Replication

To replicate the results first run the Naïve Bayes Classifier:
' python br_classification.py'

This should output:

'NB_all_results.csv' with results for all projects and Individual project results in '{project_name}_NB.csv'.

Then run the BERT model using the torch and np seeds set to 39 on lines 13.14:

'python BERT_br_classifier.py'

This should output:

'BERT_all_results.csv' with results for all projects and Individual project results in '{project_name}_BERT.csv'.

Then run the statistical analysis file:

'python statistical_tests.py'

The output of this file should state the significant statistical difference between the two models for the Accuracy, Recall, Precision and F1 Score.

The results of the models should be approximately the same and depends on hardware used. To run my model I used a RTX 3050 Laptop GPU with 4GB of VRAM.