Test 1 (Remotely Proctored)

Due Sep 30 at 12:15pm   **Points** 105   **Questions** 18

Available after Sep 30 at 11am   **Time Limit** 75 Minutes

Exam Requirements A webcam and microphone are required to take this exam.

Exam Security This exam will be monitored by **Proctorio**. **Learn more about Proctorio (https://proctorio.zendesk.com/hc/articles/200972514)** .

Instructions

Note some questions can have multiple answers.

If we have k classes the information from them will be -p1 log p1 - p2 log p2 - .... - pk log pk

where pi is the fraction of examples belonging to the ith class.

The logs are all base 2 by default so here log is implicitly $\log_2$

Attempt History

	Attempt	**Time**	**Score**
LATEST	**Attempt 1**	27 minutes	95 out of 105

> ⚠ Correct answers are hidden.

Score for this quiz: **95** out of 105

Submitted Sep 30 at 11:49am

This attempt took 27 minutes.

Question 1   **5 / 5 pts**

A regression tree may be realized as

○ a given regression function implemented in a tree

○ a decision tree with a nominal function at each leaf

◉ a decision tree with a value at each leaf

Question 2   5 / 5 pts

A decision tree will, typically, use a numeric attribute in a test

○ creating one branch for each value

◉ by picking a number and having everything less than or equal to it in one branch and everything greater than it in another branch

○ creating one branch for each value with a different class

○ none of these

Question 3   5 / 5 pts

Association rules rely on support and confidence which

○ provide only useful rules

○ are random variables

○ none of these

◉ limit the search space

Question 4   5 / 5 pts

For a nominal attribute used as a test when building a decision tree, each value usually causes a branch to be created which means

○ the nominal attribute must appear at least 2 times in the tree

○ none of these

○ you cannot have numeric attributes

◉ the nominal attribute will only be used in one test in the tree

Question 5   5 / 5 pts

The no free lunch theorem can be interpreted to mean

○
until an algorithm is shown better than another you should use it for data mining

○ decision trees are the most explainable model

○
you should always use deep neural networks for a problem because they perform the best

◉

for a given data set, we cannot tell beforehand what the best learning algorithm will be

---

Question 6

5 / 5 pts

If there are no labels on a data set what types of data mining algorithms can be applied?

☑ association rule learning

☑ clustering

☐ regression algorithms

☐ decision trees

Question 7

5 / 5 pts

The search space during training is largest for

☑ rule learners

☐ instance based learning

☐ decision trees

☐ naive bayes

Question 8   5 / 5 pts

The Laplace transform

- ◉ is used in Naive Bayes to make sure no nominal attribute has 0 occurrences for a class

- ○ for nearest neighbors

- ○ Naive bayes to get perfect probabilities

- ○ Decision trees with missing values

Question 9   5 / 5 pts

Using weather.nominal.train20 below what attribute is chosen by One-R

@relation weather.symbolic

@attribute outlook {sunny, overcast, rainy}
@attribute temperature {hot, mild, cool}
@attribute humidity {high, normal}
@attribute windy {TRUE, FALSE}
@attribute play {yes, no}

@data
sunny,   hot, high,   FALSE,no
overcast,hot, high,   FALSE,yes
rainy,   cool, normal,FALSE,yes
rainy,   cool,normal,TRUE, no
overcast,cool,normal,TRUE, yes
sunny,   mild,high,   FALSE, no
rainy,   mild,normal,FALSE,yes

sunny,   mild,normal,TRUE, yes
overcast,hot,normal,FALSE,yes
rainy,   mild,high,   TRUE,   yes

⦿

outlook -> Sunny (2 no, 1 Yes), overcast (0 no, 3yes), rainy (1 no, 3 yes) for error rate of 2/10

◯   one of temperature, humidity and windy is the best (equivalent) choice

---

◯   windy -> FALSE (2 no, 4 yes), TRUE (1 no, 3 yes), for error rate of 3/10

◯

humidity -> high (2 no, 2 yes), normal (1 no, 5 yes) -for error rate of 3/10

◯

temperature -> (1 no, 2 yes) hot, (1 no, 2 yes) cool, (1 no, 3 yes) mild for error rate of 3/ 10

Question 10   10 / 10 pts

RIPPER uses what to choose a test to add

⦿   to a rule, foil gain

◯   to add to a rule, p-n

◯   to a rule, information gain

◯   to a tree, foil gain

Incorrect

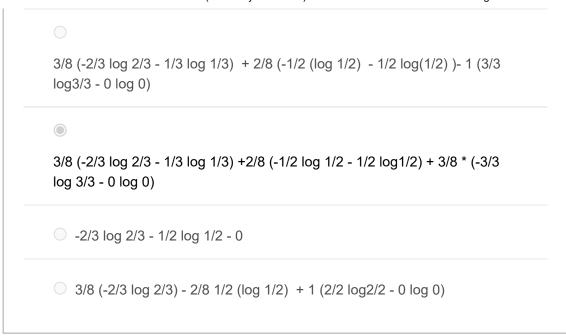## Question 11   0 / 5 pts

In association rule learning, to form a triplet (A B C) with the appropriate support you must have

⦿ at least AB and AC with the appropriate support

◯ high confidence

◯ lots of examples

◯ all pairs with the appropriate support value

Question 12   10 / 10 pts

What is the information for the test height ?

		(Class)
Height	Weight	Eating Style
Tall	normal	big
Short	heavy	big
Short	light	moderate
medium	normal	moderate
Short	heavy	moderate
Tall	light	big
medium	light	big
Tall	heavy	big

Incorrect

○

3/8 (-2/3 log 2/3 - 1/3 log 1/3) + 2/8 (-1/2 (log 1/2) - 1/2 log(1/2))- 1 (3/3 log3/3 - 0 log 0)

◉

3/8 (-2/3 log 2/3 - 1/3 log 1/3) +2/8 (-1/2 log 1/2 - 1/2 log1/2) + 3/8 * (-3/3 log 3/3 - 0 log 0)

○ -2/3 log 2/3 - 1/2 log 1/2 - 0

○ 3/8 (-2/3 log 2/3) - 2/8 1/2 (log 1/2) + 1 (2/2 log2/2 - 0 log 0)

Question 13   5 / 5 pts

Naive Bayes with 1000 attributes/features

○ have a problem with Lagrange multipliers

○ has to use laplace correction for all features

◉ will tend to have 0 or near zero probabilities

○ will work fine if there are lots of examples

Question 14   5 / 5 pts

Data mining could be defined as

○ indescribable

⦿ searching through the space of models for the best one

○ learning like a robot with reinforcement learning

○ randomly choosing models

Question 15   5 / 5 pts

RIPPER does pruning

⦿ to prevent overfitting by having more general rules

○ to get small regression models

○ to have the most specific rules possible

○ to be faster to learn

Incorrect

Question 16   0 / 5 pts

Regression solutions have been adapted to classification

⦿ only by using two classes 0 and 1 as values for regression examples

○ called logistic regression

○ never

○ via regression trees

Question 17   10 / 10 pts

With instance based learning (nearest neighbors)

○ training time is equal to testing time

◉ training time is low, but test time large for a big training set

○ test time is very slow for small data

○ training time is much slower than for RIPPER

Question 18   5 / 5 pts

What is the first test chosen for a rule for the class moderate and its Foil Gain? (equation Foil Gain = p1 (log2(p1/(p1+n1)) - log (p0/(p0+n0))

Height	Weight	(Class) Eating Style
Tall	normal	big
Short	heavy	big
Short	light	moderate
medium	normal	moderate
Short	heavy	moderate
Tall	light	big
medium	light	big
Tall	heavy	big

○ Height= Tall -> 0 (log2 0/3 - log 3/8)

○ Height = medium -> 1 (log2 1/2 - log2 3/8)

◉ Height=short -> 2 (log2 2/3 - log2 3/8)

○ Weight=heavy -> 1 (log2 1/3 - log 3/8)

Quiz Score: **95** out of 105