CSDS 440: Assignment 8

Shaochen (Henry) ZHONG, `sxz517`

Mingyang TIE, `mxt497`

Due and submitted on 10/30/2020
Fall 2020, Dr. Ray

Problem 33

The decision surface boundary is defined as $wx + b = 0$. When we modify two constants $c1$ and $c2$ that define the plus plane as $wx + b = c1$ where $c1 > 0$ and the minus plane as $wx + b = c2$ where $c2 < 0$. When $|c1| \neq |c2|$, the decision surface must be closer to either the plus or minus planes. If $|c1| > |c2|$, The decision surface boundary is closer to minus planes, because when $|c1| > |c2|$, $c1 + c2 > 0$, so $wx + b > 0$. If $|c1| < |c2|$, we can know $c1 + c2 < 0$, therefore the decision surface boundary is closer to plus planes, because $wx + b < 0$. If $|c1| = |c2|$, $wx + b = 0$, it means that the resulting decision surface is halfway between plus and minus plane, which is the general SVM case.

When we choose the decision surface boundary, it is really depend on the dataset we have. For example, if the distribution of the dataset we choose show that the negative class has less predictable data, we need to let the decision surface boundary to be closer the plus plane.

Problem 34

Assume the maximum margin classifer would be $wx + b \geq 1$ and $wx + b \leq -1$. Let two support vector $x1$ is positive, $x2$ is negative. So The maximum margin should be the projection of $x2 - x1$ onto the normal vector. Let the margin is d.

$$
\begin{aligned}
d &= \frac{w(x2-x1)}{||w||} \\
&= \frac{(1-b)-(-1-b)}{||w||} \\
&= \frac{2}{||w||}
\end{aligned}
$$

Thus according to the function, the margin of classification in an $SVM(w, b)$ is independent of b.

Problem 35

According to the definition we can get such equation:the 95intervals can be expressed:

$[(e_A - 1.96 * \sigma_A), (e_A + 1.96 * \sigma_A)]$

in the expression ,we can find these definitons :

e_A: means the average of all samples, which can be calculate by $\frac{n_1+n_2+n_3+...+n_k}{k}$, where n_k means a sample.

σ_A: means the standard devation of the samples, which can be calculate by $\sqrt{\frac{e_A(1-e_A)}{k}}$ where k is the the total number of samples, e_A is the average of all samples.

So according to the problem we should find enough big samples to sure the equation: $e_A - e_B = 0.1$ is proven ,we can get the inequality is true:

$(e_A - 1.96 * \sigma_A) > (e_B + 1.96 * \sigma_B)$

So we can change the inequality:

$$
\begin{aligned}
(e_A - 1.96 * \sigma_A) > (e_B + 1.96 * \sigma_B) &= 1.96(\sigma_B - \sigma_A) < e_A - e_B \\
&= 1.96(\sqrt{\frac{e_B(1-e_B)}{k}} + \sqrt{\frac{e_A(1-e_A)}{k}}) < e_A - e_B
\end{aligned}
$$

by the transformation we can get the relationship between (e_A, e_B) and k:

$k > (1.96(\sqrt{e_B(1-e_B)} + \sqrt{e_A(1-e_A)}))^2$

so according to the result ,we can think if the relationship between (e_A, e_B) and k is true, difference in error rates of A and B at the 95% confidence level can be proven.

Problem 36

According to the problem ,we can easily assume Professors Bobs experiment result is $[x1, x2]$ with C% confidence intervals after N times independent experiment. As the same, we can get Professors Nans experiment result is $[x3, x4]$ with C% confidence intervals after N times independent experiment. So based on the relationships . According to the definition of confidence intervals:
$[(e_A - k * \sigma_A), (e_A + k * \sigma_A)]$ where k is changed by the C% confidence.
So we can easily get the following two Professors total times with error experiment results:
Professors Bob: $E1 : N * 0.5(x1 + x2)$ and Professors Nan: $E2 : N * 0.5(x3 + x4)$
Based on the independent experiment ,we can add the different experiment results to replace the new experiment results, we can think the Professor Scoops experiment results are:
Total times experiment results: $2N$
Total times with error experiment results: $E1 + E2$
Based on this results we can calculate the results of Professor Scoop:
$e = (E1 + E2)/2N$ and $\sigma^2 = e(1 - e)/2N$ and $\sigma = \sqrt{e(1 - e)/2N}$
So with the C% confidence, the confidence interval can be expressed:
$[(e - k\sigma), (e + k\sigma)]$
Therefore $[(e - k\sigma), (e + k\sigma)]$ is the best confidence interval that Professor Scoop could report that would be consistent with Profs. Bob and Nans findings.