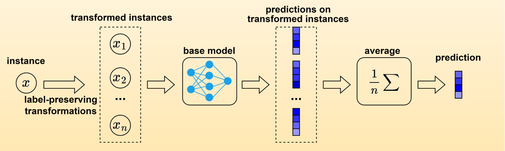
中国科学院软件研究所学术年会 暨重点实验室科技活动周 2025 第十届
学术论文

Extending Test-Time Augmentation with Metamorphic Relations for Combinatorial Problems
结合蜕变关系的测试时增强方法

卫思为，张旭东，周志阳，蔡彦

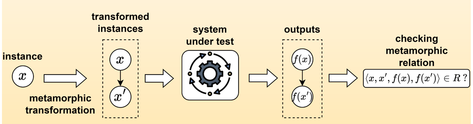
The Forty-First International Conference on Machine Learning (ICML 2024) Spotlight
联系人：蔡彦 (yancai@ios.ac.cn)

背景：测试时增强



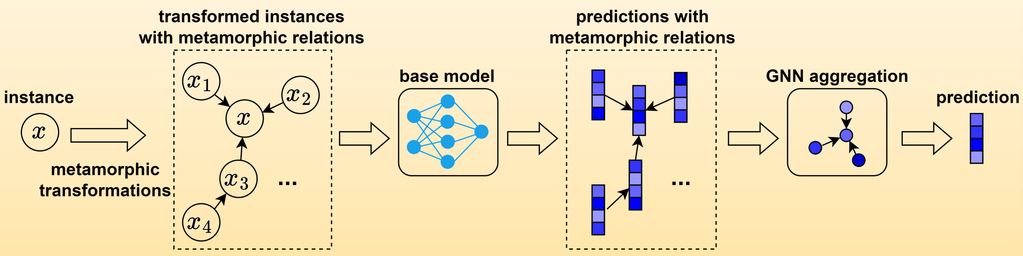
- 在模型推理时进行
- 通过平均/投票，聚合模型在等价样本上的预测
- 对标签保持变换的要求限制了其应用范围

背景：蜕变测试



- 一种软件测试方法，用于缓解测试Oracle问题
- 相比完全刻画系统行为，检测输出间关系更容易
- 多个输出间所必须满足的关系称为蜕变关系

概述



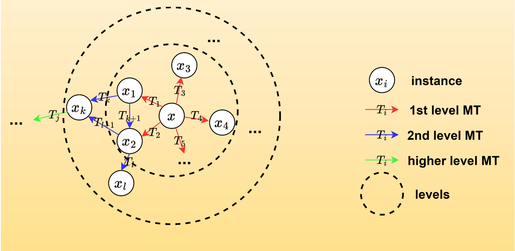
本质上推广了传统测试时增强方法，通过：

1. 使用一般的(非对称)蜕变关系
2. 使用图神经网络作为可训练的聚合模型

关键优势：

- 可以使用一般蜕变关系（一般非对称）
- 从而比传统TTA有更广泛的应用场景

构建蜕变关系图



图神经网络聚合

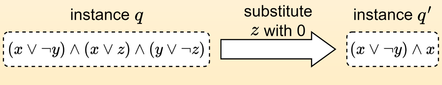
- 避免手写复杂的规则
- 可以适应不同模型的特征
- 使用MPNN聚合:

$$m_u = \frac{1}{|N(u)|} \sum_{v \in N(u)} M_t(\mathrm{h}_u^t)$$

$$\mathrm{h}_u^{t+1} = U_t(\mathrm{h}_u^t, m_u^{t+1})$$

实验评估

SAT

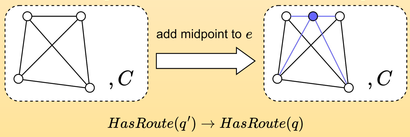


$$SAT(q') \rightarrow SAT(q)$$

Classification result for SAT in accuracy.

Dataset	NeuroSAT	MAgg(10)	MAgg(10,10)
SR(40)	0.8444	0.9548	**0.9757**
SR(80)	0.7268	0.7936	**0.8533**
SR(120)	0.6270	0.6412	**0.6643**

TSP

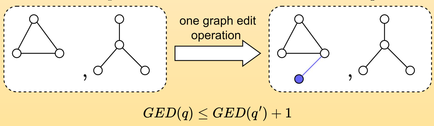


$$HasRoute(q') \rightarrow HasRoute(q)$$

Classification result for Decision TSP in accuracy.

Dataset	TSP-GNN	MAgg(10)
TSP(0.01)	0.6562	**0.6812**
TSP(0.02)	0.8101	**0.8321**

GED



$$GED(q) \leq GED(q') + 1$$

Regression result for GED in RMSE (lower is better).

Dataset	GREED	MAgg(50)
AIDS	0.7957	0.7994
LINUX	0.4151	**0.2409**
IMDB	6.7341	**6.3107**