國立成功大學

# 112學年度碩士班招生考試試題

編　號：257

系　所：數據科學研究所

科　目：計算機概論

日　期：0206

節　次：第 2 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. (45%) In the following statements, please specify if the statement is **True** or **False**. If the statement is True, explain why it is True. If it is **False**, give correct answer or explain why.

   (a) Pruning a decision tree will remove some branches. Pruning a decision tree will decrease model bias.

   (b) Pruning a decision tree will decrease model variance.

   (c) For a graph G and a node v in that graph, the DFS and BFS trees of G rooted at v always contain the same number of edges.

   (d) Prim's algorithm is a greedy algorithm but Kruskal's algorithm is not.

   (e) We cannot determine if an undirected graph $G = (V, E)$ has a cycle in $O(|V| + |E|)$ time.

   (f) Since classification is a special case of regression, logistic regression is a special case of linear regression.

   (g) The training error of k-Nearest Neighbor classifier with $k = 1$ cannot be 0.

   (h) The back-propagation algorithm learns a globally optimal neural network with hidden layers.

   (i) A classifier trained on less training data is less likely to overfit.

   (j) Dijkstra's algorithm can be used to find the shortest path between two nodes in a graph with negative edge weights.

   (k) A hash function maps each key to an array index (i.e., a hash table position). A good hash function uniformly distributes the keys among the available indices.

   (l) The core data structure of Depth-First Search is a queue.

   (m) Adding an element to a heap has worst-case time complexity $O(\log(n))$.

   (n) When optimizing a regression problem where we know that only some of our features are useful, we should use L1 regularization.

   (o) Any problem that can be solved with a greedy algorithm can also be solved with recursive algorithm.

2. (6%) Design and write an algorithm for calculating the number of paths of length k between two given vertices i and j. The graph is unweighted and you know its adjacency matrix A. Also state the runtime of your algorithm in Big-O notation and explain why your algorithm has the specified runtime.

3. (5%) Given the following four functions, please sort their time complexity in a descending order.

   $$F_1(x) = \log x^2 \qquad F_2(x) = (\log x)! \qquad F_3(x) = (x+1)! \qquad F_4(x) = \log(x!)$$

4. (6%) Give the preorder, inorder, and postorder traversal of the following tree.
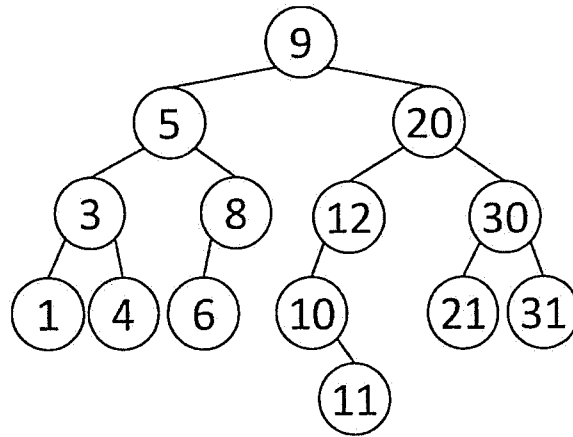
5. (5%) Please read the following Python code, and answer what will be printed out.

```python
num1, num2 = "111", "99"
ans = 0
m, n = len(num1), len(num2)
i = 0
while i < m:
    a = int(num1[m - 1 - i])
    j = 0
    while j < n:
        b = int(num2[n - 1 - j])
        ans += a * b * (10 ** (i + j))
        j += 1
    i += 1
print(ans)
```

6. (9%) Answer the following questions on operating systems.

(a) What is "context switching"? How "context switch" be processed by the operating systems?

(b) How do I/O-bound and CPU-bound programs differ?

(c) Explain the difference between internal fragmentation and external fragmentation.

7. (16%) Answer the following questions on data science.

(a) Give two supervised learning methods (e.g., decision tree, logistic regression) and explain a strategy to avoid overfitting for each of them.

(b) What is regularization? Please also give an example to explain why regularization can avoid overfitting.

(c) What are the main advantages of deep learning over traditional machine learning?

(d) What is bagging? What is boosting? Explain the differences between bagging and boosting.

8. (8%) Assume you have the following three datasets. Each dataset contains two features (x1 and x2), and a class label (cross and star). Dataset 3 contains one data point (triangle) that belongs to both cross and star classes, and is the only one overlapping point of different classes in these three datasets.

(a) By using Logistic Regression without regularization, which datasets among these three can produce 100% training accuracy? Explain your solution.

(b) By using Decision Tree (e.g., ID3), which datasets among these three can produce 100% training accuracy? Explain your solution.