AIKTC/KRRC/SoET/ACKN/QUES/2018-19/ **Date:** _____

School: SoET-CBCS **Branch:** COMP. ENGG. **SEM:** VI

To,
Exam Controller,
AIKTC, New Panvel.

Dear Sir/Madam,

Received with thanks the following **Semester/Unit Test-I/Unit Test-II (Reg./ATKT)** question papers from your exam cell:

Sr. No.	Subject Name	Subject Code	Format		No. of Copies
			SC	HC	
1	Software Engineering	CPC601		✓	02
2	System Programming & Complier Construction	CPC62		✓	02
3	Data Warehousing & Mining	CPC603		✓	02
4	Cryptography & System Security	CPC604		✓	02
5	ELECTIVE I: Machine Learning			✓	02
6					

Note: SC – Softcopy, HC - Hardcopy

(Shaheen Ansari)
Librarian, AIKTC

TE Sem-VI - Choice Based - Computers

10/5/19

Paper Subject Code: 88901 / Software Engineering

Time: 3 Hours **Marks: 80**

Note:- 1. Q1 is compulsory.
2. Solve any 3 questions from remaining questions.

Q1 Attempt any 4

a) Differentiate in between waterfall and spiral model. (5)
b) List out Requirement Elicitation Techniques. Explain any two methods. (5)
c) What is process and product metrics? Explain 3 P's of software Engineering. (5)
d) Illustrate design issues. (5)
e) Explain FTR. (5)
f) What is testing? What is the role of testing in software engineering. (5)

Q 2 a) Develop the SRS for University Management System. (10)
b) Explain the process of CMM. (10)

Q3 a) Explain Coupling and Cohesion? Explain the types of couplings with example. (10)
b) What are the testing strategies? (10)

Q4 a) Differentiate between FP based & LOC based cost estimation techniques. (10)
b) What is user interface design? Explain it with example. (10)

Q5a) What is maintenance? Explain the different types of maintenance. (10)
b) What is the use of use case diagram? Draw use case diagram for hospital management system. (10)

Q6 a) Differentiate between White Box and Black Box Testing. (10)
b) Illustrate Change Control & version control. (10)

****

(63) TE- Sem-V — Choice Based - comps 16/5/19

Paper / Subject Code: 88902 / System Programming and Compiler Construcation

(Time: 3 Hours) **Total Marks: 80**

N.B: (1) Question No. 1 is compulsory
(2) Attempt ant three questions out of remaining five questions

Q.1 (a) Differentiate between system software and application software. [05]
(b) Explain different functions of loader. [05]
(c) Explain forward reference problem and how it is handled in assembler design. [05]
(d) Explain macro and macro expansion. [05]

Q.2 (a) Find FIRST & FOLLOW for the following grammar [05]

S→Bb | Dd
B→aB | ϵ
D→cD | c

(b) Generate three address code for following code [05]

```
while(a<b) do
    if(c<d) then
        x=y+2
    else
        x=y-2
```

(c) With reference to assembler explain the following table with suitable example [10]
(i)MOT (ii)POT (iii)ST (iv)BT

Q.3 (a) Explain Synthesized and Inherited attribute with example. [10]
(b) Explain different code optimization techniques with example. [10]

Q.4 (a) Apply dead code elimination techniques for following code [05]

```
int count;
void foo( )
{
   int i;
   i=1;
   count=1;
   count=2;
   return
   count=3;
}
```

(b) Eliminate left recursion from the following grammar [05]

S→(L) | x
L→ L,S | S

(c) Explain different types of loaders in detail. [10]

Q.5 (a) Draw flowchart of a Pass-I of two pass assembler design and explain in detail. [10]

(b) Explain different features of macro with example. [10]

Q.6 (a) For the following grammar construct LL(1) parsing table and parse the string (a-a) [10]

S→F
S→(S-F)
F→a

(b) Explain different issues in code generation. [10]

****

4D8D3D396A1996D9CD52F4C3FB972F2B

TE-sem-VI- comps- Choice Based

22/5/19

Paper / Subject Code: 88903 / Data Warehouse and Mining

Time: 03 Hours **Marks: 80**

Note: 1. Question 1 is compulsory.
2. Answer any three out of remaining five questions.
3. Assume any suitable data wherever required and justify the same.

Q1 a) What are spatial data structures? Outline their importance in GIS. [5]
b) What is Metadata? Why do we need metadata when search engines like Google seem so effective? [5]
c) In real-world data, tuples with *missing values* for some attributes are a common occurrence. Describe various methods for handling this problem. [5]
d) With respect to web mining, is it possible to detect visual objects using meta-objects? [5]

Q2 a) Suppose that a data warehouse for *DB-University* consists of the four dimensions *student, course, semester,* and *instructor*, and two measures count and *avg-grade*. At the lowest conceptual level (e.g., for a given student, course, semester, and instructor combination), the *avg-grade* measure stores the actual course grade of the student. At higher conceptual levels, *avg-grade* stores the average grade for the given combination. [10]
i. Draw a *snowflake schema* diagram for the data warehouse.
ii. Starting with the base cuboid [*student, course, semester, instructor*], what specific *OLAP operations* (e.g., roll-up from *semester* to year) should you perform in order to list the average grade of CS courses for each *DB-University* student.

b) What is the relationship between data warehousing and data replication? Which form of replication (synchronous or asynchronous) is better suited for data warehousing? Why? Explain with appropriate example. [10]

Q3 a) The following table consists of training data from an employee database. The data have been generalized. For example, "31 : : : 35" for *age* represents the age range of 31 to 35. For a given row entry, *count* represents the number of data tuples having the values for *department, status, age,* and *salary* given in that row. [10]

department	status	age	salary	count
sales	senior	31 ... 35	46K ... 50K	30
sales	junior	26 ... 30	26K ... 30K	40
sales	junior	31 ... 35	31K ... 35K	40
systems	junior	21 ... 25	46K ... 50K	20
systems	senior	31 ... 35	66K ... 70K	5
systems	junior	26 ... 30	46K ... 50K	3
systems	senior	41 ... 45	66K ... 70K	3
marketing	senior	36 ... 40	46K ... 50K	10
marketing	junior	31 ... 35	41K ... 45K	4
secretary	senior	46 ... 50	36K ... 40K	4
secretary	junior	26 ... 30	26K ... 30K	6

Let *status* be the class label attribute.
i. How would you modify the basic decision tree algorithm to take into consideration the *count* of each generalized data tuple (i.e., of each row entry)?
ii. Use your algorithm to construct a decision tree from the given data.

Q3 b) Why is *tree pruning* useful in decision tree induction? What is a drawback of using a separate set of tuples to evaluate pruning? Given a decision tree, you have the option of (i) *converting* the decision tree to rules and then pruning the resulting rules, or (ii) *pruning* the decision tree and then converting the pruned tree to rules. What advantage does (i) have over (ii)? **[10]**

Q4 a) Suppose that the data mining task is to cluster points (with (x, y) representing location) into three clusters, where the points are: A_1 (2, 10), A_2 (2, 5), A_3 (8, 4), B_1 (5, 8), B_2 (7, 5), B_3 (6, 4), C_1 (1, 2), C_2 (4, 9). **[10]**
The distance function is Euclidean distance. Suppose initially we assign A_1, B_1, and C_1 as the center of each cluster, respectively. Use the *k-means* algorithm to show only (i) The three cluster centers after the first round of execution (ii) The final three clusters.

b) Briefly outline with example, how to compute the dissimilarity between objects described by the following: **[10]**

i. Nominal attributes

ii. Asymmetric binary attributes

Q5 a) Frequent pattern mining algorithms considers only distinct items in a transaction. However, multiple occurrences of an item in the same shopping basket, such as four cakes and three jugs of milk, can be important in transactional data analysis. How can one mine frequent itemsets efficiently considering multiple occurrences of items? Generate Frequent Pattern Tree for the following transaction with 30% minimum support: **[10]**

Transaction ID	Items
T1	E, A, D, B
T2	D, A, C, E, B
T3	C, A, B, E
T4	B, A, D
T5	D
T6	D, B
T7	A, D, E
T8	B, C

b) Differentiate between simple linkage, average linkage and complete linkage algorithms. Use complete linkage algorithm to find the clusters from the following dataset. **[10]**

X	4	8	15	24	24
Y	4	4	8	4	12

Q6 a) *Data quality* can be assessed in terms of several issues, including accuracy, completeness, and consistency. For each of the above three issues, discuss how data quality assessment can depend on the *intended use* of the data, giving examples. Propose two other dimensions of data quality. **[10]**

b) Present an example where data mining is crucial to the success of a business. What *data mining functionalities* does this business need (e.g., think of the kinds of patterns that could be mined)? Can such patterns be generated alternatively by data query processing or simple statistical analysis? **[10]**

***

56AB7F22531F97180656750C6B37E524

(63)

TE - Sem-V) - Choice Based — Comps

28/5/19

Paper / Subject Code: 88904 / Cryptography and System Security

[3 Hours] **[Total Marks 80]**

N. B: 1. Question No. 1 is Compulsory.
2. Solve any THREE from Question No. 2 to 6.
3. Draw neat well labeled diagram wherever necessary.

Q.1 a) Enlist security goals. Discuss their significance. (05)

b) Compare AES and DES. Which one is bit oriented? Which one is byte oriented? (05)

c) What is authentication header(AH)? How does it protect against replay attacks? (05)

d) List various Software Vulnerabilities. How vulnerabilities are exploited to launch an attack. (05)

Q.2 a) Encrypt the plaintext message "SECURITY" using affine cipher with the key pair (3, 7). Decrypt to get back original plaintext. (10)

b) Explain different types of Denial of Service attacks. (10)

Q.3 a) Users A and B use the Diffie-Hellman key exchange technique with a common prime 71 and primitive root 7. Show that 7 is primitive root of 71. If user A has private key x=5, what is A's Public Key R_1? If user B has private key y=12, what is B's public key R_2? What is the shared secret key? (10)

b) What are traditional ciphers? Discuss any one substitution and transposition cipher with example. List their merits and demerits. (10)

Q.4 a) Alice chooses public key as (7, 33) and B chooses public key as (13, 221). Calculate their private keys. A wishes to send message m=5 to B. Show the message signing and verification using RSA digital signature. (10)

b) Discuss in detail block cipher modes of operation. (10)

Q.5 a) What is the need of SSL? Explain all phases of SSL Handshake protocol in detail. (10)

b) What are the requirements of the cryptographic hash functions? Compare MD5 and SHA Hash functions. State real world applications of hash functions. (10)

Q. 6 Write short notes on any FOUR: (20)

a. Kerberos
b. Buffer Overflow
c. 3DES
d. X.509
e. IDS

——x——

(83) TE - sem - VI - Choice Based - Comps

Paper / Subject Code: 88905 / Elective - II Machine Learning

3/6/1

(3 Hours) **Total Marks: 80**

N.B. : (1) Question No. 1 is compulsory.
(2) Attempt any three questions out of remaining five.

1 (a) What is Machine Learning? How is it different than Data Mining? [05]

(b) Why Dimensionality Reduction is very Important step in Machine Learning? [05]

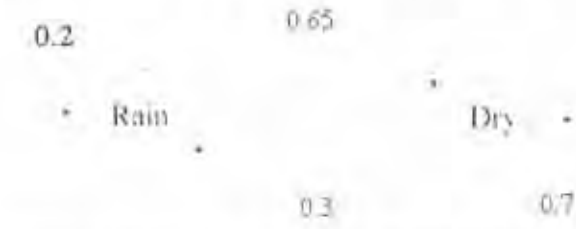
(c) Determine weights and threshold for the given data using McCulloch-Pitts neuron model. Plot all data points and show separating hyper-plane. [05]

X1	X2	D
0	0	0
0	1	0
1	0	1
1	1	0

(d) Describe Down Hill Simplex method. Why is it called Derivative Free method? [05]

2. (a) Explain the steps of developing Machine Learning applications [10]

(b) Consider Markov chain model for 'Rain' and 'Dry' is shown in following figure. [10]



Two states: 'Rain' and 'Dry'. Transition probabilities: P('Rain'|'Rain') = 0.2, P('Dry'|'Rain') = 0.65, P('Rain'|'Dry') = 0.3, P('Dry'|'Dry') = 0.7, Initial probabilities: say P('Rain') = 0.4, P('Dry') = 0.6.

Calculate a probability of a sequence of states {'Dry', 'Rain', 'Rain', 'Dry'}.

3 (a) Minimize $f(x_1, x_2) = 4x_1 - 2x_2 + 2x_1^2 + 2x_1x_2 + x_2^2$ [10]

With starting point $X_1 = \begin{Bmatrix} 0 \\ 0 \end{Bmatrix}$ using the steepest descent method.

(Perform two iterations).

3 (b) Explain following terms Initial hypothesis, Expectation step and Maximization step w.r.t E-M algorithm. Explain How Initial hypothesis converges to optimal solution? (You may explain it with an example) [10]

4. (a) Why Dimensionality reduction is an important issue? Describe the steps to reduce dimensionality using Principal Component Analysis method by clearly stating mathematical formulas used. [10]

19F968895380187F5405CE39D2FC0ADD

(b) For the following data. Calculate Gini indexes and determines which attribute is root attribute and generate two level deep decision tree. [10]

Sr. No.	Income	Defaulting	Credit Score	Location	Give Loan?
1	low	high	high	bad	no
2	low	high	high	good	no
3	high	high	high	bad	yes
4	medium	medium	high	bad	yes
5	medium	low	low	bad	no
6	medium	low	low	good	yes
7	high	low	low	good	yes
8	low	medium	high	bad	no
9	low	low	low	bad	no
10	medium	medium	low	bad	no
11	low	medium	low	good	yes
12	high	medium	high	good	yes
13	high	high	low	bad	no
14	medium	medium	high	good	yes

5. (a) Explain following terms w.r.t Bayes' theorem with proper examples. [10]

(a) Independent probabilities
(b) Dependent probabilities
(c) Conditional Probability
(d) Prior & Posterior probabilities
Define Bays theorem based on these probabilities.

(b) Draw and discuss the structure of Radial Basis Function Network. How RBFN can be used to solve non linearly separable pattern? [10]

6. Attempt any **four** [20]

(a) Illustrate Support Vector machine with neat labeled sketch and also show how to derive optimal hyper-plane?
(b) Differentiate: Derivative Based and Derivative free optimization techniques.
(c) Explain how regression problem can be solved using Steepest descent method. Write down the steps.
(d) Write Short note on ~~ICA~~ ICA and compare it with PCA.
(e) DownHill simplex method.

19F968895380187F5405CE39D2FC0ADD