Multiple-choice and short-answer questions

In this section you are asked to demonstrate your conceptual understanding of methods studied in this subject, your ability to apply them or evaluate them in the context of specific cases, and your ability to perform numeric calculations over a given set of instances.

Classify the following attributes as discrete, ordinal, or continuous:

1) Number of patients in a hospital
2) Ability to pass light, in terms of the following values: opaque, translucent and transparent

ANSWER:

Which of these is an example of a regression problem?

A. Predict which product a customer will buy, based on their search history
B. Predict whether or not a company will fail in the next year
C. Predict the age of a person from a photograph of their face
D. Predict the next word in a sentence to generate artificial text

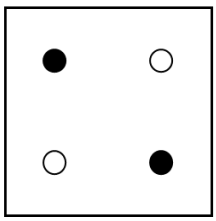
ANSWER:

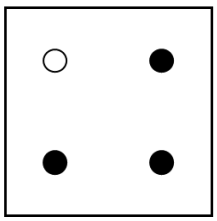
Which of these is an example of a “greedy” algorithm?

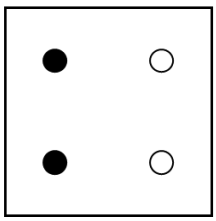
A. Support vector machine
B. Naïve Bayes
C. Multilayer perceptron
D. Expectation-maximization

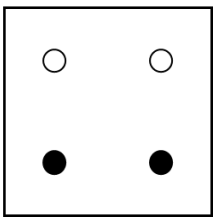
ANSWER:

Which 2D dataset is not linearly separable? The black and white dots represent instances of two different classes.

a.) 

b.) 

c.) 

d.) 

ANSWER:

An advantage of “bagging” is that it can be parallelized. How and why? (Answer in 1-2 sentences)

ANSWER:

For a “logistic regression” model trained over *n* classes, based on *k* training instances and *f* features, and tested over *t* test instances, use big-O notation to describe the approximate number of parameters in the trained model.

ANSWER:

What is the primary difference between stratified cross validation and non-stratified cross validation? (Answer in 1 sentence)

ANSWER:

What is “sum of squared errors”? Explain it in the context of a model discussed in this subject. (Answer in 1-2 sentences)

ANSWER:

Consider the dataset below:

A	B	C	Class
0	0	0	+
0	0	1	−
0	1	1	−
0	1	1	−
0	0	1	+

A	B	C	Class
1	0	1	+
1	0	1	−
1	0	1	−
1	1	1	+
1	0	1	+

1.) Estimate the conditional probabilities for P(A=1|+), P(B=1|+), and P(C=1|-) without smoothing.

ANSWER:

2.) Predict the class label for (A=0, B=1, C=0) using the given training dataset and non-smoothed probability estimates, based on the naïve Bayes approach.

ANSWER:

Consider the dataset below:

A	B	C	Class
0	0	0	+
0	0	1	−
0	1	1	−
0	1	1	−
0	0	1	+

A	B	C	Class
1	0	1	+
1	0	1	−
1	0	1	−
1	1	1	+
1	0	1	+

3.) Estimate the conditional probabilities for P(A=1|+), P(B=1|+), and P(C=1|-) using Laplacian smoothing.

ANSWER:

4.) Predict the class label for (A=0, B=1, C=0) using the given training dataset and smoothed probability estimates using Laplacian smoothing, based on the naïve Bayes approach.

ANSWER:

5.) Explain the observed differences between the predictions with and without Laplacian smoothing. (Answer in 2-3 sentences)

ANSWER: