Brake: No}) = 2/3$$

$P(\text{brake: Yes}) = 2/5$

$P(\text{brake: No}) = 3/5$.

$$1/2 \cdot 1/2 \cdot 1/2 \cdot 2/5$$

$$\leq$$

$$1/3 \cdot 2/3 \cdot 2/3 \cdot 3/5$$

$\therefore$ We choose brake: No

7) Given the following training dataset what is the precision & recall w.r.t the "Yes" class in the test dataset using the 1-nearest neighbor classifier

Feature-value	Class
1	Yes
2	No
4	Yes
5	No
7	Yes
8	No

Training data

Feature-Value	Class
1.1	Yes
2.1	Yes
5.1	No
7.1	Yes
8.1	No

Test dataset

Soln

Value	True-Label (in test data)	Predicted
1.)	Yes	Yes
2.)	Yes	No
5.)	No	No
7.)	Yes	Yes
8.)	No	No

True positives: 2

False positives: 0

False negatives: 1

Precision: 1.0

Recall: 2/3

(8) Do you think Naive Bayes is more like Bayesian learning or like Max-likelihood learning.

Soln

In general, Naive Bayes is neither Bayesian nor Max-likelihood due to the conditional independence assumption. But if we had to choose it is more like Bayesian learning since we use a prior. Interestingly, we estimate the prior using Max-likelihood learning (instead of assuming a form)

(9) In logistic regression, which of the following is true

(a) we assume that the features are independent of each other

(b) we assume that the data instances are independent of each others

Sln (b) is true while (a) is false. (b) is true because when we formulate the Conditional log-likelihood function we are implicitly assuming this.

(16). Is VC-dimension a property of i) training data or

ii) number of features or

iii) Learning algorithm.

Soln

∴ It is definitely not a function of training data.

It is inherently the property of a learning algorithm and possibly the number of features