Set 2

Claire Goeckner-Wald

October 10, 2016

Hoeffding Inequality

1) B. 0.01

See the Hoeffding code. For 100,000 repititions, the average value of v_{min} was about 0.0376.

2) D. c_1 and c_{rand}

See the Hoeffding code. For 100,000 repititions, the average value of c_1 and c_{rand} were about 0.5. This satisifies the single-bin Hoeffding Inequality because it

Error and Noise

3) E. $\lambda\mu + (1-\lambda)(1-\mu)$

Function h approximates f with error μ. But there is a $1-\lambda$ chance that f is wrong. We can determine the probability that h incorrectly approximates y with a decision tree.

4)

$$
\begin{aligned}
P(h(x) \neq y) &= \lambda\mu + (1-\lambda)(1-\mu) \\
&= \lambda\mu + (1 - \lambda - \mu + \lambda\mu) \\
&= 2\lambda\mu - \lambda - \mu + 1 \\
&= \mu(2\lambda - 1) - \lambda + 1 \\
2\lambda - 1 &= 0 \\
2\lambda &= 1 \\
\lambda &= \frac{1}{2}
\end{aligned}
$$

Thus, when $\lambda = \frac{1}{2}$, μ is irrelevant to the output of $P(h(x) \neq y)$.

Linear Regression

5) C. 0.01

See the linear regression code. For 10,000 repititions, the average in-sample error was 0.039185.

6) C. 0.01

See the linear regression code. For 10,000 repititions, the average out-sample error was 0.0484795.

7) A. 1

See the linear regression code. For 1,000 repititions, the average numbers of modified PLA iterations required is 5.334

Nonlinear Transformation

8) D. 0.5

See the non-linear transformation code. For 1,000 repititions, the average in-sample error for the non-transformed data was 0.51429.

9) A. sign$(-1 - 0.05 * x_1 + 0.08 * x_2 + 0.13 * x_1 * x_2 + 1.5 * x_1 ** 2 + 1.5 * x_2 ** 2)$

See the non-linear transformation code. For 1000 test data points, the functions a - e had accuracy rates of 0.952, 0.693, 0.677, 0.614, and 0.547, respectively. Thus, [a] most resembles the hypothesis found.

10) A. 0

See the non-linear transformation code. For 1,000 repititions, the average out-sample error for the transformed data was 0.033848.