



监督学习

- 线性回归**
 - 模型假设: $y = wx + b$
 - 代价函数: $J(w, b) = \frac{1}{2} \sum_{i=1}^n (y_i - \hat{y}_i)^2$
 - 梯度下降法: $w := w - \alpha \frac{\partial J}{\partial w}$, $b := b - \alpha \frac{\partial J}{\partial b}$
 - 正规方程: $w = (X^T X)^{-1} X^T y$
- 逻辑回归**
 - 模型假设: $\hat{y} = \sigma(wx + b)$, $\sigma(z) = \frac{1}{1 + e^{-z}}$
 - 代价函数: $J(w, b) = -\frac{1}{n} \sum_{i=1}^n [y_i \log(\hat{y}_i) + (1 - y_i) \log(1 - \hat{y}_i)]$
 - 梯度下降法: $w := w - \alpha \frac{\partial J}{\partial w}$, $b := b - \alpha \frac{\partial J}{\partial b}$
- SVM**
 - 硬间隔最大化: $\max_w \gamma$ s.t. $y_i(w \cdot x_i + b) \geq \gamma$
 - 软间隔最大化: $\min_w \frac{1}{2} \|w\|^2 + C \sum_{i=1}^n \xi_i$
- 决策树**
 - 信息熵: $H(p) = -\sum_{i=1}^n p_i \log_2 p_i$
 - 基尼指数: $G(p) = 2 \sum_{i=1}^n p_i (1 - p_i)$
 - 剪枝: 预剪枝, 后剪枝
- 集成学习**
 - Bagging: 并行训练, 降低方差
 - Boosting: 串行训练, 降低偏差
 - Random Forest: 基于Bagging的决策树集成

无监督学习

- K-Means**
 - 初始化: 随机选择K个质心
 - 迭代: 分配, 更新质心
- PCA**
 - 协方差矩阵: $\Sigma = \frac{1}{n} X^T X$
 - 特征值分解: $\Sigma = U \Lambda U^T$
- SVD**
 - 奇异值分解: $X = U \Sigma V^T$
- 关联规则**
 - 支持度: $\text{support}(X \rightarrow Y) = \frac{|X \cup Y|}{n}$
 - 置信度: $\text{confidence}(X \rightarrow Y) = \frac{|X \cup Y|}{|X|}$

强化学习

- 马尔可夫决策过程**
 - 状态: s
 - 动作: a
 - 奖励: r
 - 策略: $\pi(a|s)$
- Q-Learning**
 - 贝尔曼方程: $Q(s, a) = r + \gamma \max_{a'} Q(s', a')$

深度学习

- 神经网络**
 - 输入层, 隐藏层, 输出层
 - 激活函数: Sigmoid, Tanh, ReLU
- CNN**
 - 卷积: $z_{ij} = \sum_k \sum_l x_{kl} w_{kl}^{ij}$
 - 池化: Max Pooling, Average Pooling
- RNN**
 - 隐藏状态: h_t
 - 输入: x_t
 - 输出: y_t
- GAN**
 - 生成器: $G(z)$
 - 判别器: $D(x)$

数学基础

- 线性代数: 矩阵, 向量, 特征值
- 概率论: 贝叶斯定理, 高斯分布
- 微积分: 梯度, 泰勒展开