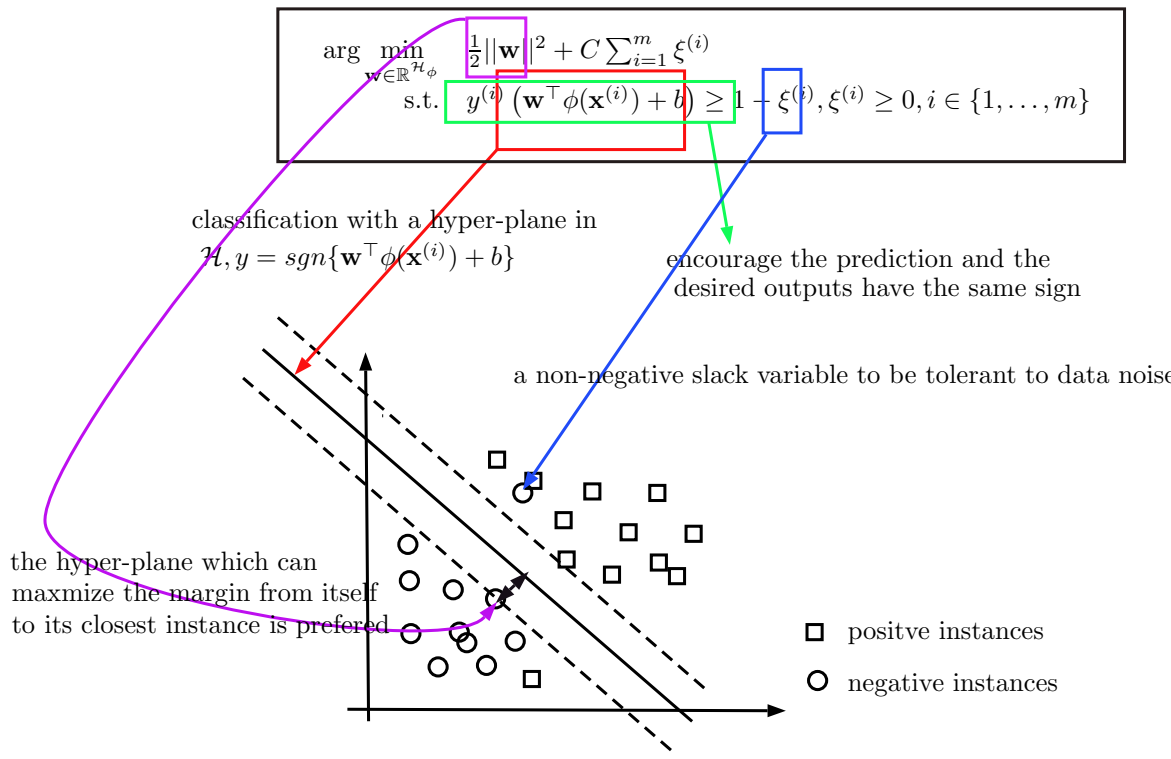
$\arg \min_{\mathbf{w} \in \mathbb{R}^{\mathcal{H}_\phi}} \frac{1}{2}||\mathbf{w}||^2 + C\sum_{i=1}^{m} \xi^{(i)}$
s.t. $y^{(i)}\left(\mathbf{w}^\top \phi(\mathbf{x}^{(i)}) + b\right) \geq 1 - \xi^{(i)}, \xi^{(i)} \geq 0, i \in \{1, \ldots, m\}$
classification with a hyper-plane in $\mathcal{H}, y = sgn\{\mathbf{w}^\top \phi(\mathbf{x}^{(i)}) + b\}$
encourage the prediction and the desired outputs have the same sign
a non-negative slack variable to be tolerant to data noise
the hyper-plane which can maxmize the margin from itself to its closest instance is prefered
positve instances
negative instances