DS&AI

Machine Learning

DPP: 1

Bayesian Learning

Q1 Consider a naive Bayes classifier with 3 boolean input variables, X_1, X_2 and X_3, and one boolean output, Y.
How many parameters must be estimated to train such a naive Bayes classifier? (you need not list them unless you wish to, just give the total)

Q2 Match the items in Column 1 with the items in Column 2 in the following table:

Column 1	Column 2
(p) Principal Component Analysis	(i) Discriminative Model
(q) Naïve Bayes Classification	(ii) Dimensionality Reduction
(r) Logistic Regression	(iii) Generative Model

(A) (p) (iii), (q) - (i), (r) - (ii)
(B) (p) (ii), (q) - (i), (r) - (iii)
(C) (p) (ii), (q) - (iii), (r) - (i)
(D) (p) (iii), (q) - (ii), (r) - (i)

Q3 Suppose you have a three class problem where class label y 0,1,2 and each training example X has 3 binary attributes X_1, X_2, X_3 0,1. How many parameters do you need to know to classify an example using the Naive Bayes classifier?
(A) 5 (B) 9
(C) 11 (D) 13

Q4 Consider the two class problem where class label y {T, F} and each training example X has 2 binary attributes X_1, X_2 {T, F}.
How many parameters will you need to know/estimate if you are to classify an example using the Naivs Bayes classifier?
(A) 5 (B) 8
(C) 3 (D) 7

Q5 Suppose that you are trying to solve a binary classification problem, and your data set has 4 at-tributes. Each attribute can take 3 possible values.
If you modeled the full joint distribution of the attributes and the class label, how many parameters would you need?

Q6 Consider a classification problem with two classes and n binary attributes.
How many would you need to learn with a Naive Bayes classifier? How many parameters would you need to learn with a Bayes optimal classifier?

Q7 Consider a classification problem with two binary features, x_1, x_2 {0,1}. Suppose P(Y = y) = 1/32, P(x_1 = 1 Y = y) = y/46, P(x_2 = 1 | Y = y) = y/62. Which class will naive Bayes classifier produce on a test item with x_1 =1 and x_2 = 0?
(A) 16 (B) 26
(C) 31 (D) 32

Q8 What is the total number of parameters needed to model P(y, x_1,x_2,x_3.....,x_d)?
• k is the number of possible values Y can take
• t is the number of value of each x_i where i ∈ {1, 2.....,d}.
(A) $k \cdot t^d - 1$
(B) $k \cdot t^{d-1}$
(C) $k \cdot (t^d - 1) + (k - 1)$
(D) (kdt)

Q9

$0A_1$	A_2	Class Label Y
True	True	+
True	True	+
True	False	–
False	False	+
False	True	–
False	False	–
False	False	–
True	False	+

What is the probability P(Y = + | A_1 = True, A_2 = True) by applying the Naïve Bayes classifier (show each step in computing the probability)? do What label should you predict for input (A_1 = True, A_2 = True) based on your result above ?

Q10 Suppose you have the following training set with three boolean input z, y and z, and a boolean output U.

x	y	z	U
1	0	0	0
0	1	1	0
0	0	1	0
1	0	0	1
0	0	1	1
0	1	0	1
1	1	0	1

Suppose you have to predict U using a naivs Bayes classifer.
After learning is complete what would be the predicted probability

$P(U = 0|x = 0, y = 1, z = 0)?$

Q11 Suppose you have the following training set with three boolean input z, y and z, and a boolean output U.

x	y	z	U
1	0	0	0
0	1	1	0
0	0	1	0
1	0	0	1
0	0	1	1
0	1	0	1
1	1	0	1

Suppose you have to predict U using a naivs Bayes classifer.
Using the probabilities obtained during the Bayes Classifier training, what would be the predicted proability $P(U = 0|x = 0)?$

Q12 Again, consider the two class problem where class label y {T, F} and each training example X has 2 binary attributes x_1, x_2 {T, F}. How many parameters will you need to estimate if we do not make the Naïve Bayes conditional independence assumption?

(A) 3 (B) 5
(C) 7 (D) 8

Q13 Consider a Bayes classifier with 3 boolean input variables, x_1, x_2, and x_3, and one boolean output Y.
How many parameters must be be estimated to train such that Bayes classifier?

Q14 Suppose that you are trying to solve a binary classification problem, and your data set has 4 attributes. Each attribute can take 3 possible values If you modeled the naive Bayes of the attributes and the class label, how many parameters would you need?

Q15 In a naive Bayes classifier, if the probability of class A given the predictor variables is 0.7 and the probability of class B given the same predictor variables is 0.3, what is the predicted class for a new data point?

(A) Class A
(B) Class B
(C) Cannot be determined
(D) Depends on the feature space

Q16 Consider a K-class dataset containing V points, where each sample is described by D continuous features. You decide to use a Gaussian Naive Bayes classifier. How many parameters need to be estimated for this model?

(A) 2DK + K (B) DK
(C) 3DK (D) 2D + K

Q17 Given the training set below with boolean input features x, y, z and a boolean output U, what would be the predicted probability P(U = 0 | x = 0, y = 1, z=0) after training a Naive Bayes classifier?

x	y	z	u
1	0	0	0
0	1	0	0
0	0	1	0
1	0	0	1
0	0	1	1
0	1	0	1
1	1	0	1

(A) 6/35 (B) 7/35
(C) 8/35 (D) 9/35

Q18 Given a dataset with K binary-valued attributes (where K > 2) for a two-class classification task, the number of parameters to be estimated for learning a naïve Bayes classifier is

(A) $2^k + 1$
(B) 2K + 1
(C) $2^{(k+1)} + 1$
(D) $K^2 + 1$

Answer Key

Q1	14~14
Q2	C
Q3	C
Q4	A
Q5	25~25
Q6	(1 + 2n)
Q7	C
Q8	D
Q9	0.142~0.142
Q10	1
Q11	0.5~0.5
Q12	B
Q13	7~7
Q14	25~25
Q15	A
Q16	A
Q17	C
Q18	B

Hints & Solutions

Note: scan the QR code to watch video solution

Q1 Text Solution:

3 dimensions X_1, X_2, X_3 Boolean binary 0, 1

clan Y Boolean 0, 1

P_{y1} P_{y2} clan probability

$X_{1=0}$

$X_{1=1}$

$X_{2=1}$

$X_{3=0}$

$X_{3=1}$

2 + 3 × 2 × 2 = 14

Q2 Text Solution:

(p) Principal Component Analysis (ii) Dimensionality Reduction

(q) Naïve Bayes Classification (iii) Generative Model

(r) Logistic Regression (i) Discriminative Model

Q3 Text Solution:

3 + 3 × 2 × 3 = 21

Q4 Text Solution:

2 + 2 × 2 × 2 = 10 5

Q5 Text Solution:

2 class

4 dimension 3 possible value

2 + 4 × 3 × 2 = 26 1 + 4 × 3 × 2 = 25

Q6 Text Solution:

2 + n × 2 × 2 = (4x + 2)

1 + (n × 2 ×1) = (1 + 2n)

Q7 Text Solution:

P(y) $P(x_1 = 1/y)$ $P(x_2 = 0/y)$ max

$$\frac{1}{32} \times \frac{y}{46}(1 - y/62)$$

$$\frac{d}{dy} = 0 \Rightarrow \frac{1}{32}\left[\frac{1}{46} - \frac{2y}{46 \times 62}\right] = 0$$

y = 31

Q8 Text Solution:

Number of parameter k + k × d × d

class probab (k + kdt)

we can neglect k ⇒kdt

Q9 Text Solution:

$$p\left(y = +/T, T\right) = \frac{P(T,T/+) \cdot P_+}{P(T,T)}$$

$$p(y/x) = \frac{p(x/y)p_y}{p_x}$$

$$\left(P_+ = \frac{4}{8} = \frac{1}{2}, P_- = \frac{4}{8} = \frac{1}{2}\right)$$

$$P(T, T/+) = P(T/+)P(T/+)$$

$$= \frac{3}{4} \times \frac{2}{4} \Rightarrow \frac{3}{8}$$

$$P(T, T/-) = P(T/-)P(T/-)$$

$$= \frac{1}{4} \times \frac{1}{4} = \frac{1}{16}$$

$$P\left(y = +/T, T\right) = \frac{3}{8} \times \frac{1}{2}$$

$$P(T, T)$$

$$P\left(y = -/T, T\right) = \frac{1}{16} \times \frac{1}{2}$$

add = 1

$$P(T, T)$$

$$\frac{3/16}{P(T,T)} + \frac{y/32}{P(T,T)} = 1$$

$$P(T, T) = 0.21875$$

$$P\left(y = +/T, T\right) = \frac{6}{7},$$

$$P\left(y = -/T, T\right) = \frac{1}{7}$$

So naive Bayes predict + Clan

Q10 Text Solution:

$P_{u=0} = 3/7$ $P_{u=1} = 4/7$

P(u = 0/x = 0, y = 1, z = 0) =

$$\frac{P(x=0, y=1, z=0/u=0)P_{u=0}}{P(x=0,y=1,z=0)}$$

$P(x=0/u=0)\ P(y=1/u=0)P(z=0/u=0)$

$\frac{2}{3} \times \frac{1}{3} \times \frac{1}{3}$

P(x = 0/u = 1) P(y = 1/u = 1) P(z=0/u = 1)

$\frac{2}{4} \times \frac{2}{4} \times \frac{3}{4}$

$P\left(u=0/x=0, y=1, z=0\right) =$

$\frac{2/63}{P(x=0,\ y=1,\ z=0)} \rightarrow \frac{8}{35}$

$P\left(u=1/x=0, y=1, z=0\right) =$

$\frac{3/28}{P(x=0,\ y=1,\ z=0)} \rightarrow \frac{27}{35}$

P(x = 0, y = 1, z = 0) = 5/36

and we anign clan 1 to the test point.

Q11 Text Solution:

$P\left(u=0/x=0\right) = \frac{P(x=0/u=0)P_{u=0}}{P_{x=0}}$

$P\left(u=1/x=0\right) = \frac{P(x=0/u=1)P_{u=1}}{P_{x=0}}$

$P(u=0/x=0) = \frac{2/7}{P(x=0)}$ add = 1

$P\left(u=1/x=0\right) = \frac{2/7}{P(x=0)}$

P(x = 0) = 4/7

P(u = 0/ x = 0) = 0.5

Q12 Text Solution:

2 + 2 × 2 × 2 = 10

⇒5

Q13 Text Solution:

2 + 3 × 2 × 2 = 14

⇒7

Q14 Text Solution:

2 + 4 × 2 × 3 = 26

⇒ 25

Q15 Text Solution:

P(A/x) = 0.7

P(B/x) = 0.3

Rule $P(C_1/x) > < P(C_2/x)$

$C_1\ C_2$

By Rule clan A

Q16 Text Solution:

- Each sample has D cont. features
- K clames

Now we break data clan wise

x^1/clan1 X^1/clan2 .... X^1/clank

μ, σ μ, σ μ, σ

Each dimension broken into K clam set and we finnd μ, σ for each set.

in 1 feature we need 2 × K parameter

in D feature we need 2 × D × K parameter

and we need K clam probabilities also = 2DK + K

Q17 Text Solution:

8/35

Q18 Text Solution:

- K binary attributes
- 2 Clanes P(y = 1) I can find $P_y = 0$
- Number of parameters = 2 + K × 2 × 2
  = 2 + 4K 1 + 2K
- $P(x_1 = 0/y)$ $P(x_1 = 1/y)$