

Ng:

监督式学习: ① 线性回归.

(已标化的数据) 线性拟合:

~~已标~~

$$h_{\theta}(x) = \theta_0 + \theta_1 x$$

minimize 误差:

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

梯度下降.

多变量(多项式回归)

a. 参数初始值

b. 学习速率 α .

② 符号(逻辑)回归: logistic Regression.

$$\text{二元分类: } h_{\theta}(x) = \frac{1}{1 + e^{-\theta^T x}}$$

它的 cost func. $\log(\dots)$

多元分类: 拆成多个二元分类

③ ~~神经网络~~ 惩罚

$$\text{过度拟合: } \text{cost} + \lambda \sum_{j=1}^n \theta_j^2$$

④ 神经网络.

cost func (惩罚), 前馈

反向传播算法

⑤ SVM. support vector machines.

折衷的 cost func.

$$h_{\theta}(x) > 0 \rightarrow 1$$

$$\leq 0 \rightarrow 0$$

C 大小重要.

鲁棒性更好.



$$\min_{\theta} \frac{1}{2} \sum_{j=1}^n \theta_j^2 = \frac{1}{2} \|\theta\|^2$$

$$p^{(1)} \cdot \| \theta \| \geq 1 \quad y^{(1)} = 1$$

$$p^{(2)} \cdot \| \theta \| \leq -1 \quad y^{(2)} = 0$$

$\therefore p^{(1)}$ 应尽可能大 θ 与边界垂直.

非线性分类 - SVM:

kernel

选标儿 l^2 构造新 feature $f_i = \exp(-\|x - l^i\|^2 / 2\sigma)$

距离 l^i 近 $f = 1$

远 $f = 0$

l^i 选择: 所有训练点集.

SVM 选 kernel. (lots of

n 大 $m \rightarrow$ logistic

n 小 m 适中 (SVM with Gau...)

n 小 m 大. 构造 feature \rightarrow logistic

非监督式学习

(未标儿的数据) 分类: ① K-Means

1. 选择 k 簇中心点. (随机)

repeat { 2. 根据距离分类

3. 中心点移动到簇内点的平均位置

使用 cost func 调试

局部最优: 多次随机初始化



林

选择分簇数

② 降维. dimensionality reduction.
feature 相关, 过多
可去除

PCA: principle Component Analysis
找一个平面使投影误差最小

a. 均值归一化 $x_j^i - \mu_j$ ~~$\frac{s_j}{n}$~~

b. feature scaling ~~$\frac{x_j^i - \mu_j}{s_j}$~~

$ND \rightarrow kD \Rightarrow$ 找 k 个向量 $(u^{(1)}, u^{(2)}, \dots, u^{(k)})$

将 data 投影到 kD 的空间.

linear regression: 点之间的误差

PCA: 点之间的垂直误差

$(x_j - \mu_j) / s_j$ $s_j =$ 标准差.

$u \rightarrow$ 降维方向 z 投影结果.

Σ data 协方差矩阵

$[U, S, V] = \text{svd}(\Sigma)$ svd 分解

U 即为 u 向量. 取前 k 个向量.

$U_{\text{reduce}} = n \times k$ $U_{\text{reduce}}^T x = z$.

low ~~dimension~~ back high dimen

$x_{\text{approx}} = U_{\text{reduce}} z$ 近似, 在线性上.



PCA ~~minimize~~ minimize $\sum_{i=1}^m \|X^i - X_{\text{approx}}^i\|^2$

使误差 ≤ 0.01 .

(3) 异常检测

正常点数据做模型

检测新样本是否异常

$P(x)$ 正常概率

应用: a. 欺诈检测, b. 制造业, c.

高斯分布 (正态分布)

估计 μ 和 σ^2

算法: data n -feature. feature, 高斯, 独立.

$P(x_i)$

① 选可能反应异常的 feature x_i

② 估计 μ_i $\sigma_i^2 \rightarrow P(x)$

③ 计算新样本的 $P(x)$ 与阈值 ϵ 做比较

train \rightarrow cross validation \rightarrow test

异常检测 vs. 监督学习

正常或异常的数据量小 \rightarrow 异常检测

画 X 的柱状图 \rightarrow 看是否正态? $\rightarrow \log(x)$ for $x^{1/2}$ or x^3

CV 错误分析 \rightarrow 挑选新 feature

多变量高斯分布 \rightarrow 协方差矩阵, 不独立.

$m \gg n$,

独立高斯分布 (常用, 可选用多 features).



论文 → 文档

申请书
合同

报告 < 西电

A. 摘要

B. 研究概

C. ①

D. ②

摘要:

摘要:

参考申请书 Copy

① 总体方案: 图. 架构

200页

一 部分

关系

新特号/问题

-

随机数

资源

摘要

② 硬件实现

摘要 (可以与实际不符)

技术

挑战

解决方案

仿真 (1/2) 中

(调整仿真参数)

技术

3GPP

1333

(要有仿真)

A图

B文章

PPT?

① 问题 (仿真) / 挑战

② 技术/方案

