Assignment 5

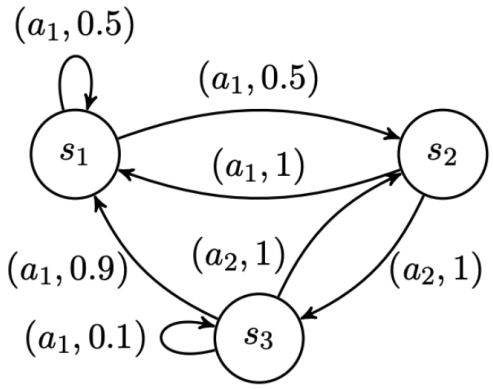
CMSC 478 – Machine Learning

April 5, 2024

Item	Summary
Assigned	Apr 4, 2023
Due	Apr 11, 11:59 PM
Topic	PCA, Reinforcement Learning
Points	70 (+30)

You are to complete this assignment on your own: that is, the code and writeup you submit must be entirely your own. However, you may discuss the assignment at a high level with other students or on the discussion board. Note at the top of your assignment who you discussed this with or what resources you used (beyond course staff, any course materials, or public Campuswire discussions).
Language and External Resources Your code must be compiled, and you will use Python. Failure to do so, or the use of any external resources, including generative AI, and other people's work, without prior written permission from the instructor, will be considered an academic integrity violation and result, at a minimum, in a 0 on this assignment.

1. Reinforcement Learning (50 points): Consider the MDP in Figure 1. There are three states: s1, s2,



(a) Transition graph

R	a_1			a_2		
	s_1	s_2	s_3	s_1	s_2	s_3
s_1	-1	0.5	—	—	—	—
s_2	10	—	—	—	—	-1
s_3	0	—	-1	—	10	—

(b) Reward table for the MDP

Figure 1: Example of a 3-state, 2-action MDP with (action, probability) on arcs

and s3. There are two actions: a1 and a2. Edges are labeled with (action, probability) pairs. For example, taking action a1 in state s1 leads back to s1 with probability 0.5 and to s2 with probability 0.5.
The table to the right gives rewards. The left column is the starting state. For each action, the rewards are given for each destination state. For example, taking action a1 in state s2 with returning to s1 yields a reward of 10.
Suppose you start with a Q-table initialized to all zeroes. Show the value of the Q-table after each of the following transitions, taken in order, with a learning rate and discount factor of 0.5. Note that the starting

Q-table for the second update is the table after the first update.

S1 $\xrightarrow{A1}$ S2
S2 $\xrightarrow{A2}$ S3

Show your work for partial credit.
What is the optimal policy for this MDP? Explain briefly why.

2. PCA (20 points):

Suppose you have a dataset with multiple features. For example, the red wine dataset. You can download the dataset from http://archive.ics.uci.edu/ml/datasets/Wine+Quality.

(a) Suppose, you used logistic regression to classify after applying PCA on the original dataset. You obtained different test accuracy for when you choose number of principal components, K=1 and K=3. Explain why that happens?

(b) How would you choose the optimal number of PCs to retain 95% variance?

Bonus Question (30 points): Write a program that classifies red wine dataset with logistic regression after applying PCA. Answer 2(b) with the results from your program.