NPTEL ONLINE CERTIFICATION COURSES

Course Name: Deep Learning

Faculty Name: Prof. P. K. Biswas

Department : E & ECE, IIT Kharagpur

Topic

Lecture 14: Linear Machine and Multiclass SVM

CONCEPTS COVERED

Concepts Covered:

- Linear Classifier
- Support Vector Machine
- Linear Machine
- Multiclass Support Vector Machine

Multiclass Problem: Linear Machine

Multiclass Problem: Linear Machine

$$f : R^D \rightarrow R^K$$

$$f(X_i, W, b) = WX_i + b = s$$

$$\begin{bmatrix} W_{11} & W_{12} & W_{13} & .. & W_{1D} \\ W_{21} & W_{22} & W_{23} & .. & W_{2D} \\ .. & .. & .. & .. & .. \\ W_{K1} & W_{K2} & W_{K3} & .. & W_{KD} \end{bmatrix} \begin{bmatrix} X_1 \\ X_2 \\ X_3 \\ .. \\ X_D \end{bmatrix} + \begin{bmatrix} b_1 \\ b_2 \\ .. \\ b_K \end{bmatrix} = \begin{bmatrix} s_1 \\ s_2 \\ .. \\ s_K \end{bmatrix}$$

Multiclass Problem: Linear Machine

$$\begin{bmatrix} 0.2 & 0.6 & -1.0 & 0.8 \\ 1.5 & 0.9 & 3.1 & 0.1 \\ 0.5 & 1.1 & 0.7 & 0.0 \\ 2.1 & 0.3 & 0.2 & 0.5 \end{bmatrix} \begin{bmatrix} 45 \\ 110 \\ 21 \\ 16 \end{bmatrix} + \begin{bmatrix} 1.1 \\ 5.3 \\ -2.1 \\ 0.6 \end{bmatrix} \Rightarrow \begin{array}{cl} \begin{bmatrix} 67.9 \\ 238.5 \\ 156.1 \\ 140.3 \end{bmatrix} & \begin{array}{l} \textit{Cat score} \\ \textit{Bird score} \\ \textit{Dog score} \\ \textit{Car score} \end{array} \end{array}$$

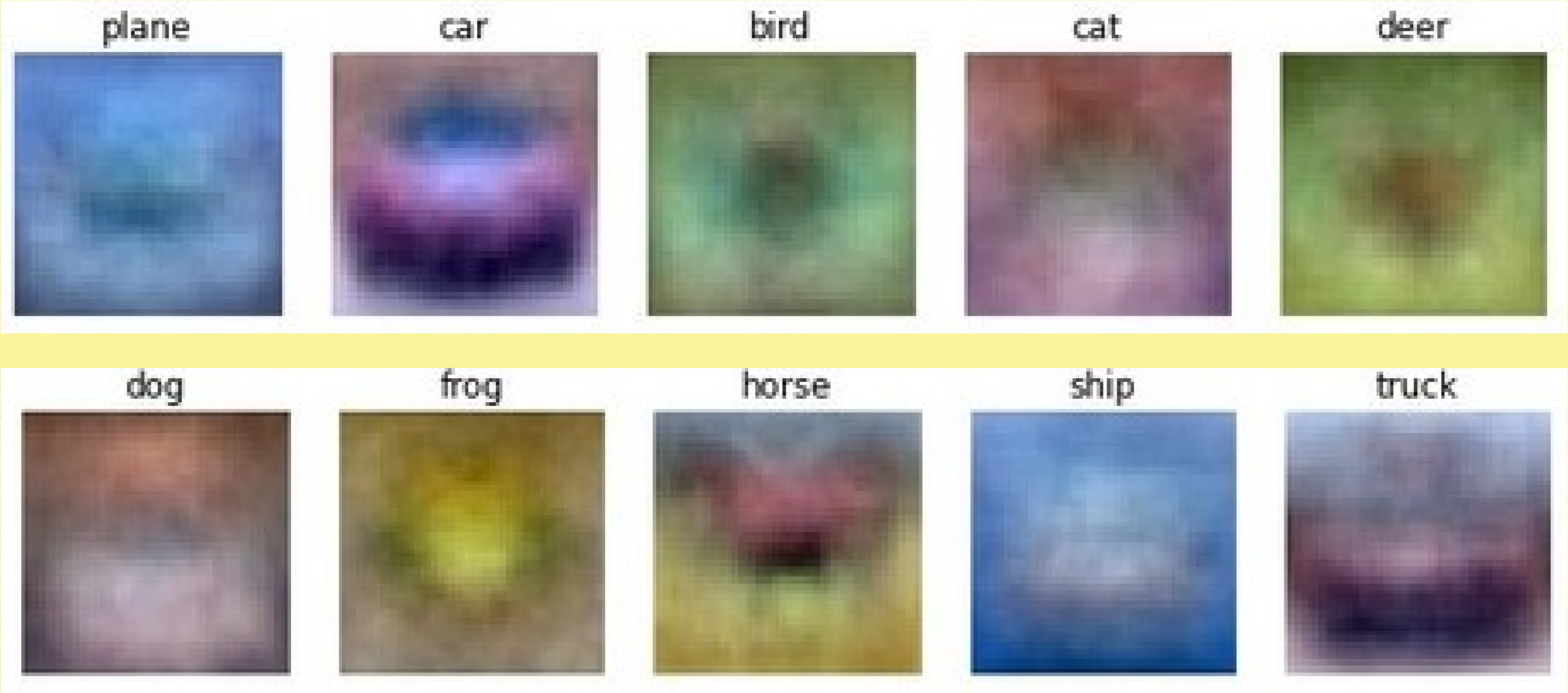
$$f(X_i, W, b)$$

Interpretation

$$\begin{bmatrix} 0.2 & 0.6 & -1.0 & 0.8 \\ 1.5 & 0.9 & 3.1 & 0.1 \\ 0.5 & 1.1 & 0.7 & 0.0 \\ 2.1 & 0.3 & 0.2 & 0.5 \end{bmatrix} \begin{bmatrix} 45 \\ 110 \\ 21 \\ 16 \end{bmatrix} + \begin{bmatrix} 1.1 \\ 5.3 \\ -2.1 \\ 0.6 \end{bmatrix} \Rightarrow \begin{array}{cc} \begin{bmatrix} 67.9 \\ 238.5 \\ 156.1 \\ 140.3 \end{bmatrix} & \begin{array}{c} \textit{Cat score} \\ \textit{Bird score} \\ \textit{Dog score} \\ \textit{Car score} \end{array} \end{array}$$

$$f(X_i, W, b)$$

Multiclass Problem: Linear Machine



Source - http://cs231n.github.io

Bias Trick

$$\begin{bmatrix} 0.2 & 0.6 & -1.0 & 0.8 & 1.1 \\ 1.5 & 0.9 & 3.1 & 0.1 & 5.3 \\ 0.5 & 1.1 & 0.7 & 0.0 & -2.1 \\ 2.1 & 0.3 & 0.2 & 0.5 & 0.1 \end{bmatrix} \begin{bmatrix} 45 \\ 110 \\ 21 \\ 16 \\ 1 \end{bmatrix} \Rightarrow \begin{array}{c} \begin{bmatrix} 67.9 \\ 238.5 \\ 156.1 \\ 140.3 \end{bmatrix} \\ f(X_i, W) \end{array} \begin{array}{l} \textit{Cat score} \\ \textit{Bird score} \\ \textit{Dog score} \\ \textit{Car score} \end{array}$$

Multiclass SVM

$$\left.\begin{aligned} s_j &= f(X_i, W)_j \\ &= WX_i \end{aligned}\right\} \rightarrow Score\ \ for\ \ j^{th}\ Class\ of\ \ i^{th}\ Vector\ (X_i, y_i)$$

$$s_{y_i} = f(X_i, W)_{y_i} \quad \rightarrow should\ be\ \max imum$$

$$s_{y_i} - s_j \geq \Delta$$

$$L_i = \sum_{j \neq y_i} \max(0, s_j - s_{y_i} + \Delta)$$

Loss Function: An Example

$$For\ some\ (X_i, y_i)\ where\ y_i = 2$$

$$s = (10\ \ 30\ \ -20\ \ 25)^t \qquad \Delta = 10$$

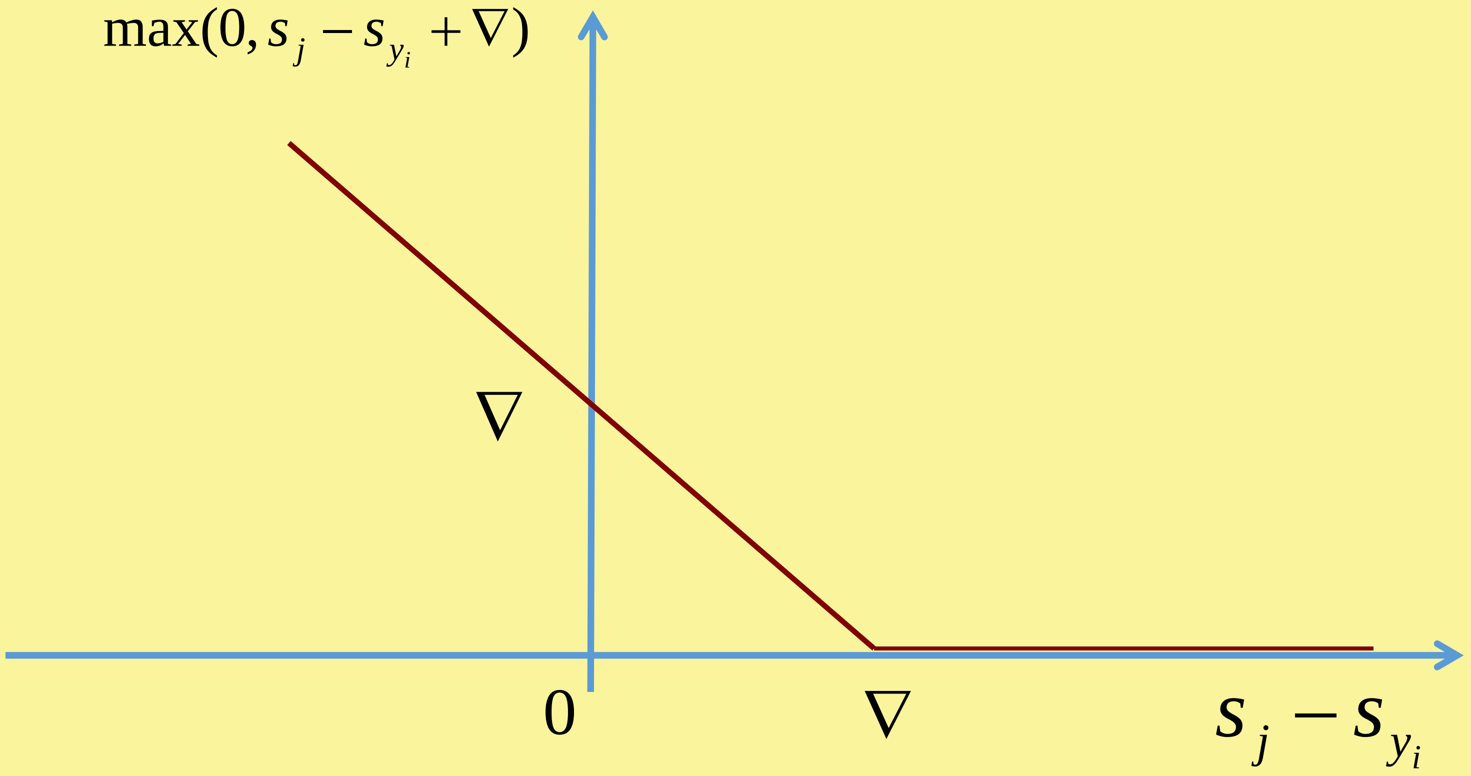
$$L_i = \sum_{j \neq y_i} \max(0, s_j - s_{y_i} + \Delta)$$

$$= \max(0,10-30+10) + \max(0,-20-30+10) + \max(0,25-30+10)$$

$$= 0 + 0 + 15$$

$$= 15$$

Hinge Loss



Regularization

$$s_j - s_{y_i} = W_j^t X_i - W_{y_i}^t X_i$$

$$\text{Scaling } W \text{ by } \lambda : W \leftarrow \lambda W$$

$$\Downarrow$$

$$s_j - s_{y_i} \leftarrow \lambda\left(s_j - s_{y_i}\right)$$

Regularization

Include a regularization term $R(W)$

$$R(W) = \lambda \sum_k \sum_l W_{kl}^2$$

$$L = \frac{1}{N} \sum_i L_i + \lambda R(W)$$

$$L = \frac{1}{N} \sum_i \sum_{j \neq y_i} [\max(0, f(X_i, W)_j - f(X_i, W)_{y_i} \ + \nabla) + \lambda \sum_k \sum_l W_{kl}^2$$

NPTEL ONLINE CERTIFICATION COURSES

Thank you