

HOMEWORK 1

1. 请列举实际中 1 个数据挖掘中有监督学习的应用，请包含以下部分:

- (1) 背景介绍（请阅读材料：如何撰写背景介绍）
- (2) 数据说明（说明因变量和自变量各是什么）
- (3) 对以上特定问题如何进行模型评估
- (4) 通过机器学习建模如何实现数据价值（无需列举具体的机器学习模型，但须说明模型如何落地，为谁服务产生价值）

2. 证明《统计学习方法》习题 1.2：

通过经验风险最小化推导极大似然估计。证明模型是条件概率分布，当损失函数是对数损失函数时，经验风险最小化等价于极大似然估计。

格式要求:

- (1) 写题目，宋体小四号，加粗居中。
- (2) 背景介绍以 4-5 个段落为宜，宋体五号单倍行距，段间距设置段前段后均为 0 行，文字两端对齐。四个部分总体：篇幅上限：4 页；篇幅下限：2 页（也就是不准写得太简略，也不准堆砌太多文字）。
- (3) 用 WORD 进行撰写，保存成 PDF 格式提交。

提交时间：3 月 9 日，晚 20:00 之前。请预留一定的时间，迟交作业扣 3 分，作业抄袭 0 分。