Method

Take benchmark dataset

Split dataset into train, val, test

Train on train and val

Evaluate on train, val, test, benchmark

Visualize result

Full data

Duplicate filtering

Full data without duplicates

Data splitting

Benchmark set

Data to split

Training

Test set

Validation set

Train set

Evaluation