

Replication Guide for Bug Report Classification Tool

Steps to Replicate the Results

1. Install Python 3.6+.
2. Install dependencies using:

```
pip install pandas numpy nltk scikit-learn xgboost
```

3. Download dataset and place it in `datasets/` folder.
4. Run the script:

```
python main.py
```

5. The script runs classification 10 times and prints average results.

Verifying Output

- Check terminal output for accuracy, precision, recall, and F1-score.
- Compare with baseline results in `report.pdf`.
- Check the saved CSV file named after the selected project for classification results.

Reproducing Statistical Analysis

- The script logs accuracy, precision, recall, and AUC across 10 runs.
- Run multiple times and use the CSV files to confirm statistical trends.

By following these steps, the reported results can be fully replicated.