## Supervised Classification

### Boosting

#### AdaBoosting

- **Weight Assignment:** $w_{i,0} = \frac{1}{n}$
- **Error Rate:** $r_j = \frac{\hat{y}_j^{(i)} \neq y^{(i)}}{\sum_{i=1}^{N} w^{(i)}}$
- Predictor Weight: $\alpha_j = \eta \ln\left(\frac{1-r_j}{r_j}\right)$

### Hard SVM

- $y(x) = w^T\phi(x) + b = 0$
- **Margin:** $\frac{1}{\|w\|}$
- **Objective:** $\min_{w,b} \frac{1}{2}\|w\|^2$
- **Discriminant Function:** $f(x) = w^T\phi(x) + b$
- **Support Vectors:** $y_i(w^T\phi(x_i) + b) = 1$
- **Polynomial Kernel:** $K(x,y) = (1 + \langle x, y\rangle)^d$
- **Gaussian RBF Kernel:** $K(x,y) = \exp\left(-\frac{\|x-y\|^2}{2\sigma^2}\right)$

### Soft SVM

- **Objective:** $\min_{w,b,\xi} \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{N}\xi_i$
- **Constraints:** $y_i(w^T\phi(x_i) + b) \geq 1 - \xi_i$
- **Slack Variables:** $\xi_i \geq 0$
- **Lagrangian:** $\mathcal{L}(w,b,\xi,\alpha,\beta) = \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{N}\xi_i - \sum_{i=1}^{N}\alpha_i(y_i(w^T\phi(x_i) + b) - 1 + \xi_i) - \sum_{i=1}^{N}\beta_i\xi_i$
- **KKT Conditions:** $\alpha_i \geq 0, \beta_i \geq 0, \alpha_i(y_i(w^T\phi(x_i) + b) - 1 + \xi_i) = 0, \beta_i\xi_i = 0$
- **Dual Problem:** $\max_\alpha \sum_{i=1}^{N}\alpha_i - \frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N}\alpha_i\alpha_j y_i y_j K(x_i, x_j)$
- **Predictor:** $f(x) = \sum_{i=1}^{N}\alpha_i y_i K(x, x_i) + b$

## Dimensionality Reduction

### Curse of Dimensionality

- **Volume:** $V_d(r) = r^d$
- **Ratio:** $ratio = \frac{V_{crust}}{V_{S_1}} = \frac{V_{S_1} - V_{crust}}{V_{S_1}}$
- **Vol Eqn:** $V = \frac{r^D \cdot \pi^{D/2}}{\rho(D/2+1)}$
- **ratio** $= 1 - (1 - \frac{\epsilon}{r})^D$

### Feature Selection

- **Embedded:** L1: L1: $\|\mathbf{w}\|_1 = \sum_{j=0}^{M}|w_j|$
- **Wrappers:** Recursive Feature Elimination using Greedy Search
- **Feature Extraction:** PCA, LDA
- **PCA:**