



中国科学院大学

University of Chinese Academy of Sciences

模式识别与机器学习

081203M04004H

Chap 9 课程作业解答

2022 年 12 月 24 号

Professor: 黄庆明



学生: 周胤昌

学号: 202228018670052

学院: 网络安全学院

所属专业: 网络安全

方向: 安全协议理论与技术

Problem 1

请给出半监督学习常用的三种假设, 并分别列举出一种依据该假设所设计的半监督学习算法.

Solution: 半监督学习常用的三种假设和主要代表算法叙述如下:

- **平滑假设:** 如果**高密度空间**中两点 x_1, x_2 距离接近, 则对应的输出 y_1, y_2 也应该接近. 主要的代表算法是**生成式模型算法**.
- **聚类假设:** 如果两点在同一个簇, 那么它们很有可能属于同一个类. 主要的代表算法是 **S³VMs 算法**.
- **流形假设:** **高维数据**大致会分布在一个低维的流形上, 并且邻近的样本拥有相似的输出. 主要的代表算法是**基于图的算法**.

Problem 2

简述直推式和归纳式半监督学习算法的异同, 并分别例举出 3 种代表性算法.

Solution: 直推式和归纳式半监督学习的**相同点**是, 他们都使用少量有标注 + 大量无标注的数据集训练. 然而**不同点**在于: 直推式半监督算法关注模型在无标注数据集上的推理预测并且可以没有显式的学习函数 f , 而归纳式半监督算法则关注于学习出一个显式函数 f 来对新的测试集进行推理. 直推式的代表性半监督算法有: TSVM 算法、基于图的算法、半监督聚类. 归纳式的代表性半监督算法有: 自训练模型、多视图学习、生成式模型.