考試科目	機器學習	系所別	人工智慧博士學位學程	命題教師	

1. Please explain the process of data preprocessing in more detail, including the reasons for such pre-processing, how to deal with the problem, and why this pre-processing can solve the problem. Please illustrate the above concepts with an actual data set. (20%)

2. When constructing a machine learning model, several different machine learning models are usually considered. Please explain how to evaluate the performance of machine learning models and the process of model selection. Please illustrate the above concepts with examples. (20%)

3. Given three classification algorithms. How can you order them from best to worst? It is noticed that you have to consider both how to assess the performance of a classification algorithm and how to design an experiment for comparing the given algorithms. (15%)

4. Consider a maze problem as shown in the following figure, where a robot is located in the maze and will be guided to learn to reach the goal. In the configuration of this problem, states are positions in the maze which the robot can visit; actions are allowable directions along which the robot can move (e.g., Up, Down, Left, Right). These four actions have their usual effects unless blocked by a wall, and the robot does not know where the internal walls are. In any given state, the robot can tell whether the state is one it has ever visited before or a new state.
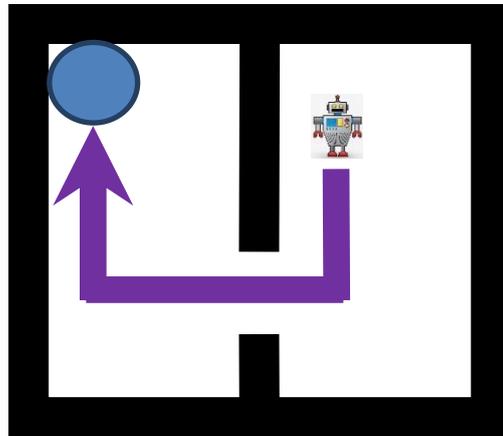
Suppose that this robot knows its initial position and where the goal is. Please describe how it can learn to reach the goal as quickly as possible, without explicit supervision of which direction to move. (15%)

5. Different machine learning methods are commonly used to solve the regression, classification, and grouping problems. Please explain the differences in theories, especially for the mathematical principles, when machine learning methods solve the regression, classification, and grouping problems. You can give examples of machine learning techniques to explain it. (15%)

6. For a specific problem, we can usually use statistical methods, machine learning methods or deep learning methods to solve it. Please explain the difference between these methods. (15%)