# 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 black background with a black square

Description automatically generated with medium confidence

1. Complete the Markov Chain transition matrix for the given problem.
2. For **one** **iteration**, calculate the value function V1(D).
3. For **one** **iteration**, calculate the value function V1 of states A, B, and C.

Q5: Given the provided data pairs where represent the true data points, and represent the predicted data points:

Answer the following questions:

1. Calculate the Mean Squared Error of this dataset
2. Calculate the Cross Entropy Loss

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

* is the starting state (S)
* is the goal state (G), when the agent arrives here, it will receive a reward +10
* and 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

Assume a deterministic policy is used.

|  |  |  |  |
| --- | --- | --- | --- |
| , S |  |  |  |
|  |  | , P |  |
|  | , P |  |  |
|  |  |  | , G |

Figure 1. the Map of Gridworld.

1. Formulate the Gridworld example as a Markov Decision Process with defining each component.
2. 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:
   1. There are two types of RL algorithms, model-based RL and model-free RL, which one is Q-learning belong to? Give the reasons.
   2. Given the update rule of Q-learning as

Assume the agent starts from the place at , it knows the action is ‘East’, and thus reaches the place, , i.e., . The initial Q-value of place is set as and assuming and learning rate , what is the Q-value update would be?