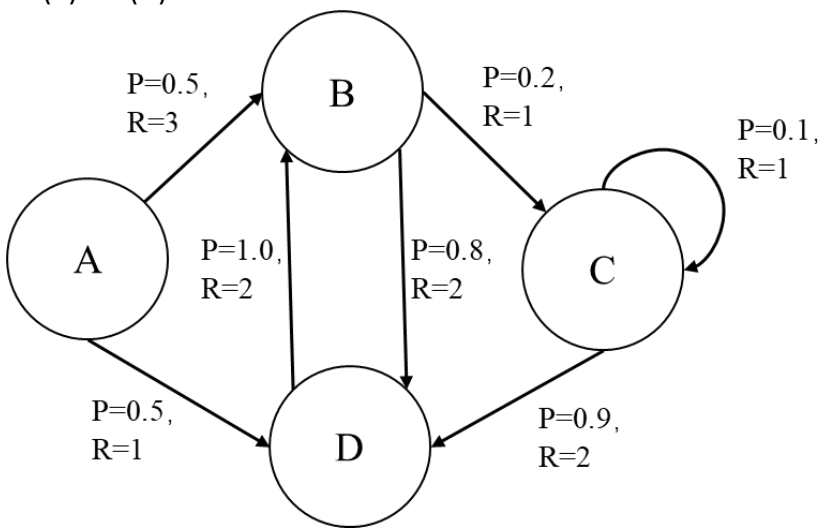
EBU4203 Introduction to AI – Week 2 Tutorial 2023

Q1: Draw a graphical representation of the Rectified Linear Unit (ReLU) activation function and provide a concise explanation of its key advantages in neural network activation functions.

Q2: Name two methods to prevent overfitting in training Deep Neural Networks (DNNs).

Q3: In the context of reinforcement learning, define the concept of a State Value Function and elaborate on its role in aiding an agent's decision-making process

Q4: Consider the following deterministic Markov Decision Process (MDP). The discount factor γ is 0.9. States are represented as A, B, C, and D. Arrows indicate state transition with corresponding actions. The action probability and immediate rewards are P and R next to the arrows, respectively. The MDP starts with an initial value function of V0(A)=V0(B)=V0(C)=V0(D)=4.



a) Complete the Markov Chain transition matrix for the given problem.
b) For **one iteration**, calculate the value function $V_1(D)$.
c) For **one iteration**, calculate the value function V_1 of states A, B, and C.

Q5: Given the provided data pairs where (y_1, y_2, y_3) represent the true data points, and $(\hat{y}_1, \hat{y}_2, \hat{y}_3)$ represent the predicted data points:

- $y_1 = [0, 0, 1], \hat{y}_1 = [0.1, 0.2, 0.7]$
- $y_2 = [0, 1, 0], \hat{y}_2 = [0.1, 0.7, 0.2]$
- $y_3 = [1, 0, 0], \hat{y}_3 = [0.3, 0.4, 0.3]$

Answer the following questions:

a) Calculate the Mean Squared Error of this dataset
b) Calculate the Cross Entropy Loss

Q6. In Figure 1, given a 4x4 Gridworld with the agent starts at the initial place s_{11} and tries to find the route to reach the goal place s_{44} with the possible movement as North, South, East and West. In the given Gridworld, s_{xy} defines different positions that the agent will pass through, where:

- $s_{11}$is the starting state (S)
- $s_{44}$is the goal state (G), when the agent arrives here, it will receive a reward +10
- s_{23} and s_{32} is a penalty state (P), when the agent arrives there, it will get a negative reward -5
- When the agent arrives all the other states, it will get a negative reward -1
- assuming a discount factor $\gamma = 0.8$

Assume a deterministic policy is used.

s_{11}, S	s_{12}	s_{13}	s_{14}
s_{21}	s_{22}	s_{23}, P	s_{24}
s_{31}	s_{32}, P	s_{33}	s_{34}
s_{41}	s_{42}	s_{43}	s_{44}, G

Figure 1. the Map of Gridworld.

a) Formulate the Gridworld example as a Markov Decision Process with defining each component.

b) Q-learning algorithm, as one of the classical reinforcement learning (RL) algorithms, is commonly used to solve this Markov Decision Process problem. Answer the following few questions about Q-learning algorithm:
   i) There are two types of RL algorithms, model-based RL and model-free RL, which one is Q-learning belong to? Give the reasons.
   ii) Given the update rule of Q-learning as

$$Q^{update}(s_t, a_t) = Q(s_t, a_t) + \alpha[r_{t+1} + \gamma max_a Q(s_{t+1}, a) - Q(s_t, a_t)]$$

Assume the agent starts from the place s_{11} at $t = 0$, it knows the action is 'East', and thus reaches the place, s_{12}, i.e., $t = 1$. The initial Q-value of place s_{12} is set as $Q(s_{12}, a) = 0.2$ and assuming $Q(s_{11}, a) = 0$ and learning rate $\alpha = 0.1$, what is the Q-value update would be?