# Exam Cheat Sheet

## Discriminative Functions for Classifiers

### Naive Bayes

$$P(y|x) = \frac{P(x|y)P(y)}{P(x)} \tag{1}$$

$$P(x|y) = \prod_{i=1}^{n} P(x_i|y) \tag{2}$$

### Fisher's Linear Discriminant Analysis

$$w = S_W^{-1}(\mu_1 - \mu_2) \tag{3}$$

$$S_W = \frac{1}{N_1}\sum_{n\in C_1}(x_n-\mu_1)(x_n-\mu_1)^T + \frac{1}{N_2}\sum_{n\in C_2}(x_n-\mu_2)(x_n-\mu_2)^T \tag{4}$$

$$S_B = (\mu_1-\mu_2)(\mu_1-\mu_2)^T \tag{5}$$

### Logistic Regression

$$\phi(x) = \frac{1}{1+e^{-x}} \tag{6}$$

$$y(x) = \phi(w^T x) \tag{7}$$

$$J(\theta) = -\frac{1}{m}[y^T \log(\sigma(X\theta)) + (1-y)^T \log(1-\sigma(X\theta))] \tag{8}$$

### Perceptron Algorithm

$$\mathcal{E}_p(\mathbf{w}, w_0) = -\sum_{n\in M} t_n(\mathbf{w}\cdot\mathbf{x}_n + w_0), \text{ where } M \text{ is misclass} \tag{9}$$

### Gradient Descent

$$\theta := \theta - \alpha\nabla_\theta J(\theta) \tag{10}$$

## Support Vector Machines

### Kernel Machines

$$K(x,x') = \exp(-\frac{\|x-x'\|^2}{2\sigma^2}) = \exp(-\gamma\|x-x'\|^2) \tag{11}$$

$$\gamma = \frac{1}{2\sigma^2} \tag{12}$$

### Hard Margin SVM

$$\min_{w,b}\frac{1}{2}\|w\|^2 \text{ s.t. } y_i(w^T x_i + b) \geq 1 \tag{13}$$

$$\mathcal{L}(\mathbf{w}, w_0, \alpha) = \frac{1}{2}\|\mathbf{w}\|^2 - \sum_{i=1}^{N}\alpha_i[t_i(\mathbf{w}^T\mathbf{x}_i + w_0) - 1] \tag{14}$$

### Soft Margin SVM

$$\mathcal{L}(\mathbf{w}, w_0, \alpha, \mu) = \frac{1}{2}\|\mathbf{w}\|^2 + C\sum_{i=1}^{N}\xi_i - \sum_{i=1}^{N}\alpha_i[t_i(\mathbf{w}\cdot\mathbf{x}_i + w_0) - 1 + \xi_i] - \sum_{i=1}^{N}\mu_i\xi_i \tag{15}$$

## Performance Metrics

### Confusion Matrix

$$\begin{bmatrix} TP & FP \\ FN & TN \end{bmatrix} \tag{16}$$

### Precision

$$\frac{TP}{TP+FP} \tag{17}$$

### Recall

$$\frac{TP}{TP+FN} \tag{18}$$

### F1 Score

$$\frac{2\times\text{Precision}\times\text{Recall}}{\text{Precision}+\text{Recall}} \tag{19}$$

### Accuracy

$$\frac{TP+TN}{TP+FP+FN+TN} \tag{20}$$

### ROC Curve

$$\text{TPR} = \frac{TP}{TP+FN} \tag{21}$$

$$\text{FPR} = \frac{FP}{FP+TN} \tag{22}$$

## Dimensionality Reduction

### PCA

$$\mathbf{X} = \mathbf{X} - \bar{\mathbf{X}} \tag{23}$$

$$\mathbf{Cov} = \frac{1}{N-1}\mathbf{X}^T\mathbf{X} \tag{24}$$

$$\mathbf{C_{ov}}\mathbf{v} = \lambda\mathbf{v} \tag{25}$$

$$C_{ov} = \mathbf{E}[\mathbf{X}\mathbf{X}^\mathbf{T}] \tag{26}$$

When $\mathbf{X}$ is mean-centered.