自动化测试-AI 测试

助教 刘佳玮 jw.liu@smail.nju.edu.cn

助教 刘关迪 liuguandi@smail.nju.edu.cn

一、研究背景

随着智能模型在各种领域扮演的角色越来越重要，智能模型的可靠性也面临着越来越高的要求，尤其是在诸如交通、医疗等安全攸关的领域中。但是，不同于由开发人员主动编写程序逻辑的传统软件，智能软件的决策逻辑是被动的从给定的数据集中学习到的，即智能软件是数据驱动决策的。因此，在智能软件的测试中，更需要注重测试用例，即测试数据的设计，研究如何通过控制测试数据的生成方向，来达成不同的测试需求，包括智能模型的鲁棒性测试、公平性测试、后门检测、可解释性分析等。如何在现有的传统软件测试基础上，针对深度学习的本质特征，对模糊测试、数据变异、蜕变关系、查分测试等一些经典的测试方法进行创新性改造，使其能够应用在智能模型的测试中，是本课程对于同学们的的基本要求。

二、课程综述

方向一、二、三围绕深度学习模型的测试，从智能模型模糊测试的角度着手，按照不同的阶段，分别关注测试数据的选择、测试数据的生成、测试结果的评估；方向四围绕深度学习框架的测试，包括深度学习框架的缺陷检测和缺陷的原因分析。在综述的完成过程中，希望同学们关注一下传统软件测试的方法是如何在智能模型测试中应用的，尤其是为了适应智能模型的特性而带来的技术挑战。

方向一：深度学习模型测试数据选择技术研究

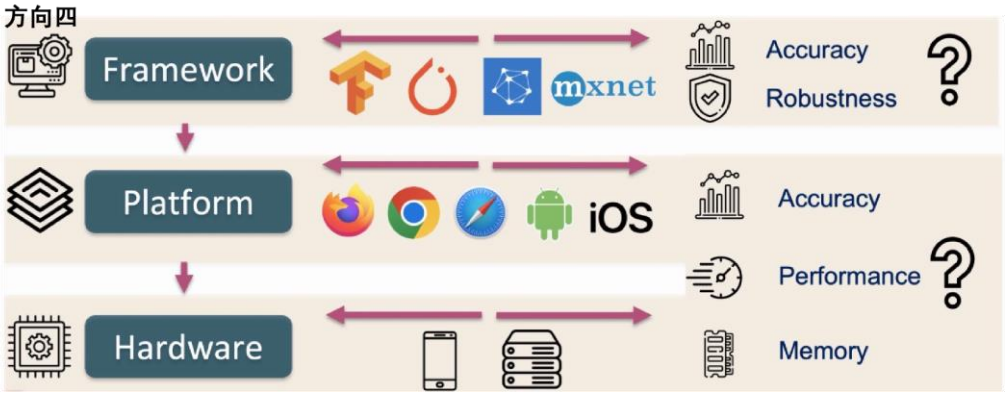
参考关键词：测试数据选择、测试数据排序、测试数据度量

方向二：深度学习模型测试数据生成技术研究

参考关键词：数据变异、蜕变关系、数据扩增

方向三：深度学习模型测试结果评估技术研究

参考关键词：鲁棒性、公平性、后门检测、可解释性分析



方向四：深度学习框架缺陷检测和原因分析技术研究

参考关键词：深度学习框架/库测试、深度学习框架/库实证研究

三、课程大作业

方向一：面向 xxx 场景的深度学习模型测试技术

在智能模型的测试中，不同场景的领域特性会对测试方法的设计带来而外的要求，例如自动驾驶场景下无意义的噪声添加并不能代表现实生活中会遇到的天气、光照等真实问题。因此，希望同学们能锻炼场景分析的能力，建立某一个特定场景下测试需求，设计并实现可用的测试方法，场景自选，可大（图像场景、点云场景等）可小（门禁人脸识别场

景、课堂教学语音生成）。本方向提供可选的待测模型和数据集，各位小组也可以根据小组的特殊场景自选其他模型和数据集。

可选待测模型：https://github.com/open-mmlab

可选数据集：https://www.kaggle.com/datasets

方向二：基于等价融合算子的深度学习框架差分测试技术

深度学习框架为深度学习模型的开发提供了必要算子的开源接口，例如 Conv1D、Conv2D 等。为了避免框架的缺陷被引入到深度学习模型中，需要对深度学习框架开展测试来发现其中的缺陷。在深度学习框架的测试中，深度学习模型是测试输入，为了检测深度学习框架是否出现了缺陷，需要利用差分测试，对比多个等价的深度学习模型来分析是否出现了不一致的异常行为。

本题目研究如何利用等价的融合算子生成等价的深度学习模型。在深度学习框架的内部，每个接口对应的算子运算可以是基础算子，例如 tf.add ，也可以由多个其他基础算子组成的复合算子，例如 tf.einsum。执行复合算子等效于执行其包含的每个基础算子。本题目提供了部分框架下示例融合算子，选择本题目的小组可以自行选择被测的深度学习框架，梳理所选框架下的尽可能完整的融合算子列表，并至少实现其中较为常用的 5 组融合算子，生成测试数据、开展深度学习框架的差分测试。

示例融合算子:

https://www.tensorflow.org/lite/models/convert/operation_fusion

https://www.mindspore.cn/tutorial/zh-CN/r0.5/advanced_use/graph_kernel_fusion.html