Quiz 9 Solution

Mar 13 2023

1 Q1

Solution: B

- A: K(K-1)/2 models are needed.
- C: For a model trained to classify K_i and K_j, we only need the data where either K_i or K_j is 1.
- D: When K is large, 1-vs-rest is preferred.

2 Q2

Solution: C
Simplify the representation

$$P = \frac{\exp^{(\theta_k - \theta_1)x}}{1 + \exp^{(\theta_2 - \theta_1)x}} \tag{1}$$

So, as long as $\theta_2 - \theta_1$ is the same, the boundary will be the same.

3 Q3

Solution: True
Refer to lecture notes S-23

4 Q4

Solution: D

- A: Exponential loss is larger in the negative region. So the punishment is higher.
- B: In Adaboost, the weak learners are not necessarily the same.
- C: The correctly-classified examples are just assigned lower weights but not zero.

5 Q5

5.1
Solution: False

Bagging is less likely to have overfit because it captures more variance in the training data.

5.2
Solution: A

$$\begin{aligned} P(correct) =& P(all\, correct) + P(only\, A\, wrong) + P(only\, B\, wrong) + P(only\, C\, wrong) \\ =& 0.5 * 0.6 * 0.7 + (1 - 0.5) * 0.6 * 0.7 + 0.5 * (1 - 0.6) * 0.7 + 0.5 * 0.6 * (1 - 0.7) \\ =& 0.65 \end{aligned}$$

6 Q6

Solution: C

- A: Not true
- B: Not guaranteed. Initialization matters.
- D: Not convex

7 Q7

Solution: A

Apply the formula

$$\begin{aligned} w_k &= \frac{\sum_n \gamma_{nk}}{\sum_k \sum_n \gamma_{nk}} \\ w_1 &= \frac{1 + 0.2}{1 + 0.2 + 1 + 0.8} = 0.4 \\ w_2 &= \frac{1 + 0.8}{1 + 0.2 + 1 + 0.8} = 0.6 \end{aligned}$$