Tribhuvan University
Institute of Science and Technology
SCHOOL OF MATHEMATICAL SCIENCES
First Assessment 2079

Subject: **Decision Analysis** *Full Marks*: 45

Course No: **MDS 606** *Pass Marks*: 22.5

Level: **MDS /II Year /III Semester** *Time*: 2hrs

Candidates are required to give their answers in their own words as far as practicable.

Attempt All Questions.

Group A [3 ×5 = 15]

1. Define Goal Programming. Distinguish between Linear programming problem and Goal programming.
2. What is conditional probability? State Bayes theorem up to 3 events.
3. A trader has two investment opportunities, A and B available to him but does not have enough capital to invest in the both. The probability of success on A is 0.70 while that on B is 0.40. Both the investments require an initial capital of Rs.20000 and both return nothing if the venture is not successful. Investment A returns Rs.30000 over cost if it is successful, whereas the successful completions of B will return Rs.50000 over cost. Using EMV criterion decide the best strategies the trader should adopt.
4. Explain the risk preferences and shape of the utility function.
5. What are the decision making under ignorance? Explain them with suitable examples.

Group B [5 × 6 = 30]

6. What are the five criteria of decision making under condition of uncertainty? Explain them with suitable examples.
7. A newspaper vendor buys a new started local paper at the rate of Rs.5 and sells it at the rate of Rs.10. The unsold papers do not have any value. The vendor knows that he cannot sell more than 20 papers in a day and the minimum sale would not be less than 18. How many papers should he buy based on (i) maximax criterion (ii) maximin criterion (iii) minimax regret criterion (iv) Laplace criterion and (v) criterion of realism if coefficient of optimism is 0.60?
8. A distribution of past daily sales of a commodity is as follows:

Daily sales (units)	1000	1200	1400	1600	1800
Probability	0.05	0.15	0.35	0.30	0.15

If selling price per unit is Rs.40 and cost price per unit is Rs.25 and salvage price per unit is Rs.5, what is

a) optimum quantity?
b) maximum expected profit?
c) expected values for perfect information?

9. There are three varieties of machines viz. A, B and C of which any one is to be purchased by a firm. The profits expected from the different machines under different markets conditions are tabulated as under:

Profit matrix (in '000 Rs.)

Market	Machine		
	A	B	C
Poor	0.50	0.00	−1.50
Fair	1.00	1.50	0.50
Good	1.50	2.50	3.50

The probabilities of the market being poor, fair, and good were estimated by the firm to be 0.3, 0.5 and 0.2 respectively. But a research group opined that these probabilities can not to be relied upon to be accurate and they assessed the following chances after a careful survey of the market conditions.

Actual state	Indicated state		
	Poor	Fair	Good
Poor	0.7	0.2	0.1
Fair	0.2	0.7	0.1
Good	0.0	0.2	0.8

From the above particulars

a) Construct the expected opportunity loss table basing on the market assessment made by the firm and determine the optimal act using EOL criterion. Also, determine the EVPI there from.

b) Construct the EOL table basing upon the information furnished by the research group and determine the net revised EOL (i.e. AMEOL) thereby.

c) State the maximum amount that should be spent after the market research.

d) Also, find the expected net gain from the sampling (ENGS), if the decision maker is prepared to obtain additional information for Rs.110.

10. Solve the following LPP graphically or simplex method:

Maximize $Z = 3x_1+9x_2$

Subject to the constraints

$x_1 + 2x_2 \leq 4$

$x_1 + 4x_2 \leq 8$

and $x_1, x_2 \geq 0$

***

Tribhuvan University
Institute of Science and Technology
SCHOOL OF MATHEMATICAL SCIENCES
First Assessment 2079

Subject: **Advanced Data Mining** *Full Marks*: 45

Course No: **MDS 602** *Pass Marks*: 22.5

Level: **MDS /II Year /III Semester** *Time*: 2hrs

Candidates are required to give their answers in their own words as far as practicable.

Attempt All Questions.

Group A [3 ×5 = 15]

1. What is data mining? Describe about Major issues in Data mining?
2. Write down the steps to find the principal component of a data set.
3. What is data warehouse? Compare OLTP and OLAP.
4. .What are categorical data? How can you handle the categorical data in association mining?
5. Explain Rule based classifier. How it is different than decision tree classifier.

Group B [5 × 6 = 30]

6. Classify the following attributes as binary, discrete, or continuous. Also classify them as qualitative (nominal or ordinal) or quantitative (interval or ratio). Some cases may have more than one interpretation, so briefly indicate your reasoning if you think there may be some ambiguity (Example: Age in years. Answer: Discrete, quantitative, ratio).
   a) Brightness as measured by a light meter.
   b) Angles as measured in degrees between 0 ◦ and 360 ◦.
   c) Bronze, Silver, and gold medals as awarded at the Olympics.
   d) Number of patients in a hospital.
   e) Ability to pass light in terms of the following values: opaque, translucent, transparent.
   f) Military rank.
7. Given the following data (in increasing order) for the attribute age: 13, 15, 16, 16, 19, 20, 20, 21, 22, 22, 25, 25, 25, 25, 30, 33, 33, 35, 35, 35, 35, 36, 40, 45, 46, 52, 70. (a) Use smoothing by bin means to smooth the above data, using a bin depth of 3. (b) How might you determine outliers in the data?
8. Consider the data set shown in Table

Transaction ID	Items Bought
1	{a, d, e}
2	{a, b, c, e}
3	{a, b, d, e}
4	{b, c, e}
5	{b, d, e}
6	{c, d}
7	{a, b, c}
8	{a, d, e}
9	{a, b, e}

Compute the support for itemsets {e}, {b, d}, and {b, d, e} by treating each transaction ID as a market basket.
Use the results in part (a) to compute the confidence for the association rules {b, d} −→ {e} and {e} −→ {b, d}. Is confidence a symmetric measure?
Repeat part (a) by treating each customer ID as a market basket. Each item should be treated as a binary variable (1 if an item appears in at least one transaction bought by the customer, and 0 otherwise.)

OR

Construct the FP tree from the above transactional data table with minimum support of 33%.

9. Explain ID3 Algorithm with an example.

10. Describe Confusion Matrix with example. Define Accuracy, Precision, TPR, TNR, FPR, FNR of the classifier model.

OR

Given the following confusion matrix, determine Accuracy, Precision, TPR, TNR, FPR, FNR of the classifier model.

N=165

	Predicted. No	Predicted: Yes	
Actual: No	50 TN	10 FP	60
Actual: Yes	5 FN	100 TP	105
	55	110	

$TPR = \frac{TP}{TP+TN} \times 1$

$TNR = \frac{TN}{TN+FN}$

$FPR =$

***

Tribhuvan University
Institute of Science and Technology
SCHOOL OF MATHEMATICAL SCIENCES
First Assessment 2079

Subject: **Monte Carlo Methods**
Course No: **MDS 607**
Level: **MDS /II Year /III Semester**

Full Marks: 45
Pass Marks: 22.5
Time: 2hrs

Candidates are required to give their answers in their own words as far as practicable.

Attempt All Questions.

Group A **[3 ×5 = 15]**

1. Name the characteristics of good random number generators.
2. Explain the meaning of "Random variables".
3. How do you estimate errors in Monte Carlo methods? Describe.
4. Write down the meaning of transition probabilities with examples.
5. What is "State space"? Explain.

Group B **[5 × 6 = 30]**

6. Discuss the congruential method to generate random numbers with its algorithm.

OR

How can one test whether random numbers are good? Explain the criteria of good random no. generators.

7. What is Monte Carlo Method? Discuss its significance.

OR

Distinguish between Simple and importance sampling in Monte Carlo methods of Integration.

8. Explain the meaning of "Markov Chain". Also explore its role in Monte Carlo methods.
9. How do you distinguish "stationary distribution" from other distributions? Explain an algorithm to get "stationary distribution" using Python code.
10. Explore the significance of "Reversible Markov chain".

***

Tribhuvan University
Institute of Science and Technology
SCHOOL OF MATHEMATICAL SCIENCES
First Assessment 2079

Full Marks: 45
Pass Marks: 22.5
Time: 2hrs

Subject: **Techniques for Big Data**
Course No: **MDS 603**
Level: **MDS /II Year /III Semester**

Candidates are required to give their answers in their own words as far as practicable.

Attempt All Questions.

Group A **[3 ×5 = 15]**

1. What do you mean by Big Data? How big the data would you consider as Big Data?
2. Briefly explain the different characteristics of Big Data.
3. What is functional Programming? What are the concepts borrowed by Map Reduce from Functional Programming?
4. When mapper worker is failed both in-progress and completed tasks are reset to idle but when Reducer worker is failed only in-progress tasks are reset to idle. What is the reason behind this?
5. What is Hadoop? Briefly explain about the history behind the development of Hadoop.

Group B **[5x6=30]**

6. What are the scopes and challenges of Big Data? Explain with suitable example.

OR

Explain about the current trend of Big Data along with its applications in the Covid 19 pandemic.

7. Explain the Distributed Execution Overview of Map Reduce.
8. Draw the block diagram to find word frequency in following text:

   Big Data Technology uses Big Data Analytics

   Big Volume of Data Gives Big Value

   Big Data has Velocity Variety and Veracity

   Big Data is really Big Technology

9. What is Hadoop Ecosystem? Explain about core components of the Hadoop Ecosystem.
10. Explain the Master/Slave Architecture in Hadoop

OR

Explain the different configuration modes to setup the Hadoop.

***

Tribhuvan University
Institute of Science and Technology
SCHOOL OF MATHEMATICAL SCIENCES
First Assessment 2079

Subject: **Research Methodology**
Course No: **MDS 601**
Level: **MDS /II Year /III Semester**

Full Marks: 45
Pass Marks: 22.5
Time: 2hrs

Candidates are required to give their answers in their own words as far as practicable.

Attempt All Questions.

Group A [3 ×5 = 15]

1. Explain the difference between research methods and Research methodology.
2. What do you mean by Research Hypothesis? Explain in brief with one case study.
3. Define Inductive and Deductive theory with suitable examples.
4. Distinguish between research hypothesis and statistical hypothesis.
5. Define literature review. What are the benefits of good literature survey?

Group B [5 × 6 = 30]

6. " The distinction between basic and applied research is that the purpose of the former is to answer a question and the purpose of the later is to solve a question". Comment on this statement.
7. What is scientific research? Describe the major steps involved in the scientific research process?
8. What library skills are required for a thorough survey of literature related to a research topic in Data science?
9. Discuss research design. Is it possible to conduct a research study without a formal research design? How?

OR

Define descriptive research design? What are the characteristics of descriptive research design?

10. Describe a research design for testing the following hypothesis. "Male managers will have more access to critical information than female managers in the same ranks".

OR

What are the elements of research design? Point out the initial questions in developing a research design.

***