السؤال الأول: (20 درجة)

شبكة عصبية (موضحة في الشكل 1) تخضع لمرحلة تدريب باستخدام خوارزمية الانتشار العكسي (Backpropagation). إذا كانت بيانات التدريب المحدودة هي: مدخلات (x_1=0.18, x_2=0.54) وقيم المخرجات المستهدفة (y_d=0.9)، والأوزان الحالية لكلتا الطبقتين كانت كما يلي:

Hidden layer:

$$W = \begin{bmatrix} 0.9 & 1.35 \\ 0.45 & 1.8 \end{bmatrix} \qquad b = \begin{bmatrix} 0.9 \\ -0.45 \end{bmatrix}$$

Output layer:

$$W = [2.7 \quad 1.8] \qquad b = [-3.6]$$

تم تعديل معمارية (Topology) الشبكة العصبية لتصبح كما بالشكل 2 والأوزان الحالية للطبقتين كانت كالتالي:

Hidden layer:

$$W = \begin{bmatrix} 0.9 & 1.35 \\ 0.45 & 1.8 \\ 2.25 & 2.7 \end{bmatrix} \qquad b = \begin{bmatrix} 0.9 \\ -0.45 \\ 1.35 \end{bmatrix}$$

Output layer:

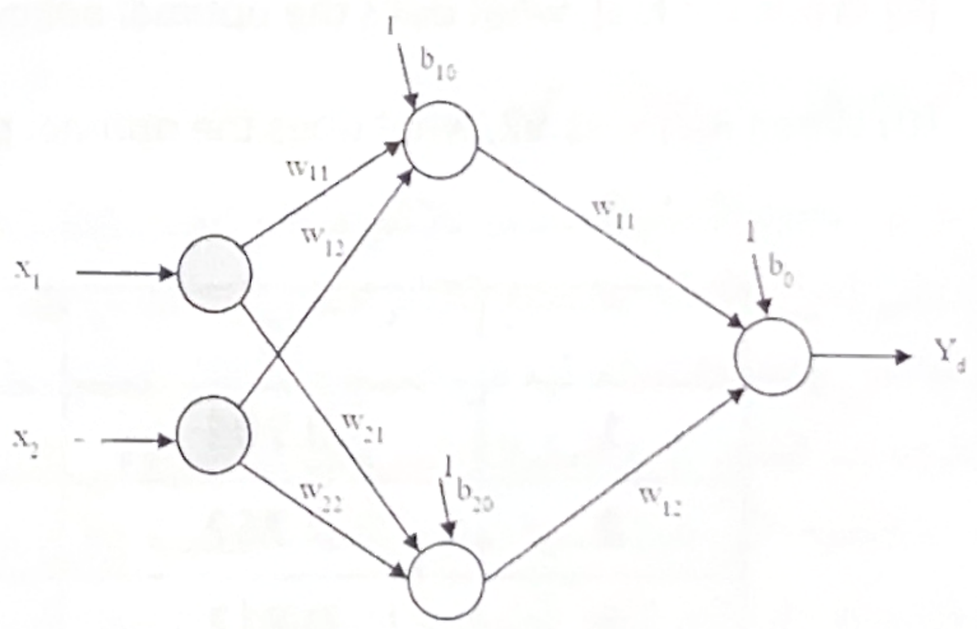
$$W = [2.7 \quad 1.8 \quad 0.9] \qquad b = [-3.6]$$

أجب عن الأسئلة التالية:

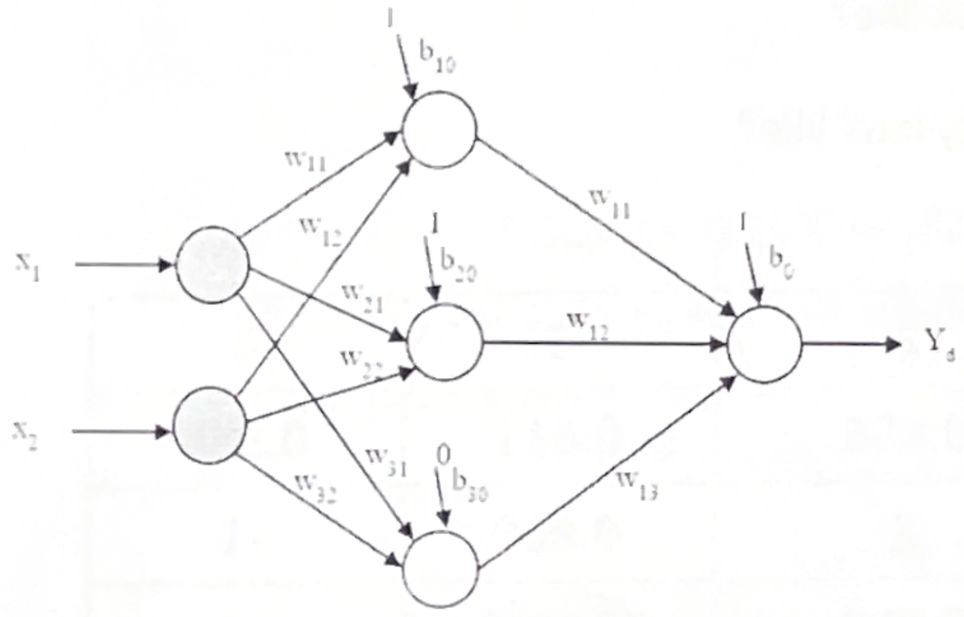
(أ) احسب مخرجات الشبكتين وقارن بين النتائج من حيث دقة الوصول إلى الحل.

(ب) حدد القيم الجديدة للأوزان بعد إجراء دورة تدريب واحدة (One iteration) للشبكة في الشكل 1، مع الأخذ في الإعتبار بأن معامل التعليم هو 0.1 ودالة التنشيط المستخدمة هي "Sigmoid".

(ج) هل أظهرت النتائج في الفقرة (ب) تحسنًا في دقة الحل بعد التدريب؟



الشكل 1



الشكل 2

Q2 (20 marks):

Assuming a 3x4 grid world problem:

- Let s_{ij} denote the position in row i and column j.
- s_{11} is the initial state.
- There is a wall at s_{22}
- s_{24} and s_{34} are goal states
- The robot escapes the world upon reaching either goal state.
- There are four actions: up, down, left, and right.
- Each action is possible in every state.

The transition model $P(s'|s,a)$:

- An action achieves its intended effect with probability 0.7.
- An action leads to a 90-degree left turn with probability 0.1.
- An action leads to a 90-degree right turn with probability 0.2.
- If the robot bumps into a wall, it stays in the same square.

The reward function $R(s)$ is the reward of entering state s:

- $R(s_{24})=-1$
- $R(s_{34})= 1$
- Otherwise, $R(s)=-0.04$

(A) The robot is in s_{21} and tries to move to the right. What is the probability that the robot stays in s_{21}?

(B) What is the optimal action for state s_{33}?

(C) When 0 < R(s), what does the optimal policy look like?

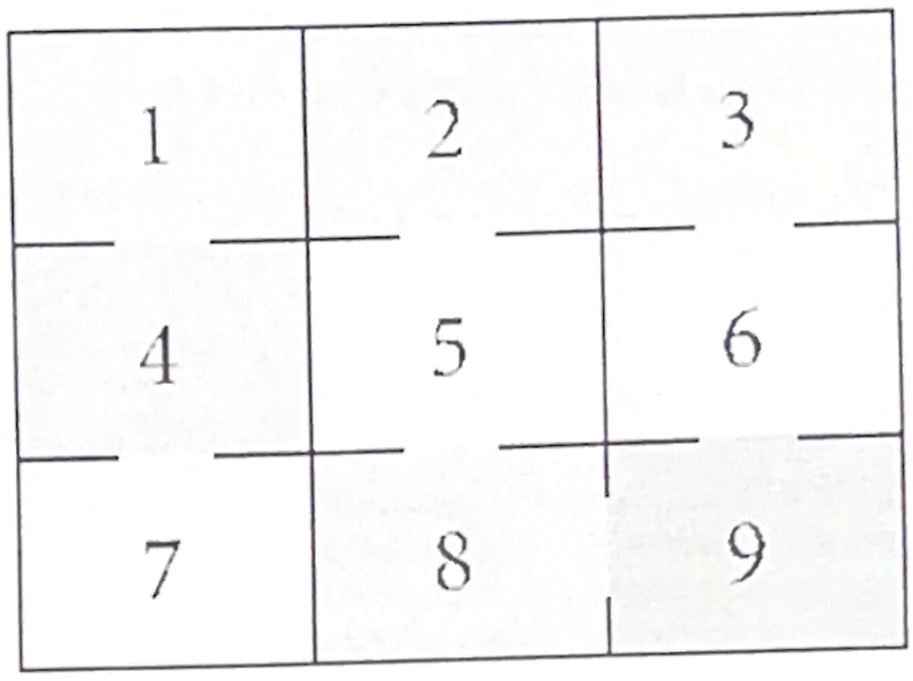
(D) When R(s) < −1.92, what does the optimal policy look like?

	1	2	3	4
1	0.705	0.655	0.611	0.388
2	0.762	X	0.660	-1
3	0.812	0.868	0.918	+1

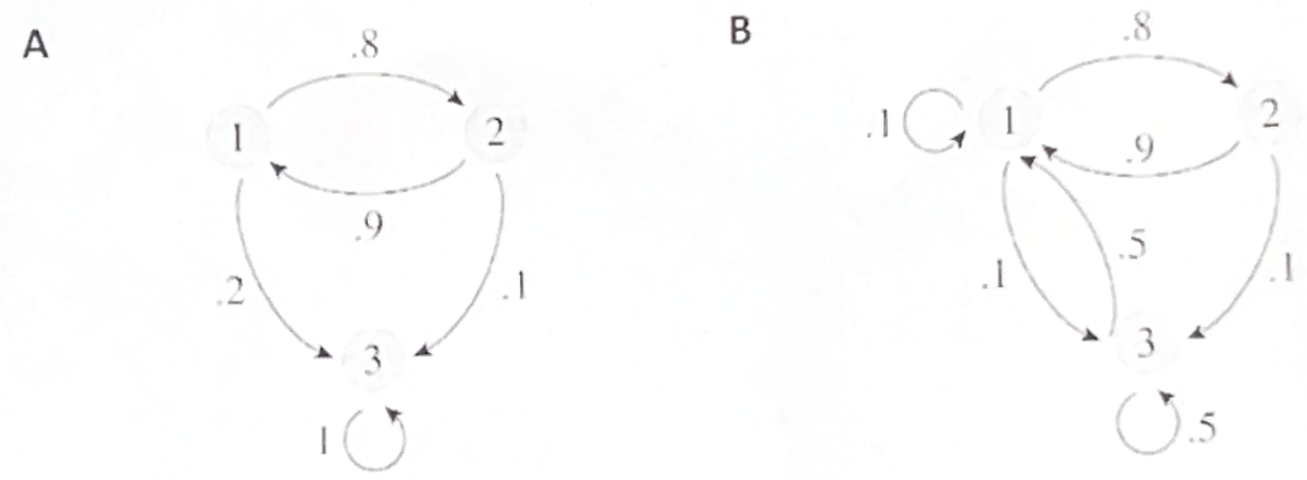
Q3 (5 marks): Consider the maze with 9 rooms shown in the figure. The system çonsists of the maze and a robot. We assume that the following learning pattern exists: If the robot is in room 1, 2, 3, 4, or 5, it either remains in that room or it moves to any room that the maze permits, each having the same probability; if it is in room 8, it moves directly to room 9; if it is in room 6, 7, or 9, it remains in that room.

(A) Explain why the above experiment is a Markov chain.

(B) Construct the transition matrix P.



Q4 (5 marks): Write down the transition matrix associated with each state transition diagram.



السؤال الرابع: (10 درجات)

(أ) ما هو مفهوم التعلم الخاضع للإشراف؟ وكيف يختلف عن الأنواع الأخرى من التعلم؟
(ب) متى يكون التعلم الخاضع للإشراف خيارًا أفضل مقارنة بالتعلم غير الخاضع للإشراف والعكس؟
(ج) ما هي التطبيقات العملية في الصناعات التي تعتمد على التعلم الخاضع وغير الخاضع للإشراف؟
(د) كيف يتم تحديد نوع المشكلة (تصنيف أو تنبؤ) في التعلم الخاضع للإشراف؟
(هـ) ما التقنية التي تستخدمها نماذج اللغة الضخمة LLM لإنتاج محتوى؟
(و) هل يمكن لنماذج اللغة الضخمة أن تستبدل الإنسان في بعض الأعمال؟ لماذا أو لماذا لا؟

Good Luck