Model	Predicted value for $(\overline{\mathbf{W}}, \overline{\mathbf{X_i}})$
[1] Perceptron (Smoothed surrogate)	$\hat{y}_i = sign(\overline{W} \cdot \overline{X_i})$
[2] Widrow-Hoff / Fisher (Least-Squares)	$\hat{y}_i = \overline{W} \cdot \overline{X_i}$
[3] Logistic Regression	$\hat{y}_i = 1/(1 + \exp[-\overline{W} \cdot \overline{X_i}])$
[4] Support Vector Machine (Hinge)	$\hat{y}_i = \overline{W} \cdot \overline{X_i}$
[5] Support Vector Machine (Hinton's L_2-Loss)	

Model	Loss function for $(\overline{\mathbf{X_i}}, \mathbf{y_i})$	Gradient of the Loss Function
[1]	$L_i = \max\{0, -y_i \cdot (\overline{W} \cdot \overline{X_i})\}$	$\frac{\partial L_i}{\partial \overline{W}} = -y_i \overline{X_i}[I(y_i \hat{y}_i < 0)]$
[2]	$L_i = (y_i - \overline{W} \cdot \overline{X_i})^2 = \{1 - y_i \cdot (\overline{W} \cdot \overline{X_i})\}^2$	$\frac{\partial L_i}{\partial \overline{W}} = -2(y_i - \overline{W} \cdot \overline{X_i}) \cdot \overline{X_i}$
[3]	$L_i = \log(1 + \exp[-y_i \cdot (\overline{W} \cdot \overline{X_i})])$	$\frac{\partial L_i}{\partial \overline{W}} = -\frac{y_i \overline{X_i}}{1 + e^{y_i \overline{W} \cdot \overline{X_i}}}$
[4]	$L_i = \max\{0, 1 - y_i \cdot (\overline{W} \cdot \overline{X_i})\}$	$\frac{\partial L_i}{\partial \overline{W}} = -y_i \overline{X_i}[I(y_i \hat{y}_i < 1)]$
[5]	$L_i = [\max\{0, 1 - y_i \cdot (\overline{W} \cdot \overline{X_i})\}]^2$	

- **Update of the Weights** for α (learning rate) and λ (regularization parameter):

$$\overline{W} \Leftarrow \overline{W}(1 - \alpha\lambda) - \alpha \frac{\partial L_i}{\partial \overline{W}}$$