

survival8

Data Science Course

Learn Data Science in a Classroom. Get placed in top companies with Placement Assistance.

Great Learning Chennai

BITS WILP Data Mining Mid-Sem Exam 2017-H2 (Make Up)

[Index](#) [Subjects ▼](#) [Mail Us](#)

BITS WILP Data Mining Mid-Sem Exam 2017-H2 (Make Up)
 BITS WILP Database Design And Applications Mid-Sem 2017-H1 (Regular)
 Work Integrated Learning Programmes Division
 First Semester 2017-18
 Mid Semester Test
 EC-2 Make Up
 Course No: IS ZG415
 Course Title: Data Mining
 Date of exam: 14/Oct/2017 (AN)

Pages

- [Postings Index](#)
- [Index of BITS WILP Exam Papers and Content](#)
- [Index of Lessons in Technology](#)
- [Index of Guest Interviews](#)
- [Downloads](#)
- [Book Requests](#)

Blog Archive

- ▼ [2020](#) (31)
 - ▼ [May](#) (1)
 - [Covid-19 and response of IT companies \(by Divjot S...](#)
 - ▶ [April](#) (6)
 - ▶ [March](#) (12)
 - ▶ [February](#) (6)
 - ▶ [January](#) (6)
- ▶ [2019](#) (48)
- ▶ [2018](#) (31)
- ▶ [2017](#) (15)
- ▶ [2016](#) (6)

Popular Posts



You Are a Badass. How to stop doubting your greatness and start living an awesome life (Jen Sincero, 2013)

INTRODUCTION The language used in the book extremely funny and Jen Sincero still makes sure that she m...

[Covid-19 and response of IT companies \(by Divjot Singh\)](#)

As the Covid-19 pandemic ravages the world, many domains like airlines, tourism and services...

[Innovation to beat the Coronavirus \(Covid19\)](#)

Coronavirus' Exponential growth and decline In the first phase of the pandemic, we saw a...

[Download fiction books \(March 2018\)](#)
 Download fiction books for free: Link for Google Dr...

[Life Lessons By Steve Jobs](#)

Birla Institute of Technology & Science, Pilani
Work-Integrated Learning Programmes Division
First Semester 2017-2018

Mid-Semester Test
(EC-2 Make-up)

Course No. : IS ZC415
 Course Title : DATA MINING
 Nature of Exam : Closed Book
 Weightage : 30%
 Duration : 2 Hours
 Date of Exam : 14/10/2017 (AN)

No. of Pages = 2
 No. of Questions = 5

Note:

- Please follow all the *Instructions to Candidates* given on the cover page of the answer book.
- All parts of a question should be answered consecutively. Each answer should start from a fresh page.
- Assumptions made if any, should be stated clearly at the beginning of your answer.

Q.1. The following data is of diameter of a random sample of 11 trees.
 11.4, 13.3, 18.3, 2.7, 26.0, 43.6, 32.6, 52.2, 47.2, 69.3, 2.3

- (a) Find five-number summary and Inter Quartile Range (IQR). [3]
 (b) Draw boxplot for the observations. Show all labels and outliers, if any. [3]

Q.2. You are given 10 training samples. They are divided into four classes: CLASS-A, CLASS-B, CLASS-C, and CLASS-D. One sample belongs to CLASS-A, two belong to CLASS-B, three belong to CLASS-C, and four to CLASS-D. Use the following log₂ table to answer the questions.

p	log(p)
0.1	-3.32
0.2	-2.32
0.3	-1.74
0.4	-1.32
0.5	-1.00
0.6	-0.74

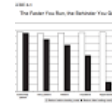
- (a) What is the total entropy of the samples? [2]
 (b) What is the Gini index of the samples? [2]
- Q.3. Imagine that you have created a large and complex decision tree from a training dataset that contain many attributes.
- (a) Your tree has a high accuracy on the training set, but the accuracy on the test set is very low. Explain why the accuracy on the test data could be so much lower than on the training data. [3]
 (b) You apply pruning methods and create a series of successively smaller subtrees from the full decision tree. As you continue to prune more, the test accuracy goes up in the beginning, but then starts to drop again. Explain the increase and then decrease of test accuracy in terms of specific terminology and concepts. [3]

Steve Jobs' last words will change your views on life. The billionaire passed away at the ...

Effects of news and world events on Nifty50 and stock market
 Day: 10th Aug 2017
 Sensex tanks 267 points. Nifty hits one-month low. 1. Market outlook: ...



Why Bill Gates would raise chickens
 I'm excited about the poverty-fighting power of poultry. If you were living on \$2 a day, wh...



Intelligent investor (Ben Graham & Jason Zweig, 4e)
 Reading from "A Note About Benjamin Graham by Jason Zweig" Here are Graham...



The Essays Of Warren Buffett (Lessons For Corporate America)
INTRODUCTION Buffett has applied the traditional principles as chief executive officer of Berkshi...

How To Talk TO Anyone (92 Little Tricks For Big Success In Relationships, by Leil Lowndes) - Book Summary
 There are two kinds of people in this life: Those who walk into a room and say, "Well, here I..."



max bupa health insurance

Know More

*T&C Apply



About Me



Ashish Jain
 View my complete profile

- Q.4. (a) The table below is given to you as training dataset where each sample has a labelled category either + or -. Each feature can take one of three nominal values: A, B, or C.

F1	F2	F3	Category
A	C	A	+
C	A	C	+
A	A	C	-
B	C	A	-
C	C	B	-

Use Naïve Bayes classifier to classify the following test example into proper category:

F1 = A, F2 = C, F3 = B

[4]

- (b) Given the following data set, *Electricity use* is the class to be predicted. Would Naive Bayes be effective for developing a classifier for the following data set? Why or why not?

Temp	Season	Electricity use
Below average	Winter	High
Above average	Winter	Low
Below average	Summer	Low
Above average	Summer	High

[2]

- Q.5 (a) After mining a transaction database for frequent itemsets, there is only one frequent itemset of size 10. Let N be the total number of frequent itemsets of all sizes (including the one of size 10). What is the minimal value of N?

[3]

- Q.5 (b) Consider the transaction database below:

TID	Items
1	1, 1, 3, 6
2	1, 2, 3, 6, 6
3	1, 4, 5, 5
4	2, 3, 3, 6
5	1, 3, 5
6	2, 3, 4, 5, 6
7	1, 2, 2, 2
8	1, 2, 4, 6

Assuming that minimum support is 50%, find all frequent itemsets by using apriori algorithm.

[5]

No comments:

Post a Comment

Enter your comment...



Comment as:

Narendran (Go ▼)

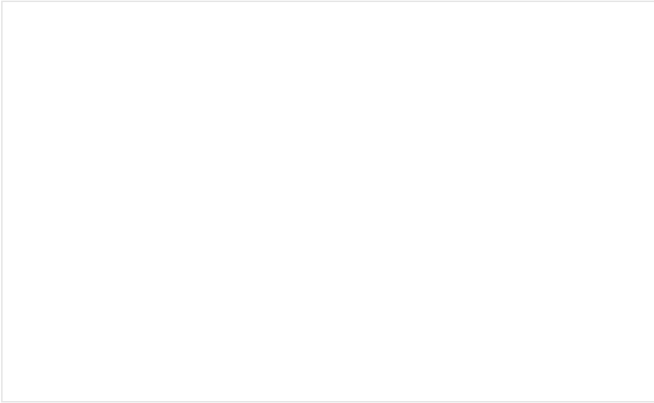
Sign out

Publish

Preview

☐ Notify me


Subscribe to: [Posts \(Atom\)](#)



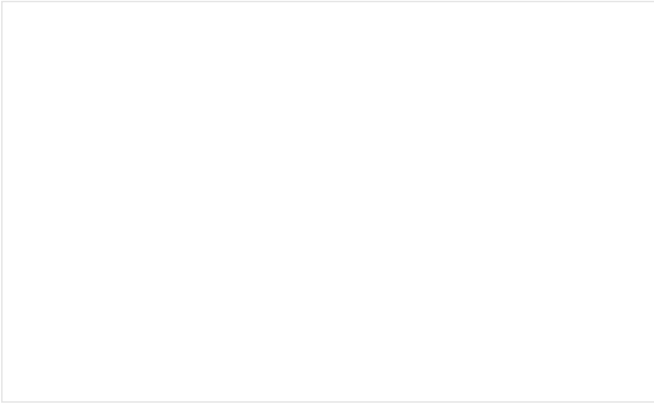
Data Science Course

Learn Data Science in a Classroom. Get placed in top companies with Placement Assistance.

greatlearning.in



Followers



Aakash Digital - JEE

Test Prep. Anywhere

Real time learning with immediate doubt clarification from qualified faculties.

digital.aakash.ac.in

OPEN

