亞洲大學博士班資格考考試試題

文件編號：AC-0054
機密等級：2 限閱
版本：1

考試日期： 110 年 2 月 1 日

科目	機器學習	系所別	人工智慧博士學位學程	命題教師	

1. For each of the following tasks, identify which type of learning is involved (supervised, unsupervised, or reinforcement) and the training data to be used. If a task can fit more than one type, explain how and describe the training data for each type.
Recommending a book to a user in an online bookstore.
Playing tic-tac-toe
Categorizing movies into difference types
Learning to play music.

2. Please explain the concept of VC dimension., and show the VC dimension of the triangle model is 7. Consider the "triangle" learning model, where $h : \mathbb{R}^2 \rightarrow \{-1, +1\}$ and $h(x) = +1$ if x lies within an arbitrarily chosen triangle in the plane and −1 otherwise.

3. Describe the main idea of logistic regression, and use a practical example to illustrate your description.

4. Illustrate the Analysis Process such as Supervised Learning (e.g. spam classification, credit card fraud detection)

5. Please explain how to use various validation metrics (e.g. confusion matrix, ROC curve, AUC, Accuracy, Precision, Recall & F1 Score) to evaluate Machine Learning models.

6. You are assigned the task to build a tool that can distinguish junk email. What is in a junk e-mail that lets you know it is junk? How can the computer detect junk through a syntactic analysis? What would you like the computer to do if it detects a junk e-mail—delete it automatically, move it to a different file, or just highlight it on the screen? Discuss the pros and cons of the potential approaches in a comparative manner.

7. 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 a machine learning experiment for comparing the given algorithms.