

# 机器学习天问

Fduzjrqlw

## 目录

<b>1 机器学习理论</b>	<b>3</b>
1.1 简述机器学习问题. . . . .	3
1.2 误差分解公式. . . . .	3
1.3 风险 (risk), 经验风险 (empirical risk) 的含义. . . . .	3
1.4 过拟合, 欠拟合和误差的关系. . . . .	3
<b>2 数据预处理</b>	<b>3</b>
2.1 归一化 (normalization) 和标准化 (standardization), 以及它们的相同点和不同点. . . . .	3
2.2 为什么要使用归一化/标准化? . . . . .	3
2.3 什么时候用归一化, 什么时候用标准化? . . . . .	3
2.4 一定要归一化吗? 举出一些不需要归一化的例子. . . . .	3
<b>3 特征工程</b>	<b>3</b>
3.1 特征抽取的目的. . . . .	3
3.2 独热编码. . . . .	3
<b>4 正则化</b>	<b>3</b>
4.1 做正则化的原因. . . . .	3
4.2 模型复杂程度的评价指标. . . . .	3
4.3 线性回归需要对偏差项 (bias term) 做正则吗? . . . . .	4
4.4 限制经验风险最小化 (Constrained ERM) v.s. 惩罚项经验风险最小化 (Penalized ERM) 的等价性. . . . .	4
4.5 $L_1$ 正则的作用和原因. . . . .	4
4.6 $L_2$ 正则的作用和原因. . . . .	5

4.7	为什么要引入 Elastic Net 中的 $L_1$ 和 $L_2$ 的组合正则. . . . .	5
<b>5</b>	<b>优化算法</b>	<b>5</b>
5.1	GD,SGD 与 mini_batch GD 之间的区别和联系. . . . .	5
5.2	什么时候用 SGD?mini_batch GD 中 batch_size 的选择. . . .	5
5.3	梯度下降算法的收敛率 (convergence rate). . . . .	5
5.4	SGD 的使用技巧. . . . .	6
5.5	SGD 的理论依据. . . . .	6
5.6	在线学习的动机和方法. . . . .	6
5.7	次梯度和次梯度下降算法. . . . .	6
<b>6</b>	<b>模型</b>	<b>6</b>
6.1	SVM . . . . .	6
6.1.1	什么是 SVM 的退化? . . . . .	6
6.1.2	如何构造新的输入数据, 使得一个线形可分的 SVM 问题退化? . . . . .	6