

基于知识萃取的深度学习模型的可解释测试研究

可解释测试覆盖指标研究

具有语义逻辑的测试覆盖目标的抽取

受限场景下的模型决策逻辑覆盖的模拟

深度学习模型测试行为的分析（包含白盒和黑盒测试场景）

样本表示关键特征抽取

各层判别决策能力分析

决策路径覆盖指标设计

可解释测试集生成方法研究

不同测试数据的检测能力分析与提炼

大规模无标注数据的代表数据选择

可能导致模型错误行为的边界数据选择