Experiment Replication Guide

This document provides detailed step-by-step instructions for reproducing all experimental results presented in our paper "Enhanced Bayesian Optimization for Automated System Configuration Tuning."

1. Environment Setup

1.1 Prerequisites

Python 3.8-3.10

8GB RAM minimum (16GB recommended for LLVM dataset)

Approximately 2GB free disk space

1.2 Installation

git clone https://github.com/sgjqk/ISE_Coursework.git

pip install -r requirements.txt

2. Dataset Preparation

All datasets are included in the repository under the datasets/ directory.

ls datasets/

7z.csv Apache.csv brotli.csv LLVM.csv PostgreSQL.csv spear.csv storm.csv x264.csv

3. Running Baseline (Random Search)

3.1 Execute Random Search Algorithm

Run random search with 20 independent runs, 100 evaluations per run

python main_random.py

3.2 Generate Baseline Visualizations

Create visualization of random search results

python visualize_multiple_runs.py

4. Running Enhanced TPE Method

4.1 Execute Enhanced TPE Algorithm

Run enhanced TPE with 20 independent runs, 300 evaluations per run

python main.py

4.2 Generate TPE Visualizations

Create visualization of TPE results

python visualize_tpe_search_results.py

5. Statistical Comparison

5.1 Compare Random Search vs. TPE

Run statistical comparison

python stats_test.py

6. Additional Visualization Options

6.1 For Specific Systems Only

Visualize multiple runs for specific system

python visualize_multiple_runs.py --dataset storm

6.2 For TPE Results Only

Visualize TPE results for specific system

python visualize_tpe_search_results.py --dataset PostgreSQL

7. Verification Process

7.1 Verify Performance Improvements

Check stats_test.py output:

All p-values should be < 0.05 (statistically significant)

Performance improvements should match Table 1:

storm: ~100% improvement

spear: ~50% improvement

PostgreSQL: ~0.05% improvement

Examine visualization results:

Compare files in visualization_results_multi/ with those in tpe_visualization_multi/

TPE convergence curves should show faster convergence

TPE box plots should show lower median values for all systems

Storm should show 0 values for all TPE runs

7.2 Verify File Integrity

Check number of result files

ls -l random_search/ | grep -c ".csv" # Should show ~160 files (8 systems × 20 runs)

ls -l tpe_results_improved/ | grep -c ".csv" # Should show ~160 files (8 systems × 20 runs)