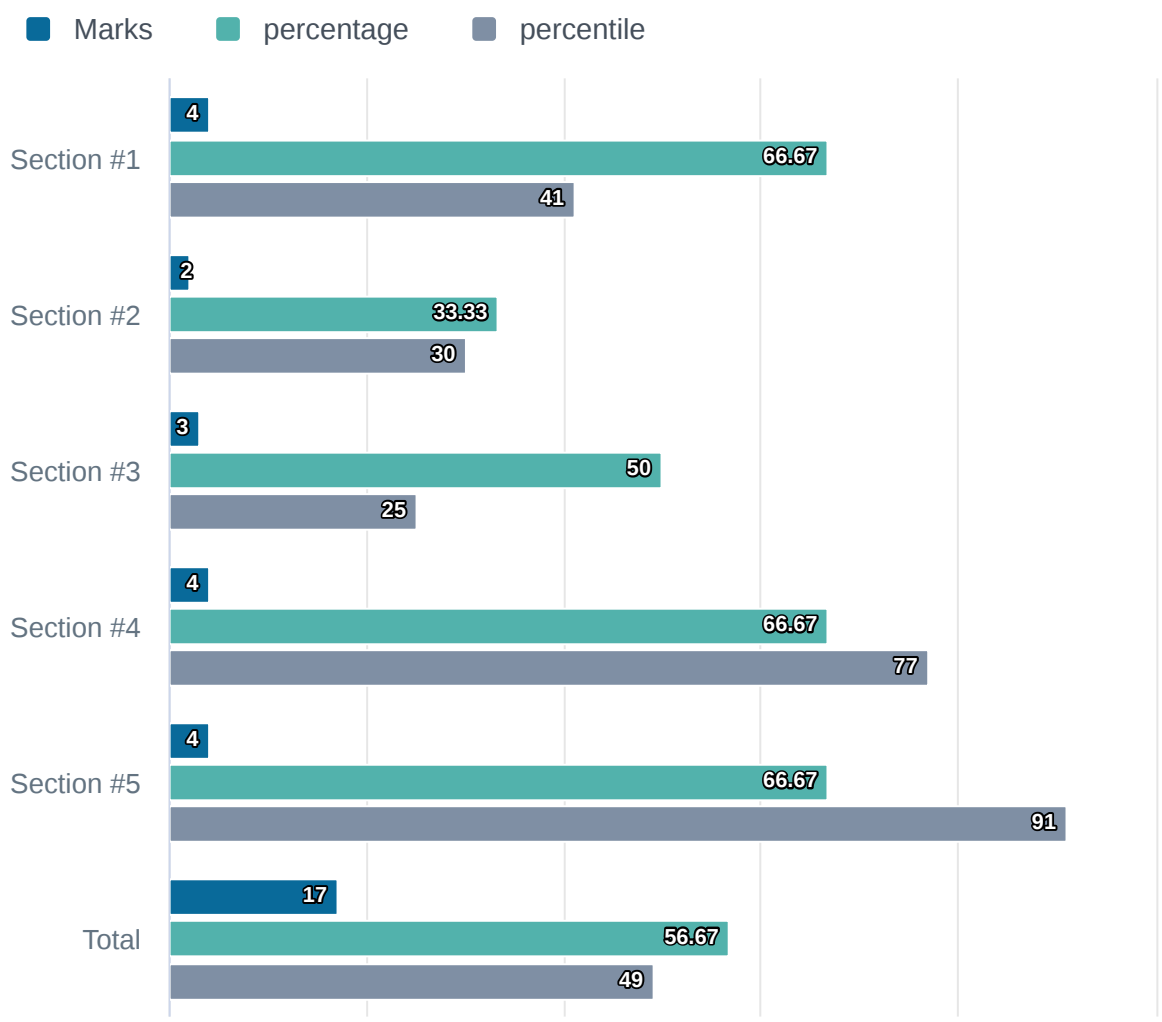
MTAIML_CLUS_BATCH2_SEM 2_1-2023_EC2R_AIMLCZG512_DEEP REINFORCEMENT LEARNING_Online

GULAM SARWAR | 27 Jan 2024

Overall Summary

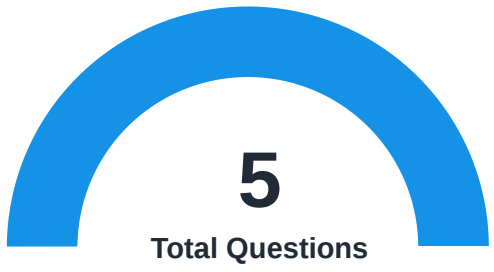
17 Marks Scored out of 30	56.67 % 49.12 percentile out of 226 Test Takers	2h 14m 56s Time taken of 2hr 25mins

Marks Scored



Attempt Summary

Distribution of questions attempted in a total of 5 question(s).



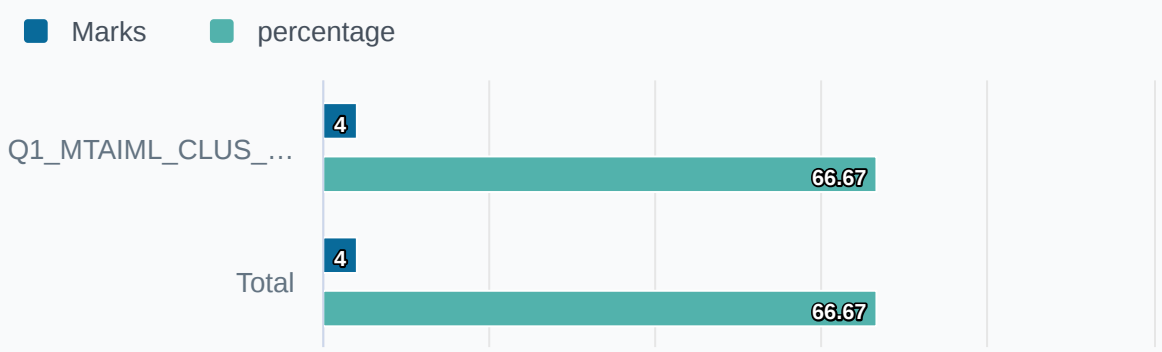
This shows the correctness of questions attempted by the test taker

Correct	0 Ques	0/0 Marks
Incorrect	0 Ques	0/0 Marks
Partially Correct	5 Ques	17/30 Marks
Not Attempted	0 Ques	0/0 Marks

Section-Wise Details

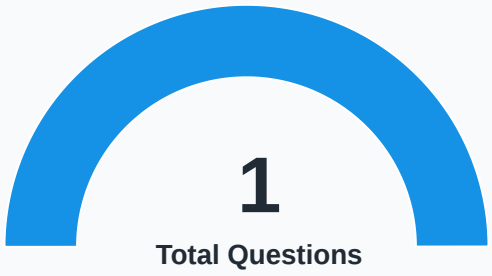
Section 1	question(s)	Time taken	Marks Scored
Section #1	1 Q.	35m 16s (Untimed)	4 / 6

Marks Scored



Attempt Summary

Distribution of questions attempted in a total of 1 question(s).



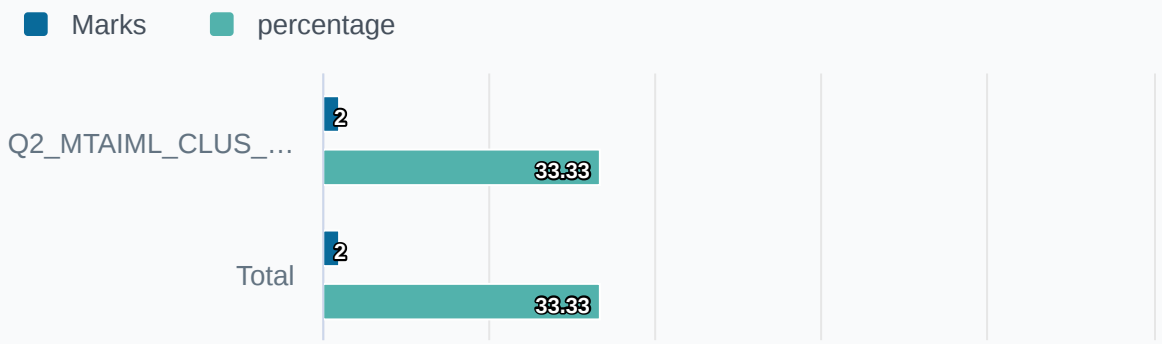
This shows the correctness of questions attempted by the test taker

Partially Correct	1 Ques	4/6 Marks

Section-Wise Details

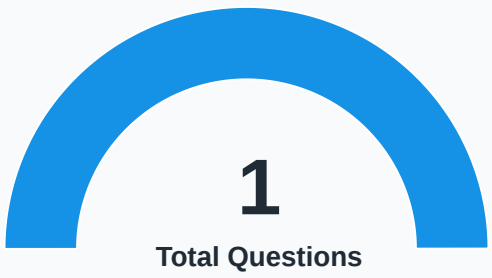
Section 2	question(s)	Time taken	Marks Scored
Section #2	1 Q.	43m 22s (Untimed)	2 / 6

Marks Scored



Attempt Summary

Distribution of questions attempted in a total of 1 question(s).



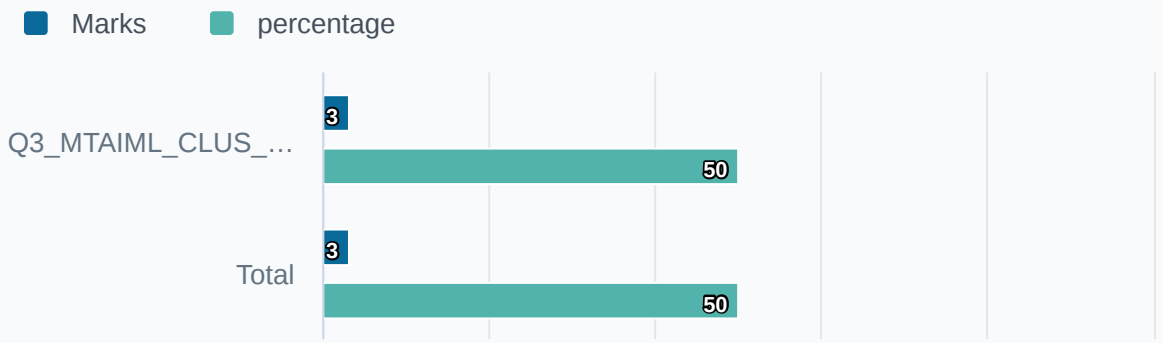
This shows the correctness of questions attempted by the test taker

Partially Correct	1 Ques	2/6 Marks

Section-Wise Details

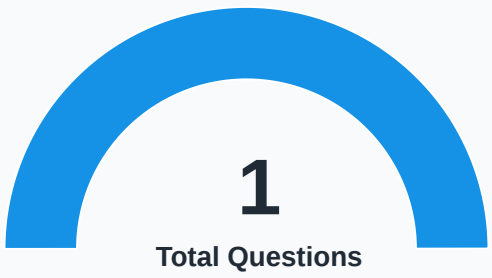
Section	question(s)	Time taken	Marks Scored
Section #3	1 Q.	22m 20s (Untimed)	3 / 6

Marks Scored



Attempt Summary

Distribution of questions attempted in a total of 1 question(s).



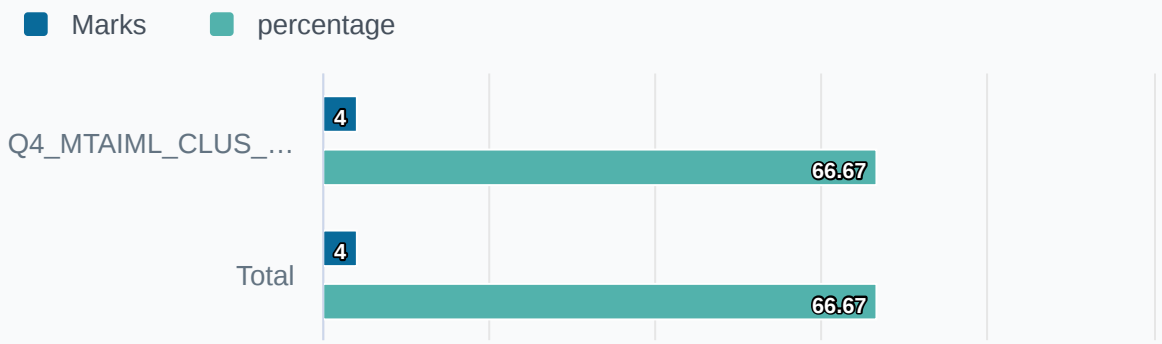
This shows the correctness of questions attempted by the test taker

Partially Correct	1 Ques	3/6 Marks

Section-Wise Details

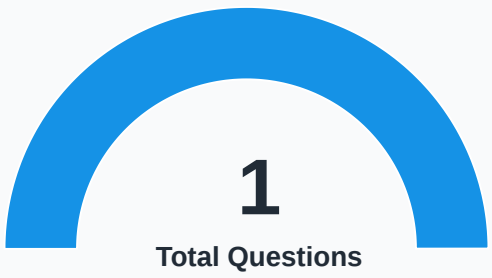
Section 4	question(s)	Time taken	Marks Scored
Section #4	1 Q.	21m 14s (Untimed)	4 / 6

Marks Scored



Attempt Summary

Distribution of questions attempted in a total of 1 question(s).



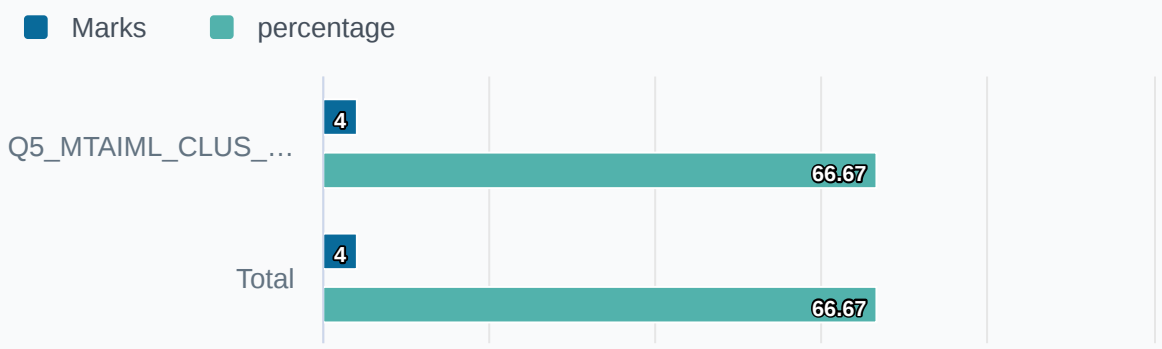
This shows the correctness of questions attempted by the test taker

Partially Correct	1 Ques	4/6 Marks

Section-Wise Details

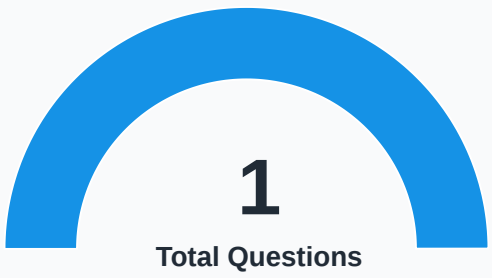
Section 5	question(s)	Time taken	Marks Scored
Section #5	1 Q.	12m 42s (Untimed)	4 / 6

Marks Scored



Attempt Summary

Distribution of questions attempted in a total of 1 question(s).



This shows the correctness of questions attempted by the test taker

Partially Correct	1 Ques	4/6 Marks

Question-Wise Details

Section 1
Section #1

1	35m 16s	4/6
question(s)	Time taken	Marks Scored

Q. 2

Question 1

Mark for Re-evaluation | Time taken: 35m 16s | Marks Scored: 4/6

Please write your answer on a piece of paper and scan and upload the handwritten answers using the QR Code available in the Scan and Upload Section of this exam

Consider tic-tac-toe to answer this question. Assume that states are numbered from S_1 to S_n.

(a) List the four elements of reinforcement learning and write one well-articulated formal statement explaining the role of each element. **[2 M]**

(b) Write the temporal difference rule for learning each state's value. **[0.5 M]**. Explain various elements and the workings of this rule. **[1.5 M]**

(c) Let the value of the current state be 4.5, and all its possible successor/predecessor states have a value of 2.7. Use 0.9 to be the parameter value for any parameter you need to use to solve this problem. Given this, Revise the estimate of the value of the current state using your answer to (b). Explain your answer **[2 M]**

[2.0 + 2.0+ 2.0 = 6 Marks]

Response:

Words : 0

Answer Image 1706333521922.png
1.58 MB Click to Download

Answer Image 1706333555889.png
1.47 MB Click to Download

128152855__0_1708068487503__DRL q4 ans - GULAM SARWAR.pdf
1.67 MB Click to Download

ⓘ Warning: The uploaded files are not verified to be safe. Please access them carefully.

File Upload Settings for this question:

Maximum number of files allowed	20
Maximum individual File Size Allowed	30 MB
File Types Allowed	All File Types Allowed by Mettl ⓘ

Max marks: **6**

EVALUATOR'S NAME	MARKS	EVALUATION COMMENTS	LAST MODIFIED ON
Pooja Harde	4.0	part b) and c) partially solved. no explanations	08:32 PM Feb 8, 2024

Question-Wise Details

Section 2
Section #2

1	43m 22s	2/6
question(s)	Time taken	Marks Scored

Q. 2

Question 1

Mark for Re-evaluation | Time taken: 43m 22s | Marks Scored: 2/6

Please write your answer on a piece of paper and scan and upload the handwritten answers using the QR Code available in the Scan and Upload Section of this exam

(a) Write any one use-case (application) which can be modeled as a multi-armed bandit for its solution. Your answer should have all the elements that identify the use case and can be modeled as a multi-armed bandit problem. **[2 M]**
(b) Write down a bandit **algorithm** that is □-greedy. [Note: only a formal algorithm pseudo-code will be accepted as an answer.] **[2 M]**. Comment on the efficiency of this algorithm. **[0.5 M]**.
(c) How does this algorithm balance exploration vs exploitation trade-off? Explain. **[1.5 M]**.

[2.0 + 2.5+ 1.5 = 6 Marks]

Response:

Words : 0

Answer Image 1706333666985.png
1.62 MB Click to Download

Answer Image 1706333691404.png
1.69 MB Click to Download

Answer Image 1706333718589.png
1.64 MB Click to Download

Answer Image 1706333756928.png
1.53 MB Click to Download

128152855__0_1708068487503__DRL q4 ans - GULAM SARWAR.pdf
1.67 MB Click to Download

Warning: The uploaded files are not verified to be safe. Please access them carefully.

File Upload Settings for this question:

Maximum number of files allowed	20
Maximum individual File Size Allowed	30 MB
File Types Allowed	All File Types Allowed by Mettl

Max marks: **6**

EVALUATOR'S NAME	MARKS	EVALUATION COMMENTS	LAST MODIFIED ON
Karthika	2.0	Generic answer is given. Use case based explanation is missing.	09:54 PM Feb 4, 2024

Question-Wise Details

Section 3
Section #3

1	22m 20s	3/6
question(s)	Time taken	Marks Scored

Q. 2

Question 1

Mark for Re-evaluation | Time taken: 22m 20s | Marks Scored: 3/6

Please write your answer on a piece of paper and scan and upload the handwritten answers using the QR Code available in the Scan and Upload Section of this exam

(a) Devise two example tasks of your own that fit into the MDP framework, identifying for each its states, actions, and rewards. Make the examples as different from each other as possible. **[2 M]**

(b) In the following two cases you will find the reward sequences (Starting from R1) is given.

You should find the Return (G_0 and G_1) in each case. Assume the γ = 0.8 in both cases.

(i) An episodic task with T=5. The rewards received are $R_1 = 1$, $R_2 = 2$, $R_3 = 6$, $R_4 = 3$, and $R_5 = 2$ **[1.5 M]**

(ii) A continuing task with $R_1 = 2$, followed by an infinite sequence of 0's. **[1.5 M]**

(c) Write down the complete expression for either $v_\pi(s)$ or $q_\pi(s,a)$. **[1.0M]**

[2.0 + 3.0+ 1.0 = 6 Marks]

Response:

Words : 0

Answer Image 1706333811786.png
1.57 MB Click to Download

Answer Image 1706333835947.png
1.46 MB Click to Download

128152855__0_1708068487503__DRL q4 ans - GULAM SARWAR.pdf
1.67 MB Click to Download

ⓘ Warning: The uploaded files are not verified to be safe. Please access them carefully.

File Upload Settings for this question:

Maximum number of files allowed	20
Maximum individual File Size Allowed	30 MB
File Types Allowed	All File Types Allowed by Mettl

Max marks: **6**

EVALUATOR'S NAME	MARKS	EVALUATION COMMENTS	LAST MODIFIED ON
Rebecca Sandra Paul	3.0	3b is incorrect	12:20 PM Feb 11, 2024

Question-Wise Details

Section 4
Section #4

1	21m 14s	4/6
question(s)	Time taken	Marks Scored

Q. 2

Question 1

Mark for Re-evaluation | Time taken: 21m 14s | Marks Scored: 4/6

Please answer this question only by typing your answers in the typing space available for this question

(a) What do you perform in Policy Evaluation and Policy Improvement? How are they useful in estimating optimal policy? [Answer must be phrased using formal, unambiguous statement] **[2 M]**

(b) Provide a high-level algorithm for policy iteration. Use the computation of either action or state values. Make necessary assumptions. **[2 M]**

(c) Explain the characteristics of reinforcement learning problems for which a solution using dynamic programming is appropriate. Provide any two examples of problems. **[2 M]**

[2.0 + 2.0+ 2.0 = 6 Marks]

Response:

Words : 0

128152855__0_1708068487503__DRL q4 ans - GULAM SARWAR.pdf
1.67 MB Click to Download

Max marks: **6**

EVALUATOR'S NAME	MARKS	EVALUATION COMMENTS	LAST MODIFIED ON
Anusha P	4.0	a. uses not specified b. assumptions not given	09:54 PM Feb 20, 2024

Question-Wise Details

Section 5
Section #5

1	12m 42s	4/6
question(s)	Time taken	Marks Scored

Q. 2

Question 1

Mark for Re-evaluation | Time taken: 12m 42s | Marks Scored: 4/6

Please write your answer on a piece of paper and scan and upload the handwritten answers using the QR Code available in the Scan and Upload Section of this exam

(a) Assume you have data in the form of just the following 5 complete episodes. Non-terminal States are labeled A and B, the numbers in the episodes denote Rewards, and all states end in a terminal state T.

- A 2 A 6 B 1 B 0 T
- A 3 B 2 A 4 B 2 B 0 T
- A 0 B 2 A 4 B 4 B 2 B 0 T
- B 8 B 0 T

Estimate the value of states of A & B. Show the steps. **[2 M]**

(b) What techniques monte-carlo algorithms use to ensure the exploration vs. exploitation trade-off is balanced. Write one line for each of the techniques stating their use. **[2 M]**

(c) How are □-soft and □-greedy policies are related? A small example (a state with 4 actions, with one of the actions being greedy) illustrates the working of □-greedy action selection. Assume suitable values for each action to help your explanation. **[2 M]**

[2.0 + 2.0+ 2.0 = 6 Marks]

Response:

Words : 0

Answer Image 1706334307830.png
1.47 MB Click to Download

Answer Image 1706334637442.png
1.73 MB Click to Download

Answer Image 1706334696096.png
1.66 MB Click to Download

128152855__0_1708068487503__DRL q4 ans - GULAM SARWAR.pdf
1.67 MB Click to Download

Warning: The uploaded files are not verified to be safe. Please access them carefully.

File Upload Settings for this question:

Maximum number of files allowed	20
Maximum individual File Size Allowed	30 MB

File Types Allowed	All File Types Allowed by Mettl

Max marks: **6**

EVALUATOR'S NAME	MARKS	EVALUATION COMMENTS	LAST MODIFIED ON
Manjula AK	4.0	problem not written	08:47 PM Feb 23, 2024

About the Report

This Report is generated electronically on the basis of the inputs received from the assessment takers. This Report including the AI flags that are generated in case of availing of proctoring services, should not be solely used/relied on for making any business, selection, entrance, or employment-related decisions. Mettl accepts no liability from the use of or any action taken or refrained from or for any and all business decisions taken as a result of or reliance upon anything, including, without limitation, information, advice, or AI flags contained in this Report or sources of information used or referred to in this Report.