

机器学习实验流程

- 1 标记数据集记作 D 。
- 2 将 D 固定均分为 k 个子集（一般 $k=10$ 或 5 ），记作 $F_1 \cdots F_k$ 。
- 3 对于 i 从 1 到 k ：
 - 3.1 取 F_i ，记作 $TEST_i$ 。
 - 3.2 取其它 $k-1$ 个 F 的并集，记作 TV_i 。
 - 3.3 将 TV_i 按 9:1 比例固定划分为 $TRAIN_i$ 和 $VALID_i$ 。
 - 3.4 根据 grid search，确定超参组合的候选集 THETA。
 - 3.5 对于 THETA 中的每种超参组合 θ ，在 $TRAIN_i$ 上训练，在 $VALID_i$ 上验证，验证效果最好的一种超参组合记作 $\theta_{i,1}$ 。
 - 3.6 对于 j 从 1 到 R （一般 $R=3$ ）：
 - 3.6.1 将 TV_i 按 9:1 比例随机划分为 $TRAIN_{i,j}$ 和 $VALID_{i,j}$ 。
 - 3.6.2 采用超参组合 $\theta_{i,1}$ ，在 $TRAIN_{i,j}$ 上训练，每个 epoch 在 $VALID_{i,j}$ 上验证，如果验证效果连续 n 个 epochs 都未提升，训练即可 early stop。
 - 3.6.3 在 $TEST_i$ 上测试，结果记作 $result_{i,j}$ 。
 - 3.7 取 $result_{i,1} \cdots result_{i,R}$ 的均值，记作 $result_i$ 。
- 4 取 $result_1 \cdots result_k$ 的均值，记作 $result$ 。
- 5 报告 $result$ 和标准差。

备注：

- 1 所有“划分”都应当是 stratified（保持 class 比例）。
- 2 所有“固定”划分应当预先存储，供所有待比较的方法共享。
所有“随机”划分不应预先存储，每次重新随机。

参考文献：

A Fair Comparison of Graph Neural Networks for Graph Classification (ICLR 2019)