Haonan Hu

4/1/2021

2863545

1. Assume you have an Ensemble ML model of 6 base-learners that each outputs the following values for a test sample: 0.2, 0.3, 0.7, 0.2, 0.9, and 0.4.

> a) If you assign the following weights to the 6 base-learners, respectively, 0.12, 0.18, 0.21, 0.18, 0.17, and 0.14, what is the fused output of the ensemble using a weighted average?
>
> Fusion = $0.2 * 0.12 + 0.3 * 0.18 + 0.7 * 0.21 + 0.2 * 0.18 + 0.9 * 0.17 + 0.4 * 0.14 = 0.47$
>
> b) What is the fused output using a median?
>
> 0.2, 0.2, 0.3, 0.4, 0.7. 0.9
>
> Median = 0.3, 0.4
>
> Fusion = $(0.3 + 0.4) / 2 = 0.35$

2. Suppose you have 4 classes, A, B, C, and D. List the pairs of classes the base-learner, binary-classification models will be trained with, if you use the Method 2 of Pairwise Coupling.

Base learner: (d1, d2, d3, d4, d5, d6)

Binary-classification:

d1: A and B   d2: A and C   d3: A and D

d4: B and C   d5: B and D   d6: C and D

3. Using the Method 2 of Pairwise Coupling:

- Assume for a test sample, the outputs

  for the base-learners (d1, d2, and d3) are (assume the left class is assigned to 1 and the right class to 0):

- d1: A and B=0.6   p1(A) = 0.6 p1(B) = 0.4
- d2: A and C=0.7   p2(A) = 0.7 p2(C) = 0.3
- d3: B and C=0.5.   p3(B) = 0.5 p3(C) = 0.5
     a) What is p(A), p(B), and p(C)?

     P(A) = p1(A) * p2(A) = 0.6 * 0.7 = **0.42**

     P(B) = p1(B) * p3(B) = 0.4 * 0.5 = 0.2

     P(C)= p2(C) * p3(C)= 0.3 * 0.5 = 0.15

     b) Which class will the ensemble classify the test sample as?

     Class A, because P(A) is highest among 3

4. Assume you are using four Neural Networks as the base-learners for a Bagging Ensemble and that the training set (X) contains 10 samples labeled 1-10. If N is 5, create a subset of training samples for each the base-learners. Make sure at least one of the subsets demonstrates "with replacement".

Sample1: 1, **2, 3,** 4, 5

Sample2: **2, 3,** 4, 6, 9

Sample3: 1, 6, 7, 8, 9

Samlpe4: 2, 3, 5, 7, 8

5. Using the original Boosting algorithm, assume:

- X1 = {a, b, c, d, e}
- X2 = {f, g, h, i, j}
- X3 = {k, l, m, n, o}

- In Iteration1, that d1 misclassified{i, j} of X2

Give a data set for training d2 in Iteration 2.

d2 = {i, j, g, h}