

Classifier: SVM vs Logistic Regression

Learning Type: **Supervised**, Task: **Classification**, Algorithm: **SVM and Logistic Regression**

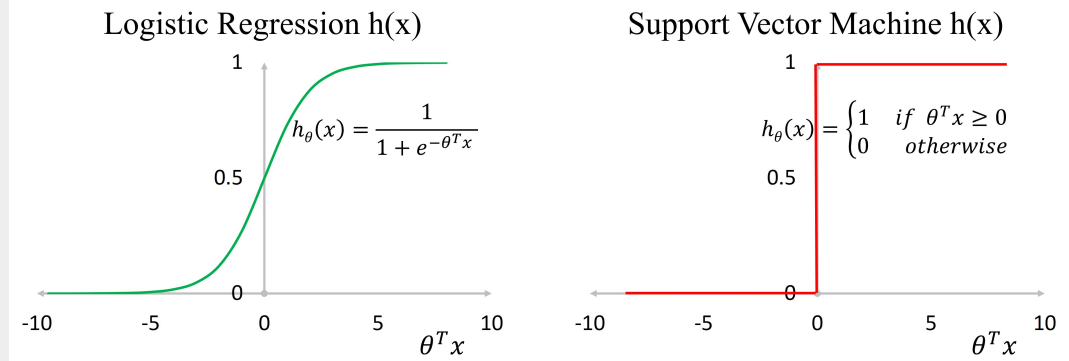
Model $h(x)$:

Logistic Regression (sigmoid):

$$h_{\theta}(x)_{LR} = \frac{1}{1+e^{-\theta^T x}} = \frac{e^{\theta^T x}}{e^{\theta^T x} + 1}$$

Support Vector Machine:

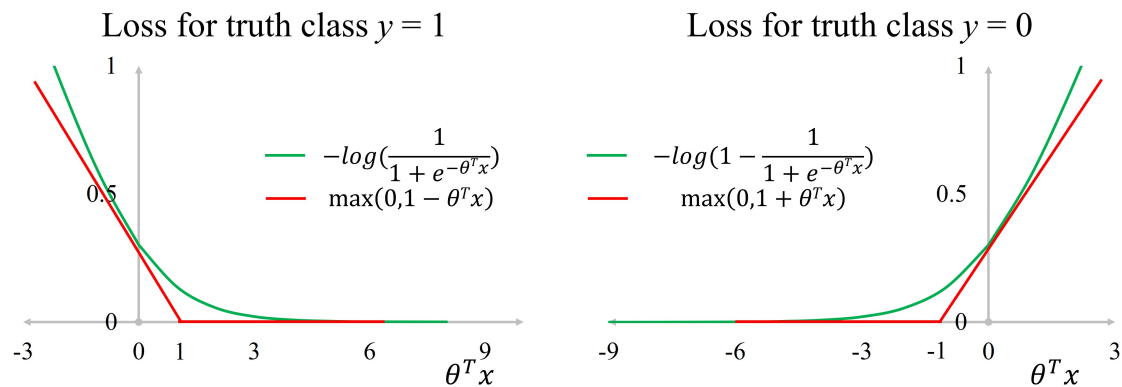
$$h_{\theta}(x)_{SVM} = \begin{cases} 1 & \text{if } \theta^T x \geq 0 \\ 0 & \text{otherwise} \end{cases}$$



Evaluation Criteria: Cost Function

Cross Entropy or Log Loss for LR: $J(\theta) = -\frac{1}{m} [\sum_{i=1}^m y^{(i)} \log h_{\theta}(x^{(i)}) + (1 - y^{(i)}) \log (1 - h_{\theta}(x^{(i)}))] + \frac{\lambda}{2m} \sum_{j=1}^n \theta_j^2$

Hinge Loss for SVM: $J(\theta) = C[\sum_{i=1}^m y^{(i)} \max(0, 1 - \theta^T x) + (1 - y^{(i)}) \max(0, 1 + \theta^T x)] + \frac{1}{2} \sum_{j=1}^n \theta_j^2$
 where m = no. of examples; n = no. of features; X = dataset and θ = model weights



More Details about SVM and Logistic Regression

	Logistic Regression	Support Vector Machine
Regularization	$\lambda \uparrow$ Overfitting \downarrow	$C \uparrow$ Overfitting \uparrow $C = \frac{1}{\lambda}$
MultiClass	One-vs-all	One-vs-all
NonLinear Dataset	Nonlinear transformation of $\theta^T x$	Kernel Function to convert X to higher dimension and use linear hyperplane
Class Margin	Effect with λ	$C \uparrow$ Loss \uparrow Margin \downarrow Hard SVM $C \downarrow$ Loss \downarrow Margin \uparrow Soft SVM

