



## Continuous Assessment Test – I

Course Code	19DS910	Course Name	Web Analytics		
Degree	M.Sc.	Programme	Data Science	Semester	IX
Date	08/08/2023	Duration	90 Minutes	Max. Marks	50
Faculty-in-Charge	Mr.B.Ramprakash				

Answer All Questions

Part A		05x2=10Marks	CO
A1.	What advantages does Web analytics offer to businesses and website owners?		CO1
A2.	How do Universal Analytics and Google Analytics 4 differ from each other?		CO4
A3.	Define: Churn probability.		CO1
A4.	What is the role of a web log server, and how does it contribute to website management and analysis?		CO4
A5.	What is Referrer log and how does it play a crucial role in tracking website visitors?		CO1
Part B		02 x 5=10 Marks	
B1.	What does a web log server contain, and how does it capture and store information about website visitors' interactions and activities?		CO1
B2.	What are the key features and functionalities offered by Google Analytics 4, and how do they help website owners and marketers gain insights into user behaviour, track website performance, and optimize online strategies?		CO4
Part C		03x10= 30Marks	CO
C1.	<p>You are a data analyst working for an e-commerce website. Your company recently implemented clickstream data tracking to monitor user interactions on the website. The clickstream data records various events, including page views, clicks, and interactions with product recommendations.</p> <p>Based on the clickstream data collected from the website, analyse and identify the most frequently visited product categories during the last month. Provide insights on user preferences and potential strategies to optimize the user experience for those categories. Additionally, highlight any significant drop-off points in the clickstream data that may indicate areas for improvement in the website's navigation or design. (10)</p>		CO2
OR			

C2.	You are a web analyst working for a content-based website that publishes articles, videos, and info graphics on various topics. Your company recently started using clickstream data to understand user behaviour on the website. The clickstream data includes information on user sessions, page views, time spent on each page, and click-through rates on recommended content.	CO2
C3.	Analyse the clickstream data for the past week and identify the most engaging content types based on the average time users spend on each page. Provide recommendations on how the website can leverage this information to create more compelling content and improve user engagement. Additionally, investigate the click-through rates on recommended content and propose strategies to optimize the content recommendation system to increase user interaction with related articles, videos, and infographics.	CO3
<b>OR</b>		
C4.	You are a web server administrator responsible for managing the performance of a high-traffic e-commerce website. The website experiences periodic slowdowns and occasional downtime during peak hours, resulting in dissatisfied users and potential revenue loss. You want to identify and resolve the performance bottlenecks to ensure a smooth and reliable user experience.	CO3
C5.	For a popular social media platform you are a web master .This site experiences a significant increase in user activity during specific events, such as product launches, major announcements, or trending topics. The website's performance has been impacted during these peak traffic events, leading to slow loading times, delayed response, and occasional server crashes. This has resulted in user frustration and negative feedback. Identify a holistic approach to optimize web server performance during peak traffic events. The proposed solutions should help enhance user experience, maintain website stability, and uphold the reputation of the social media platform during critical moments of high user activity.	CO3
C6.	Examine the effectiveness of these data sources in measuring website traffic and user engagement for the e-commerce company based on the available web log server data, clickstream data, and other web channel data. Use the below scenario to understand the existing problem better.	CO5
<b>OR</b>		
C6.	Evaluate the current content recommendation system's performance and its impact on user engagement. Propose key performance indicators (KPIs) to measure the success of the personalized content recommendation system. Using the available web log server data, clickstream data, and other web channel data, assess the effectiveness of these data sources in understanding user preferences and interactions with the news platform. Use the below scenario to improve your understanding.	CO5
Consider you a data scientist working for a popular online news platform that attracts a large audience and generates significant web traffic. The platform relies on web log server data, clickstream data, and other web channel data to understand user behaviour and improve content recommendations. The platform is considering implementing a new personalized content recommendation system to enhance user engagement and retention.		

\*\*\*\*\*



Continuous Assessment Test – I

Course Code	19DS920	Course Name	Natural Language Processing		
Degree	M.Sc.	Programme	Data Science	Semester	IX
Date	8/08/2023	Duration	90 Minutes	Max. Marks	50
Faculty-in-Charge	Dr. P. Sharmila				

Answer All Questions

Part A		5 x 2 = 10	COs
A1.	What is bound morpheme?		CO1
A2.	Differentiate Stemming and Lemmatizing with an example		CO3
A3.	Construct morphological tree for “irreplaceableity” (hint: irreplaceability)		CO1
A4.	What is orthographic rules? Give an example		CO1
A5.	Create a regular expression for a word that matches both uppercase and lowercase letters  [a-zA-Z]		CO2
Part B		3 x 5 = 15	COs
B1.	Construct Finite State Transducer (FST) for both regular and irregular noun pluralization for the given word Book → Books and Leaf → Leaves		CO1
B2.	Find all possible different meanings for the following sentences A. “Can you please bring me the orange bowl” B. “I saw bats”		CO2
B3.	List the types of Ambiguities with an example		CO2
Part C		1x10=10, 1x15=15	COs
C1.	Compute: A) Normalized Hamming distance score (5marks) B) Jaccard similarity score (5marks) between the sequences. Text Sequence 1: "The quick brown fox jumps over the lazy dog." Text Sequence 2: "The quick brown fox jumps over the lazy cat."		CO3
OR			
C2.	Demonstrate how the Levenshtein distance may be used to compare the similarities between the strings "kitten" and "sitting". (10)		CO3
C3.	Compute tf-idf using Bag of Words without any NLP based preprocessing and find the similarity using cosine similarity Text Sequence 1: "The quick brown fox jumps over the lazy dog." Text Sequence 2: "The quick brown fox jumps over the lazy cat." (15)		CO3
OR			
C4.	Explain co-occurrence matrix of window size 3 for the given sentences "I love to read books." "Books are a great source of knowledge." Find the similarity using cosine similarity between "books and knowledge". (15)		CO3

\*\*\*\*\*



**THIAGARAJAR COLLEGE OF ENGINEERING, MADURAI 625 015.**  
**Department of Applied Mathematics and Computational Science**  
**Continuous Assessment Test – I**

Course Code	19DS930	Course Name	Computer Vision		
Degree	M.Sc.	Programme	M.Sc. Data Science	Semester	IX
Date	09/08/2023	Duration	Slot 2 - 90 minutes	Max. Marks	50
Faculty-in-Charge		Prof S.T. Padmapriya			

**Answer All Questions**

Part A		$5 \times 3 = 15$	CO
A1	What is digital image processing?		CO1
A2	Write about the factors that affect the quality of images during image acquisition.		CO1
A3	List out the different types of image transformations.		CO1
A4	Define image segmentation. List out the types of segmentation.		CO2
A5	Differentiate image enhancement and image restoration.		CO1
Part B		$4 \times 5 = 20$	
B1	Tabulate the changes in coordinates during translation, reflection and rotation.		CO1
B2	Explain the salt and pepper noise and the filters to remove these noises from the images without affecting the quality of the image.		CO1
B3	“RGGROW algorithm is not good enough for texture based segmentation” – Justify this statement and explain the step-by-step process of the algorithm.		CO2
B4	Explain the step by step process to apply Fourier transformation on images.		CO1
Part C		$1 \times 5 + 1 \times 10 = 15$	
C1	Do histogram equalization on the following image which has 8 discrete pixel levels (0 - 7), transforming it into a histogram equalized image also with 8 discrete grey levels in the range (0-7).		CO3
	<pre> 1 1 1 1 1 1 1 1 0 2 5 5 5 5 2 0 0 3 2 6 7 2 3 0 0 3 3 2 2 3 3 0 0 2 3 2 2 3 3 0 0 3 2 4 4 2 4 0 0 2 6 4 4 4 2 0 1 1 1 1 1 1 1 1 </pre>		
	(OR)		
C2	An input image $g(x, y)$ is blurred by convolution by a blurring function $h(x, y)$ and then has random noise $n(x, y)$ added. Give a mathematical expression for the resulting image $f(x, y)$ . Describe how to restore the image using the Inverse Fourier filter, i.e. how to estimate $g(x, y)$ given $f(x, y)$ and knowing $h(x, y)$ . Give an expression for the noise distribution $n'(x, y)$ in the restored image.		CO3
C3	Given the histogram of an image. How will you segment your Region of Interest in that image using histogram-based clustering? Explain the steps involved in the above process.		CO3
	(OR)		
C4	How will you fix the optimal number of clusters to be identified in an image in K-Means algorithm? Is it possible to define the optimal number of clusters manually? What are the methods that are used for fixing the number of clusters?		CO3

\*\*\*\*\*



**THIAGARAJAR COLLEGE OF ENGINEERING, MADURAI 625 015.**  
**Department of Applied Mathematics and Computational Science**  
**Continuous Assessment Test – I**

Course Code	19DSPG0	Course Name	Soft Computing		
Degree	M.Sc.	Programme	Data Science (5 yrs. Integrated)	Semester	IX
Date	10/08/2023	Duration	90 minutes	Max. Marks	50
Faculty-in-Charge		Dr. ANITHA D			

**Answer All Questions**

**Part A**

$5 \times 2 = 10$

CO

CO1

CO4

CO4

CO1

CO2

A1. List down the different knowledge representation techniques

A2. Define heuristic search.

A3. Write down the pseudocode for Generate and Test search technique

A4. Identify the data structure used in Breadth first search and state the reason for the same.

A5. Present the diagram showing fuzzy inference system architecture.

**Part B**

$3 \times 5 = 15$

CO1

CO1

CO2

B1. Present the Depth First search algorithm with an example graph and give the time complexity

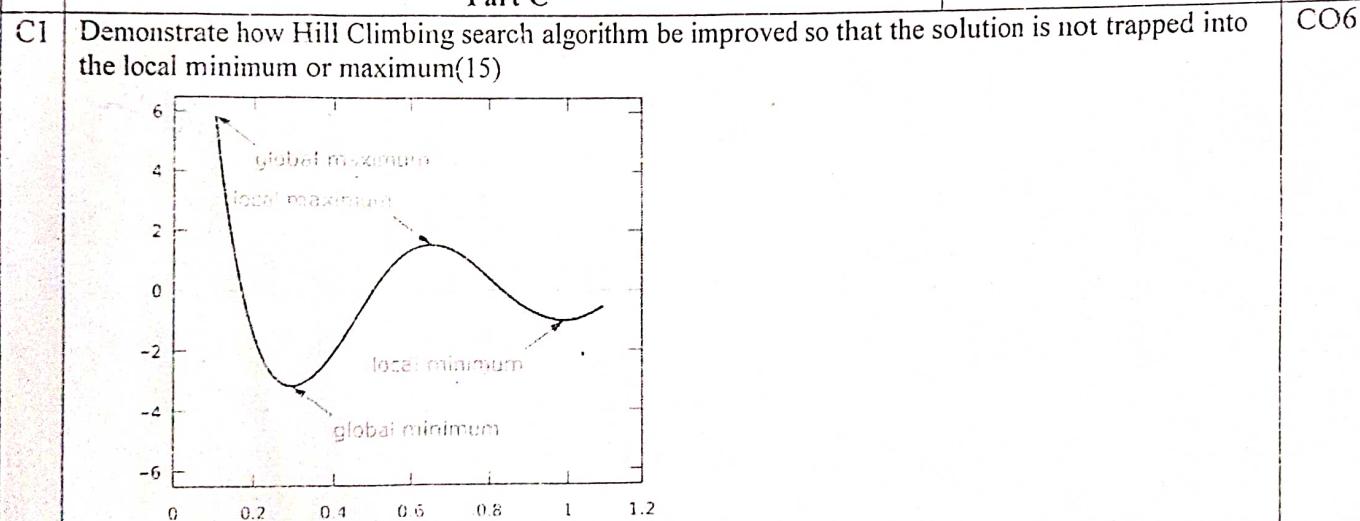
B2. Formulate an algorithm to find the cost function of a placement of queen in an NQueen's problem

B3. Explain any two membership functions that helps in fuzzification with examples for each

**Part C**

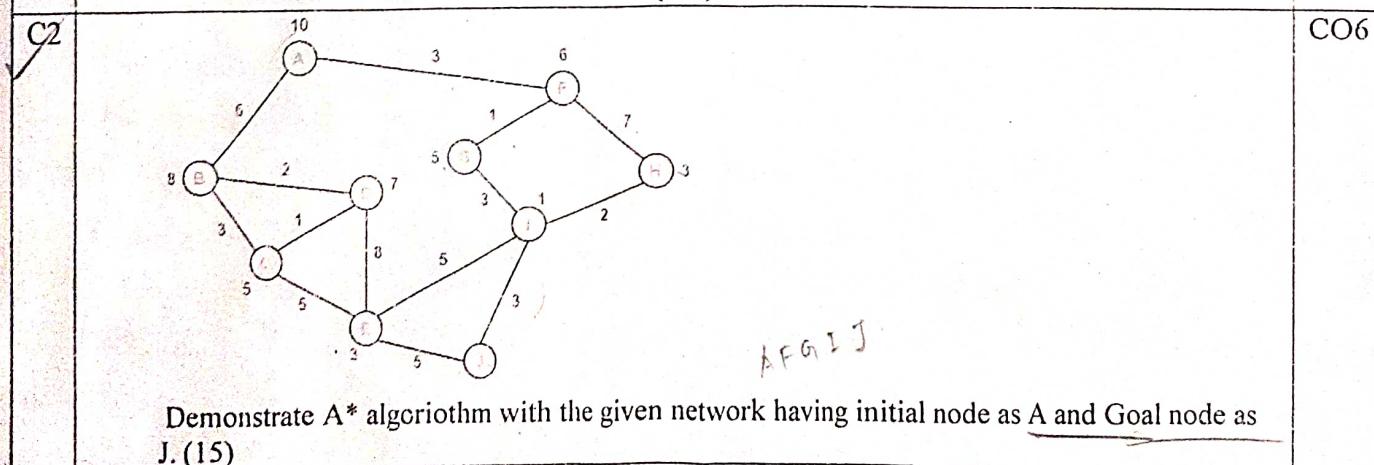
$1 \times 15 + 1 \times 10 = 25$

CO6



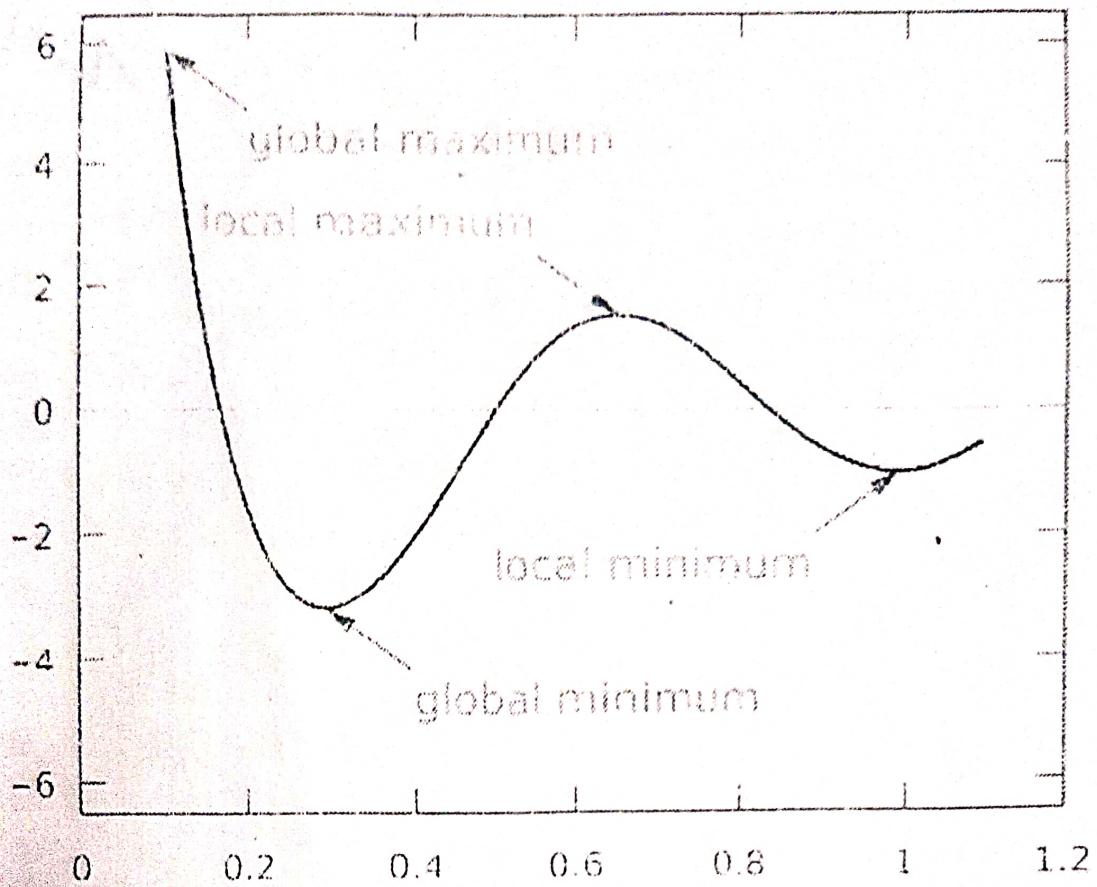
(OR)

CO6



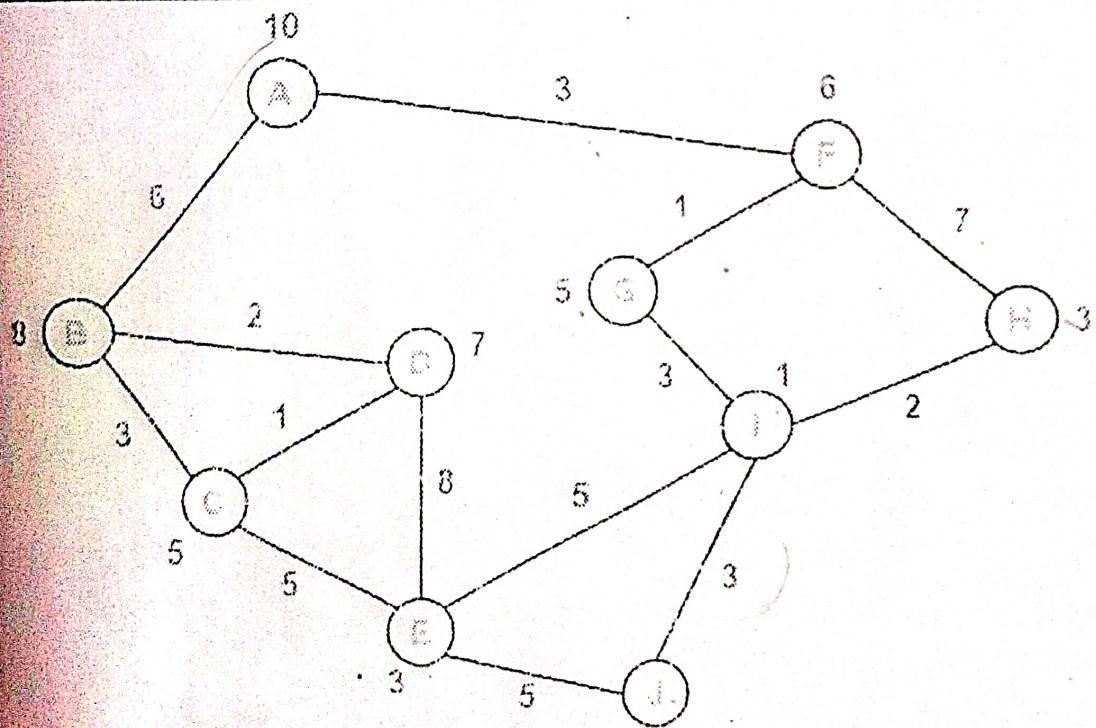
C1

Demonstrate how Hill Climbing search algorithm be improved to escape from the local minimum or maximum(15)



(OR)

C2

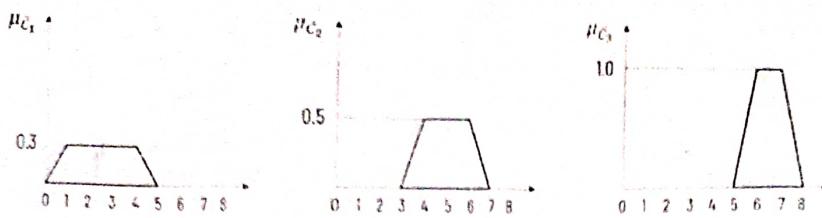


Demonstrate A\* algorioothm with the given network having J. (15)

C3

Consider the three fuzzy output sets, find the crisp value of the same with Center of sums and weighted average method. Present the need of this process with an example scenario (10)

CO4



(OR)

C4

Suppose you are a soils engineer. You wish to track the movement of soil particles under strain in an experimental apparatus that allows viewing of the soil motion. You are building pattern recognition software to allow a computer to monitor and detect the motions. You define two fuzzy sets on a universe of nondimensional particle areas,  $X = [0, 1, 2, 3, 4]$ :  $A \sim$  is a fuzzy set whose elements belong to the occlusion, and  $B \sim$  is a fuzzy set whose elements belong to inadequate segmentation. Let

$$A = \left\{ \frac{0.1}{0} + \frac{0.4}{1} + \frac{1}{2} + \frac{0.3}{3} + \frac{0.2}{4} \right\}$$

$$B = \left\{ \frac{0.2}{0} + \frac{0.5}{1} + \frac{1}{2} + \frac{0.4}{3} + \frac{0.1}{4} \right\}$$

Find the following:

(a)  $A \cup B$  *met*

(b)  $A \cap B$  *min*

(c)  $\bar{A}$

(d)  $\bar{B}$

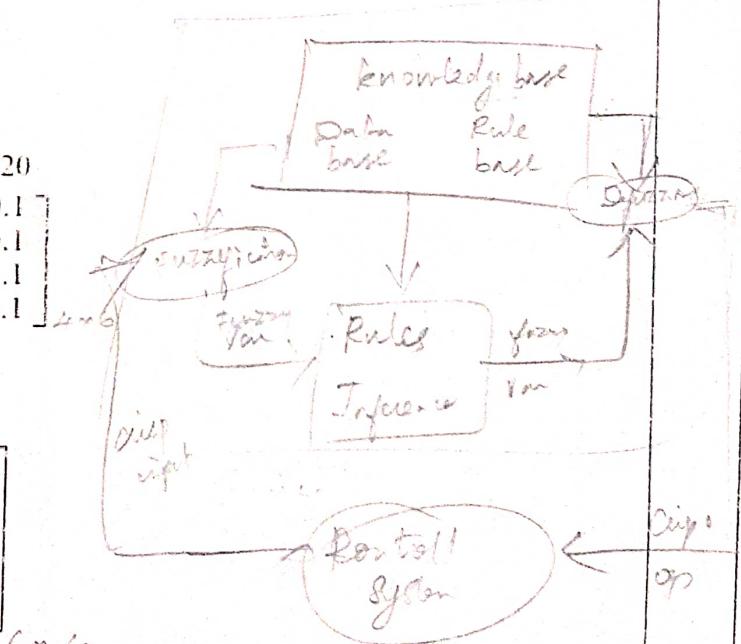
(e)  $\overline{A \cap B} = \bar{A} \cup \bar{B}$

(f)  $\overline{A \cup B} = \bar{A} \cap \bar{B}$  (6)

$$R = \begin{bmatrix} & 20 & 40 & 60 & 80 & 100 & 120 \\ 30 & 0.2 & 0.3 & 0.3 & 0.3 & 0.3 & 0.1 \\ 60 & 0.2 & 0.4 & 0.6 & 0.7 & 0.7 & 0.1 \\ 100 & 0.2 & 0.4 & 0.6 & 0.8 & 1 & 0.1 \\ 120 & 0.2 & 0.2 & 0.2 & 0.2 & 0.2 & 0.1 \end{bmatrix}$$

$$S = \begin{bmatrix} & 500 & 1000 & 1500 & 1800 \\ 20 & 0.2 & 0.2 & 0.2 & 0.15 \\ 40 & 0.33 & 0.4 & 0.4 & 0.15 \\ 60 & 0.33 & 0.6 & 0.6 & 0.15 \\ 80 & 0.33 & 0.67 & 0.8 & 0.15 \\ 100 & 0.33 & 0.67 & 1 & 0.15 \\ 120 & 0.1 & 0.1 & 0.1 & 0.1 \end{bmatrix}$$

Find min max composition (4)



6 x 4



## Continuous Assessment Test - I

Course Code	19DSPE0	Course Name	Marketing Analytics		
Degree	M.Sc	Programme	M.Sc DS	Semester	IX
Date	11/08/2023	Duration	90 Minutes	Max. Marks	50
Faculty-in-Charge		Dr.T.Chandrasekharan			

Answer All Questions

Part A		2 x 2.5 = 5
A1.	Define Model and Metrics. CO1	
A2.	List the factors driving adoption of Marketing Analytics. CO1	

Part B		3 x 5 = 15
B1.	Summarize the advantages of Marketing Analytics. CO1	
B2.	Define Hierarchy of Market definitions. CO1	
B3.	How do you estimate the Market size? CO1	

Part C		2 x 15 = 30
C1	a) How do you express models in different styles? Illustrate with example. (10) CO2 b) Illustrate the metrics family and dashboard for your organization.(5)  (OR)	
C2	a) Identify examples of descriptive, predictive, and normative models used by your organization. (10) CO2 b) What is market segmentation, Show its advantages.(5)	
C3	Illustrate in detail about Market Sizing. .(15) CO2  (OR)	
C4	Explain the two Market force analysis tool with a neat sketch. .(15)CO2	



THIAGARAJAR COLLEGE OF ENGINEERING, MADURAI 625 015  
DEPARTMENT OF APPLIED MATHEMATICS AND COMPUTATIONAL SCIENCE

Continuous Assessment Test – II

Course Code	19DS910	Course Name	Web Analytics		
Degree	M.Sc.	Programme	Data Science	Semester	IX
Date	12/09/2023	Duration	90 Minutes	Max. Marks	50
Faculty-in-Charge	Mr.B.Ramprakash				

Answer All Questions

Part A		05x2=10Marks	CO
A1.	Explain the primary goal of conducting a heuristic evaluation in web analytics.		CO1
A2.	Define: Bounce Rate .		CO2
A3.	Describe advantages and limitations of using surveys as a data collection method in web analytics.		CO2
A4.	What is the meaning of user retention when considering web analytics, and why is it a crucial metric for websites?		CO4
A5.	Explain the term "site visits" means in the context of web analytics, and briefly describe its significance as a metric for websites.		CO5
Part B		02 x 5=10 Marks	
B1.	Examine the essential factors to take into account and recommended methods when creating a survey aimed at gaining insights through web analytics.		CO2
B2.	Explain how acquisition and conversion are measured in web analytics and their role in understanding user behaviour and website performance.		CO6
Part C		02x15= 30Marks	CO
C1.	As a web analyst for a movie streaming website, your company is launching a new feature that enables users to create wish lists of movies. Your task is to design a survey to gather insights into user preferences and the feature's effectiveness. Outline a comprehensive survey plan for the scenario.		CO5

Contd...

**OR**

C2.	You are the web analyst for an online fitness platform that recently introduced a personalized workout recommendation feature. The management wants to assess how users are engaging with this feature and whether it positively impacts their fitness journey. Design a survey plan to collect relevant data and insights.	CO5
-----	---	-----

C3.	As the web analytics lead for a popular news website experiencing a decline in user engagement and ad revenue. The executive team is concerned and wants you to propose a strategy to improve the situation. Design a comprehensive plan for applying KPIs to diagnose the website's issues and formulate a strategy for recovery	CO6
-----	---	-----

**OR**

C4.	The e-commerce company has recently launched a new marketing campaign and wants to evaluate its success using KPIs. You are the web analytics manager for an e-commerce company that sells fashion products online. You have access to the website's analytics data. Please create a comprehensive plan to apply relevant KPIs to assess the impact of the marketing campaign.	CO6
-----	--	-----

**OR**

\*\*\*\*\*



THIAGARAJAR COLLEGE OF ENGINEERING, MADURAI 625 015  
DEPARTMENT OF APPLIED MATHEMATICS AND COMPUTATIONAL SCIENCE

Continuous Assessment Test – II

Course Code	19DS920	Course Name	Natural Language Processing		
Degree	M.Sc.	Programme	Data Science	Semester	IX
Date	12/09/2023	Duration	90 Minutes	Max. Marks	50
Faculty-in-Charge	Dr.P.Sharmila				

Answer All Questions

Part A		4 x 2 = 8 Marks
A1.	Differentiate open class words and closed class words with examples	CO2
A2.	State the advantages of bottom-up chart parser compared to top-down parsing	CO3
A3.	Differentiate Constituency parsing and Dependency Parsing	CO3
A4.	Discuss how Markov model applied in google search engine	CO3
Part B		2 x 6 = 12 Marks
B1.	Explain transition-based dependency parser with example.	CO3
B2.	Explain Attachment ambiguity and Coordination Ambiguity with an proper example	CO3
Part C		1 x 15 = 15 Marks
C1.	Determine the POS tag using Viterbi algorithm for the sentence "Will will google campus" (7 marks) Given: Nitish loves campus N V N Can Nitish google campus M N V N Will Ankita google campus M N V N Ankita loves will N V N Will loves google N V N Construct look up table (2 marks), transition matrix (3 marks) and emission matrix (3 marks) for the given data	1 x 15 = 15 Marks CO3
	OR	
	Contd...	

C2.	Parse the given sentence "Book the flight through Houston" using probabilistic CKY parser	CO3
	$S \rightarrow NP VP \quad 0.8$ $S \rightarrow X1 VP \quad 0.1$ $X1 \rightarrow Aux NP \quad 1.0$ $S \rightarrow book \mid include \mid prefer \quad 0.01 \quad 0.004 \quad 0.006$ $S \rightarrow Verb NP \quad 0.05$ $S \rightarrow VP PP \quad 0.03$ $NP \rightarrow I \mid he \mid she \mid me \quad 0.1 \quad 0.02 \quad 0.02 \quad 0.06$ $NP \rightarrow Houston \mid NWA \quad 0.16 \quad .04$ $Det \rightarrow the \mid a \mid an \quad 0.6 \quad 0.1 \quad 0.05$ $NP \rightarrow Det Nominal \quad 0.6$ $Nominal \rightarrow book \mid flight \mid meal \mid money \quad 0.03 \quad 0.15 \quad 0.06 \quad 0.06$ $Nominal \rightarrow Nominal Nominal \quad 0.2$ $Nominal \rightarrow Nominal PP \quad 0.5$ $Verb \rightarrow book \mid include \mid prefer \quad 0.5 \quad 0.04 \quad 0.06$ $VP \rightarrow Verb NP \quad 0.5$ $VP \rightarrow VP PP \quad 0.3$ $Prep \rightarrow through \mid to \mid from \quad 0.2 \quad 0.3 \quad 0.3$ $PP \rightarrow Prep NP \quad 1.0$	

OR

C3.	Generate any two possible sequences with maximum of 3 sentences using <u>Markov Model</u> from the given data.	CO4
	<p>"I love Markov chains."</p> <p>"Markov chains love me."</p> <p>"Markov is I."</p> <p>a. Look-Up Table (4 marks)</p> <p>b. Transition Table (5 marks)</p> <p>c. Sentence Generation (6 marks)</p>	

OR

C4.	Construct Dependency Parsing for the paraphrased sentences given below:	CO4
	<p>a. I have a strong affinity for consuming coffee. (5 marks)</p> <p>b. Coffee is one of my favorite drinks. (5 marks)</p> <p>c. Discuss the result in terms of similarities among the three sentences (5 marks)</p>	

\*\*\*\*\*



Course Code	19DS930	Course Name	Computer Vision	Semester	IX
Degree	M.Sc.	Programme	M.Sc. Data Science	Max. Marks	50
Date	13/09/2023	Duration	Slot 1 - 90 minutes		
Faculty-in-Charge		Prof S.T. Padmapriya			

**Answer All Questions**

Part A		5 x 2 = 10	CO
A1	What are the three types of discontinuity in digital image?		CO2
A2	What is the advantage of using sobel operator?		CO2
A3	Give the properties of the second derivative around an edge.		CO2
A4	What is meant by object point and background point?		CO2
A5	Define object detection.		CO2

Part B		4 x 5 = 20	
B1	Write short notes on Canny Edge Detection method.		CO2
B2	Explain the types of Graph-Cut segmentation methods with an example.		CO2
B3	Differentiate LoG and DoG edge detection methods using appropriate formulations.		CO2
B4	Detail the process of texture based segmentation.		CO2

Part C		2x10= 20	
C1	Apply region growing on the following image with seed point as 6 and threshold value as 3. Show the resultant segmented image. (10)		CO4

*After connectivity*

1 iteration

5 6 6 7 6 7 6 6  
 6 7 6 7 5 5 4 7  
 6 6 4 4 3 2 5 6  
 5 4 5 4 2 3 4 6  
 0 3 2 3 3 2 4 7  
 0 0 0 0 2 2 5 6  
 1 1 0 1 0 3 4 4  
 1 0 1 0 2 3 5 4

**(OR)**

C2	Apply 3 X 3 low pass averaging mask, high pass averaging mask and median filter on a given image patch by considering two cases. i. Ignore border and modify 3 X 3 sub image (5) ii. Zero padding. (5)	CO4
----	--	-----

The image patch is

10 25 30 24 15  
 15 20 19 21 18  
 12 30 25 30 22  
 17 27 28 26 18  
 20 24 21 25 10

C3	Given set of points. Use Hough Transform to join these points. (10) A(1,4) B(2,3) C(3,1) D(4,1) E(5,0)	CO3
----	---	-----

**(OR)**

C4	Given matrix Ix and Iy. Ix:	CO3
----	--------------------------------	-----

1 2 3 4

1	1	3	6	5
2	12	2	11	3
3	9	2	4	10
4	3	7	1	2
5	7	2	7	8

2 ah 4 ah ✓

Iy:

1	7	3	1	7
2	9	2	5	3
3	7	1	2	2
4	5	9	3	3
5	4	3	1	2

3 ah 2 ah

Find whether the point (3,3) is corner or not using Harris Corner detection. The threshold value T = 5000. (10)

\*\*\*\*\*

R.

19S035



**THIAGARAJAR COLLEGE OF ENGINEERING, MADURAI 625 015.**  
**Department of Applied Mathematics and Computational Science**  
**Continuous Assessment Test – II**

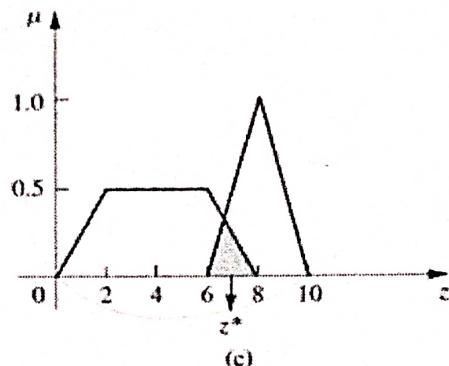
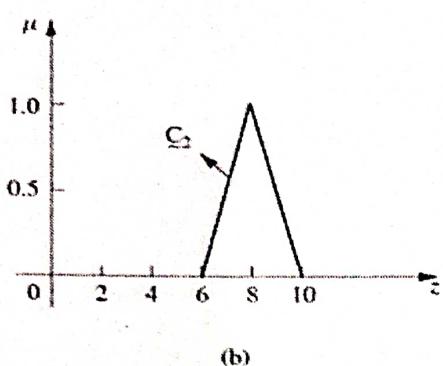
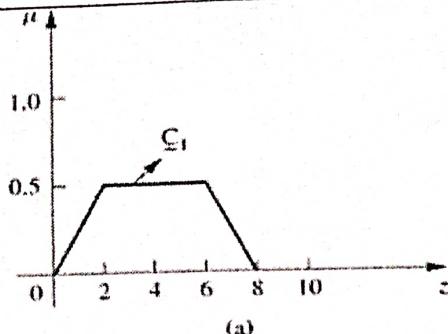
Course Code	19DSPG0	Course Name	Soft Computing		
Degree	M.Sc.	Programme	Data Science (5 yrs. Integrated)	Semester	IX
Date	14/09/2023	Duration	90 minutes		Max. Marks
Faculty-in-Charge		<b>Dr. ANITHA D</b>			

**Answer All Questions**

Part A		$4 \times 5 = 20$	CO
A1	Describe the steps of fuzzy reasoning with fired rules.		CO2
A2	Give two functions for dilation and concentration each. What is intensification?		CO2
A3	Give a brief note on Sugeno fuzzy models.		CO2
A4	Explain the extension principle with examples.		CO2

Part B		$2 \times 15 = 30$																	
B1	Consider a fuzzy logic controller is used to control the speed of a motor by changing its input voltage (V) according to two input variables; speed (SP), and speed change rate SC. Let the fuzzy set of SP be {Slow (S), Normal (N), Fast (F)}, and the fuzzy set for SC be {Low (L), Medium (M), High (H)}, and for the control action be {Slow Down (DN), No Change (NC), Speed Up (Up)}, where, (SP $\in$ [500, 1000]), (SC $\in$ [0, 10]), and (V $\in$ [2, 3]) with step = 0.1. The given table shows the rules. Explain the triggering of these rules with suitable input variables and appropriate membership functions for fuzzification and functions for defuzzification.		CO6																
	<p style="text-align: center;"><b>Table 2</b></p> <table border="1" style="margin-left: auto; margin-right: auto;"> <tr> <td><math>\wedge</math></td><td>S</td><td>N</td><td>F</td></tr> <tr> <td>L</td><td>Up</td><td>NC</td><td>NC</td></tr> <tr> <td>M</td><td>Up</td><td>NC</td><td>NC</td></tr> <tr> <td>H</td><td>NC</td><td>DN</td><td>DN</td></tr> </table>			$\wedge$	S	N	F	L	Up	NC	NC	M	Up	NC	NC	H	NC	DN	DN
$\wedge$	S	N	F																
L	Up	NC	NC																
M	Up	NC	NC																
H	NC	DN	DN																
	(OR)																		
B2	Consider a Project management application that accepts two inputs : Project funding and Project Staffing that gives an output as profitable. Design a fuzzy logic decision system to find the percentage of profit.		CO6																

B3



The following image shows the Center of Area defuzzification method for the *Steering Angle z* output linguistic variable of a vehicle maneuvering fuzzy system. Demonstrate how the Center of Gravity method reach the given defuzzified result. Also find the same with Weighted Average method.

(OR)

B4

CO4

$$\tilde{A} = \text{temperature of input steam is hot} = \left\{ \frac{0}{175} + \frac{0.7}{180} + \frac{1}{185} + \frac{0.4}{190} \right\}$$

$$\tilde{B} = \text{separation of mixture is good} = \left\{ \frac{0}{89} + \frac{0.5}{92} + \frac{0.8}{95} + \frac{1}{98} \right\}$$

$$\tilde{A}' = \left\{ \frac{1}{170} + \frac{0.8}{175} + \frac{0.5}{180} + \frac{0.2}{185} \right\}$$

- Find the new  $B'$  given  $A'$  (8)
- What happens if the temperature of the input steam is very hot? (7)

$A' \text{ OR } B' \text{ or } \tilde{A}' \cup (\tilde{A} \times \tilde{B})$

$(\tilde{A} \times \tilde{B}) \cup (\tilde{A}' \times \tilde{B}')$



THIAGARAJAR COLLEGE OF ENGINEERING, MADURAI 625 015.

Department of AMCS (Data Science)

Continuous Assessment Test – II

Course Code	19DSPE0	Course Name	Marketing Analytics		
Degree	M.Sc DS	Programme	M.Sc DS	Semester	IX
Date	15/09/2023	Duration	90 Minutes	Max. Marks	50
Faculty-in-Charge		Dr.T.Chandrasekharan			

Answer All Questions

Part A		2 x 2.5 = 5
A1.	Show the benefits of Competitive Analysis. CO2	
A2.	How Analytics approach is used to Organizational strategy decision. CO3	

Part B		2 x 5 = 10
B1.	What are the different methods to market to target segments CO3	
B2.	Summarize the Core Competency based on organizational functions. CO4	

Part C		2 x 17.5 = 35
C1	a) How do you perform causal Analysis Forecasting? (10) CO3 b) Illustrate the potential outcomes of the strategic scenario in detail.(7.5) (OR)	
C2	a) Give an explanation of the financial and non-financial measures that a business uses. (10) CO3 b) Describe how the Trial rate Forecasting Method works.(7.5)	
C3	Provide a concrete illustration of the competitive analysis process with your own example. (17.5) CO4 (OR)	
C4	Provide a comprehensive breakdown of the Strategic Decision Models.(17.5) CO4	



19S035

Continuous Assessment Test – III

Course Code	19DS910	Course Name	Web Analytics		
Degree	M.Sc.	Programme	Data Science	Semester	IX
Date	17/10/2023	Duration	90 Minutes	Max. Marks	50
Faculty-in-Charge	Mr.B.Ramprakash				

Answer All Questions

Part A		02x2.5=5Marks	CO
A1.	Give the importance of user behaviour analysis in web analytics and with real-world example.		CO1
A2.	Describe Omniture features and advantages for web analytics and its role in aiding businesses in tracking.		CO4
Part B		01 x 5=5 Marks	
B1.	Explain the significance of the key features in Google Analytics.		CO
Part C		02x12.5= 25Marks and 01x15=15marks	CO
C1.	As a web analytics manager for a multimedia-rich website, explain the significance of tracking non-HTML files in Google Analytics. Discuss tracking methods for non-HTML content, provide implementation guidance, apply the resulting data for content insights, and offer recommendations for optimizing user engagement and content strategy.		CO6
OR			
C2.	You are an analyst for an online publication, explain the significance of tracking outbound links' impact on user engagement. Apply the importance of outbound link tracking in Google Analytics, detail tracking methods, and offer recommendations for enhancing user engagement via optimized outbound linking strategies.		CO6

C3.	For a non-profit organization with a donation website. Conduct qualitative interviews with past donors to understand their motivations for contributing and their impressions of the donation process. Apply the qualitative data to identify key factors influencing donor decisions and propose strategies to increase donations.	CO2
OR		
C4.	As a digital marketing analyst for an online travel agency. Your task is to optimize advertising campaigns to maximize the return on investment (ROI). Analyze click-through rate (CTR), cost per click (CPC), and conversion rate data for various advertising channels. Use quantitative analysis to identify the most cost-effective channels and propose strategies to allocate the advertising budget for maximum ROI.	CO2
OR		
C5.	As a web analyst for an online news website tasked with improving user engagement. Analyze visitor clicks and web channel data to identify popular content types and assess the effectiveness of various traffic sources. Provide insights on content optimization and channel targeting to increase user engagement and readership, emphasizing the practicality and potential impact of your recommendations.	CO5
C6.	You're a web analytics consultant for a financial institution tasked with analyzing web channel data for online banking services while ensuring data privacy and security compliance. Analyze user behaviour patterns, assess privacy and security implications, evaluate institutional policies, and provide recommendations for optimizing data analysis while safeguarding customer data and complying with regulations.	CO5

\*\*\*\*\*



**Continuous Assessment Test – III**

I b - 7

Course Code	19DS920	Course Name	Natural Language Processing		
Degree	M.Sc.	Programme	Data Science	Semester	IX
Date	17/10/2023	Duration	90 Minutes	Max. Marks	50
Faculty-in-Charge		Dr. P Sharmila			

**Answer All Questions**

Part A		$5 \times 2 = 10$ Marks
A1.	What is Wordnet?	CO5
A2.	What is Coreference and Anaphora in linguistics?	CO5
A3.	List the challenges of AI's Chat bot for Language Translation	CO5
A4.	Why do we need semantic role labelling when there's already parsing?	CO6
A5.	What is subjectivity and polarity in sentiment analysis	CO6
Part B		$2 \times 5 = 10$ Marks
B1.	Explain the dictionary based approach – LESK algorithm for WSD	CO5
B2.	What is the difference between pragmatics and discourse analysis?	CO5
Part C		$1 \times 8 = 8, 1 \times 12 = 12, 1 \times 10 = 10$
C1.	Identify Thematic roles for the following sentences: (8 marks) “John planted the tree. The water in the pot boiled quickly. Mary listened to the music coming from the park. John made Peter sweep the garage”	CO5
OR		
C2.	Explain how unsupervised learning algorithm using concept clustering to overcome Word Sense Disambiguation. (8 marks)	CO5
C3.	a) Explain and differentiate the challenges in semantic analysis with examples (5 marks) b) Discuss Aspect-based Sentiment Analysis for the given Sentence: (7 marks) “Our stay at the hotel was nice. The staff was friendly and the rooms were clean, but our beds were quite uncomfortable.”	CO6
OR		
Cont...		

What is PNL in supervised Resolution classifier and how it is applied for the sentence: (5 marks)

CO6

- C4. "John took his license when he was 18.  
He passed his exam at his first attempt."  
Also apply Hobbs algorithm for the above sentences (7 marks)

Why Pragmatics Matters in Multilingual Chat bots? (3 marks)

CO4

C5. Discuss both generative and retrieval-based chat-bots with appropriate scenarios for both open and closed domains. (7 marks)

OR

Discuss Named Entity in Resume Filtering in terms of (10 marks)

CO4

- C6
- a. Ambiguity and Abbreviations
  - b. OoV words
  - c. Spelling Variations

\*\*\*\*\*



**THIAGARAJAR COLLEGE OF ENGINEERING, MADURAI 625 015.**  
**Department of Applied Mathematics and Computational Science**  
**Continuous Assessment Test – III**

Course Code	19DSPG0	Course Name	Soft Computing		
Degree	M.Sc.	Programme	Data Science (5 yrs. Integrated)	Semester	IX
Date	19/10/2023	Duration	90 minutes		
Faculty-in-Charge	Dr. ANITHA D				

**Answer All Questions**

**Part A**

**5 x 2 = 10**

**CO**

- |     |   |     |
|-----|---|-----|
| A1. | List any two short comings of traditional optimization methods.                     | CO5 |
| A2. | Match any four genetic characteristics with that of the genetic algorithm features. | CO1 |
| A3. | Enumerate any four methods of genetic representations.                              | CO1 |
| A4. | List any four stopping criteria for the Genetic algorithm off spring generation.    | CO3 |
| A5. | Define Elitism in Genetic Algorithm. Present its significance.                      | CO3 |

**Part B**

**2x5 = 10**

**CO**

- |    |   |     |
|----|---|-----|
| B1 | Explain how Travelling salesman problem be approached with Genetic algorithm.                     | CO5 |
| B2 | Write the algorithm for Boltzmann selection and Rank selection methods of selecting best parents. | CO3 |

**Part C**

**1x 20 + 1 x 10 = 30**

**CO5**

- |    |  |     |
|----|--|-----|
| C1 | Perform three generations of binary coded genetic algorithm to solve the following optimization problem. Maximize $f(x) = x^3 - 5x^2$ , $0 \leq x \leq 63$ , x is an integer.<br>Use Roulette wheel selection, single point crossover, value summation mutation method and population size of six (20) | CO5 |
|----|--|-----|

**(OR)**

- |    |  |     |
|----|--|-----|
| C2 | Perform three generations of binary coded genetic algorithm to solve the following optimization problem. Maximize $f(x) = x^2 - 5x$ , $0 \leq x \leq 63$ , x is an integer<br>Use Tournament selection, multi point crossover, parity summation mutation and population size of six (20) | CO5 |
|----|--|-----|

- |    |   |     |
|----|---|-----|
| C3 | Formulate a genetic algorithm solution to 0/1 knapsack problem. Demonstrate the generation of offsprings with appropriate genetic operators (10)<br><b>(OR)</b> | CO6 |
|----|---|-----|

- |    |   |    |   |   |   |   |   |   |    |   |   |   |   |   |   |    |   |   |   |   |   |   |    |   |   |   |   |   |   |     |
|----|---|----|---|---|---|---|---|---|----|---|---|---|---|---|---|----|---|---|---|---|---|---|----|---|---|---|---|---|---|-----|
| C4 | <table border="1"> <tr> <td>x1</td> <td>1</td><td>0</td><td>1</td><td>0</td><td>0</td><td>0</td> </tr> <tr> <td>x2</td> <td>0</td><td>0</td><td>1</td><td>0</td><td>0</td><td>1</td> </tr> <tr> <td>x3</td> <td>1</td><td>0</td><td>0</td><td>0</td><td>0</td><td>1</td> </tr> <tr> <td>x4</td> <td>0</td><td>0</td><td>1</td><td>0</td><td>1</td><td>0</td> </tr> </table> <p>Genotype</p> <p>Chromosome</p> <p>Population</p> <p>With the given population assuming your own objective function, demonstrate the application of three mutation operators and three different cross-over operators. (10)</p> | x1 | 1 | 0 | 1 | 0 | 0 | 0 | x2 | 0 | 0 | 1 | 0 | 0 | 1 | x3 | 1 | 0 | 0 | 0 | 0 | 1 | x4 | 0 | 0 | 1 | 0 | 1 | 0 | CO6 |
| x1 | 1   | 0  | 1 | 0 | 0 | 0 |   |   |    |   |   |   |   |   |   |    |   |   |   |   |   |   |    |   |   |   |   |   |   |     |
| x2 | 0   | 0  | 1 | 0 | 0 | 1 |   |   |    |   |   |   |   |   |   |    |   |   |   |   |   |   |    |   |   |   |   |   |   |     |
| x3 | 1   | 0  | 0 | 0 | 0 | 1 |   |   |    |   |   |   |   |   |   |    |   |   |   |   |   |   |    |   |   |   |   |   |   |     |
| x4 | 0   | 0  | 1 | 0 | 1 | 0 |   |   |    |   |   |   |   |   |   |    |   |   |   |   |   |   |    |   |   |   |   |   |   |     |



19 So 38

**THIAGARAJAR COLLEGE OF ENGINEERING, MADURAI 625 015.**  
**Department of Applied Mathematics and Computational Science**  
**Continuous Assessment Test – III**

Course Code	19DS930	Course Name	Computer Vision		
Degree	M.Sc.	Programme	M.Sc. Data Science	Semester	IX
Date	18/10/2023	Duration	Slot 1 - 90 minutes	Max. Marks	50
Faculty-in-Charge Prof S.T. Padmapriya					

**Answer All Questions**

**Part A**

$5 \times 2 = 10$  CO

A1	What is the primary difference between LDA and ICA?	CO6
A2	What is the role of the 'k' parameter in K-Nearest Neighbors?	CO6
A3	What distinguishes model-based object recognition methods from appearance-based methods?	CO4
A4	What is the main challenge in determining the number of clusters in a dataset for clustering algorithms like k-means?	CO6
A5	What does the term 'centroid' refer to in the context of clustering?	CO6

**Part B**

$4 \times 5 = 20$

B1	Explain the fundamental differences between structural, model-based, appearance-based, and shape-based methods in object recognition.	CO4
B2	Explain the Bag of Words (BoW) model in the context of object recognition.	CO4
B3	What are the advantages and limitations of the Bayes classifier compared to the KNN algorithm in terms of classification tasks?	CO6
B4	Explain the primary objective of dimensionality reduction techniques such as PCA, LDA, and ICA.	CO6

**Part C**

$2 \times 10 = 20$

C1	Can you provide practical examples or scenarios where using Artificial Neural Networks may not be the most suitable approach for image-related tasks? Explain the limitations of ANNs in these cases and suggest alternative techniques or models that could be more effective.	CO5
----	---	-----

**(OR)**

C2	Imagine you are tasked with diagnosing skin diseases from images using an Artificial Neural Network model. How would you design the ANN architecture, and what preprocessing steps would you apply to the image dataset to optimize the model's performance for this specific computer vision problem?	CO5
----	--	-----

C3	For a facial recognition system in a large organization, discuss the supervised and semi supervised techniques you would employ to improve accuracy while considering data privacy concerns. How can labeled and unlabeled data be used effectively in this context?	CO6
----	--	-----

**(OR)**

C4	Use the k-means algorithm and Euclidean distance to cluster the following 8 examples into 3 clusters: <u>A1=(2,10)</u> , <u>A2=(2,5)</u> , <u>A3=(8,4)</u> , <u>A4=(5,8)</u> , <u>A5=(7,5)</u> , <u>A6=(6,4)</u> , <u>A7=(1,2)</u> , <u>A8=(4,9)</u> . Suppose that the initial seeds (centers of each cluster) are A1, A4 and A7. Run the k-means algorithm and show: a) The new clusters (i.e. the examples belonging to each cluster) (3) b) The centroids of the new clusters. (2) d) How many more iterations are needed to converge? Show the iterations. (5)	CO6
----	--	-----

\*\*\*\*\*



## Department of AMCS (Data Science)

### Continuous Assessment Test – III

Course Code	19DSPE0	Course Name	Marketing Analytics		
Degree	M.Sc DS	Programme	M.Sc DS	Semester	IX
Date	20/10/2023	Duration	90 Minutes	Max. Marks	50
Faculty-in-Charge	Dr.T.Chandrakumar				

Answer All Questions

Part A		$2 \times 2.5 = 5$
A1	Identify the important features of Conjoint analysis.	CO4
A2	Name the different types of Pricing techniques.	CO5

Part B		$2 \times 5 = 10$
B1.	Show the drawbacks associated with Decision Tree models.	CO5
B2.	Give an example for Tiered Pricing.	CO6

Part C		$2 \times 17.5 = 35$
C1	a) Show the three step process for selecting a new retail store location in detail with example (10) b) How is the internet changing the distribution of products? (7.5)  (OR)	CO5
C2	a) In the Ecommerce Sales model, there are a lot of distinct inputs and outputs, but how do they all function together? (10) b) Provide an overview of the several popular ways of promotion, including some instances. 7.5)	CO5
C3	Provide a concrete illustration of the promotional metrics for Traditional and Social media. (17.5)  (OR)	CO6
C4	Illustrate with an example of three popular models to assess the impact of pricing on organizational goals. (17.5)	CO6

193035

Name of the Candidate:

Reg. No.:

**M.Sc., DEGREE - 19DS910  
NOVEMBER 2023- EXAMINATIONS  
BRANCH : DATA SCIENCE  
WEB ANALYTICS**

Duration : 3 Hours

Maximum: 100 Marks

**Answer All Questions**

**PART - A**

**(10 x 2 = 20 )**

CO	Marks
CO1	(2)
CO1	(2)

- A1.** What is the significance of bounce rate in web analytics? **CO1** (2)
- A2.** How does Google Analytics facilitate the identification of the outbound links that receive the highest click-through rates on a website? **CO1** (2)
- A3.** What type of data can be captured using page tags in web analytics? **CO1** (2)
- A4.** How can clickstream data help in improving website usability? **CO2** (2)
- A5.** What metric in web analytics is commonly used to measure user retention over time? **CO2** (2)
- A6.** What is the primary differentiation between first-party tracking and third-party tracking within the domain of web analytics? **CO3** (2)
- A7.** What information regarding user behavior can be extracted through the analysis of site visits in web analytics? **CO3** (2)
- A8.** What does the term "conversion rate" measure in web analytics, and how is it calculated? **CO4** (2)
- A9.** How is click tracking done in web analytics, and what data does it collect? **CO4** (2)
- A10.** Give key features of Google Analytics that allows you to track the sources of website traffic. **CO6** (2)

**PART - B**

**(4 x 5 = 20 )**

CO	Marks
CO1	(5)

- B1.** Give primary heuristics frequently applied in heuristic evaluations for web analytics and their role in identifying usability problems on a website. **CO1** (5)
- B2.** How metrics like sessions, page views, and bounce rates in Google Analytics are utilized to assess user behavior on a website. **CO2** (5)
- B3.** Explain the purpose and methodology of survey-based web analytics and highlight its distinctions from traditional tools like Google Analytics. **CO4** (5)
- B4.** Explore the importance of establishing goals and conversions in Google Analytics and optimizing website conversion rates. **CO4** (5)

**PART - C** ( 60 Marks)

- C1.** As a web analyst at a streaming service. Examine clickstream data from the last six months to determine the most-watched genres of content, the peak hours of user activity, and which devices users prefer for streaming. Suggest strategies to enhance content offerings and optimize user experience. **CO2** (12.5)

**(OR)**

- C2.** You work as a web analyst for a job search website. Analyze the clickstream data from the past year to understand user behavior and preferences. Propose strategies for enhancing job recommendations, increasing user engagement, and ultimately improving the job search experience based on your analysis. **CO2** (12.5)

- C3.** For a leading online travel agency, you are tasked with optimizing the website's performance during peak traffic periods, such as holiday seasons and special promotions. The site has experienced slowdowns and occasional outages during these high-traffic events, resulting in user dissatisfaction. Propose a comprehensive strategy to enhance web server performance, ensuring a smooth user experience, site stability, and maintaining the agency's reputation during critical high-traffic periods. **CO3** (12.5)

**(OR)**

**CONTD.,**

- C4.** Imagine you are the web analytics manager for an online booking platform specializing in event tickets. Evaluate the effectiveness of the recent email marketing campaign promoting a major concert. Analyze the campaign data, identify opportunities for improvement, and propose data-driven strategies to enhance email marketing efforts for better ticket sales and user engagement. **CO3 (12.5)**
- C5.** You are a data analyst for a music streaming service with a diverse user base. Evaluate the impact of the platform's music recommendation algorithms on user engagement and retention. Suggest KPIs to quantify the success of the recommendation system. Analyze user data, such as listening history and user-generated playlists, to improve music suggestions. **CO5 (12.5)**
- C6.** You work as a web analytics manager for a subscription-based streaming service. The company wants to assess the effectiveness of its content recommendation algorithm. Select KPIs to measure recommendation success and retention. Explain data collection processes and how these KPIs will help refine content suggestions. **CO5 (12.5)**
- C7.** As a web analyst for an e-learning platform, the company wants to gather insights on the effectiveness of their course content. Create a survey plan to assess user satisfaction with course materials, instructional design, and overall learning experience. Define the survey objectives, audience, data collection methods, question types, and explain how the survey data will drive content improvements. **CO5 (10)**
- C8.** As a web analyst for a travel booking website. The management is interested in understanding the impact of recent changes to the booking process on user satisfaction. Develop a survey plan to assess user satisfaction and collect feedback about the booking experience. Describe the survey objectives; audience, data collection methods, question types, and how the findings will guide further improvements. **CO5 (10)**
- C9.** As an analyst for an online publication, you're tasked with tracking outbound links to enhance user engagement. Explain the significance of monitoring outbound links in Google Analytics, detailing the tracking methods and demonstrating their application. Provide recommendations on how optimized outbound linking strategies can positively impact user engagement for your online publication. **CO6 (12.5)**
- C10.** You're a web analytics manager for an online gaming platform. The company is launching a referral program to increase user acquisition. Develop a comprehensive plan for using KPIs to measure the success of the referral program. Describe the specific KPIs, how you will collect and analyze the data, and how these KPIs will help in gauging the program's effectiveness. **CO6 (12.5)**

\*\*\*\*\*

Duration : 3 Hours

M.Sc. DEGREE - NOVEMBER 2023 - EXAMINATIONS  
BRANCH : DATA SCIENCE  
NATURAL LANGUAGE PROCESSING

Reg. No.: 19S035

Answer All Questions  
**PART - A**

Maximum: 100 Marks

- A1. List the differences for syntax level analysis and semantic level analysis (5 x 2 = 10) CO Marks  
A2. What is bound morpheme and free morpheme? CO1 (2)  
A3. What is orthographic rules? Give an example CO1 (2)  
A4. Differentiate open class words and closed class words with examples CO1 (2)  
A5. What is WordNet? CO2 (2)

- B1. Explain the different types of Ambiguities with an example (6 x 5 = 30) CO Marks  
B2. Compare syntactic parsing and semantic parsing with an example CO1 (5)  
B3. Differentiate Constituency parsing and Dependency Parsing CO2 (5)  
B4. Explain transition-based dependency parser with example. CO2 (5)  
B5. Explain Attachment and Coordination Ambiguity with suitable example CO3 (5)  
B6. What is the difference between pragmatics and discourse analysis? CO4 (5)

- C1. a) Explain how distance measure is applied in text similarity (60 Marks) CO Marks  
b) Demonstrate how the Levenshtein Distance may be used to compare the CO3 (3)  
similarities between the strings "kitten" and "sitting". CO3 (5)  
c) Compare the output with Jaccard Similarity Score and Normalized Hamming CO3 (7)  
Distance score

- C2. Identify the POS tag using Viterbi algorithm for the sentence (OR) CO3 (15)  
"Will will google campus"  
Construct look up table, transition matrix and emission matrix for the given data  
Given:

Nitish loves campus  
Can Nitish google campus  
Will Ankita google campus  
Ankita loves will  
Will loves google

- C3. Compute Tf-Idf using Bag of Words with all required preprocessing and find the CO4 (15)  
similarity using cosine similarity

S1 = "I want to start learning to charge something in life"  
S2 = "Reading something about life no one else knows"  
S3 = "Never stop learning"

(OR)

- C4. Generate any two possible sequences with maximum of 3 sentences using CO4 (15)  
Markov Model from the given data.

"Word embeddings are a type of word representation"  
"word embedding to learn a better quality of word"  
"Word embedding is a mapping of words into vectors"  
"Text embeddings, also known as word embeddings"

- a. Look-Up Table  
b. Transition Table  
c. Sentence Generation

CONTD.,

<b>C5.</b>	a) Why Pragmatics Matters in Multilingual Chat bots? b) Discuss both generative and retrieval-based chat-bots with appropriate scenarios for both open and closed domains.	(OR)	<b>CO4</b>	<b>(3)</b>
<b>C6.</b>	Explain co-occurrence matrix of window size 3 for the given sentences "I love to read books." "Books are a great source of knowledge." Find the similarity using cosine similarity between books and knowledge		<b>CO4</b>	<b>(10)</b>
<b>C7.</b>	a) Why do we need semantic role labelling when there is already parsing? b) Identify Thematic roles for the following sentences: "John planted the tree. The water in the pot boiled quickly. Mary listened to the music coming from the park. John made Peter sweep the garage"	(OR)	<b>CO5</b> <b>CO5</b>	<b>(2)</b> <b>(8)</b>
<b>C8.</b>	a) Explain the dictionary based approach – LESK algorithm for WSD b) Explain how unsupervised learning algorithm using concept clustering to overcome Word Sense Disambiguation.		<b>CO5</b> <b>CO5</b>	<b>(4)</b> <b>(6)</b>
<b>C9.</b>	a) Explain and differentiate the challenges in semantic sentiment analysis with examples b) Discuss Aspect-based Sentiment Analysis for the given Sentence: "Our stay at the hotel was nice. The staff was friendly and the rooms were clean, but our beds were quite uncomfortable."	(OR)	<b>CO6</b> <b>CO6</b>	<b>(3)</b> <b>(7)</b>
<b>C10.</b>	a) Discuss Named Entity in Resume Filtering in terms of a. Ambiguity and Abbreviations b. OoV words c. Spelling Variations b) Apply Hobbs algorithm for the sentences to replace pronouns "John took his license when he was 18. He passed his exam at his first attempt."		<b>CO6</b> <b>CO6</b>	<b>(5)</b> <b>(5)</b>

\*\*\*\*\*

Duration : 3 Hours

Answer All Questions  
PART - A

	$(10 \times 2 = 20)$	CO	Marks
A1.	Define pixel.	CO1	(2)
A2.	What is an orthogonal transformation in image processing?	CO1	(2)
A3.	Define convolution in the context of image processing.	CO1	(2)
A4.	How histogram equalization can be useful in image enhancement?	CO1	(2)
A5.	What are the key steps involved in the Canny edge detection algorithm?	CO2	(2)
A6.	What is the purpose of using Harris edge detector in images?	CO2	(2)
A7.	Differentiate between object detection and object recognition in computer vision.	CO2	(2)
A8.	What is PCA?	CO2	(2)
A9.	Give an application where activity recognition using computer vision can be beneficial.	CO4	(2)
A10.	Define CBIR and CBVR.	CO4	(2)

PART - B

$(5 \times 8 = 40)$

	$(5 \times 8 = 40)$	CO	Marks
B1.	Compare and contrast orthogonal and affine transformations in the context of image processing. Provide examples where each type of transformation is useful.	CO1	(8)
B2.	Describe common methods and approaches used for texture segmentation.	CO2	(8)
B3.	Explain the principles behind graph-cut segmentation and mean-shift segmentation methods.	CO2	(8)
B4.	Compare region growing and edge-based segmentation approaches and provide examples of scenarios where one method is preferred over the other.	CO2	(8)
B5.	Define unsupervised learning and discuss its main objectives and applications. How does it differ from supervised learning?	CO6	(8)

PART - C

$(4 \times 10=40)$

	$(4 \times 10=40)$	CO	Marks
C1.	Consider the image given below.	CO3	(10)

Image:

30 30 35 40 30  
30 32 31 33 35  
34 33 33 34 34  
32 33 33 31 30  
30 31 30 30 30

Perform region growing with a seed intensity value of 32 and a threshold of 2. Implement the region growing algorithm, and list the coordinates of all the pixels that belong to the segmented region.

(OR)

C2.	Consider a gray scale image with pixel intensity values ranging from 0 to 255:	CO3	(10)
-----	--	-----	------

Image:

100 102 105 200 198  
103 105 109 199 201  
104 107 108 198 203  
110 112 111 202 205  
115 113 114 203 207

Perform K-Means clustering to segment this image into three clusters. Start by initializing three cluster centroids as 105, 198 and 205. Perform two iterations of the K-Means algorithm, updating the cluster assignments and centroids. Provide the updated cluster assignments and centroids after this iteration.

C3.	In the context of scene recognition in robotics, explain how the Bag of Words model can be applied to identify different scenes in an environment. Describe the feature extraction and learning steps involved in this recognition process.	CO4	(10)
-----	---	-----	------

(OR)

CONTD.,

- C4.** Compare and contrast structural-based and appearance-based methods for object recognition. Provide a scenario where structural-based recognition would be more suitable and another scenario where appearance-based recognition would excel. Justify your choices. **CO4 (10)**
- C5.** Given a 5x5 gray scale image: **CO6 (10)**
- Image:**  
 10 20 10 30 40  
 50 60 70 80 90  
 30 20 10 20 30  
 70 80 90 80 70  
 50 40 30 20 10
- Apply a 3x3 convolution with the following kernel:  
**Kernel:**  
 -1 0 1  
 -1 0 1  
 -1 0 1
- Perform the convolution operation, and provide the resulting 3x3 output matrix with explanation.
- (OR)**
- C6.** Given a gray scale image: **CO6 (10)**
- Image:**  
 10 20 30 40 50  
 60 70 80 90 100  
 110 120 130 140 150  
 160 170 180 190 200  
 210 220 230 240 250
- Apply a 3x3 mean filter (averaging filter) to the image with zero padding. Calculate the new pixel values for each pixel in the output image.
- C7.** Describe the KNN algorithm and how it can be applied to image matching. What are the main steps involved in using KNN for finding similar images in a dataset? Explain it with an example. **CO5 (10)**
- C8.** Provide a simple example where using a KNN classifier might be more suitable than a Naive Bayes classifier. What characteristics of the data make KNN a reasonable choice in this scenario? **CO5 (10)**
- \*\*\*\*\*

19S025

Name of the Candidate:

Reg. No.:

**19DSPG0**  
**M.S.C. DEGREE - NOVEMBER 2023 -EXAMINATIONS**  
**BRANCH : DATA SCIENCE**  
**SOFT COMPUTING**

Duration : 3Hours

Maximum: 100 Marks

**SPECIAL INSTRUCTIONS:** Scientific Calculators are permitted  
 Answer All Questions

**PART - A****(10 x 2 = 20)****CO Marks**

- A1. Define State space search. Give an example. CO1 (2)  
 A2. Give an example of semantic representation CO1 (2)  
 A3. Give the space complexity and time complexity of breadth first search. CO1 (2)  
 A4. Write how Depth Limited DFS be extended to iterative deepening search. CO1 (2)  
 A5. Present any four differences between Mamdani and Sugeno Fuzzy inference systems CO2 (2)  
 A6. Draw the fuzzy inference system architecture CO2 (2)  
 A7. Present any four membership functions in fuzzification. CO2 (2)  
 A8. List any eight optimization techniques. CO5 (2)  
 A9. Write short notes on evolutionary algorithms. CO5 (2)  
 A10. What is the necessity of Elitism in Genetic algorithms? How is it performed? CO3 (2)

**PART - B****(4 x 5 = 20)****CO Marks**

- B1. Demonstrate Depth Limited DFS with an example CO1 (5)  
 B2. Demonstrate heuristic search with an example cost function CO1 (5)  
 B3. Explain the role of linguistic variables in fuzzy logic and explain any three linguistic hedges. CO2 (5)  
 B4. Explain any three mutation strategies for generating offspring in a genetic algorithm CO3 (5)

**PART - C****(60 marks )****CO Marks**

- C1. Given a population set for genetic algorithm, demonstrate any two selection operators, two cross over operators and two mutation operators CO5 (10)

Fitness	Initial Population
22	101010100011110101
9	110011010101011100
8	111110101111010101
70	111001111100001001
19	11001101010101011100
48	101110101111001001
23	110011010101011100
38	111001111100001001

**(OR)**

- C2. Develop a genetic representation for a Travelling Salesman problem. Explain your representation and demonstrate the generation of offsprings with appropriate genetic operators for a Travelling Salesman problem with a specification of stopping criteria CO5 (10)

- C3. Perform three generations of binary coded genetic algorithm to solve the following optimization problem. Maximize  $f(x) = \sin x$ ,  $0 \leq x \leq 32$ ,  $x$  is an integer. Use Tournament selection, multi point crossover, interchanging mutation operator and population size of four CO3 (15)

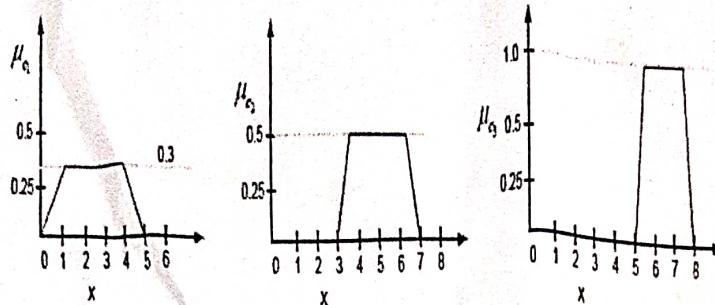
**(OR)**

- C4. With a maximizing function of  $x_1^2 + x_2^2$ , Give a binary code genetic algorithm solution with 8 bits of population string in which first 4 bits represents  $x_1$  and the last 4 bits represent  $x_2$ . CO3 (15)

**CONTD.,**

C5.

CO4 (15)



Consider the application of three rules in a fuzzy inference system. Find the crisp value after defuzzification of the system with Center of Gravity method, Weighted average method, Mean of Maxima and Last of Maxima method.

(OR)

- C6. a. Identify the inputs, rules and output needed for selecting project managers from a pool of project managers using a fuzzy inference system CO4 (5)
- b. You are asked to develop a controller to regulate the temperature of a room. Knowledge of the system allows you to construct a simple rule of thumb: when the temperature (measured in Fahrenheit) is HOT then cool room down by turning the fan at the fast speed (measured in rpm), or, expressed in rule form, IF temperature is HOT, THEN fan should turn FAST. CO4 (10)

$$\tilde{H} = \text{"hot"} = \left\{ \frac{0}{60} + \frac{0.1}{70} + \frac{0.7}{80} + \frac{0.9}{90} + \frac{1}{100} \right\}$$

$$\tilde{F} = \text{"fast"} = \left\{ \frac{0}{0} + \frac{0.2}{1} + \frac{0.5}{2} + \frac{0.9}{3} + \frac{1}{4} \right\}$$

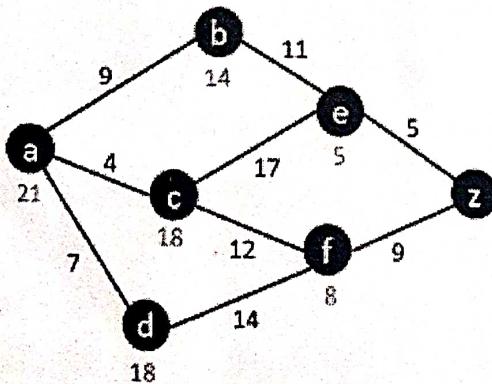
- (a) From these two fuzzy sets construct a relation for the rule using classical implication.  
 (b) Suppose a new rule uses a slightly different temperature, say "moderately hot," use max-product composition, find the resulting fuzzy fan speed.

- C7. a. "When the temperature is higher, more likely the algorithm accepts a worse solution". CO6 (10)  
 Present how Simulated Annealing algorithm stands true for the given statement  
 b. Demonstrate how iterative deepening algorithm works better than depth first search. CO6 (10)

(OR)

C8.

CO6 (20)



Given a graph of a network, a need of travelling from node a to node z has occurred. Present a comparison study between the functionality of A\* algorithm and any other search algorithm in finding a minimum path from node a to node z.

\*\*\*\*\*

Name of the Candidate:

Reg. No.:

5505

**19DSPE0**  
**M.Sc DEGREE–NOVEMBER 2023 –EXAMINATIONS**  
**BRANCH : DATA SCIENCE**  
**MARKETING ANALYTICS**

Duration : 3Hours

Maximum: 100 Marks

Answer All Questions

PART – A

(5 x 2 = 10)

CO Marks

- A1.** Define Marketing Analytics. CO1 (2)  
**A2.** Name any four typical applications for Market Sizing. CO1 (2)  
**A3.** What are some inherent limitations of segmentation? CO1 (2)  
**A4.** Provide an overview of several different competitive information sources. CO1 (2)  
**A5.** What are the factors used in selecting forecasting methods? CO1 (2)

PART – B

(4 x 5 = 20)

CO Marks

- B1.** Identify specific factors driving adoption of marketing analytics in an organization. CO1 (5)  
**B2.** Perform a SWOT analysis for your own organization and at least two of your principal competitors. CO2 (5)  
**B3.** What is AHP? Show its usage. CO3 (5)  
**B4.** Summarize the typical distribution costs. CO4 (5)

PART – C

(70 mark)

CO Marks

- C1.** How do you conduct a PESTLE Market analysis for an organization's market? Of the six categories of trends, which has the largest potential to make significant change? CO2 (12.5)  
**(OR)**
- C2.** Illustrate Trial Rate Forecasting method with example. CO2 (12.5)
- C3.** Explain the product and service metrics used for tracking sales, profitability. CO3 (12.5)  
**(OR)**
- C4.** Give a detailed explanation of Creaming pricing and Demand based pricing with example. CO3 (12.5)
- C5.** What type of pricing assessment technique does your organization use to evaluate pricing for proposed new products and services? Illustrate. CO4 (12.5)  
**(OR)**
- C6.** Describe the distribution channel member evaluation and selection model. CO4 (12.5)
- C7.** Give a detailed study on promotion metrics for Traditional media. CO5 (12.5)  
**(OR)**
- C8.** Examine the various sales metrics to understand the effectiveness of marketing and sales efforts. CO5 (12.5)
- C9.** **a)** Identify the forces acting on your organization's market using the Porter five forces framework. How will they change the market over time? CO6 (10)  
**b)** Prepare a QSPM decision model for entry into a new market. CO6 (10)  
**(OR)**
- C10.** **a)** Prepare a Monte Carlo analysis for a planned growth in an existing market. Create scenarios for a weak, typical and strong market. CO6 (10)  
**b)** Show how Target-return Pricing and Tiered pricing will work for marketing a product. CO6 (10)

\*\*\*\*\*