Experiment Replication Guide

1. Data Acquisition

The required dataset links are provided in the GitHub repository. Please download and place the data in the data/ directory.

2. Environment Configuration

Ensure that Jupyter Notebook is installed along with all dependencies listed in requirements.pdf.

3. Steps to Replicate the Experiment in Jupyter Notebook

1. **Open the Jupyter Notebook**

jupyter notebook

2. **Load and Preprocess Data**

   Run the corresponding notebook cells that handle data preprocessing.

3. **Train the Model**

   Execute the training cells to train the SVM model with SMOTE.

4. **Perform Model Evaluation**

   Run the evaluation cells to generate performance metrics (Precision, Recall, F1-score).

4. Results Verification

Experiment results, including evaluation metrics and confusion matrices, are displayed within the notebook. The results can be verified directly in the output of the notebook cells. Additionally, to save the results, execute the notebook cell that exports the results to the results/ directory.

To manually check the saved results, use:

cat results/experiment_summary.txt

To visualize the results, execute the notebook cell responsible for plotting performance graphs.

5. Expected Outcomes

The replication should yield classification accuracy similar to that reported in the project (F1-score around 0.88). If significant deviations occur, ensure that the dataset and preprocessing steps are correctly executed.