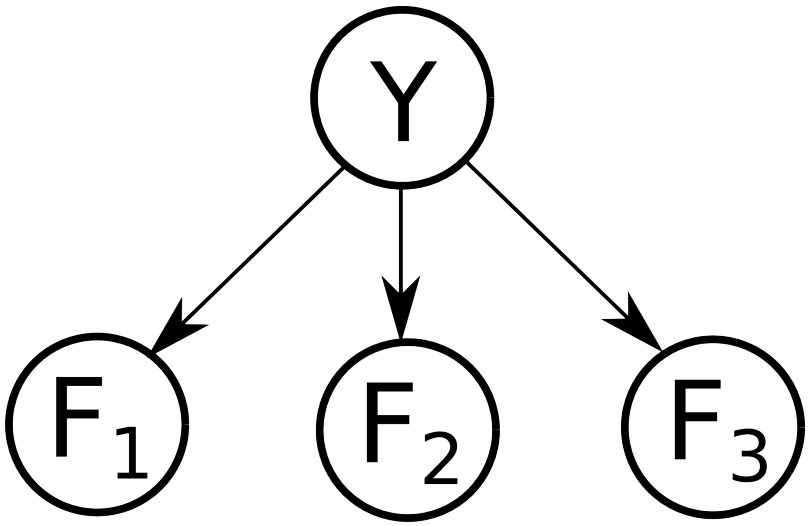
复查测验提交: Homework 6

用户	信息科学与技术学院 何耀宇
课程	人工智能I
测试	Homework 6
已开始	24-12-27 上午3:15
已提交	24-12-27 下午4:51
截止日期	24-12-27 下午11:59
状态	已完成
尝试分数	得 114 分，满分 124 分
已用时间	13 小时 36 分钟
显示的结果	所有答案, 已提交的答案, 正确答案

问题 1

得 18 分，满分 18 分

In this question, we will train a Naive Bayes classifier to predict class labels Y as a function of input features F_i . (**Keep 3 decimal places**)



We are given the following 15 training points:

F_1	0	1	1	1	0	1	1	1	1	1	1	0	1	1	1
F_2	0	0	1	0	1	0	1	1	1	1	1	1	1	0	1
F_3	1	0	1	1	1	1	1	1	1	1	1	1	0	1	0
Y	A	A	B	B	B	B	B	B	B	B	B	C	C	C	C

What is the maximum likelihood estimate of the prior P(Y)?

Y	P(Y)
A	**[q1.1]**
B	**[q1.2]**
C	**[q1.3]**

What are the maximum likelihood estimates of the conditional probability distributions? Fill in the tables below (the second and third are done for you).

F_1	Y	$P(F_1 \mid Y)$
0	A	**[q1.4]**
1	A	**[q1.5]**
0	B	**[q1.6]**
1	B	**[q1.7]**
0	C	**[q1.8]**
1	C	**[q1.9]**

F_2	Y	$P(F_2 \mid Y)$
0	A	1.000
1	A	0.000
0	B	0.222
1	B	0.778
0	C	0.250
1	C	0.750

F_3	Y	$P(F_3 \mid Y)$
0	A	0.500
1	A	0.500
0	B	0.000
1	B	1.000
0	C	0.500
1	C	0.500

q1.1 的指定答案: 0.133

q1.2 的指定答案: 0.600

q1.3 的指定答案: 0.267

q1.4 的指定答案: 0.500

q1.5 的指定答案: 0.500

q1.6 的指定答案: 0.111

q1.7 的指定答案: 0.889

q1.8 的指定答案: 0.250

q1.9 的指定答案: 0.750

q1.1 的正确答案:		
评估方式	**正确答案**	**区分大小写**
完全匹配	0.133	
q1.2 的正确答案:		
评估方式	**正确答案**	**区分大小写**
完全匹配	0.6	
完全匹配	0.600	
q1.3 的正确答案:		
评估方式	**正确答案**	**区分大小写**
完全匹配	0.267	
q1.4 的正确答案:		
评估方式	**正确答案**	**区分大小写**
完全匹配	0.5	
完全匹配	0.500	
q1.5 的正确答案:		
评估方式	**正确答案**	**区分大小写**
完全匹配	0.5	
完全匹配	0.500	
q1.6 的正确答案:		
评估方式	**正确答案**	**区分大小写**
完全匹配	0.111	
q1.7 的正确答案:		
评估方式	**正确答案**	**区分大小写**

完全匹配	0.889	
q1.8 的正确答案:		
评估方式	**正确答案**	**区分大小写**
完全匹配	0.25	
完全匹配	0.250	
q1.9 的正确答案:		
评估方式	**正确答案**	**区分大小写**
完全匹配	0.75	
完全匹配	0.750	

问题 2

得 14 分，满分 14 分

Following question 1, Now consider a new data point ($F_1 = 0, F_2 = 0, F_3 = 1$).

Use your classifier to determine the joint probability of causes Y and this new data point, along with the posterior probability of Y given the new data: (**Keep 3 decimal places**)

Y	$P(Y, F_1 = 0, F_2 = 0, F_3 = 1)$
A	**[q2.1]**
B	**[q2.2]**
C	**[q2.3]**

Y	$P(Y \mid F_1 = 0, F_2 = 0, F_3 = 1)$
A	**[q2.4]**
B	**[q2.5]**
C	**[q2.6]**

What label does your classifier give to the new data point? (Break ties alphabetically). Enter capital letters only

[q2.7]

The training data is repeated here for your convenience:

F_1	0	1	1	1	0	1	1	1	1	1	1	0	1	1	1
F_2	0	0	1	0	1	0	1	1	1	1	1	1	1	0	1
F_3	1	0	1	1	1	1	1	1	1	1	1	1	0	1	0
Y	A	A	B	B	B	B	B	B	B	B	B	C	C	C	C

q2.1 的指定答案: 0.033

q2.2 的指定答案: 0.015

q2.3 的指定答案: 0.008

q2.4 的指定答案: 0.589

q2.5 的指定答案: 0.268

q2.6 的指定答案: 0.143

q2.7 的指定答案: A

q2.1 的正确答案:

评估方式	正确答案	区分大小写
完全匹配	0.033	

q2.2 的正确答案:

评估方式	正确答案	区分大小写
完全匹配	0.015	

q2.3 的正确答案:

评估方式	正确答案	区分大小写
完全匹配	0.008	

q2.4 的正确答案:

评估方式	正确答案	区分大小写
完全匹配	0.589	
完全匹配	0.59	
完全匹配	0.590	

q2.5 的正确答案:

评估方式	正确答案	区分大小写
完全匹配	0.268	
完全匹配	0.262	

q2.6 的正确答案:

评估方式	正确答案	区分大小写
完全匹配	0.143	
完全匹配	0.148	
q2.7 的正确答案：		
评估方式	**正确答案**	**区分大小写**
完全匹配	A	

问题 3

得 18 分，满分 18 分

Following the previous questions, now use Laplace Smoothing with strength k = 3 to estimate the prior P(Y) for the same data. (**Keep 3 decimal places**)

Y	P(Y)
A	**[q3.1]**
B	**[q3.2]**
C	**[q3.3]**

Use Laplace Smoothing with strength k = 3 to estimate the conditional probability distributions below (again, the second two are done for you).

F_1	Y	$P(F_1 \mid Y)$
0	A	**[q3.4]**
1	A	**[q3.5]**
0	B	**[q3.6]**
1	B	**[q3.7]**
0	C	**[q3.8]**
1	C	**[q3.9]**

F_2	Y	$P(F_2 \mid Y)$
0	A	0.625

1	A	0.375
0	B	0.333
1	B	0.667
0	C	0.400
1	C	0.600

F_3	Y	$P(F_3 \mid Y)$
0	A	0.500
1	A	0.500
0	B	0.200
1	B	0.800
0	C	0.500
1	C	0.500

q3.1 的指定答案: 0.208

q3.2 的指定答案: 0.500

q3.3 的指定答案: 0.292

q3.4 的指定答案: 0.500

q3.5 的指定答案: 0.500

q3.6 的指定答案: 0.267

q3.7 的指定答案: 0.733

q3.8 的指定答案: 0.400

q3.9 的指定答案: 0.600

q3.1 的正确答案:

评估方式	正确答案	区分大小写
完全匹配	0.208	

q3.2 的正确答案:

评估方式	正确答案	区分大小写
完全匹配	0.500	
完全匹配	0.5	

q3.3 的正确答案:		
评估方式	**正确答案**	**区分大小写**
完全匹配	0.292	
q3.4 的正确答案:		
评估方式	**正确答案**	**区分大小写**
完全匹配	0.5	
完全匹配	0.500	
q3.5 的正确答案:		
评估方式	**正确答案**	**区分大小写**
完全匹配	0.5	
完全匹配	0.500	
q3.6 的正确答案:		
评估方式	**正确答案**	**区分大小写**
完全匹配	0.267	
q3.7 的正确答案:		
评估方式	**正确答案**	**区分大小写**
完全匹配	0.733	
q3.8 的正确答案:		
评估方式	**正确答案**	**区分大小写**
完全匹配	0.4	
完全匹配	0.400	
q3.9 的正确答案:		
评估方式	**正确答案**	**区分大小写**
完全匹配	0.6	
完全匹配	0.600	

问题 4

得 14 分，满分 14 分

Now consider again the new data point $F_1 = 0, F_2 = 0, F_3 = 1$. Use the Laplace-Smoothed version of your classifier to determine the joint probability of causes Y and this new data point, along with the posterior probability of Y given the new data: (**Keep 3 decimal places**)

Y	$P(Y, F_1 = 0, F_2 = 0, F_3 = 1)$
A	**[q4.1]**

B	**[q4.2]**
C	**[q4.3]**

Y	$P(Y \mid F_1 = 0, F_2 = 0, F_3 = 1)$
A	**[q4.4]**
B	**[q4.5]**
C	**[q4.6]**

What label does your (Laplace-Smoothed) classifier give to the new data point? (Break ties alphabetically). Enter a single capital letter.

[q4.7]

q4.1 的指定答案: 0.033

q4.2 的指定答案: 0.036

q4.3 的指定答案: 0.023

q4.4 的指定答案: 0.359

q4.5 的指定答案: 0.391

q4.6 的指定答案: 0.250

q4.7 的指定答案: B

q4.1 的正确答案:		
评估方式	**正确答案**	**区分大小写**
完全匹配	0.033	
q4.2 的正确答案:		
评估方式	**正确答案**	**区分大小写**
完全匹配	0.036	
q4.3 的正确答案:		
评估方式	**正确答案**	**区分大小写**
完全匹配	0.023	
q4.4 的正确答案:		
评估方式	**正确答案**	**区分大小写**
完全匹配	0.359	
完全匹配	0.356	
q4.5 的正确答案:		
评估方式	**正确答案**	**区分大小写**
完全匹配	0.391	

完全匹配	0.389	
q4.6 的正确答案:		
评估方式	**正确答案**	**区分大小写**
完全匹配	0.25	
完全匹配	0.250	
完全匹配	0.255	
q4.7 的正确答案:		
评估方式	**正确答案**	**区分大小写**
完全匹配	B	

问题 5

得 10 分，满分 10 分

Select all correct statements.

所选答案:
- You have a lot of training data, but it is **very noisy and far from** the true data distribution (test data distribution), it is very likely to overfit the training data if you train a model to fit it.
- You have **too** few training samples (but close to the true data distribution), and you train a very expressive model to fit it, it is very likely to overfit.
- You have inadequate training data (and close to the true data distribution) , and you train an expressive model to fit it. Applying regularization techniques will make the model less likely to overfit the training data.

答案:
- You have a lot of training data, but it is **very noisy and far from** the true data distribution (test data distribution), it is very likely to overfit the training data if you train a model to fit it.
- You have **too** few training samples (but close to the true data distribution), and you train a very expressive model to fit it, it is very likely to overfit.
- You have inadequate training data (and close to the true data distribution) , and you train an expressive model to fit it. Applying regularization techniques will make the model less likely to overfit the training data.

问题 6

得 10 分，满分 10 分

In which model is Laplace smoothing commonly applied to solve the "zero probability" problem?

所选答案: Naive Bayes Classifier

答案: Linear Regression

Naive Bayes Classifier

K-Nearest Neighbors

Support Vector Machine

问题 7

得 10 分，满分 10 分

Which factors affect the position of the decision boundary of a linear classifier?

所选答案: The weight vector W

The bias term b

The correlation between features

答案: The weight vector W

The bias term b

The order of the training samples

The correlation between features

问题 8

得 10 分，满分 10 分

What are the characteristics of Mini-batch Gradient Descent?

所选答案: Uses a small subset of samples to compute the gradient in each iteration

Faster than Batch Gradient Descent and more suitable for large-scale datasets

答案: Uses a small subset of samples to compute the gradient in each iteration

Uses the entire dataset to compute the gradient in each iteration

Convergence speed is generally slower than Stochastic Gradient Descent

Faster than Batch Gradient Descent and more suitable for large-scale datasets

问题 9

得 10 分，满分 10 分

The K-means algorithm:

所选答案: None of the above

答案: Requires the dimension of the feature space to be no bigger than the number of samples

Has the smallest value of the objective function when K = 1

Converges to the global optimum if and only if the initial means are chosen as some of the samples themselves

None of the above

问题 10

得 0 分，满分 10 分

Which of the following statements about the EM algorithm is **incorrect**?

所选答案：
- ✗ In the E-step, the expected value of the posterior distribution is computed.
- ✓ The EM algorithm always converges to the global optimum.

答案：
The EM algorithm does not require gradient computation.

In the E-step, the expected value of the posterior distribution is computed.

In the M-step, the marginal likelihood function is maximized.

- ✓ The EM algorithm always converges to the global optimum.

2025年1月8日 星期三 下午03时40分08秒 CST

← 确定