机器学习导论第四次作业

韦俊林 (201928016029023)

2020 年 5 月 15 日

证明简答题

1. 试列举几种常见的半监督学习方法。比较有监督学习、无监督学习、半监督学习、主动学习以及强化学习的异同；

   答：常见的半监督学习有基于高斯混合模型的生成式方法、半监督支持向量机 (Semi-Supervised Support Vector Machine,S3VM)、图半监督学习、协同训练以及半监督聚类等。

   监督学习：在监督学习中输入的每一个数据集都有一个对应的标签，如对防垃圾邮件系统中“垃圾邮件”“非垃圾邮件”，对手写数字识别中的“1”，“2”，“3”等。在建立预测模型的时候，监督式学习建立一个学习过程，将预测结果与“训练数据”的实际结果进行比较，不断的调整预测模型，直到模型的预测结果达到一个预期的准确率。监督式学习的常见应用场景如分类问题和回归问题。常见算法有逻辑回归 (Logistic Regression) 和反向传递神经网络 (Back Propagation Neural Network)；

   无监督学习：无监督学习与监督学习刚好相反，输入的数据集仅提供相应的特征，却没有对应的标签。常见的算法有密度估计 (Density Estimation)、聚类 (Clustering)；

   半监督学习：介于监督学习与无监督学习之间，在输入的数据集中，只给定少量有标签的样本，而更多的是无标签的样本。半监督学习研究的是如何让学习器基于少量的标记样本，自动地利用未标记样本来提升学习性能；

   主动学习：属于半监督学习的一类，首先利用少量有标记的样本训练得到一个模型，然后用该模型选择最有用的未标记样本，并交由专家进行标记，接着将这些专家标记的样本加入训练集重新训练一个模型。因此，主动学习是引入了额外的专家知识，通过与外界的交互来将部分未标记的样本转变为有标记的样本；

   强化学习：在这种学习模式下，输入数据作为对模型的反馈，不像监督模型那样，输入数据仅仅是作为一个检查模型对错的方式，在强化学习下，输入数据直接反馈到模型，模型必须对此立刻作出调整。常见的应用场景包括动态系统以及机器人控制等。常见算法包括 Q-Learning 以及时间差学习 (Temporal difference learning)；

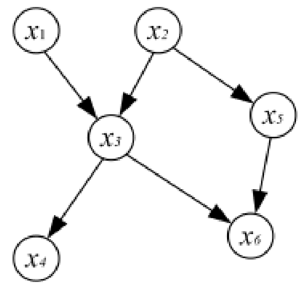
2. 试给出协同训练的方法步骤；

   答：首先，在每个视图上基于有标记样本分别训练一个分类器；

   然后，让每个分类器分别去挑选自己“最优把握的”未标记样本并赋予伪标记。并将伪标记样本提供给另一个分类器作为新的有标记样本用于训练更新；

   这样不断迭代上述步骤（“互相学习，共同进步”），直到两个分类器都不再发生变化，或者达到预定的迭代轮次结束。

3. 针对如下概率图模型，



a) 试写出如下有向图模型对应的联合概率分布函数；

b) 试根据 D-separation 定理，判断随机变量 x_1 与 x_6 是否独立？

c) 试根据 D-separation 定理，判断随机变量 x_1 与 x_5 是否独立？

d) 试根据 D-separation 定理，判断随机变量 x_1 与 x_5 在给定 x_4 时是否条件独立？

答：

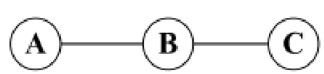
a) $P(x_1)P(x_2)P(x_3|x_1,x_2)P(x_4|x_3)P(x_5|x_2),P(x_6|x_3,x_5)$

b) 图中 x_1 到 x_6 其中一条路径 $x_1 \to x_3 \to x_6$，节点 x_3 是“head-to-tail”类型，这里 x_3 不作为观察点，因此 x_1 与 x_6 不是 x_3 条件独立，所以 x_1 与 x_6 不独立。

c) 图中 x_1 到 x_5 有两条路径分别是 $x_1 \to x_3 \leftarrow x_2 \to x_5$ 和 $x_1 \to x_3 \to x_6 \leftarrow x_5$。在第一条路径中 x_3 是“head-to-head”，x_3 不作为观察点，所以 x_1 与 x_5 是 x_3 下条件独立。在第二条路径中 x_6 是“head-to-head”，x_6 不作为观察点，所以 x_1 与 x_5 是 x_6 下条件独立。因此，从 x_1 到 x_5 上任意路径都经过碰撞点，所以 x_1 与 x_5 独立。

d) 图中 x_1 到 x_5 的一条路径 $x_1 \to x_3 \leftarrow x_2 \to x_5$，在 x_4 给定的情况下不会被 x_2 与 x_3 阻塞，所以 x_1 与 x_5 在给定 x_4 时不条件独立。

4. 试指出如下无向概率图模型的所有最大团，写出对应的联合概率分布函数，并利用概率基本公式证明：A 和 C 在给定 B 的条件下独立，即 $P(A,C|B) = P(A|B)P(C|B)$。



答：该无向图有两个最大团，分别是 $\{A,B\}$ 和 $\{B,C\}$，从而对应的联概率分布函数为：

$$P(A,B,C) = \frac{1}{Z}\psi_{AB}(A,B)\psi_{BC}(B,C)$$

其中 Z 为归范化因子，$\psi_C(C)$ 为势函数；

基于条件概率可得：

$$
\begin{aligned}
P(A,C|B) &= \frac{P(A,B,C)}{P(B)} = \frac{P(A,B,C)}{\sum_{A'}\sum_{C'}P(A',B,C')} \\
&= \frac{\frac{1}{Z}\psi_{AB}(A,B)\psi_{BC}(B,C)}{\sum_{A'}\sum_{C'}\frac{1}{Z}\psi_{AB}(A',B)\psi_{BC}(B,C')} \\
&= \frac{\psi_{AB}(A,B)}{\sum_{A'}\psi_{AB}(A',B)} \cdot \frac{\psi_{BC}(B,C)}{\sum_{C'}\psi_{BC}(B,C')} \\
P(A|B) &= \frac{P(A,B)}{P(B)} = \frac{\sum_{C'}P(A,B,C')}{\sum_{A'}\sum_{C'}P(A',B,C')} \\
&= \frac{\sum_{C'}\frac{1}{Z}\psi_{AB}(A,B)\psi_{BC}(B,C')}{\sum_{A'}\sum_{C'}\frac{1}{Z}\psi_{AB}(A',B)\psi_{BC}(B,C')} \\
&= \frac{\psi_{AB}(A,B)}{\sum_{A'}\psi_{AB}(A',B)}
\end{aligned}
$$

综上所述，证得 $P(A,C|B) = P(A|B)P(C|B)$。