1. 线性回归 → 回归任务

学习 $f(x_i)=wx+b$ 使得 $f(x_i)\simeq y_i$ 用均方误差最小化来求 $w$ 和 $b$

$$\frac{\partial E(w,b)}{\partial w}=2\left(w\sum x_i^2-\sum(y_i-b)x_i\right) \Rightarrow w=\frac{\sum y_i(x_i-\bar{x})}{\sum x_i^2-\frac{1}{m}(\sum x_i)^2}$$ closed-form 的解

$$\frac{\partial E(w,b)}{\partial b}=2\left(mb-\sum(y_i-wx_i)\right) \qquad b=\frac{1}{m}\sum(y_i-wx_i)$$

更一般来说 $\frac{\partial E}{\partial \hat{w}}=2X^T(X\hat{w}-y)$

$\hat{w}^*=(X^TX)^{-1}Xy$

$y=g^{-1}(w^Tx+b)$ 广义线性模型

$g(y)=w^Tx+b$

例: $\ln y=w^Tx+b$

$y=e^{w^Tx+b}$

2. logistic regression

$\ln\frac{y}{1-y}=w^Tx+b$ $\quad \frac{y}{1-y}$ 称为 odds $\quad y$ 表示视样本 $x$ 为正样本的几率

用极大似然法估计参数

优点: 无需假设数据分布

可得到预测概率

sigmoid ⇒ 任意阶可导

3. Linear Discriminant Analysis LDA 线性判别分析 → 监督降维技术

将样例投影到直线上 使同类样例投影点尽可能接近

def: 两类样本的中心在直线上的投影 $w^T\mu_0$, $w^T\mu_1$; 两类样本协方差 $w^T\Sigma_0 w$, $w^T\Sigma_1 w$

欲使同类样本接近 异类样本远离 $w^T\Sigma_0 w+w^T\Sigma_1 w\downarrow$ $\|w^T\mu_0-w^T\mu_1\|_2^2$ 尽可能大 ⇒ $J=\frac{\|w^T\mu_0-w^T\mu_1\|_2^2}{w^T\Sigma_0 w+w^T\Sigma_1 w}$

$=\frac{w^T(\mu_0-\mu_1)(\mu_0-\mu_1)^Tw}{w^T(\Sigma_0+\Sigma_1)w}$

类内散度矩阵 $S_w=\Sigma_0+\Sigma_1=\sum_{x\in X_0}(x-\mu_0)(x-\mu_0)^T+\sum_{x\in X_1}(x-\mu_1)(x-\mu_1)^T$

$S_b=(\mu_0-\mu_1)(\mu_0-\mu_1)^T$

$=\frac{w^TS_bw}{w^TS_ww}$ ⇒ LDA 最大化目标

$\Rightarrow \min\ -w^TS_bw$

s.t. $w^TS_ww=1$

$\Rightarrow S_bw=\lambda S_ww=\lambda(\mu_0-\mu_1)$

$w=S_b^{-1}(\mu_0-\mu_1)$

多分类学习 {
- OvO $N(N-1)/2$ 两两配对
- OvR
- MvM

特例

类别不均衡 {
- 理论解决方法: $\frac{y}{1-y}>1$ 视为正例 ⇒ $\frac{y}{1-y}>\frac{m^+}{m^-}$ 视为正例: 再缩放 代价敏感学习
- 实际操作方法: {
  - 欠采样: 去除反例 EasyEnsemble
  - 过采样: 增加正例 SMOTE
  - 阈值移动 threshold-moving