CSE 415, Autumn 2025

Assignment 6

Last name:_____ First name:_____ UWNetID:_____

Due Monday night December 1 via Gradescope at 11:59 PM. (Note: Because of the tight schedule at the end of the quarter, there will be no grace days on this assignment, and two days only to turn in an assignment late – for a 25 percent penalty per day.) You may turn in either of the following types of PDFs: (1) dcans of these pages that include your answers (handwriting is OK, if it's clear), or (2) documents you create with the answers, saved as PDFs. When you upload to Gradescope, you'll be prompted to identify where in your document your answer to each question lies.

Perform and provide answers for each of the seven exercises. Each TA on the staff has contributed or helped on two of the questions. These are intended to take 15-40 minutes each if you know how to do them. Each is worth between 15 points and 20 points. The total possible points for completing the exercises is 150 (not including the extra credit in the formatting bonus, described below).

If any corrections have to be made to this assignment, these will be posted in ED.

This is an individual-work assignment. Do not collaborate on this assignment. Do not use AI systems, such as ChatGPT, to produce any of your answers.

Prepare your answers in a neat, easy-to-read PDF. Our grading rubric will be set up such that when a question is not easily readable or not correctly tagged or with pages repeated or out of order, then points will be be deducted. However, if all answers are clearly presented, in proper order, and tagged correctly when submitted to Gradescope, we will award a 5-point bonus. (Restated, one could lose the bonus due to poor photo contrast, handwriting that graders have difficulty reading, or mess-ups with the tagging of the answers when submitting to Gradescope.)

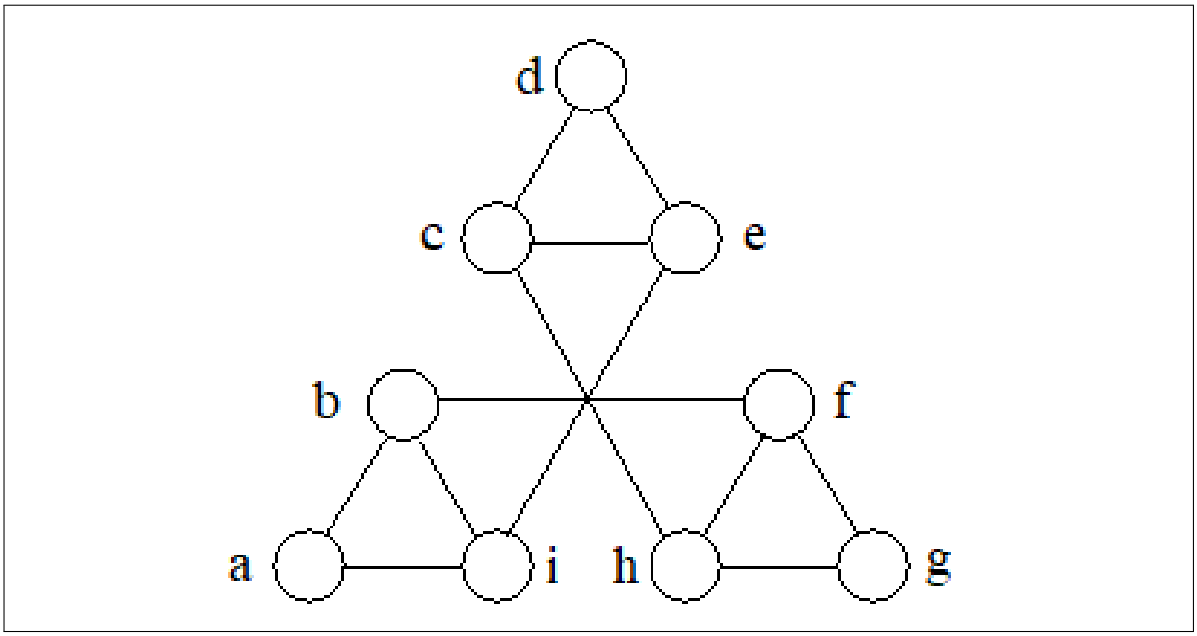
If you choose to typeset your answers in Latex using the template file for this document, please put your answers in blue while leaving the original text black. Using Latex is not required for the bonus points, but it might help in achieving a neat, easy-to-read PDF.

---

1 MDP Basics (Eric)

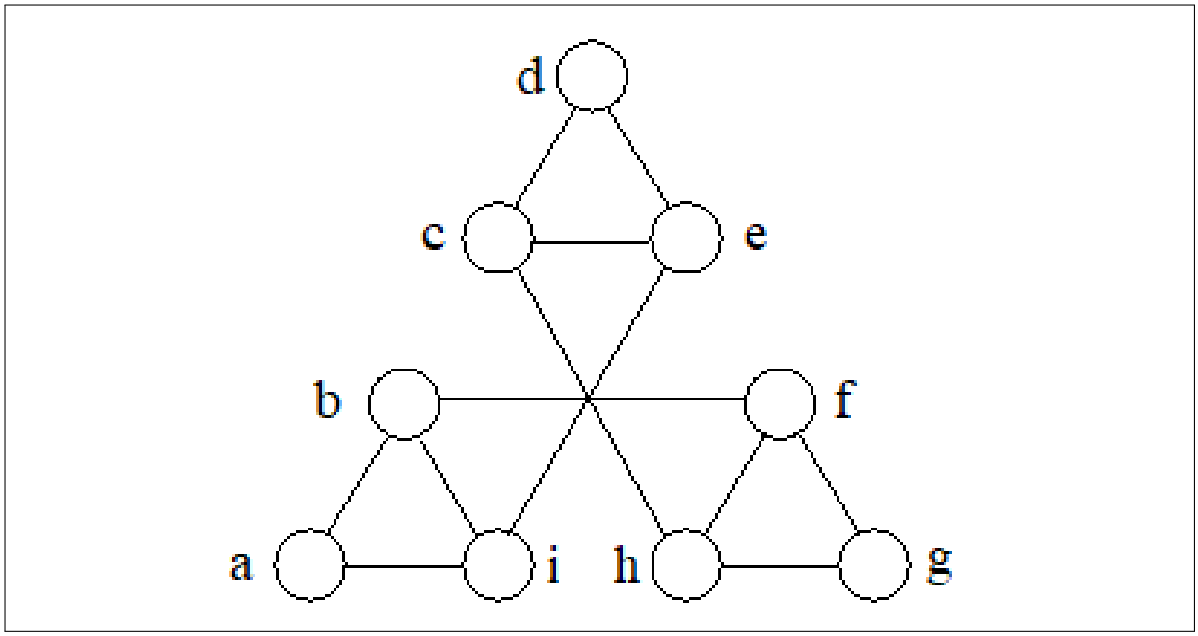
(20 points) This exercise covers basic properties of an MDP.

(a) (4 pts) Suppose that the transition function of TOH2-World is deterministic. Also, suppose that all transitions have a reward of -2, except the transition from g (the "goal") to the Terminal state, which has a reward of 50. Next assume that an agent, finding itself in an arbitrary state of this world, not using any discounting, seeks to maximize its total reward in whatever time remains in its episode. (Say it takes one second per transition in this world.) Assuming it acts optimally, there will be a specific sum of rewards that it can get from each state. Write down that value in each state below.



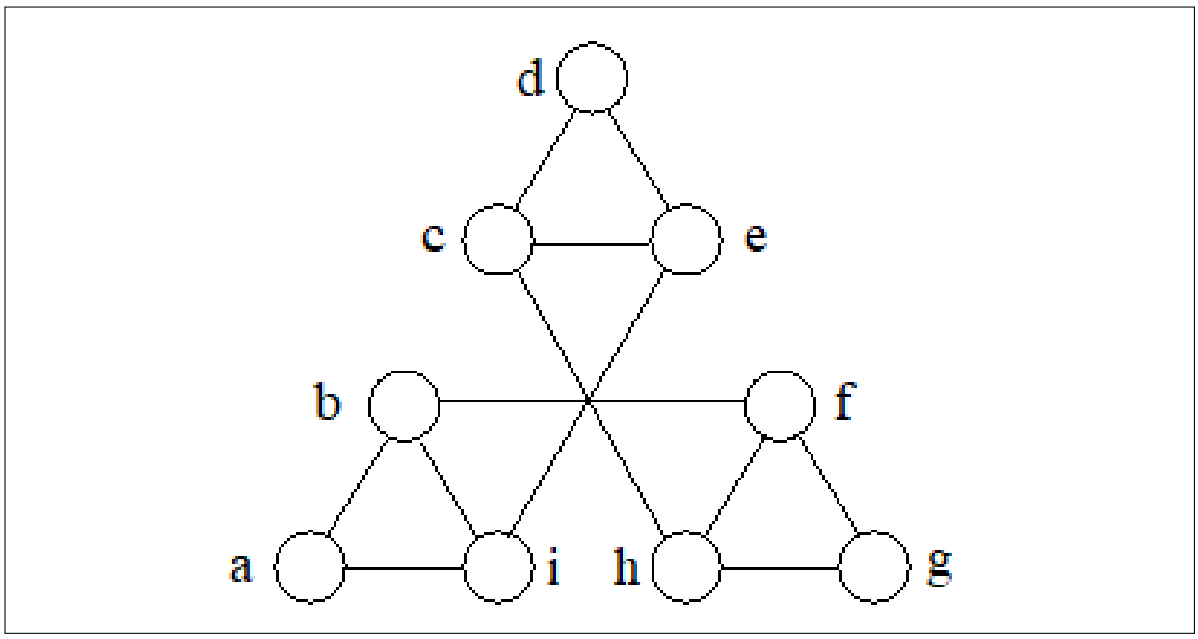
(b) (4 pts.) Drawing on the same diagram, use arrowheads on some of the edges to indicate an optimal policy.

(c) (3 pts.) Suppose state i is a secondary goal node. When the agent is at i, the only legal action is Exit, which transitions the agent to the Terminal state and yields a reward of 20. Using the copy of the graph below, write the value of each state assuming a time horizon of 2 moves. Except for the addition of the new goal assume that the reward function is as before. (and the main goal is still the main goal).



(d) (3 pts.) Also, indicate the optimal policy using arrowheads on this graph.

(e) (3 pts.) Now assume the agent uses discounting of future rewards, with $\gamma = 0.75$. Show the value of each node now. (Settings in part c still works here)



(f) (3 pts.) Indicate the optimal policy in this case using arrowheads on this graph.

2 Basic Q-Learning in an MDP (Sahil)

(15 points) Use the Markov Decision Process (MDP) diagram below to help you as you answer the following question on Q-Learning. The states are A, B, C, D, E, F, G, and an unshown Terminal state.

A	B (Exit -8)	C	D	E	F (Exit $+12$)	G

Use Q-Learning to estimate the true Q values of this MDP given the observed state transitions below. For each state transition, indicate which Q value is changing by giving the state and action associated with that value and giving its *new* value. You may wish to keep your own diagram of the current values while you work through the given transitions.

You may assume that at states $\{A, C, D, E, G\}$ there are two allowed actions, {Left, Right}. At states $\{B, F\}$, there are two allowed actions: {Left, Exit}. At the Terminal state there are no valid actions. When the Exit action is taken from B or F, the agent transitions directly to the Terminal state and then the episode ends.

All Q values are initialized to 0. Use a discount factor of $\gamma = 0.9$ and a learning rate of $\alpha = 0.5$. Make sure to process the transitions in the order provided. Use the standard Q-Learning update rule:

$$Q(s,a) \leftarrow Q(s,a) + \alpha \left[r + \gamma \max_{a'} Q(s',a') - Q(s,a) \right].$$

State Transitions				New Q Values		
State (s)	Action (a)	New State (s')	Reward (r)	State	Action	$Q(s,a)$
E	Left	D	-1			
D	Left	C	-1			
C	Right	D	-1			
D	Right	E	-1			
E	Right	F	-1			
F	Exit	Terminal	$+12$			
A	Right	B	-1			
B	Exit	Terminal	-8			

3 Joint Distributions and Inference (Emilia)

(15 points) Let C represent the proposition that it is cloudy in Seattle. Let R represent the proposition that it is raining in Seattle. Consider the table given below.

C	R	$P(C, R)$
cloudy	*rain*	0.73
cloudy	*sun*	0.14
clear	*rain*	0.02
clear	*sun*	0.11

(a) (2 points) Compute the marginal distribution $P(C)$ and express it as a table.

(b) (2 points) Similarly, compute the marginal distribution $P(R)$ and express it as a table.

(c) (2 points) Compute the conditional distribution $P(R|C = \textit{cloudy})$ and express it as a table. Show your work/calculations.

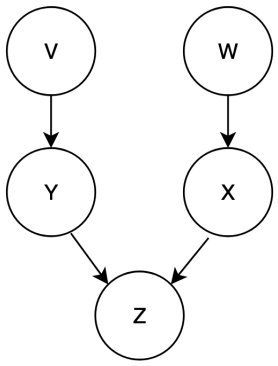
(d) (2 points) Compute the conditional distribution $P(C|R = \textit{sun})$ and express it as a table. Show your work/calculations.

(e) (3 points) Is it true that $C \perp\!\!\!\perp R$? (i.e., are they statistically independent?) Explain your reasoning.

(f) (4 points) Suppose you decide to track additional weather patterns of Seattle such as temperature (hot/cold), humidity (humid/dry), and wind (windy/calm) denoted as the random variables T, W, H respectively. Is it possible to compute $P(C, R, T, W, H)$ as a product of five terms? If so, show your work. What assumptions need to be made, if any? Otherwise, explain why it is not possible.

4 Bayes Net Structure and Meaning (Krish)

(20 points) Consider a Bayes net whose graph is shown below.



Random variable V has a domain with two values $\{v_1, v_2\}$, the domain for W has two values $\{w_1, w_2\}$; the domain for X has three values: $\{x_1, x_2, x_3\}$; Y's domain has three values: $\{y_1, y_2, y_3\}$; and Z's domain has two values: $\{z_1, z_2\}$.

(a) (6 points) Give a formula for the joint distribution of all four random variables, in terms of the marginals (e.g., $P(W = w_i)$), and conditionals that must be part of the Bayes net (e.g., $P(Z = z_m | X = X_j, Y = y_k)$).

(b) (2 points) How many probability values belong in the (full) joint distribution table for this set of random variables?

(c) (4 points) For each random variable: give the number of probability values in its marginal (for W) or conditional distribution table (for the others).

V: W: X:

Y: X:

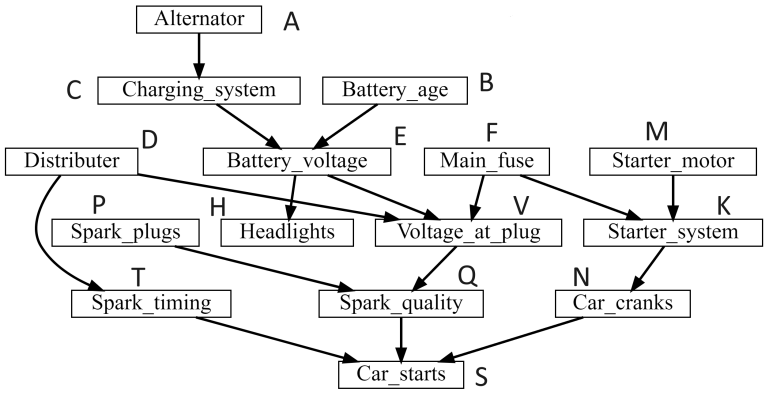
(d) (8 points) For each random variable, give the number of *non-redundant* probability values in its table from (c). Note: If a joint probability distribution has n rows, and their probabilities are $p_0, p_1, \ldots, p_{n_1}$, then we know $\sum_{i=0}^{n-1} p_i = 1$ and any $n - 1$ of them are non-redundant, because the remaining value can be computed from them by subtracting their sum from 1. In a conditional distribution, there are groups of probabilities that sum to 1, and within each group, one is "redundant" while the others are "non-redundant.")

V: W: X:

Y: Z:

5 D-Separation (Krish)

(20 points) Consider the automobile engine-diagnosis Bayes Net below.



(a) (1 point) Assuming no observations have been made, are F and M independent?

(b) (1 point) From the diagram can we assume that, are M and N independent?

(c) (5 points) List all "undirected" paths from C to S.

(d) (3 points) Which of the above are active paths?

(e) (2 points) Is it guaranteed that $A \perp\!\!\!\perp S|Q, C$? If not, give an active path from A to S.

(f) (2 points) Is it guaranteed that $B \perp\!\!\!\perp S|V, K, T$? If not, give an active path from B to S.

(g) (2 points) Is it guaranteed that $P \perp\!\!\!\perp B|S$? If not, give an active path from P to B.

(h) (2 points) Is it guaranteed that $A \perp\!\!\!\perp M|F, V, S$? If not, give an appropriate active path.

(i) (2 points) What is the longest loop-free undirected path you can find in this graph? What nodes would need to be observed to make it an active path?

6 Choosing Hidden Sequences (Eric)

(20 points) Consider the Jones family – Mom, Dad, Jim, Sally. Jim and Sally are twin siblings, and both 16 years old. Once a month, the family goes out for ethnic food – it's either Thai or Indian. Jim has his list of favorite restaurants, and Sally has her own list of her favorites. Each month, Dad uses one of the lists to choose what type of food the family will eat during their outing. He tends to stick with the same list more often than switch, and he tends to slightly favor Jim's list, maybe because of his own preferences.

With the current list, Dad selects a restaurant at random, assuming the restaurants on the list are equally likely to be picked. But they don't necessarily go to that restaurant. If the restaurant is Indian, Jim gets to pick any Indian restaurant from his list, and they go there. If the restaurant is Thai, then Sally gets to pick any Thai restaurant from her list, and they go there.

```
Jim's list:
  My-Pad-Thai (Thai), Punjabi Kitchen (Indian), Delhi Curry (Indian).
Sally's list:
  Bangkok Bites (Thai), Mango Mansion (Thai),
  Siam Shack (Thai), Delhi Curry (Indian).
```

Then after Dad makes his random selection from the current list, he uses some kind of spinning device that we don't understand, and he determines whether to switch lists for next time according to the following transition probabilities.

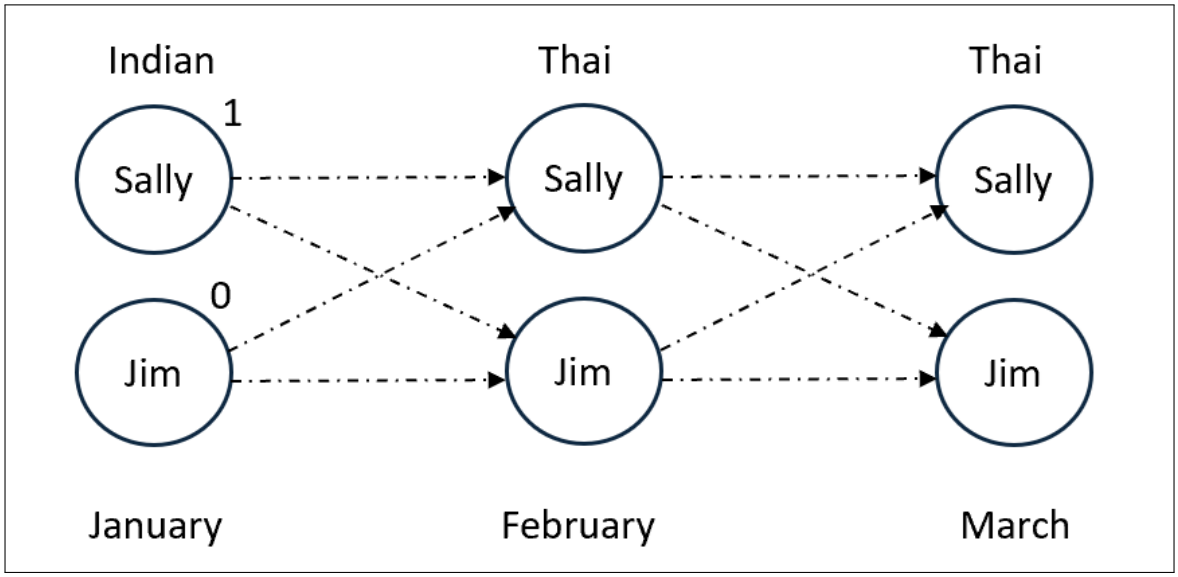
If the current list is Jim's, the probably of staying with Jim's is 0.6, and switching to Sally's is 0.4. If the current list is Sally's the probably of staying with Sally's is 0.7, and switching to Jim's is 0.3.

(a) (2 points) Assume Dad starts the year using Sally's list for January's outing. He's public about that, but in subsequent months, he does not tell anyone which list he is using.

One possible 3-month sequence is list uses is: Sally's, Sally's, Jim's (for January, February, March). What is the probability of that sequence of list usages, without considering the ethnicities of the restaurants they visited?

(b) (2 points) Suppose that quarter of the year they end up going to an Indian restaurant in January, a Thai restaurant in February, and a Thai restaurant in March. Taking that additional information into consideration. what is the probability of that same list-usage sequence and seeing that sequence of cuisine choices?
(i.e., compute P(X_1 =Sally, X_2 =Sally, X_3 =Jim, E_1 =Indian, E_2 =Thai, E_3 =Thai).

(c) (8 points) What is the most likely sequence of list usages by Dad given that same restaurant ethnicity sequence? (Indian, Thai, Thai). Use the diagram template below to create a trellis diagram for this problem. If the probability coming into a dashed arrow is 0, leave it dashed. Otherwise, write over the dashed line to make it a more solid arrow. Before identifying the most likely list-usage sequence, compute, at each node, the probability of reaching that node along a most likely path from the Sally-in-January node. The probability value 1 is provided on that starting node, since it's a given in this exercise. Naturally, the probability of getting to the Jim-in-January node is 0, because we're given that Sally's list is the one being used in the first month (January). Use the Viterbi algorithm to get these probabilities at the other 4 nodes of the diagram. Finally, highlight the most probable path that starts at Sally-in-January and ends at one of the nodes in the March column. (Hint: the HMM worksheet involves a similar computation.) Show the factors you are multiplying on the appropriate edges of the trellis. For the March column, at least, use a calculator to show the two probabilities to 4 decimal places.



(d) (4 points) Compute the stationary probability of Jim's list vs Sally's list being used. (Assume that when the sequence extends beyond 12 months, Dad does NOT automatically go back to Sally's list each January. In other words, the Markov Model represented by the transition CPT is not limited to 12 time steps.)

Show the needed equations before you solve them, and show the steps you use in solving them.

(e) (4 points) Compute the marginal probabilities of the family's having Indian and Thai food on their outings, corresponding to the stationary probabilities of list usages.

7 Perceptrons (Sahil)

(20 points) You are planning an exciting hiking trip across California and need to decide which trails to include in your itinerary. You classify trails into two categories: Must-Hike Trails and Skip Trails. A trail is labeled as Must-Hike (+1) if it is highly attractive for you, while a trail is labeled as Skip (−1) if it does not meet your preferences. You decide to use a perceptron model to classify trails based on their features.

Dataset Features and Specification:
Scenery: Represented as poor = −1, decent = 0, amazing = 2.
Crowd Level: Represented as empty = −1, moderate = 0, crowded = 1.
Driving Distance: Represented as far = −2, near = 2.

Perceptron Parameters

- The perceptron has a weight vector, $\mathbf{w}$, with four components: bias term, scenery, crowd level, and driving distance.
- Initial weights: $\mathbf{w} = [0, -1, 0, 0]$.
- Threshold: 0 (i.e., predict +1 if $\mathbf{w}^\top \mathbf{x} \geq 0$, and −1 otherwise).
- Learning rate: 1.

Each input example $\mathbf{x}$ is represented as a 4-dimensional vector
$[1, \text{Scenery}, \text{Crowd Level}, \text{Driving Distance}]$, where the first component is the bias input.

Example Number	Scenery	Crowd Level	Driving Distance	Label
1	amazing	crowded	near	Must-Hike
2	decent	crowded	far	Skip
3	amazing	empty	far	Skip
4	decent	moderate	near	Must-Hike
5	poor	empty	near	???

Table 1: Hiking Trail Dataset

Assume the standard perceptron update rule: If the prediction is incorrect:

$$\mathbf{w} \leftarrow \mathbf{w} + \eta \cdot y \cdot \mathbf{x}$$

where $y \in \{+1, -1\}$ is the true label and η is the learning rate. If the prediction is correct, no update is made.

(a) (4 points) What would be the updated weights $\mathbf{w}$ after processing example number 1?

(b) (3 points) What would be the updated weights $\mathbf{w}$ after processing example number 2?

(c) (3 points) What would be the updated weights $\mathbf{w}$ after processing example number 3?

(d) (3 points) What would be the updated weights $\mathbf{w}$ after processing example number 4?

(e) (3 points) What is your prediction for the label of example number 5 based on the final weights?

(f) (4 points) Is convergence guaranteed for this perceptron and data set? Why or why not?

8 AI and the Potential for Harm (Emilia)

(20 points)

Two short articles published earlier this year by MIT News are (1) ***Explained: Generative AI's environmental impact*** and (2) ***Q&A: The climate impact of generative AI***. They cover a topic that has been getting more attention recently as companies push to build more data centers — what are the environmental effects of large-scale AI? Please read through these articles. In forming your responses to this question, keep in mind the information you gained from the article(s).

You are welcome to seek out additional sources when completing your responses. If you do so, please credit/cite these sources and provide a link to the sources you used.

(a) What do you currently use AI for? If you don't directly use AI yourself, you may comment on actions you take where you see AI being applied without you specifically asking for it (e.g., search summaries, customer service "conversations", face-tracking).

(b) Do you ever consider the environmental impacts before you use an AI tool? If so, what are your main concerns?

(c) What are some of the potential real/claimed benefits of building a new data center?

(d) What are some of the drawbacks?

(e) Consider the following applications of AI: to learn new material or to complete homework assignments (Learning/HW), to better focus advertising, to streamline traveling by using facial photographs rather than human agents (Travel), to develop a new medication to treat a new pathogen (Medicine), for national defense or use in war (War), to study climate change or other large social issues (Social), to produce artwork for movies or other media (Art), for company if you are lonely or bored (Companionship), for entertainment to make a game more exciting or engaging. (Entertainment), or to enable robotic house servants (Robots). Complete the following table, stating whether you consider the application a "Good" or "Bad" use of AI or whether "It Depends" and commenting on the reasons for your classification. Try to consider both positives and negatives in your responses. Your reasons do NOT need to be limited to the environmental impacts of AI.

Please enter your responses into the table below.

Application	Good / Bad / Depends	Comment
Learning/HW		
Travel		
Medicine		
War		
Social		
Art		
Companionship		
Entertainment		
Robots		