

用户验收测试计划

1. 测试目标

- 验证 TimeMixer 模型能否有效地预测 HVAC 系统的故障。这包括预测故障发生的 时间、类型，以及 严重程度。
- 验证 TimeMixer 模型在实际 HVAC 系统运行环境中的性能表现。这包括模型的 预测精度、稳定性、实时性，以及对不同运行条件的 适应性。
- 验证 TimeMixer 模型的易用性。这包括模型的 部署、配置、数据输入、结果解读，以及对用户的友好程度。

2. 测试范围

2.1 功能测试

- 数据集准备:
 - 验证 ttl 文件能否被正确解析，并转换为 TimeMixer 模型可接受的输入格式。
 - 验证数据清洗和预处理步骤的有效性，例如缺失值处理、异常值处理、数据标准化等。
 - 验证数据集中是否包含足够的故障样本，以支持模型的训练和评估。
- 故障预测性能:
 - 使用历史 HVAC 运行数据，训练 TimeMixer 模型，并用测试集评估其预测性能。
 - 评估指标应包括但不限于：
 - 准确率 (Accuracy): 预测正确的故障样本占总样本的比例。
 - 精确率 (Precision): 预测为故障的样本中，实际为故障的样本比例。
 - 召回率 (Recall): 实际为故障的样本中，被正确预测为故障的样本比例。
 - F1 分数 (F1 Score): 精确率和召回率的调和平均值。
 - 提前预警时间 (Lead Time): 预测的故障发生时间与实际故障发生时间之间的间隔。
- 不同故障类型和严重程度的预测:
 - 验证 TimeMixer 模型能否准确预测不同类型的 HVAC 故障，例如传感器故障、制冷剂泄漏、压缩机故障等。
 - 验证 TimeMixer 模型能否区分不同严重程度的故障，例如轻微故障、中等故障、严重故障等。
- 实时性:
 - 验证 TimeMixer 模型的预测速度是否满足 HVAC 系统实时监控的需求。
 - 测量模型的预测延迟，并与系统可接受的延迟进行比较。
- 适应性:
 - 验证 TimeMixer 模型对不同运行条件的适应性，例如不同的环境温度、负荷水平、设备型号等。
 - 使用来自不同运行条件的数据集，评估模型的预测性能。

2.2 非功能测试

- **易用性:**
 - 评估 TimeMixer 模型的部署和配置过程是否简单易懂。
 - 评估用户界面是否友好，用户能否方便地输入数据、查看预测结果，并理解模型的输出。
- **可维护性:**
 - 评估 TimeMixer 模型的代码是否易于理解和维护，例如代码结构、注释、文档等。
 - 评估模型是否易于更新和扩展，例如添加新的数据源、新的故障类型等。
- **安全性:**
 - 评估 TimeMixer 模型的数据安全性和隐私保护措施。
 - 验证模型是否能够抵御潜在的安全威胁，例如数据泄露、恶意攻击等。

3. 测试方法

- **模拟测试:** 使用历史 HVAC 运行数据，模拟真实的系统运行环境，对 TimeMixer 模型进行测试。
- **现场测试:** 将 TimeMixer 模型部署到真实的 HVAC 系统中，进行现场测试，收集实际运行数据，评估模型的性能表现。
- **用户访谈:** 与 HVAC 系统的操作人员和维护人员进行访谈，收集他们对 TimeMixer 模型的反馈意见，例如易用性、实用性等。

4. 测试环境

- 操作系统: Linux, macOS, Windows
- 硬件平台: CPU, GPU (MPS, CUDA), 多GPU (根据实际部署环境选择)
- 软件环境: Python, PyTorch, pandas, numpy, matplotlib, rdflib
- HVAC 系统模拟器或真实的 HVAC 系统

5. 测试用例

5.1 数据集准备

测试用例	测试步骤	预期结果
1.1	解析 Brick 模型 ttl 文件	ttl 文件被正确解析，没有错误信息。
1.2	提取时间序列数据	从 ttl 文件中提取出 HVAC 运行数据的时间序列，数据类型和维度正确。
1.3	数据清洗和预处理	缺失值、异常值被正确处理，数据被标准化到合适的范围。
1.4	故障样本数量	数据集中包含足够的故障样本，以支持模型的训练和评估。

Table 1

5.2 故障预测性能

测试用例	测试步骤	预期结果
2.1	训练 TimeMixer 模型	模型训练成功，没有错误信息。
2.2	评估预测精度	模型在测试集上的准确率、精确率、召回率、F1 分数达到预设阈值。
2.3	预测提前预警时间	模型预测的故障发生时间与实际故障发生时间之间的间隔达到预设要求。

Table 2

5.3 不同故障类型和严重程度的预测

测试用例	测试步骤	预期结果
3.1	预测不同类型的故障	模型能够准确预测不同类型的 HVAC 故障，例如传感器故障、制冷剂泄漏、压缩机故障等。
3.2	区分不同严重程度的故障	模型能够区分不同严重程度的故障，例如轻微故障、中等故障、严重故障等。

Table 3

5.4 实时性

测试用例	测试步骤	预期结果
4.1	测量预测延迟	模型的预测延迟小于系统可接受的延迟。

Table 4

5.5 适应性

测试用例	测试步骤	预期结果
5.1	使用不同运行条件的数据集	模型在不同运行条件下的预测性能满足预设要求。

Table 5

5.6 易用性

测试用例	测试步骤	预期结果
6.1	部署和配置模型	模型部署和配置过程简单易懂。

测试用例	测试步骤	预期结果
6.2	使用用户界面	用户界面友好，用户能够方便地输入数据、查看预测结果，并理解模型的输出。

Table 6

5.7 可维护性

测试用例	测试步骤	预期结果
7.1	代码审查	模型代码结构清晰，注释完整，文档详细。
7.2	更新和扩展模型	模型易于更新和扩展，例如添加新的数据源、新的故障类型等。

Table 7

5.8 安全性

测试用例	测试步骤	预期结果
8.1	数据安全和隐私保护	模型的数据安全性和隐私保护措施符合相关标准和规范。
8.2	安全性测试	模型能够抵御潜在的安全威胁，例如数据泄露、恶意攻击等。

Table 8

6. 测试进度

- 制定测试用例：1天
- 搭建测试环境：2天
- 执行功能测试：5天
- 执行非功能测试：2天
- 分析测试结果：1天
- 编写测试报告：1天

7. 测试交付物

- 测试计划
- 测试用例
- 测试报告
- TimeMixer 模型部署文件 (如果适用)

8. 测试责任人

- 开发团队负责编写测试代码、准备测试数据集、部署模型，并执行测试。
- 测试团队负责设计测试用例、搭建测试环境、分析测试结果，并编写测试报告。

- HVAC 领域专家负责提供故障预测相关的专业知识，并评估模型的预测结果。

9. 风险

- 测试时间不足：需要合理安排测试进度，并根据实际情况进行调整。
- 硬件资源不足：需要提前申请足够的硬件资源，以保证测试顺利进行。
- 测试用例覆盖不全：需要尽可能全面地设计测试用例，以减少漏测的可能性。
- 故障样本不足：需要收集足够的故障样本，以确保模型的训练和评估效果。