HW4: Support Vector Machine
Advanced Machine Learning

Francisco Calderon

2/22/2017

1.0 Toy SVM

Marginals

$x^2 = 1$

$x^2 = 3$

Decision Boundary

$x^2 = 2$

a) Points with Lagrangian coefficient = 0

$\alpha_3 = 0$ $\alpha_6 = 0$ $\alpha_8 = 0$

$\alpha_4 = 0$ $\alpha_7 = 0$

$$w_1 = \alpha_1(1)(3) + \alpha_2(3)(1) + \alpha_5(1)(-1)$$
$$w_2 = \alpha_1(1)(1) + \alpha_2(-1)(1) + \alpha_5(-1)(0)$$

$$1 = 3\alpha_1 + 3\alpha_2 - \alpha_5)$$
$$+\ 0 = \alpha_1 + \alpha_2 - \alpha_5)3$$
$$\frac{1}{2} = \frac{2\alpha_5}{2}$$
$$\alpha_5 = \frac{1}{2}$$

$$0 = \alpha_1 - \alpha_2$$
$$+\ \frac{1}{2} = \alpha_1 + \alpha_2$$
$$\frac{1}{2} = \frac{2\alpha_1}{2}$$

$$\frac{1}{4} = \alpha_2 \qquad \frac{1}{4} = \alpha_1$$

b) Find w and w_0

$(3,1)$ x_1: $w_0 + 3w_1 + w_2 = 1$ $\quad -2+3$

$(3,-1)$ x_2: $w_0 + 3w_1 - w_2 = 1$ $\quad -2+3$

$(1,0)$ x_5: $w_0 + w_1 = -1$

$w_2 = 0$

$$w_0 + 3w_1 = 1$$
$$-(w_0 + w_1 = -1)$$
$$\frac{2w_1}{2} = \frac{2}{2}$$
$$w_1 = 1$$

$$w_0 + 2 = -1$$
$$-2 \quad -2$$
$$w_0 = -2$$

$w = (1, 0)$

$w_0 = -2$

c) The support vectors are

$x_1 = (3,1)$

$x_2 = (3,-1)$

$x_5 = (1,0)$

d) $$\text{Sign}(w_0 + WX) = \text{Sign}\left(w_0 + \sum_{i=1}^{N} \alpha_i y_i x_i \cdot x\right)$$
$$= \text{Sign}\left(w_0 + \sum_{i=1}^{N} \alpha_i y_i k_i\right)$$

$$\text{sign}\left(-2 + \left(\frac{1}{4} x_1^T x + \frac{1}{4} x_2^T x - \frac{1}{2} x_5^T x\right)\right)$$

c) $$\text{sign}\left(-2 + \left(\frac{1}{4}[3,1]\begin{bmatrix}0.5\\0.5\end{bmatrix} + \frac{1}{4}[3,-1]\begin{bmatrix}0.5\\0.5\end{bmatrix} - \frac{1}{2}[1,0]\begin{bmatrix}0.5\\0.5\end{bmatrix}\right)\right)$$
$$\text{sign}\left(-2 + \left(\frac{1}{2} + \cancel{\frac{1}{4} + -\frac{1}{4}}\right)\right)$$
$$\text{sign}(-1.5) = -1$$

2.0 Kernel SVM

Estimator	Accuracy	Best Parameters
SVM-Gaussian	83.5%	{C: 6.1, gamma: 0.1}
SVM-Quadratic	84.5%	{C: 1.1, gamma: 0.1}
KNN	82.%	{n_neighbors:10}

Table 1: Interpolating the Optimal Learning Rate

Table 1 shows the results of three models that were trained using `RandomizedSearchCV`. As one can see, the Quadratic Support Vector Machine performs the best using a C parameter of 6.1 and gamma of 0.1. The Support Vector Machine with a Gaussian kernel performed admirably with an accuracy of 83.5%. Lastly, although KNN had the lowest accuracy, it was faster to train and predict with KNN than either of the support vector machines. This may be a more important metric for a practical application.