

survival8

## Data Science Course

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

Great Learning Chennai

### BITS WILP Advanced Data Mining Mid-Sem Exam (Regular) 2017-H2

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

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

#### Mid-Semester Test (EC-2 Regular)

Course No. : SS ZG548  
 Course Title : ADVANCED DATA MINING  
 Nature of Exam : Closed Book  
 Weightage : 35%  
 Duration : 2 Hours  
 Date of Exam : 24/09/2017 (FN)  
 No of pages: 2  
 No of questions: 4  
 Note:

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

Q1. Explain the following concepts with respect to data mining. [2 + 2 + 2 = 6]

- a. Difference between error and noise in data
- b. Predictive and Descriptive data mining tasks with example
- c. Overfitting and regularization.

Q2. Describe incremental association mining setting. How fast update algorithm (FUP) differs from fast update 2 algorithm (FUP2). Explain how one can maintain certain data structures required by FUP2 in an incremental way. [1 + 1 + 2 = 4]

Q3. Consider association rule mining for incremental databases. What are the main advantages of using Compact Pattern Stream tree (CPS Tree) over FP-Tree. Assuming a pan contains two transactions (that is restructuring happens after arrival of two transactions). And each window to contain two pans. Consider step by step arrival of following six transactions

1. A, B, C, D, E
2. B, C, E, F
3. E, C, D, F
4. B, F, C, A
5. D, E, C
6. F, B, D

#### 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**  
 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...

Starting from a null CPS tree, draw all intermediate CPS tree on incremental arrival of every transaction. Clearly explain modifications to the CPS tree in each step.  $[2 + 8 = 10]$

Q4. Consider the problem of clustering an evolving database. Assume database updates happens at a regular interval. Let initially there be 15 datum points P1, P2, ..., P15 spread in 2D space. Distances between every pair of datum point is provided in Table-1. A Density-Based Spatial Clustering algorithm DBSCAN is applied with parameters **Eps=31** and **MinPts=3**.  $[3 + 1 + 3 + 3 + 1 + 4 = 15]$

- Identify Core, border and noise points
  - How many clusters are there?
  - Determine the cluster membership for each of the core or border point
- Consider the arrival of five more data points P16, P17, P18, P19, P20 in the database. Now we have 20 datum points. Refer Table-1 to get the distance between any two datum point of the database. Apply incremental DBSCAN on updated database to report the following.
- Identify Core, border and noise points
  - How many clusters are there?
  - Determine the cluster membership for each of the core or border point


Table 1.: Distance between every pair of datum points is given below

	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11	P12	P13	P14	P15	P16	P17	P18	P19	P20
P1	0	30	11	14	67	10	64	50	81	69	60	98	80	41	14	60	56	40	60	53
P2	30	0	25	40	44	40	46	60	60	55	62	76	65	18	42	30	36	46	68	67
P3	11	25	0	15	57	18	53	42	70	58	50	88	69	40	25	55	46	30	52	47
P4	14	40	15	0	69	10	64	40	82	67	50	100	78	53	20	70	58	31	50	41
P5	67	44	57	69	0	75	11	60	15	21	52	32	26	55	81	46	11	50	61	70
P6	10	40	18	10	75	0	70	50	88	75	60	106	86	50	10	70	64	41	60	50
P7	64	46	53	64	11	70	0	50	18	11	41	36	20	60	78	53	10	41	50	60
P8	50	60	42	40	60	50	50	0	65	46	14	82	53	78	60	86	50	14	10	10
P9	81	60	70	82	15	88	18	65	0	20	55	18	18	71	95	60	25	58	65	75
P10	69	55	58	67	21	75	11	46	20	0	35	36	11	70	83	65	20	40	45	55
P11	60	62	50	50	52	60	41	14	55	35	0	70	41	80	70	84	44	20	10	22
P12	98	76	88	100	32	106	36	82	18	36	70	0	30	85	113	70	42	76	80	92
P13	80	65	69	78	26	86	20	53	18	11	41	30	0	80	94	72	30	50	50	63
P14	41	18	40	53	55	50	60	78	71	70	80	85	80	0	50	22	50	64	86	84
P15	14	42	25	20	81	10	78	60	95	83	70	113	94	50	0	70	70	50	70	60
P16	60	30	55	70	46	70	53	86	60	65	84	70	72	22	70	0	44	72	92	94
P17	56	36	46	58	11	64	10	50	25	20	44	42	30	50	70	44	0	40	53	60
P18	40	46	30	31	50	41	41	14	58	40	20	76	50	64	50	72	40	0	22	22
P19	60	68	52	50	61	60	50	10	65	45	10	80	50	86	70	92	53	22	0	14
P20	53	67	47	41	70	50	60	10	75	55	22	92	63	84	60	94	60	22	14	0

\*\*\*\*\*

No comments:

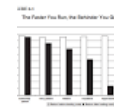
Post a Comment


Comment as: Narendran (Go ▾)

☐ Notify me

[Home](#)

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



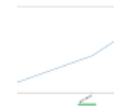
Intelligent investor (Ben Graham & Jason Zweig, 4e)

Reading from "A Note About Benjamin Graham by Jason Zweig" Here

are Graham...

Life Lessons By Steve Jobs

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

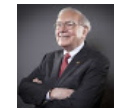
Day: 10th Aug 2017  
Sensex tanks 267 points.  
Nifty hits one-month

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...



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...


#ExamFromHome

LAST DAY TODAY



NMIMS-NPAT  
Admission Test for NMIMS Prog

APPLY NOW

About Me



Ashish Jain

[View my complete profile](#)

## Data Science Course

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

Great Learning Chennai

Followers

Followers (0)

Follow

## Database For 1.1mn Companies

275+ Industries

Access database of over 1.1M Indian companies  
Financial data. Private & public companies.

[info.emis.com](http://info.emis.com)

OPEN