Final Exam

- 题目数量：7

- 1: 判断题（about 7小题）
  - (　) Both Principle Component Analysis and Spectral Clustering need to perform eigen-decomposition on certain matrices, the size of these two matrices are the same.

© Deng Cai, College of Computer Science, Zhejiang University

2: 多选题 （about 8小题）

- (6) Which of the following statements are true? (　　)
  - **A.** Since *k-means* is an unsupervised learning algorithm, it cannot overfit the data, and thus it is always better to have as large a number of clusters as is computationally feasible.
  - **B.** If we are worried about *k-means* getting stuck in bad local optima, one way to ameliorate (reduce) this problem is if we try using multiple random initializations.
  - **C.** A good way of initializing k-means is setting　.
  - **D.** For some datasets, the "right" or "correct" value of k (the number of clusters) can be ambiguous and hard even for a human expert looking carefully at the data to decide.

© Deng Cai, College of Computer Science, Zhejiang University

3: 简答题 （about 4小题）

- (3) Write down the objective functions of PCA (There are two forms of the objective goals, and you should write down both of them), explain the meaning of these goals.

© Deng Cai, College of Computer Science, Zhejiang University

4~7 问答题

- Give the data, ask you to design a Naïve Classifier.
- Give the data, ask you to build a decision tree.

© Deng Cai, College of Computer Science, Zhejiang University