

Sample Midterm 1

Due: March 20th, 2018

There are six problems, choose one to count as extra credit by circling its number. The extra credit problem is worth $1/10$ of a problem.

You may use a calculator, but no notes or web search.

Do not turn this page until we give the signal.

Name: _____

Problem 1 (algorithms)

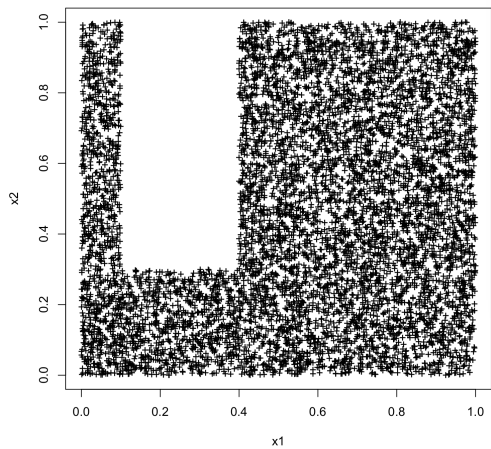
For this problem, we are looking for responses that both indicate your assessment as to a possible accuracy change and your understanding of the algorithm that led to this assessment. Answers should be two or three sentences long and focus on the relevant and important issue.

a. We have trained Naive Bayes (binary vector input, binary label) on a data set. Then, we create a new data set that is identical to the original but it duplicates one of the existing attributes and retrain. What would you expect to happen to the testing performance of the algorithm (assuming the new feature is available)?

b. We have trained 1-Nearest Neighbor (continuous vector input, binary label) on a data set. Then, we create a new data set that is identical to the original but includes a new attribute that is identical to the label and retrain. What would you expect to happen to the testing performance of the algorithm (assuming the new feature is available)?

Problem 2 (classifiers)

We trained a decision tree on a data set with a two-dimensional continuous input space. Here is a sample of the points the resulting classifier assigns positive labels to. Draw the corresponding decision tree. Make sure to label all branches, nodes, and leaves.



Problem 3 (representation)

a. In a two-bit input space, a binary Naïve Bayes classifier can be represented by b , the prior probability that the output class is 1 ($\Pr(y = 1)$), and b^{iy} , the probability that an example with output class y has bit i set to 1 ($\Pr(x_i = 1|y)$). Define b , b^{10} , b^{11} , b^{20} , and b^{21} so the resulting classifier represents a conjunction—it assigns an output of 0 to inputs 00, 01, and 10, and an output of 1 to 11.

b. Will the following training set produce the correct classifier? For smoothing purposes, any 0/1 probability is replaced by 0.1/0.9.

x_1	x_2	y	x_1	x_2	y
0	0	0	1	0	0
0	1	0	1	0	0
1	0	0	0	1	0
1	1	1	0	1	0
1	0	0	0	1	0

Problem 4 (VC dimension)

Consider the hypothesis class of decision trees of depth at most d that categorize continuous inputs of the form $a \in \mathbb{R}$ to one of two binary classes $\{0, 1\}$.

What is the VC dimension of this hypothesis class? Express in terms of d . Provide a complete proof.

Problem 5 (loss)

Here's how Wikipedia defines the **Sleeping Beauty Problem**:

Sleeping Beauty volunteers to undergo the following experiment and is told all of the following details: On Sunday she will be put to sleep. Once or twice, during the experiment, Beauty will be awakened, interviewed, and put back to sleep with an amnesia-inducing drug that makes her forget that awakening. A fair coin will be tossed to determine which experimental procedure to undertake:

- *If the coin comes up heads, Beauty will be awakened and interviewed on Monday only.*
- *If the coin comes up tails, she will be awakened and interviewed on Monday and Tuesday.*

In either case, she will be awakened on Wednesday without interview and the experiment ends.

Any time Sleeping Beauty is awakened and interviewed she will not be able to tell which day it is or whether she has been awakened before. During the interview Beauty is asked: "What is your credence now for the proposition that the coin landed heads?"

We will attack this problem by defining a loss function and asking Sleeping Beauty to declare her "credence" the coin landed heads as by reporting the value that minimizes the loss. Note that she can't distinguish the different awakenings, so she must report one value, which we'll call p , no matter when she is asked to report her credence.

a. Each time she is awakened, her loss is increased by p if the coin is tails and $1 - p$ if the coin is heads. What should she report to minimize her loss and why?

b. Each time she is awakened, her loss is increased by p^2 if the coin is tails and $(1 - p)^2$ if the coin is heads. What should she report to minimize her loss and why?

Problem 6 (optimizers)

Consider (again) the class of signed intervals, $H = \{h_{a,b,s} : a \leq b, s \in \{-1, 1\}\}$ where:

$$h_{a,b,s} = \begin{cases} s & \text{if } x \in [a, b] \\ -s & \text{if } x \notin [a, b] \end{cases}$$

a. Explain how to compute the ERM for this class in the realizable case. State the computational complexity of the algorithm in the context of a data set of size m .

b. Explain how to compute the ERM for this class in the agnostic case. State the computational complexity of the algorithm in the context of a data set of size m .