0 Instructions

Homework is due Tuesday, February 20, 2024 at 23:59pm Central Time. Please refer to `https://courses.grainger.illinois.edu/cs446/sp2024/homework/hw/index.html` for course policy on homeworks and submission instructions.

1 Soft-margin SVM: 4pts

The Lagrangian form of the soft-margin SVM is given by

$$L(\omega, b, \xi, \alpha, \beta) = \frac{1}{2}||\omega||^2 + C\sum_{i=1}^{n}\xi_i - \sum_{i=1}^{n}\alpha_i[y_i(\omega^T x_i + b) - 1 + \xi_i] - \sum_{i=1}^{n}\beta_i\xi_i$$

The dual form of the problem is then given by

$$D(\alpha, \beta) = \min_{\omega, b, \xi} L(\omega, b, \xi, \alpha, \beta)$$

Because the problem is convex, we know that the maximum of the dual is the minimum of the primal. The solution to the dual occurs when the gradients of the Lagrangian are 0, i.e.

$$\begin{aligned}
\nabla_\omega L(\omega, b, \xi, \alpha, \beta) &= \omega - \sum_{i=1}^{n}\alpha_i y_i x_i = 0 \\
\nabla_b L(\omega, b, \xi, \alpha, \beta) &= -\sum_{i=1}^{n}\alpha_i y_i = 0 \\
\nabla_\xi L(\omega, b, \xi, \alpha, \beta) &= C - \alpha_i - \beta_i = 0
\end{aligned}$$

Substitue these back into the Lagrangian, we have

$$\begin{aligned}
D(\alpha, \beta) &= \frac{1}{2}||\sum_{i=1}^{n}\alpha_i y_i x_i||^2 + C\sum_{i=1}^{n}\xi_i - \sum_{i=1}^{n}\alpha_i[y_i(\sum_{j=1}^{n}\alpha_j y_j x_j^T x_i + b) - 1 + \xi_i] - \sum_{i=1}^{n}\beta_i\xi_i \\
&= \sum_{i=1}^{n}\alpha_i - \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}\alpha_i\alpha_j y_i y_j x_i^T x_j
\end{aligned}$$

subject to

$$
\begin{aligned}
\alpha_i &\geq 0 \\
\beta_i &\geq 0 \\
\alpha_i + \beta_i &= C \\
\sum_{i=1}^{n} \alpha_i y_i &= 0
\end{aligned}
$$

2 SVM, RBF Kernel and Nearest Neighbor: 6pts

2.1

$$\begin{aligned}\hat{\omega} &= \sum_{i=1}^{N} \hat{\alpha}_i y_i x_i \\ f(x) &= (\sum_{i=1}^{N} \hat{\alpha}_i y_i x_i)^T x\end{aligned}$$

2.2

$$\begin{aligned}\hat{\omega} &= \sum_{i=1}^{N} \hat{\alpha}_i y_i \phi(x_i) \\ f(x) &= (\sum_{i=1}^{N} \hat{\alpha}_i y_i \phi(x_i))^T \phi(x) \\ &= \sum_{i=1}^{N} \hat{\alpha}_i y_i K(x_i, x)\end{aligned}$$

2.3

$$\begin{aligned}\lim_{\delta\to 0} \frac{\sum_{i=1}^{N} \hat{\alpha} y_i e^{-\frac{||x_i - x||^2}{2\delta^2}}}{e^{-\frac{\rho^2}{2\delta^2}}} &= \lim_{\delta\to 0} \sum_{i=1}^{S} \hat{\alpha} y_i e^{-\frac{||x_i - x||^2}{2\delta^2}} \\ &= \lim_{\delta\to 0} \sum_{i=1}^{T} \hat{\alpha} y_i e^{-\frac{||x_i - x||^2 - \phi^2}{2\delta^2}} + \lim_{\delta\to 0} \sum_{i=1}^{S/T} \hat{\alpha} y_i e^{-\frac{||x_i - x||^2 - \phi^2}{2\delta^2}} \\ &= \sum_{i=1}^{T} \hat{\alpha} y_i + 0 \\ &= \sum_{i=1}^{T} \hat{\alpha} y_i\end{aligned}$$

3 Decision Tree and Adaboost: 12 pts

4 Learning Theory: 14pts

5 Coding: SVM, 4pts