Exam 1

Started: Oct 21 at 5:56pm

Quiz Instructions

This Exam is Open Resources, but Individual.

- You may use any course materials (slides, textbooks, Piazza in read-only mode, etc.) at your disposal
  - You may not consult with other humans, enrolled in this course or not
- You may use electronic resources (calculators, python, etc.) to perform operations
  - All operations on this exam should be completable with a 4-function calculator and scratch paper

You will submit your “proof of work” as an external file to <Link Here> on Canvas to be considered for partial credit.

- This can be a PDF, DOC, PPT, Image of scratch paper, etc.

The exam will not give feedback on results until the conclusion of everyone's exam time. In addition, the Canvas grading tool can only effectively mark "correct vs. incorrect" and will not be able to initially factor in partial credit in many cases.

- This exam is scored out of 100 points. The points in the Canvas Quiz "activity" may not accurately reflect the ground truth of points given to each problem. Trust the [X Points] given to each problem in the text for accurate representations of the breakdown of point values.
- Your initial score and feedback on the full exam should be considered a lower bound before the application of partial credit done via manual grading.
- You can submit any numeric answers as reduced fractions or as decimals with at least two digits past the decimal if rounded.

Question 1 — 0 pts

[25 Points, 5 Points each] Match the following definitions to their terms. Each term will be used only once, but not all terms will be used.

The inability of the model to capture the complexity of the underlying phenomenon, relying instead upon the constraints of the model itself.	[Choose]
The inability of the model to generalize from train to test, relying instead on the idiosyncrasies of the training data.	[Choose]
Reducing a numeric or categorical variable to a smaller set of options to reduce complexity.	[Choose]
Transforming a single feature into a series of binary features.	[Choose]
A tool to visualize the performance of a classifier.	[Choose]

Question 2 — 0 pts

[5 Points] Which of the following methods is an ensemble classification technique that iteratively creates weak classifiers and weights each individual classifier’s importance according to how accurately it predicts the outcomes, then increases the importance of difficult-to-classify points on the following iteration?

- ○ Decision Trees
- ○ AdaBoost
- ○ Support Vector Machines
- ○ Random Forests

Question 3 — 0 pts

[5 Points] In a linear regression problem with 5 features, we test out two types of feature regularization to penalize our weights. For the first regularization, we get w_1 = [0.5, 1.3, 0.002, 3.9, 0.01]. For the second regularization, we get w_2 = [0.48, 0.8, 0, 4, 0]. Which of the regularizations is more likely to be using Lasso?:

- ○ Second regularization generating w_2
- ○ First regularization generating w_1

Question 4 — 0 pts

[5 Points] Name one benefit of Boosting over Bagging.

p

0 words

Consider the following confusion matrix for a trained and tested classifier:

	Predicted as	**True**	**False**		*Row Total*
Ground Truth					
True		40	20		*60*
False		10	30		*40*
Column Total		**50**	**50**		***100***

Question 5

0 pts

[5 Points] What is the Accuracy of your classifier?

Question 6

0 pts

[2.5 Points] Are there more False Positives or False Negatives in your results?

- ○ False Negatives
- ○ False Positives

Question 7

0 pts

[5 Points] What is the best Baseline Accuracy we could generate for this classifier?

You are working with the following series of 5 samples as a Training Set for a K-Nearest-Neighbors Classifier of odd human measurements to classify CU Alum status.

Name (SampleID)	X_1: Vertical Jump (in)	X_2: Hair Length (mm)	X_3: Pupillary Distance (mm)	Y: CU Alum
Alice	15	600	60	No
Brandin	15	250	55	Yes
Callie	20	10	60	No
David	60	100	70	Yes
Elanor	10	90	75	Yes

Consider the new incoming Test case below.

Name (SampleID)	X_1: Vertical Jump (in)	X_2: Hair Length (mm)	X_3: Pupillary Distance (mm)	Y: CU Alum
Fahmid	20	200	70	???

Question 8 0 pts

[5 points] Given *no adjustments* (i.e. no scaling or normalization), which sample is the closest neighbor to Fahmid? (use the Manhattan distance for ease of calculation)

- ○ Callie
- ○ David
- ○ Brandin
- ○ Elanor
- ○ Alice

Question 9 0 pts

[5 Points] Using Min-Max Scaling on the Training Data, what is your scaled value for Fahmid's Hair Length?

You are developing a Decision Tree classifier to determine whether a user has watched "Machine Learning Adventures" by Dr. Quigley based on binary {Yes, No} survey responses (features) within your online movie distribution platform.

X_1: Likes Action	**X_2: Likes Educational**	**Y: Watched "MLA"**
Yes	No	No
Yes	Yes	Yes
Yes	No	No
No	Yes	No
Yes	No	Yes

Use the Misclassification Error = min(p, 1-p) as your measure of Impurity

Question 10 0 pts

[5 Points] What is the Information Gain for a root node split on X_1: Likes Action?

Question 11 0 pts

[5 Points] What is the Information Gain for a root node split on X_2: Likes Educational?

Question 12 0 pts

[2.5 Points] Which feature is a better split for our root node?

- ○ X_1: Likes Action
- ○ X_2: Likes Educational

Question 13 1 pts

[2.5 Points] If we were to create a full depth tree, can we ever get 100% accuracy on our training set (i.e. will every leaf have 0 Impurity)?

- ○ No
- ○ Yes

You are developing a Naive Bayes classifier to predict whether a racecar driver has won a race so far this year (Winner = {Yes, No}) based on the results from driver skill tests (Skills = {Low, Medium, High}, Reflexes = {Fast, Slow}). The training set consists of 5 samples:

Driver #	X_1: Skills Results	X_2: Reflexes Results	Y: Winner Status
1	High	Fast	Yes
2	Medium	Slow	No
3	Medium	Fast	No
4	Medium	Slow	Yes
5	Medium	Fast	No

Consider the new incoming driver test case below. *Note: We are not incorporating any smoothing or other adjustments into the main problem.*

Driver #	X_1: Skills Results	X_2: Reflexes Results	Y: Winner Status
6	Medium	Slow	???

Question 14 0 pts

[5 Points] What is the Class-Conditional of Winner = Yes for our Driver 6?

Question 15 0 pts

[5 Points] What is the Class-Conditional of Winner = No for our Driver 6?

Question 16 0 pts

[5 Points] What is the Prior Probability of Winner = Yes for our Driver 6?

Question 17 0 pts

[5 Points] What is the Prior Probability of Winner = No for our Driver 6?

Question 18 0 pts

[2.5 Points] What classification would your Naive Bayes Classifier give for Driver 6?

- ○ Winner = No
- ○ Winner = Yes

Question 19 0 pts

[BONUS 10 Points] What is the Class-Conditional Probability of Winner = Yes for our Driver 6 if we incorporate Laplace (i.e. add-1) Smoothing?

Not saved Submit Quiz