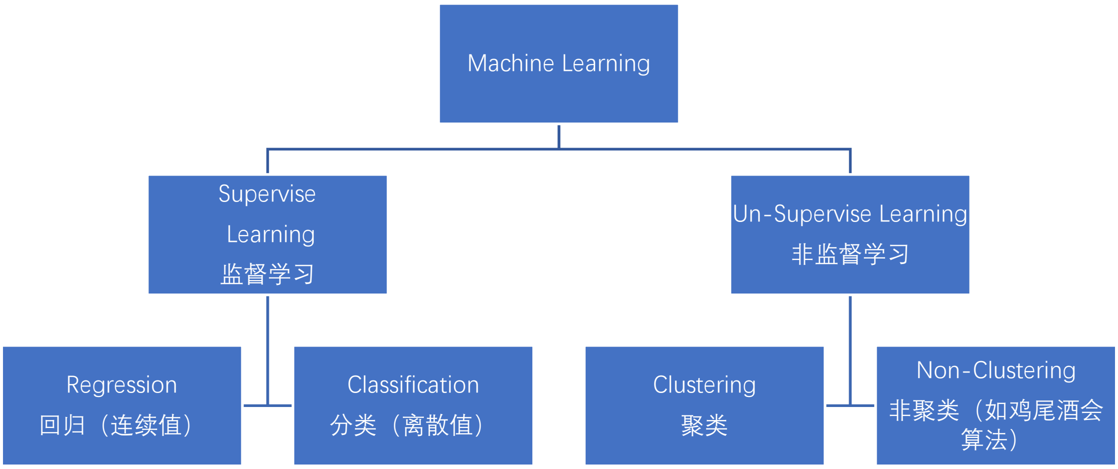
Week 1 笔记

by 骆剑 2017/5/21

Introduction



Supervise Learning 监督学习

数据有标签，模型用来预测输入对应的输出

Regression 回归：预测连续的输出值

Classification 分类：预测离散的输出值

UnSupervise Learning 非监督学习

数据没有标签，模型用来学习数据之间的关系

Linear Regression （线性回归）

Basic Concept（基本概念）：

1.Hypothesis: $h_\theta(x) = \theta_0 + \theta_1 x$

(假设函数)

2.Parameters: θ_0 ； θ_1

3.Cost Function: $J(\theta_0, \theta_1) = \frac{1}{2m}\sum_{i=1}^{m}\left(h_\theta\left(x^{(i)}\right) - y^{(i)}\right)^2$

(损失函数)

4.Goal: $\underset{\theta_0,\theta_1}{\text{minimize}}\ J(\theta_0, \theta_1)$

【注】：式 3 中的 $x^{(i)}$ $y^{(i)}$ 表示第 i 个 example 样本的 x 和 y

样本总数为 m

Gradient descent algorithm（梯度下降算法）

Repeat until convergence （循环直至收敛）

{ //同步更新，赋值(3)(4)必须在计算(1)(2)之后，设 tmp

$$tmp0 := \theta_0 - \alpha \cdot \frac{\partial}{\partial \theta_0} J(\theta_0, \theta_1) \qquad (1)$$

$$tmp1 := \theta_1 - \alpha \cdot \frac{\partial}{\partial \theta_1} J(\theta_0, \theta_1) \qquad (2)$$

$$\theta_0 := tmp0 \qquad (3)$$

$$\theta_1 := tmp1 \qquad (4)$$

}

推导（一）

$$tmp0 := \theta_0 - \alpha \cdot \frac{\partial}{\partial \theta_0} J(\theta_0, \theta_1)$$

$$:= \theta_0 - \alpha \cdot \frac{1}{m} \sum_{i=1}^{m} \left(h_\theta\left(x^{(i)}\right) - y^{(i)}\right)$$

推导（二）

$$tmp1 := \theta_1 - \alpha \cdot \frac{\partial}{\partial \theta_1} J(\theta_0, \theta_1)$$

$$:= \theta_1 - \alpha \cdot \frac{1}{m} \sum_{i=1}^{m} \left[\left(h_\theta\left(x^{(i)}\right) - y^{(i)}\right) \cdot x^{(i)} \right]$$

偏导推导过程（一）

$$\frac{\partial}{\partial\theta_0}J(\theta_0,\theta_1) = \frac{\partial}{\partial\theta_0}\{\frac{1}{2m}\sum_{i=1}^{m}\left(h_\theta\left(x^{(i)}\right)-y^{(i)}\right)^2\}$$

$$= \frac{\partial}{\partial\theta_0}\{\frac{1}{2m}\sum_{i=1}^{m}\left(\theta_0+\theta_1 x^{(i)}-y^{(i)}\right)^2\}$$

$$= \frac{1}{m}\sum_{i=1}^{m}\left(\theta_0+\theta_1 x^{(i)}-y^{(i)}\right)$$

$$= \frac{1}{m}\sum_{i=1}^{m}\left(h_\theta\left(x^{(i)}\right)-y^{(i)}\right)$$

偏导推导过程（二）

$$\frac{\partial}{\partial\theta_1}J(\theta_0,\theta_1) = \frac{\partial}{\partial\theta_1}\{\frac{1}{2m}\sum_{i=1}^{m}\left(h_\theta\left(x^{(i)}\right)-y^{(i)}\right)^2\}$$

$$= \frac{\partial}{\partial\theta_1}\{\frac{1}{2m}\sum_{i=1}^{m}\left(\theta_0+\theta_1 x^{(i)}-y^{(i)}\right)^2\}$$

$$= \frac{1}{m}\sum_{i=1}^{m}\left(\theta_0+\theta_1 x^{(i)}-y^{(i)}\right)$$

$$= \frac{1}{m}\sum_{i=1}^{m}\left[\ \left(h_\theta\left(x^{(i)}\right)-y^{(i)}\right)\cdot x^{(i)}\ \right]$$