COMP7015 Artificial Intelligence **Semester 1, 2024-2025**

Written Assignment 2

Instructor: Dr. Kejing Yin due on Dec. 9, 2024 (11:59am)

Submission instruction: *Put your answers to all questions to one single PDF file, make sure that your answers are readable. Name the file as "*`wa2_StudentID_YourFullName.pdf`*" and upload it to Moodle submission box.*

Problem 1 Computational Graphs and Backpropagation (25 marks)

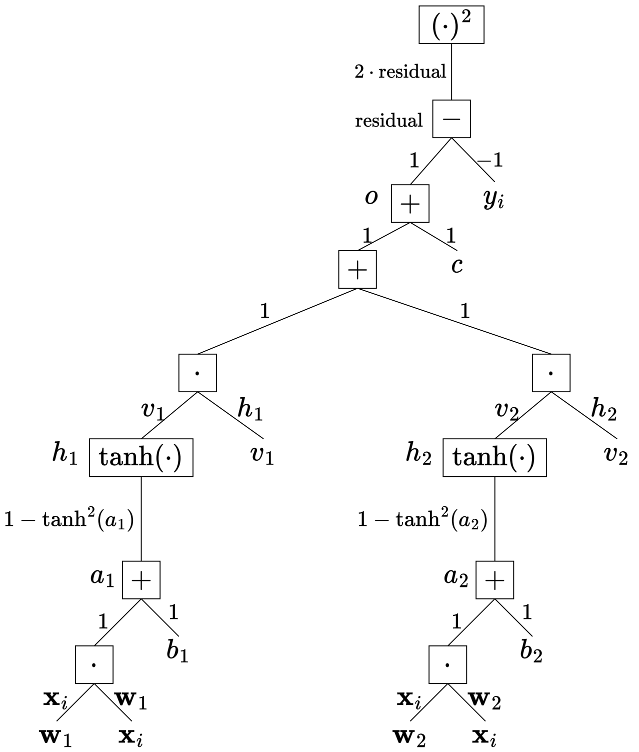
Given a dataset $\mathcal{D} = \{(\mathbf{x}_1, y_1), \ldots, (\mathbf{x}_n, y_n)\}$, consider a neural network, which computes the loss function for the i-th data sample using the following formulas:

$$\begin{aligned}
\mathbf{a}_i =& \mathbf{b} + \mathbf{W}\mathbf{x}_i, \\
\mathbf{h}_i =& \tanh(\mathbf{a}_i), \\
o_i =& \mathbf{v}^\top \mathbf{h}_i + c, \\
\ell_i =& (o_t - y_t)^2,
\end{aligned}$$

where $\tanh(z) = \frac{e^z - e^{-z}}{e^z + e^{-z}}$. For simplicity, let's assume the input data is three dimensional (*i.e.*, $\mathbf{x} \in \mathbb{R}^3$), the hidden size is 2 (*i.e.*, $\mathbf{h} \in \mathbb{R}^2$) in this problem.

(a) (12 marks) Construct the computational graph for ℓ_i.

Solution: The computational graph is given by:



(b) (5 marks) What are the model parameters to be learned?

Solution: The model parameters to be learned are $\mathbf{b}$, $\mathbf{W}$, $\mathbf{v}$, and c.

(c) (8 marks) Given the following model parameters and data, compute the loss function in forward pass and compute the gradient with respect to all model parameters in backward pass.

$$\begin{aligned}\mathbf{b} =& [0.1, 0]^\top \\ \mathbf{W} =& \begin{bmatrix} 0.8 & -0.3 & 0.5 \\ -1 & 0.1 & 0.9 \end{bmatrix} \\ c =& 0.7 \\ \mathbf{v} =& [0.5, 0.8]^\top \\ \mathcal{D} =& \{([1, -1, 1]^\top, 1.5)\}\end{aligned}$$

Solution: (1) The forward pass:

$$\begin{aligned}h_1 =& \tanh\left(\mathbf{w}_1^\top \mathbf{x}_1 + b_1\right) = \tanh(1.7) = 0.9354 \\ h_2 =& \tanh\left(\mathbf{w}_2^\top \mathbf{x}_1 + b_2\right) = \tanh(-0.2) = -0.1974 \\ o =& (h_1 v_1 + h_2 v_2) + c = 1.0098 \\ \ell =& (o - y)^2 = 0.2403\end{aligned}$$

(2) The backward pass:

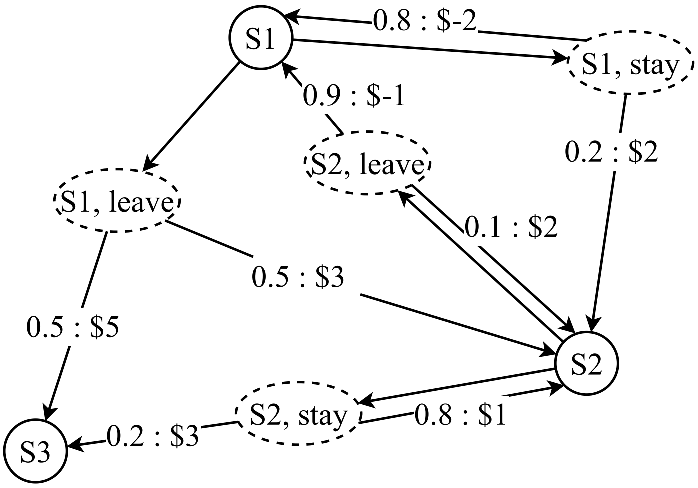
$$\begin{aligned}\frac{\partial \ell}{\partial c} =& 2 \cdot (o-y) \cdot 1 \cdot 1 = -0.9804 \\ \frac{\partial \ell}{\partial v_1} =& 2 \cdot (o-y) \cdot 1 \cdot 1 \cdot 1 \cdot h_{11} = -0.9171 \\ \frac{\partial \ell}{\partial v_2} =& 2 \cdot (o-y) \cdot 1 \cdot 1 \cdot 1 \cdot h_{12} = 0.1935 \\ \frac{\partial \ell}{\partial b_1} =& 2 \cdot (o-y) \cdot 1 \cdot 1 \cdot 1 \cdot v_1 \cdot (1 - \tanh^2(a_1)) \cdot 1 = -0.0613 \\ \frac{\partial \ell}{\partial b_2} =& 2 \cdot (o-y) \cdot 1 \cdot 1 \cdot 1 \cdot v_2 \cdot (1 - \tanh^2(a_2)) \cdot 1 = -0.7538 \\ \frac{\partial \ell}{\partial \mathbf{w}_1} =& 2 \cdot (o-y) \cdot 1 \cdot 1 \cdot 1 \cdot v_1 \cdot (1 - \tanh^2(a_1)) \cdot 1 \cdot 1 \cdot \mathbf{x}_1 = [-0.0613, 0.0613, -0.0613] \\ \frac{\partial \ell}{\partial \mathbf{w}_2} =& 2 \cdot (o-y) \cdot 1 \cdot 1 \cdot 1 \cdot v_2 \cdot (1 - \tanh^2(a_2)) \cdot 1 \cdot 1 \cdot \mathbf{x}_1 = [-0.7538, 0.7538, -0.7538]\end{aligned}$$

Therefore, the gradient with respect to the model parameters are:

$$\begin{aligned}\nabla_{\mathbf{b}} \ell =& [-0.0613, -0.7538]^\top \\ \nabla_{\mathbf{W}} \ell =& \begin{bmatrix} -0.0613 & 0.0613 & -0.0613 \\ -0.7538 & 0.7538 & -0.7538 \end{bmatrix} \\ \nabla_{\mathbf{v}} \ell =& [-0.9171, 0.1935]^\top \\ \nabla_c \ell =& -0.9804\end{aligned}$$

Problem 2 Markov Decision Process (25 marks)

The following figure describes a Markov decision process.



(a) (6 marks) Define the transition function based on the figure above. Namely, list out the values of $\mathrm{T}(s, a, s')$ for all possible combinations of s, a, and s'.

Solution: The transition function is defined as below.

s	a	s'	$\mathrm{T}(s, a, s')$
S1	stay	S1	0.8
S1	stay	S2	0.2
S1	stay	S3	0
S1	leave	S1	0
S1	leave	S2	0.5
S1	leave	S3	0.5
S2	stay	S1	0
S2	stay	S2	0.8
S2	stay	S3	0.2
S2	leave	S1	0.9
S2	leave	S2	0.1
S2	leave	S3	0

Note: The rows with zero transition probability can be omitted.

(b) (6 marks) Similarly, define the reward function based on the figure above. Namely, list out the values of $\mathrm{Reward}(s, a, s')$ for all possible combinations of s, a, and s'.

Solution: The reward function is defined as below.

s	a	s'	$\mathrm{Reward}(s, a, s')$
S1	stay	S1	$-2
S1	stay	S2	$2
S1	leave	S2	$3
S1	leave	S3	$5
S2	stay	S2	$1
S2	stay	S3	$3
S2	leave	S1	$-1
S2	leave	S2	$2

(c) (13 marks) Given a policy "$\pi(S1) = $ leave; $\pi(S2) = $ stay", use iterative algorithm to compute $V_\pi(S1)$ and $V_\pi(S2)$. Give the estimation after three iterations.

Solution: The steps of running the iterative algorithm are as below.
1) We initialized $V_\pi^{(0)}(S1) = 0, V_\pi^{(0)}(S2) = 0, V_\pi^{(0)}(S3) = 0$, $S3$ is the end state, so $V_\pi(S3)$ does not need to updated. Then, we follow the following iterative update rule.

$$V_\pi^{(t)}(s) \leftarrow \sum_{s'} T\left(s, \pi(s), s'\right)\left[\text{Reward}\left(s, \pi(s), s'\right) + \gamma V_\pi^{(t-1)}\left(s'\right)\right]$$

2) $t = 1$:

$$V_\pi^{(1)}(S1) \leftarrow 0.5 \times (3 + 1 \times 0) + 0.5 \times (5 + 1 \times 0) = 4$$

$$V_\pi^{(1)}(S2) \leftarrow 0.8 \times (1 + 1 \times 0) + 0.2 \times (3 + 1 \times 0) = 1.4$$

2) $t = 2$:

$$V_\pi^{(2)}(S1) \leftarrow 0.5 \times (3 + 1 \times 1.4) + 0.5 \times (5 + 1 \times 0) = 4.7$$

$$V_\pi^{(2)}(S2) \leftarrow 0.8 \times (1 + 1 \times 1.4) + 0.2 \times (3 + 1 \times 0) = 2.52$$

2) $t = 3$:

$$V_\pi^{(3)}(S1) \leftarrow 0.5 \times (3 + 1 \times 2.52) + 0.5 \times (5 + 1 \times 0) = 5.26$$

$$V_\pi^{(3)}(S2) \leftarrow 0.8 \times (1 + 1 \times 2.52) + 0.2 \times (3 + 1 \times 0) = 3.416$$

Problem 3 Q-Learning (25 marks)

An agent lives in a 2×2 grid, and is using Q-learning to learn a policy. The state is the index of square that the agent locates. The agent will get reward $r = 10$ when it lands in State 3. There are no other rewards or penalties.

1	2
3 reward: 10	4

The agent has four possible actions, which are MoveEast, MoveWest, MoveNorth, and MoveSouth. The Q-table is initialized to all zeros. The agent starts at Square 1 and observe the following episodes:

Episode	State s	Action a	Reward r	Resulting State s'
#1	1	MoveSouth	(i)	3
#2	3	MoveEast	(ii)	4
#3	4	MoveNorth	(iii)	2
#4	2	MoveWest	(iv)	1
#5	1	MoveEast	(v)	3

After each episode, the Q-table is updated correspondingly.

(a) (10 marks) Determine the values of (i–v) in the table above based on the descriptions of the grid world.

Solution: The values of (i-v) are 10, 0, 0, 0, and 10, respectively.

(b) (15 marks) List out the nonzero entries of the Q-table after all updates. Show your steps of making updates to the Q-table.

Solution: The Q-table is initialized to all zeros as follows.

	MoveSouth	MoveEast	MoveNorth	MoveWest
1	0	0		
2	0			0
3		0	0	
4			0	0

In this example solution, we use $\eta = 0.9$ and $\gamma = 1$. It's ok to use other values as long as you explicitly state them in your solutions.

1) $s = 1$, $a =$MoveSouth, $s' = 3$, $r = 10$.

$$Q(1, \text{MoveSouth}) \leftarrow Q(s,a) + \eta \left(r + \left[\gamma \max_{a'} Q(s',a') \right] - Q(s,a) \right) = 0 + 0.9 \times (10 + 0 - 0) = 9$$

	MoveSouth	MoveEast	MoveNorth	MoveWest
1	9	0		
2	0			0
3		0	0	
4			0	0

2) $s = 3$, $a =$MoveEast, $s' = 4$, $r = 0$.

$$Q(3, \text{MoveEast}) \leftarrow Q(s,a) + \eta \left(r + \left[\gamma \max_{a'} Q(s',a') \right] - Q(s,a) \right) = 0 + 0.9 \times (0 + 0 - 0) = 0$$

3) $s = 4$, $a =$MoveNorth, $s' = 2$, $r = 0$.

$$Q(4, \text{MoveNorth}) \leftarrow Q(s,a) + \eta \left(r + \left[\gamma \max_{a'} Q(s',a') \right] - Q(s,a) \right) = 0 + 0.9 \times (0 + 0 - 0) = 0$$

4) $s = 2$, $a =$MoveWest, $s' = 1$, $r = 0$.

$$Q(2, \text{MoveWest}) \leftarrow Q(s,a) + \eta \left(r + \left[\gamma \max_{a'} Q(s',a') \right] - Q(s,a) \right) = 0 + 0.9 \times (0 + 1 \times 9 - 0) = 8.1$$

5) $s = 1$, $a =$MoveEast, $s' = 3$, $r = 10$.

$$Q(1, \text{MoveEast}) \leftarrow Q(s,a) + \eta \left(r + \left[\gamma \max_{a'} Q(s',a') \right] - Q(s,a) \right) = 0 + 0.9 \times (10 + 1 \times 0 - 0) = 9$$

	MoveSouth	MoveEast	MoveNorth	MoveWest
1	9	9		
2	0			8.1
3		0	0	
4			0	0

Therefore, at the end of this phase, the nonzero entries of the Q-table are:

$$Q(1, \text{MoveSouth}) = 9, \quad Q(1, \text{MoveEast}) = 9, \quad Q(2, \text{MoveWest}) = 8.1$$

Problem 4 Naive Bayes Classifier (25 marks)

In a clinical study about seasonal flu, data were collected from 90 people. The following table shows the number of flu patient and healthy people having fever and sore throat, respectively.

Flu			Healthy		
Fever	Sore throat	count	Fever	Sore throat	count
Yes	Yes	11	Yes	Yes	2
Yes	No	5	Yes	No	3
No	Yes	8	No	Yes	8
No	No	2	No	No	51
	Total:	26		Total:	64

(a) (15 marks) Consider a person with a sore throat but no fever. What would be the prediction generated by a naïve Bayes classifier trained using the data above? Show the steps of getting your answer.

Solution: Step 1. The prior probabilities are given by:

$$p(\text{Flu}) = 0.29 \text{ and } p(\text{Healthy}) = 0.71.$$

Step 2. The conditional probability of each feature is given by:

$$\begin{aligned}
p(\text{Fever} = \text{No}|\text{Flu}) &= 10/26 = 0.38, \\
p(\text{Fever} = \text{No}|\text{Health}) &= 59/64 = 0.92, \\
p(\text{Sore throat} = \text{Yes}|\text{Flu}) &= 19/26 = 0.73, \\
p(\text{Sore throat} = \text{Yes}|\text{Health}) &= 10/64 = 0.16,
\end{aligned}$$

Step 3. Hence, for the person with sore throat but no fever, we have

$$p(\text{Flu})p(\text{Fever} = \text{No}|\text{Flu})p(\text{Sore throat} = \text{Yes}|\text{Flu}) \approx 0.08$$

$$p(\text{Healthy})p(\text{Fever} = \text{No}|\text{Health})p(\text{Sore throat} = \text{Yes}|\text{Health}) \approx 0.10.$$

Since $0.1 > 0.08$, the naïve Bayes classifier will classify this person as Healthy.

(b) (10 marks) Name at least two advantages and two disadvantages of the naïve Bayes classifier.

Solution: The advantages: (1) Naïve Bayes classifier is very fast. (2) the probability used in naïve Bayes classifier can easily be updated when new training data comes in; therefore, it is very easy to achieve incremental learning. (3) It is easy to implement.

The disadvantage: (1) It assumes that all features are independent, which may not align with real datasets. (2) zero frequency problem: if a categorical feature has a value in test set that is not observed in training set, naïve Bayes classifier will assign a zero probability and cannot make predictions.