

Replication Instructions

Steps to Replicate the Results

1. Use the provided `bug_reports.csv` in the `data/` folder.
2. Ensure all dependencies are installed (see `requirements.pdf`).
3. Run the scripts in the following order:

- (a) Preprocess the dataset:

```
python src/preprocess.py
```

- (b) Balance the dataset:

```
python src/balance_data.py
```

- (c) Train the model:

```
python src/train.py
```

- (d) Evaluate the model:

```
python src/evaluate.py
```

4. Outputs produced:

- `svm_model.pkl`
- `tfidf_vectorizer.pkl`
- `classification_report.txt`
- `confusion_matrix.png`

Running these steps on the same dataset should yield results consistent with the original experiment.