

After running the manual, three metrics and the statistic test results for each dataset will be directly displayed in the terminal.

Baseline:

```
C:\Users\TlANS\Desktop\ISE作业\lab1>python br_classification.py
[nltk_data] Downloading package stopwords to
[nltk_data] C:\Users\TlANS\AppData\Roaming\nltk_data...
[nltk_data] Package stopwords is already up-to-date!
=== Naive Bayes + TF-IDF Results ===
Average Precision: 0.6120
Average Recall: 0.7575
Average F1 score: 0.5624
=== Naive Bayes + TF-IDF Results ===
Average Precision: 0.6373
Average Recall: 0.7112
Average F1 score: 0.5388
=== Naive Bayes + TF-IDF Results ===
Average Precision: 0.6300
Average Recall: 0.6936
Average F1 score: 0.5412
=== Naive Bayes + TF-IDF Results ===
Average Precision: 0.6005
Average Recall: 0.7364
Average F1 score: 0.5159
=== Naive Bayes + TF-IDF Results ===
Average Precision: 0.5698
Average Recall: 0.6668
Average F1 score: 0.4611
```

Scheme 1:

```
C:\Windows\System32\cmd.exe
[nltk_data] Package stopwords is already up-to-date!
2025-04-04 12:01:06.679156: I tensorflow/core/util/port.cc:153] oneDNN custom operations are on. You may see slightly different numerical results due to floating-point round-off errors from dif
ferent computation orders. To turn them off, set the environment variable 'TF_ENABLE_ONEDNN_OPTS=0'.
2025-04-04 12:01:06.970155: I tensorflow/core/util/port.cc:153] oneDNN custom operations are on. You may see slightly different numerical results due to floating-point round-off errors from dif
ferent computation orders. To turn them off, set the environment variable 'TF_ENABLE_ONEDNN_OPTS=0'.
=== MLP + TinyBERT Experiment Results ===
Project: pytorch
Average Precision: 0.7191
Average Recall: 0.6703
Average F1 score: 0.6842
--- Statistical Test Results ---
Precision: t-statistic = 12.9388, p-value = 0.0000
Recall: t-statistic = -8.5269, p-value = 0.0000
F1 Score: t-statistic = 12.2897, p-value = 0.0000
=== MLP + TinyBERT Experiment Results ===
Project: tensorflow
Average Precision: 0.7626
Average Recall: 0.6878
Average F1 score: 0.7101
--- Statistical Test Results ---
Precision: t-statistic = 19.8206, p-value = 0.0000
Recall: t-statistic = -3.2463, p-value = 0.0029
F1 Score: t-statistic = 18.8983, p-value = 0.0000
=== MLP + TinyBERT Experiment Results ===
Project: keras
Average Precision: 0.7664
Average Recall: 0.6966
Average F1 score: 0.7170
--- Statistical Test Results ---
Precision: t-statistic = 22.7855, p-value = 0.0000
Recall: t-statistic = 0.3722, p-value = 0.7124
F1 Score: t-statistic = 22.8648, p-value = 0.0000
=== MLP + TinyBERT Experiment Results ===
Project: incubator-mxnet
Average Precision: 0.7522
Average Recall: 0.7121
Average F1 score: 0.7233
--- Statistical Test Results ---
Precision: t-statistic = 15.0153, p-value = 0.0000
Recall: t-statistic = -1.9415, p-value = 0.0620
F1 Score: t-statistic = 20.1757, p-value = 0.0000
=== MLP + TinyBERT Experiment Results ===
Project: caffe
Average Precision: 0.7054
Average Recall: 0.6165
Average F1 score: 0.6373
--- Statistical Test Results ---
Precision: t-statistic = 5.5877, p-value = 0.0000
Recall: t-statistic = -3.7591, p-value = 0.0000
F1 Score: t-statistic = 11.0510, p-value = 0.0000
```

Scheme 2:

```
C:\Windows\System32\cmd.exe
[nltk_data] C:\Users\TIAN\S\AppData\Roaming\nltk_data...
[nltk_data] Package stopwords is already up-to-date!
=== SVM + TF-IDF Experiment Results ===
Project:      pytorch
Average Precision: 0.6804
Average Recall: 0.7547
Average F1 score: 0.7817

--- Statistical Test Results ---
Precision: t-statistic = 14.6478, p-value = 0.0000
Recall:    t-statistic = -0.3366, p-value = 0.7388
F1 Score:  t-statistic = 22.3492, p-value = 0.0000
=== SVM + TF-IDF Experiment Results ===
Project:      tensorflow
Average Precision: 0.7847
Average Recall: 0.8395
Average F1 score: 0.7873

--- Statistical Test Results ---
Precision: t-statistic = 41.9981, p-value = 0.0000
Recall:    t-statistic = 22.5097, p-value = 0.0000
F1 Score:  t-statistic = 54.6136, p-value = 0.0000
=== SVM + TF-IDF Experiment Results ===
Project:      keras
Average Precision: 0.7862
Average Recall: 0.8828
Average F1 score: 0.7920

--- Statistical Test Results ---
Precision: t-statistic = 24.4758, p-value = 0.0000
Recall:    t-statistic = 14.3188, p-value = 0.0000
F1 Score:  t-statistic = 34.7912, p-value = 0.0000
=== SVM + TF-IDF Experiment Results ===
Project:      incubator-mxnet
Average Precision: 0.7511
Average Recall: 0.7837
Average F1 score: 0.7641

--- Statistical Test Results ---
Precision: t-statistic = 22.7974, p-value = 0.0000
Recall:    t-statistic = 4.5083, p-value = 0.0001
F1 Score:  t-statistic = 26.7978, p-value = 0.0000
=== SVM + TF-IDF Experiment Results ===
Project:      caffe
Average Precision: 0.6612
Average Recall: 0.6162
Average F1 score: 0.6297

--- Statistical Test Results ---
Precision: t-statistic = 4.6627, p-value = 0.0001
Recall:    t-statistic = -3.6875, p-value = 0.0009
F1 Score:  t-statistic = 10.6790, p-value = 0.0000
```