Concept Quiz Over Week 3 Material

Due Oct 22 at 11:59pm **Points** 1 **Questions** 9
Available Oct 18 at 12am - Oct 22 at 11:59pm **Time Limit** None

Score for this survey: **1** out of 1
Submitted Oct 22 at 7:51pm
This attempt took 1,630 minutes.

Question 1

Match the algorithm with its description.

ou Answered	**Logistic Regression**	Linear classifier with a (
ou Answered	**Perceptron**	Linear classifier with nc
ou Answered	**k Nearest Neighbor**	Exemplar classifier witl
ou Answered	**Linear Regression**	Linear regression mode

Logistic regression and the perceptron are both linear classifiers, but the perceptron has no probabilistic interpretation. k Nearest Neighbors can form arbitrary decision boundaries based on the examples it uses to compute neighbors, but lacks any direct probabilistic interpretation. Linear regression is obviously a linear regression model and can be interpreted as a linear model with Gaussian noise.

Question 2

The standard perceptron learning algorithm is guaranteed to converge even for non-linearly separable data.

True

ou Answered

False

False. It will continue to oscillate as zero error will never be reached.

Question 3

Match the term to its definition.

ou Answered **False Positive** Negative example that

ou Answered **True Positive** Positive example that c

ou Answered **False Negative** Positive example that c

ou Answered **True Negative** Negative example that

False positive -> Negative example that our model predicts as positive

True positive -> Positive example that our models predicts as positive

False negative -> Positive example that our model predicts as negative

True Negative -> Negative example that our model predicts as negative

Question 4

What are recall and precision?

Your Answer:

Recall is the fraction of my model predicted as positive

Precision is Of things your model predicts as positives, what fraction are correct?

Recall is the fraction of positive examples correctly predicted as positive. Precision is the fraction of positive predictions that are actually positive examples.

Question 5

Bayes error is the error due to ______________.

- training a model from a finite dataset.
- the difficulty of optimization.

ou Answered

inherit uncertainty in the problem.

our model not being expressive enough.

inherit uncertainty in the problem.

If two examples look identical except they have different outputs, no model can get them both correct.

Question 6

Both discriminative and generative classifiers make decisions according to P(y|x) (where y is the output and x is the input). However, discriminative model learn _________, while generative models learn _____________.

ou Answered

P(y|x), P(y|x) & P(y)

P(y|x) directly; P(x|y) and P(y) and apply Bayes Theorem

Generative models make use of Bayes Theorem and learn make decisions according to argmax_y P(x|y)P(y)

Question 7

Given random variables X, Y, and Z -- what does it mean for X and Y to be conditionally independent given Z?

(Hint: Review the height/vocabulary/age example in lecture)

Your Answer:

P(X,Y|Z) = P(X)P(Y|Z)

P(X,Y|Z) = P(X|Z)

P(Y,X|Z) = P(Y|Z)

Conditioned on the value of Z, X and Y do not provide additional information about each other. That is to say P(X,Y|Z) =P(X|Z)P(Y|Z).

Question 8

The Naive Bayes assumption is that __________.

ou Answered

- the input features are conditionally independent given the class label.
- the input features are distributed according to a Normal distribution.
- the input features are independent.
- that P(y|x) is proportional to P(x|y)P(y).

the input features are conditionally independent given the class label.

Question 9

The Naïve Bayes assumption generally reduces the number of parameters we have to learn in our model. This is because we don't need to learn the full conditional joint distribution $P(x_1, x_2, \ldots, x_d|y)$ and can instead learn conditionals for individual input dimensions $P(x_i|y) \quad \forall i$.

ou Answered

- True

False

True! The conditional independence assumption in Naïve Bayes lets us assume $P(x_1, x_2, \ldots, x_d|y) = P(x_1|y) * P(x_2|y) * \ldots * P(x_d|y)$ so we can then just learn the individual conditional (right) rather than the full conditional joint (left).

Survey Score: **1** out of 1