CMPT 423 Final Exam—Winter 2025

Deadline: April 15 at 2pm CST on Canvas

Name: Enter your name here.
NSID: Enter your NSID here.

Submission instructions

To complete this take-home exam, enter your answers below each question by replacing the "ENTER YOUR ANSWER HERE AND REMOVE THIS LINE" section. Generate a .pdf file using this Latex template and submit the .pdf file as part of the Canvas Assignment "CMPT 423 Final Exam".

Make sure to submit early and to submit often on Canvas. As a first step, you should enter your name and NSID into this template, generate the .pdf file and submit this final without answers on Canvas. As you work through this exam, you can always submit your answers by resubmitting on Canvas.

1 Short answer questions

Answer each question in at most five sentences.

1.1 Inductive bias [grade weight: 10%]

What is an inductive bias in a machine learning model? Provide an explanation using an example.

Answer

INSERT YOUR ANSWER HERE AND REMOVE THIS LINE.

1.2 Logistic regression [grade weight: 10%]

What is the purpose of the sigmoid function in logistic regression, and how does it help in modelling binary data?

Answer

INSERT YOUR ANSWER HERE AND REMOVE THIS LINE.

1.3 Generalization [grade weight: 10%]

What is the purpose of cross-validation in machine learning, and how does it help prevent overfitting?

Answer

INSERT YOUR ANSWER HERE AND REMOVE THIS LINE.

1.4 Reinforcement learning [grade weight: 10%]

Suppose you train a RL system, for example DQN, to play an Atari video game. Instead of converging to a high-performing policy, DQN is stuck looping back and forth between a small number of states resulting in few rewards and poor performance. Describe one technique to address this issue to ensure that DQN converges to an optimal policy that maximizes reward in the game.

Answer

INSERT YOUR ANSWER HERE AND REMOVE THIS LINE.

2 Maximum likelihood method [grade weight: 20%]

Suppose you are given a dataset of scalar values

$$\mathcal{D} = \{x_1, ..., x_n | x_i \in \mathbb{R}\}.$$

Assume that each $x_i \in \mathcal{D}$ is sampled i.i.d. from a normal distribution with mean μ and standard deviation one. Derive the maximum likelihood estimate $\widehat{\mu}$ for the parameter μ given the dataset $\mathcal{D}$.

Answer

INSERT YOUR ANSWER HERE AND REMOVE THIS LINE.

3 Language models [grade weight: 20%]

Suppose you use the maximum likelihood method train a language model to perform list sorting. This can be accomplished by training the model with the following prompt-response pairs

```
Prompt:   "[ 5, 3, 6, 0 ]"
Response: "[ 0, 3, 5, 6 ]"
```

How would you create a token dataset to train a model that can sort lists and respond as in the example above? Considering the different sequence models discussed in class, which architecture would you use and why?

Answer

INSERT YOUR ANSWER HERE AND REMOVE THIS LINE.

4 Markov Decision Processes

Consider the following MDP with five states and an action space $\mathcal{A} = \{a, b\}$:

s_1 $a, p = 1,$ $r = 0.1$ s_0 b $p = 1 - \alpha$ $r = 0$ s_2 $a, b, p = 1,$ $r = -1$ s_3

$p = \alpha, r = 1$ s_4
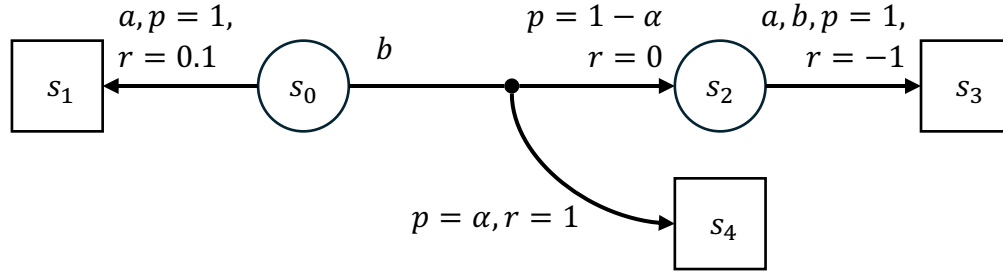
When selecting the action a in state s_0 a reward of 0.1 is given to the agent. When selecting the action b in state s_1, the agent transitions to state s_4 with probability α and a reward of 1 is given. With probability $1 - \alpha$, the agent transitions to state s_2 and no reward is given. From state s_2 the agent can select actions a or b again, but both actions cause a deterministic transition to state s_3 for which a reward of -1 is received. The states s_1, s_4, and s_3 are all absorbing, meaning selecting any action at these states leads to a reward of zero and a transition back into the same state.

Depending on the setting of α, the action a or b leads to a higher γ-discounted return. In this questions you will use Q-value functions, to derive a bound for α where action b is optimal in state s_0 and leads to the highest γ-discounted return.

Part A [grade weight: 6%]

What is the Q-value $Q(s_0, a)$ of selecting action a in state s_0 for a discount factor γ?

Answer

INSERT YOUR ANSWER HERE AND REMOVE THIS LINE.

Part B [grade weight: 6%]

What is the Q-value $Q(s_0, b)$ of selecting action b in state s_0 for a discount factor γ?

Answer

INSERT YOUR ANSWER HERE AND REMOVE THIS LINE.

Part C [grade weight: 8%]

For which setting of α is selecting action b optimal?

Answer

INSERT YOUR ANSWER HERE AND REMOVE THIS LINE.