FBA | QUANTITATIVE FINANCE RESEARCH GROUP

Financial Machine Learning

Homework 3

Due at 07:00 pm (Korea Standard Time) on Sunday, August 21.

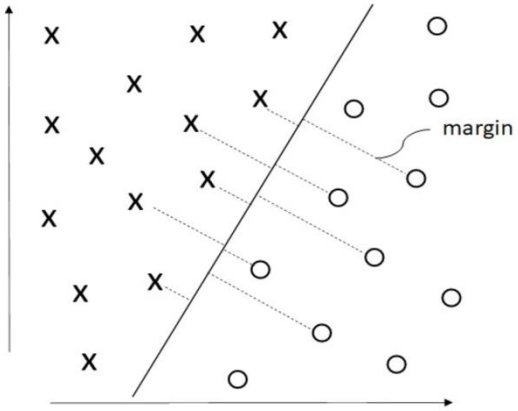
Submit one file: written solutions with executable Python code

Problem 1. Text book: Hands-on Machine Learning. Submit .ipynb file.

(a) Practice all the codes in the Text book Chapter 5. And show that they work well.

Problem 2. Maximum Margin Classification. Capture your answer and put in your .ipynb file.

Assume that we have $n = 1, \dots, N$ input data. Let` say $\boldsymbol{x_n}$ (2-dimensional vectors) is the n-th input data, and $t_n \in \{1, -1\}$ is the n-th labeled value $n = 1, \dots, N$. Also, we denote $\boldsymbol{w} = (w_1, w_2, \dots, w_N)\ as\ weight$ and b_0 (scalar value) as a bias.



(a) Derive equation of Decision Boundary

(b) Derive equation of distance between Boundary and data set (margin)

Assume $\kappa = \min_n (\text{margin of n} - \text{th})$, which is scalar value.

(c) Derive κ (Note that we have two labels, and each label is 1 or -1)

(d) For better classification, should we maximize κ or minimize κ?. Then why?

(e) Find the conditions of parameters W, b to optimize κ using Lagrange Multipliers $a = (a_1, a_2, \dots, a_N)$; Hint: $L(\boldsymbol{w}, b, \boldsymbol{a}) = \frac{1}{2}\boldsymbol{w^T w} - \sum_{n=1}^{N} a_n(t_n(\boldsymbol{w^T x_n} + b) - 1)$

Problem 3. Dual Representation

(a) Find $\tilde{L}(a)\ with\ only$ Lagrange multipliers a $= (a_1, a_2, \dots, a_N)$ by substituting and erasing $\boldsymbol{w}$

(b) Derive equation of distance between Boundary and data set with only Lagrange multipliers a $= (a_1, a_2, \dots, a_N)$ by substituting and erasing $\boldsymbol{w}$