Q1 Exam regulations

0 Points

BBM406 Honor Code

I promise that, for the BBM406 Midterm Exam

- All my exam work will be done entirely by myself, with no help from others;
- I will not communicate with anybody except the proctors during the exam;
- I will not consult any people or sources other than my printed/handwritten course notes, slides and the reference books listed in the course webpage;
- I will not provide any information about the exam's contents to other students until the exam deadline; and
- I will turn on my camera on Zoom session during the whole exam period.

Understanding this, I pledge my honor that I will not violate this Honor Code during the exam. I certify that all solutions will be entirely my own, that I will not consult people or sources other than those permitted, and that I will not share information with others during the exam.

Do NOT sign nor take this exam if you do not agree with this.

Signature (Specify your name and surname as your signature)

Mehmet Taha USTA MTUSTA

Q2 True/False Questions

20 Points

Question 1. True/False Questions [20 POINTS]

Please state whether the following statements are true or not. *To get points from these questions, you have to provide your reasonings for the answers you give and they should be correct.*

(a) [3 POINTS] Using linear regression, you can learn the maximum likelihood model parameters (α) for the model $y_i = \log(x_1^{\alpha_1} e^{\alpha_2}) + \epsilon_i$ where $\epsilon_i \sim N(0, \sigma^2)$ denotes iid noise.

(b) [3 POINTS] Using linear regression, you can learn the maximum likelihood model parameters (α) for the model $y_i = \log(\alpha_1 x_1 x_2^3) + \epsilon_i$ where $\epsilon_i \sim N(0, \sigma^2)$ denotes iid noise.

(c) [3 POINTS] You are assigned to solve a binary classification problem. You trained a classifier on the training data you collected and obtained a low training error. Claiming that you solved the problem could be problematic.

(d) [3 POINTS] You are assigned to solve a binary classification problem where one class is very rare (e.g., detecting fraud transactions). The dataset for the problem consists of 100 positive examples and 4.000 negative examples. Achieving 98% classification accuracy on this task can be considered as a success.

(e) [4 POINTS] A neural network with one hidden layer is less likely to overfit if it has fewer units in the hidden layer.

(f) [4 POINTS] You figure out that the data collected for the binary classification problem that you are interested in are obtained via some noisy sensors. Hence, you perform a feature selection procedure on the full data and reduced the large feature set to a smaller set. As usual, you then split the data into two, as training and test splits, and train several models on the training split using different hyperparameter settings. Finally, you show the best test error that they achieved to your friends just to impress them. They reply that these results are problematic.

Please attach your solution to Question 1 (True/False Questions) as a pdf file using the following link:

IMG_20210421_094908.jpg Download

Q3 Naïve Bayes Classifier

20 Points

Question 2. Naïve Bayes Classifier [20 POINTS]

Suppose you are asked to solve a binary classification problem where each training sample has binary features which are denoted by X_1, $X_2 \in \{0,1\}$, respectively. Assume that these features, X_1 and X_2, are conditionally independent given the class label Y, the class priors are defined as P(Y=0)=P(Y=1)=0.5, and the conditional probabilities are given as follows:

$P(X_1 \mid Y)$	X_1=0	X_1=1
Y=0	0.6	0.4
Y=1	0.3	0.7

$P(X_2 \mid Y)$	X_2=0	X_2=1
Y=0	0.9	0.1
Y=1	0.4	0.6

(a) [3 POINTS] Please fill in the following table and specify the predicted class labels for all the four possible configurations of X_1 and X_2.

X_1	X_2	$P(X_1,X_2,Y=0)$	$P(X_1,X_2,Y=1)$	Predicted class label
0	0			
0	1			
1	0			
1	1			

One can define the expected error rate as the probability which a classifier gives a misprediction for an observation. While estimating the error rates asked in the following questions, you can assume that you have infinite training data used to train the classifiers.

(b) [3 POINTS] Please estimate the expected error rate of the naïve Bayes classifier that predicts the class labels Y given X_1 only. Similarly, estimate the expected error rate of the naïve Bayes classifier that predicts the class labels Y given X_2 only. Which one of the classifiers you prefer considering these error rates?

(c) [4 POINTS] Now, estimate the expected error rate of the naïve Bayes classifier that predicts the class labels Y given both of the features $\{X_1, X_2\}$.

(d) [4 POINTS] Suppose that we introduce a new feature X_3, which is a direct copy of X_1. What will be the new expected error rate of the naïve Bayes classifier that predicts Y given all these attributes $\{X_1, X_2, X_3\}$?

(e) [6 POINTS] Explain the reason behind the different behaviors of the naïve Bayes classifiers in part (c) and part (d). What happens if you train a logistic regression model, do you observe a similar issue? If not, why?

Please attach your solution to Question 2 (Naïve Bayes Classifier) as a pdf file using the following link:

IMG_20210421_094927.jpg Download

Midterm Exam - Part 1

GRADED

STUDENT
MEHMET TAHA USTA

TOTAL POINTS
8 / 40 pts

QUESTION 1
Exam regulations **0** / 0 pts

QUESTION 2
True/False Questions **6** / 20 pts

QUESTION 3
Naïve Bayes Classifier **2** / 20 pts