Formula Sheet, CS 440/ECE 448 Exam 1

Probability

$$P(X=x)=\Pr(X=x)\quad \ldots \quad \text{or} \quad \ldots \quad P(X=x)=\frac{d}{dx}\Pr(X\le x)$$

$$P(X,Y)=P(X|Y)P(Y)$$

$$E[f(X,Y)]=\sum_{x,y} f(x,y)P(X=x,Y=y)$$

Decision Theory

$$f(x)=\underset{y}{\operatorname{argmax}}\, P(Y=y|X=x),\quad \text{Bayes Error Rate}=\sum_x P(X=x)\min_y P(Y\ne y|X=x)$$

$$\text{Precision}=P(Y=1|f(X)=1)=\frac{TP}{TP+FP}$$

$$\text{Recall}=\text{Sensitivity}=P(f(X)=1|Y=1)=\frac{TP}{TP+FN}$$

$$\text{Selectivity}=P(f(X)=0|Y=0)=\frac{TN}{TN+FP}$$

Naïve Bayes

$$f(x)\approx \operatorname{argmax}_y\left(\log P(Y=y)+\sum_{i=1}^n \log P(W=w_i|Y=y)\right)$$

$$P(W=w_i|Y=y)=\frac{k+\text{Count}(w_i,y)}{k+\sum_{v\in V}(k+\text{Count}(v,y))}$$

Hidden Markov Model

$$v_1(j)=\pi(j)b_j(\boldsymbol{x}_t)$$

$$v_t(j)=\max_i v_{t-1}(i)\,a_{i,j}b_j(\boldsymbol{x}_t),\qquad \psi_t(j)=\underset{i}{\operatorname{argmax}}\, v_{t-1}(i)\,a_{i,j}b_j(\boldsymbol{x}_t)$$

$$y^*(T)=\underset{i}{\operatorname{argmax}}\, v_T(i)\,,\qquad y^*(t)=\psi_{t+1}\big(y^*(t+1)\big)$$

Fairness

- Demographic Parity: $P(f(X)|A=1)=P(f(X)|A=0)$
- Equal Odds: $P(f(X)|Y,A=1)=P(f(X)|Y,A=0)$
- Predictive Parity: $P(Y|f(X),A=1)=P(Y|f(X),A=0)$

Learning

$$\mathcal{R}=\mathrm{E}[\ell(Y,f(X))]$$

$$\mathcal{R}_{\text{emp}}=\frac{1}{n}\sum_{i=1}^n \ell(y_i,f(x_i))$$

Linear Regression

$$f(\boldsymbol{x})=\boldsymbol{w}^T\boldsymbol{x}+b$$

$$\mathcal{L}=\frac{1}{n}\sum_{i=1}^n \mathcal{L}_i\,,\quad \mathcal{L}_i=\frac{1}{2}\epsilon_i^2,\quad \epsilon_i=f(\boldsymbol{x}_i)-y_i$$

$$\boldsymbol{w}\leftarrow \boldsymbol{w}-\eta\frac{\partial\mathcal{L}}{\partial\boldsymbol{w}},\quad \frac{\partial\mathcal{L}}{\partial\boldsymbol{w}}=\frac{1}{n}\sum_{i=1}^n \epsilon_i\boldsymbol{x}_i$$

Perceptron

$$f(\boldsymbol{x}) = \operatorname{argmax} \boldsymbol{W}\boldsymbol{x} + \boldsymbol{b}$$

$$\boldsymbol{w}_c \leftarrow \begin{cases} \boldsymbol{w}_c - \eta\boldsymbol{x} & c = \operatorname{argmax} \boldsymbol{W}\boldsymbol{x} + \boldsymbol{b} \\ \boldsymbol{w}_c + \eta\boldsymbol{x} & c = y \\ \boldsymbol{w}_c & \text{otherwise} \end{cases}$$

Softmax & Sigmoid

$$f_c(\boldsymbol{x}) = \frac{\exp(\boldsymbol{w}_c^T\boldsymbol{x} + b_c)}{\sum_{k=1}^{v}\exp(\boldsymbol{w}_k^T\boldsymbol{x} + b_k)} \approx \Pr(Y = c|\boldsymbol{x})$$

$$\sigma(\boldsymbol{w}^T\boldsymbol{x} + b) = \frac{1}{1 + e^{-(\boldsymbol{w}^T\boldsymbol{x}+b)}} \approx \Pr(Y = 1|\boldsymbol{x})$$

$$\mathcal{L} = -\ln f_y(\boldsymbol{x}), \quad \frac{\partial\mathcal{L}}{\partial f_c(\boldsymbol{x})} = \begin{cases} -\frac{1}{f_c(\boldsymbol{x})} & c = y \\ 0 & \text{otherwise} \end{cases}$$

$$\boldsymbol{w}_c \leftarrow \boldsymbol{w}_c - \eta\frac{\partial\mathcal{L}}{\partial\boldsymbol{w}_c} = \boldsymbol{w}_c - \eta\epsilon_c\boldsymbol{x}, \quad \epsilon_c = \begin{cases} f_c(x_i) - 1 & c = y \\ f_c(x_i) - 0 & \text{otherwise} \end{cases}$$

Multi-Layer

$$f_k = \operatorname*{softmax}_k \boldsymbol{z}^{(2)}, \qquad z_k^{(2)} = b_k^{(2)} + \sum_{j=1}^{n} w_{k,j}^{(2)} h_j, \qquad h_j = \operatorname{ReLU}\left(b_j^{(1)} + \sum_{i=1}^{d} w_{j,i}^{(1)} x_i\right)$$

$$\frac{\partial\mathcal{L}}{\partial w_{j,i}^{(1)}} = \sum_{k=1}^{v}\left(\frac{\partial\mathcal{L}}{\partial z_k^{(2)}}\right)\left(\frac{\partial z_k^{(2)}}{\partial h_j}\right)\left(\frac{\partial h_j}{\partial w_{j,i}^{(1)}}\right) = \sum_{k=1}^{v}\left(f_k - 1_{y=k}\right)w_{k,j}^{(2)}1_{h_j>0}x_i$$

Image Formation & Processing

$$\frac{x'}{f} = -\frac{x}{z}, \quad \frac{y'}{f} = -\frac{y}{z}$$

$$h_x'(x', y') = \frac{(h(x'+1, y') - h(x'-1, y'))}{2}, \qquad h_y'(x', y') = \frac{(h(x', y'+1) - h(x', y'-1))}{2}$$

ConvNets

$$y[k, l] = \sum_i\sum_j x[k-i, l-j]h[i, j], \qquad \frac{d\mathcal{L}}{dh[i, j]} = \sum_k\sum_l \frac{d\mathcal{L}}{dy[k, l]}\frac{dy[k, l]}{dh[i, j]}$$

$$z[m, n] = \max_{\substack{(m-1)p+1 \le k \le mp, \\ (n-1)p+1 \le l \le np}} y[k, l], \quad \frac{d\mathcal{L}}{dy[k, l]} = \begin{cases} \frac{d\mathcal{L}}{dz[m, n]} & \text{if } y[k, l] = \max_{\substack{(m-1)p+1 \le i \le mp, \\ (n-1)p+1 \le j \le np}} y[i, j] \\ 0 & \text{otherwise} \end{cases}$$