Question 1: Simple Optimizations [8]
6 Marks to find the $x$ values and 2 marks for finding the lowest $f'(x)$

Question 2: 3-NN Classification [20]
a) 5 marks for correctly classifying each point
b) 5 marks in total

Question 3: From Rules to Decision Trees [12]
7 points for correctly building the tree for $class_1$ and 5 points $class_2$
4 marks will be deducted if a correct but not the shortest tree is generated.

Question 4: Pruning a Decision Tree by Hand [20]
The marks are distributed in the question itself

Question 5: Programming Decision Tree [40]
The score of this question mainly consists of three part. For students who have submitted their own programs they will get $X = 5$ points. For students whose codes are able to compile, they can get another $Y = 5$ points. For the remaining $Z = 40 - X - Y = 30$ points, we will test their programs on our designed datasets and compare their outputs with the standard ones. We have prepared 5 test cases in total and for each test case there are 5 testing points corresponding to 5 different outputs. Both different test cases and the test points have their own weight.

The information of 5 test cases is listed below:

- Case 1: # of training = 30, # of tuning = # of testing = 7. This small trivial dataset sets all class labels to be 1, so the right decision tree will only contain the root node. Weight = 2
- Case 2: # of training = 150, # of tuning = # of testing = 33. This dataset set class labels solely rely on attribute A5, i.e. label = 0 when A5 = 0 or 2, and label = 1 when A5 = 1, 3, so the right decision will have only one level and four leaf nodes. Weight = 2
- Case 3: # of training = 300, # of tuning = # of testing = 66. This is the normal and primary test case, so it takes the highest weight. Weight = 3
- Case 4: same as case 3 except for randomly flipping with ratio of 5% the class label of samples to add up some noise. Weight = 1
• Case 5: same as case 3 except for adding two extra useless dimension whose values are all 1 and 2 respectively. Weight = 1

Here, case 1-2 are trivial cases to test the basic logic of the program. Case 3 is the primary test case for judging the correctness of building and pruning decision tree. As for case 4 and 5, they are used to further judging students’ program with some border conditions, but since they are relatively less important, so they take much smaller weights.

The information of 5 test points within each test case is shown as following:

• 0: print the information gain for each attribute at the root node. Weight 1.
• 1: create a decision tree from a training set, print the tree. Weight 3.
• 2: create a decision tree from a training set, print the classifications for instances in a test set. Weight 3.
• 3: create a decision tree from a training set then prune, print the tree. Weight 2.
• 4: create a decision tree from a training set then prune, print the classifications for instances in a test set. Weight 2.

Both test case weights and test points weights will be normalized before computing the score. Finally the score will be calculated based on the following formula:

\[
\text{score} = X \cdot I(\text{submitted}) + Y \cdot I(\text{compiled}) + Z \cdot \sum_{i=1}^{5} \sum_{j=1}^{5} w^{i}_{\text{case}} \cdot w^{j}_{\text{point}} \cdot I(\text{output}_{ij} = std_{ij})
\]