國立成功大學

112學年度碩士班招生考試試題

編　號：188

系　所：電腦與通信工程研究所

科　目：人工智慧概論

日　期：0206

節　次：第 1 節

備　註：不可使用計算機

※ 考生請注意：本試題不可使用計算機。　請於答案卷(卡)作答，於本試題紙上作答者，不予計分。

1. (40%) True or False Questions. Explain your reasons.

   (a) We can get multiple local optimum solutions if we solve a linear regression problem by minimizing the sum of squared errors using gradient descent.

   (b) When a decision tree is grown to full depth, it is more likely to fit the noise in the data

   (c) When the hypothesis space is richer, over fitting is more likely.

   (d) When the feature space is larger, over fitting is more likely.

   (e) We can use gradient descent to learn a Gaussian Mixture Model.

   (f) As the number of training examples goes to infinity, your model trained on that data will have Lower variance.

   (g) As the number of training examples goes to infinity, your model trained on that data will have Lower bias.

   (h) Suppose you are given an EM algorithm that finds maximum likelihood estimates for a model with latent variables. You are asked to modify the algorithm so that it finds MAP estimates instead. You need to modify the Expectation step.

2. (20%) Assume we have a set of data from patients who have visited NCKU hospital during the year 2021. A set of features (e.g., temperature, height) have been also extracted for each patient. Our goal is to decide whether a new visiting patient has any of diabetes, heart disease, or Alzheimer (a patient can have one or more of these diseases).

   (a) We have decided to use a neural network to solve this problem. We have two choices: either to train a separate neural network for each of the diseases or to train a single neural network with one output neuron for each disease, but with a shared hidden layer. Which method do you prefer? Justify your answer.

   (b) Some patient features are expensive to collect (e.g., brain scans) whereas others are not (e.g., temperature). Therefore, we have decided to first ask our classification algorithm to predict whether a patient has a disease, and if the classifier is 80% confident that the patient has a disease, then we will do additional examinations to collect additional patient features. In this case, which classification methods do you recommend: neural networks, decision tree, or naive Bayes? Justify your answer in one or two sentences

3. (20%) Explain the principle of the gradient descent algorithm. Accompany your explanation with a diagram. Explain the use of all the terms and constants that you introduce and comment on the range of values that they can take.

4. (20%) For each of the listed descriptions below, answer whether the experimental set up is ok or problematic. If you think it is problematic, briefly state all the problems with their approach:

   (a) A project team performed a feature selection procedure on the full data and reduced their large feature set to a smaller set. Then they split the data into test and training portions. They built their model on training data using several different model settings, and report the best test error they achieved.

   (b) A project team split their data into training and test. Using their training data and cross-validation, they chose the best parameter setting. They built a model using these parameters and their training data, and then report their error on test data.