Motilal Nehru National Institute of Technology, Allahabad
School of Management Studies
B. Tech 8th Semester, Even Semester, Mid semester Examination, 2018-19
MS1888- Cross cultural Management

Time: 1 ½ hours Max marks:20

Attempt all questions.

Q.1 Write short notes on any five: 2x5

a) Systems theory of management
b) Taylor's rational model of management
c) Globalization
d) Ethnocentrism
e) Cultural relativism
f) Cultural determinism
g) Theory X

Q.2. Attempt any one part: 5x2

a) "Culture is like an ice-berg." Throw light on this statement and define the concepts of implicit and explicit cultures.
b) Define Rousseau's ring. Elucidate on the multilayered model of culture as given by Rousseau.

OR

c) Define and explain the elements of culture.
d) "Culture is the software of mind." Throw light on this statement of Hofstede and explain his concept of culture.

Motilal Nehru National Institute of Technology Allahabad
Department of Computer Science & Engineering
Mid Semester Examination 2018-19
B. Tech Information Technology VIII Semester
Subject Code/Name: **CS 1803 – Privacy Preserving Publishing**

Duration: **1.5 hours** Full Marks: **20**

NOTE: All questions are compulsory. Be specific and to the point in your answers. Make assumptions wherever necessary and quote it.

Q1. Describe the general process of data collection and publishing. Provide the supporting block diagram. (2)

Q2. What is privacy preserving data publishing? Explain with a scenario in which we need to publish sensitive data and possible privacy threat to it. (3)

Q3. Explain following attack models with suitable examples and their possible countermeasures. (4)

a) Record Linkage Model
b) Attribute Linkage Model
c) Table Linkage Model
d) Probabilistic Linkage Model

Q4. a) What is a linking attack? Describe with an example. Explain, how the *k*-anonymity technique can prevent it? (2)

Q4. b) Apply generalization and suppression operators to generate a 4-anonymous table from the given raw table. Treat *Condition* as a sensitive attribute. (1)

S. No.	Zip Code	Age	Nationality	Condition
1	13053	28	Indian	Heart Disease
2	13068	29	Russian	Heart Disease
3	13061	36	Japanese	Cancer
4	13054	39	Russian	Cancer

Q5. What is the entropy *l*-diversity? Calculate the entropy of the distribution of sensitive attributes in the table given in **Q4. b**. Convert the same table into a 2-diverse table. (3)

Q6. What is the homogeneity attack over *k*-anonymity technique? What techniques should be used to prevent it, show with an example. (2)

Q7. Describe the Earth Mover Distance in details? Calculate the Earth mover Distance of the distribution set **P= (3, 2, 0, 2, 2)** and **Q = (4, 3, 3, 3, 4)**. (3)

Motilal Nehru National Institute of Technology Allahabad
Department of Computer Science & Engineering
Mid-Semester Examination (Even Semester 2018-2019)

CS-1805: PATTERN RECOGNITION
B.Tech. (CS+IT) VIII Semester

Time : 90 Min. **M. M. : 20**

Note : All questions are compulsory.

Q1. Consider the multivariate normal density for a d dimensional sample $\mathbf{x}$ as $p(\mathbf{x})\sim\mathbf{N}(\boldsymbol{\mu},\boldsymbol{\Sigma})$, where $\boldsymbol{\mu}=[\mu_1,\mu_2,\dots,\mu_d]^t$ and $\boldsymbol{\Sigma}=diagonal(\sigma_1^2,\sigma_2^2,\dots,\sigma_d^2)$.

a) State the Mahalanobis distance from $\mathbf{x}$ to $\boldsymbol{\mu}$.

b) Show that the evidence **[4]**

$$p(\mathbf{x})=\frac{1}{\prod_{i=1}^{d}\sqrt{2\pi}\sigma_i}\exp\left[-\frac{1}{2}\sum_{i=1}^{d}\left(\frac{x_i-\mu_i}{\sigma_i}\right)^2\right]$$

Q2. Consider the following samples $\{(2,0)^t,(6,0)^t,(4,1)^t,(4,-1)^t\}$ in class ω_1 and $\{(-3,1)^t,(-3,-1)^t,(-5,0)^t\}$ in class ω_2 to be drawn from two different Gaussian functions. Take the decision for the sample $(0,0)^t$ using Bayesian classifier when $P(\omega_1)=P(\omega_2)=\frac{1}{2}$. **[5]**

Q3. For the two class problem with the class labels ω_1 and ω_2, let $P(\omega_1)$ and $P(\omega_2)$ be their prior probabilities. The likelihood functions for these classes are normal distributions with the same covariances but different means: $\mathbf{N}_1(\boldsymbol{\mu}_1,\boldsymbol{\Sigma})$ for class ω_1 and $\mathbf{N}_2(\boldsymbol{\mu}_2,\boldsymbol{\Sigma})$ for class ω_2. Prove that the Bayes decision boundary between class ω_1 and ω_2 is of the form $\mathbf{w}^t(\mathbf{x}-\mathbf{x}_0)=0$. State the value of $\mathbf{w}$ and $\mathbf{x}_0$. ***[5]***

Q4. In many pattern classification problems one has the option either to assign the pattern to one of c classes, or to reject it as being unrecognizable. If the cost for rejects is not too high, rejection may be a desirable action. Let

$$\lambda(\alpha_i|\omega_j)=\begin{cases}0 & if\ i=j \quad where\ i,j\in\{1,2,\dots,c\}\\ \lambda_r & if\ i=c+1\\ \lambda_s & otherwise\end{cases}$$

where λ_r is the loss incurred for choosing the $(c+1)$-th action which is rejection, and λ_s is the loss incurred for making a substitution error.

a) Show that the minimum risk is obtained if we decide ω_i if $p(\omega_i|\mathbf{x})\geq p(\omega_j|\mathbf{x})$ for all j and if $p(\omega_i|\mathbf{x})\geq 1-\frac{\lambda_r}{\lambda_s}$, and reject otherwise.

b) Show that the discriminant functions for the above shall be **[6]**

$$g_i(\mathbf{x})=\begin{cases}p(\mathbf{x}|\omega_i).P(\omega_i) & if\quad i=1,2,\dots,c\\ \frac{\lambda_s-\lambda_r}{\lambda_s}\sum_{j=1}^{c}p(\mathbf{x}|\omega_j).P(\omega_j) & if\quad i=c+1\end{cases}$$

Motilal Nehru National Institute of Technology Allahabad
Department of Computer Science & Engineering

B. Tech 8th Semester (Information Technology)
Mid Semester Examination 2018-19
Subject Code/Name: CS 1804 / Research Trends in IT

Duration: **90 Minutes** Max. Marks: **20**

NOTE: Attempt all sections.

[SECTION –A]

1. Explain the importance and steps of research process **[3]**
2. Why do we use stratified sampling? A population is divided into three strata using N1 = 5000, N2 = 2000 and N3 = 3000. Standard deviations of each strata is 15, 18, and 5 respectively. How can we draw a sample of size n = 84 from the above three strata using proportional as well as disproportional allocation? Explain Briefly. **[4]**

[SECTION –B]

3. (a) Explain how Blockchain technology can bring improvements in any organization? **[2]**
   (b) Consider a scenario where a group of 6 people (with different roles) are working on a project. Each member can make changes according to his/her responsibilities. However the changes can be published through quorum. For this scenario **[3]**
   I. Design the Blockchain.
   II. Explain about the approval criteria.
   III. Propose the criteria for reward point.
4. Explain how Fog Computing Technology is different from cloud computing? Also, highlight certain research domains in which Fog Computing is being successfully utilized. **[3]**
5. Explain the following terms: **[5]**
   I. Data Virtualization
   II. IOT in healthcare.
   III. Cloud Telephony
   IV. Grid Computing
   V. Telemedicine

Mid Semester Examination, Feburary-2019
Department of Computer Science & Engineering,
Motilal Nehru National Institute of Technology Allahabad,
B.Tech VIII Semester
"Computer Science & Engineering and Information Technology"

Subject: Web Mining **Paper code: CS-1810**
Duration: 90 Minutes **Max. Marks: 20**

Note: **Attempt all questions. Make assumptions wherever necessary and quote it. For solving numerical part must write all used formula.**

Q1. What is web mining? Give the taxonomy of web mining with issues in each of them. How data mining techniques are used in web mining? Give the example of each technique. **[4]**

Q2. A corpus contains the following five documents:

D1	To be or not to be, this is the question!
D2	I have a pair of problems for you to solve today.
D3	It's a long way to Tipperary, it's a long way to go.
D4	I've been walking a long way to be here with you today.
D5	I am not able to question these orders.

The indexing system only considers nouns, adjectives, pronouns, adverbs and verbs. All forms are converted to singular, verbs are converted to the infinitive tense, removes all punctuation marks and translates all letters to uppercase. Conjunctions, prepositions, articles and exclamations are discarded as well. Multiple occurrences of the same term within a document are not counted.

For instance, the phrase

Hey, it's not too late to solve these exercises!

becomes

IT BE NOT TOO LATE SOLVE THIS EXERCISE

a) What is the minimum dimension (number of coordinates) of the TFIDF vector space for this collection of documents? Calculate the tf-idf matrix for all documents. **[2]**
b) Fill the 5×5 matrix of Jaccard coefficients between all pairs of documents. **[2]**
c) Apply an agglomerative clustering procedure to the collection. as a measure of similarity between two clusters C1 and C2, consider the highest similarity between D1 and D2, with D1 $\in$ C1 and D2 $\in$ C2. **[2]**
d) Draw the resulting dendrogram. **[2]**

Q3. Explain the different IR queries with help of example. What is the use of Inverted index? Consider the three documents below: **[4]**

D1	Web mining is useful.
D2	Usage mining applications.
D3	Web stricture mining studies the web hyperlink structure.

Create the inverted indices and construct the inverted index for above three documents.

Q4. What is the sentiment analysis? Explain different field of sentiment analysis. Define infrequent aspect extraction, explicit and implicit aspects and basic rules of opinions with help of example. **[4]**