File - apomakrismeno_test

```
1 C:\Users\User\Desktop\ergasia\.venv\Scripts\python.exe C:\Users\User\Desktop\ergasia\apomakrismeno_test.py
2 Loading and Preprocessing Data (Brute Force)...
3 Creating Shingles (Brute Force)...
4 Starting Brute Force Jaccard computation...
5 Finished Brute Force Jaccard computation.
6 Brute Force total time: 9403.71 sec
7
8
9 Loading and preprocessing data (LSH) with threshold=0.5, num_perm=128...
10 Generating MinHash signatures (LSH)...
11 Building LSH index...
12 Finding nearest neighbors...
13 Nearest neighbor search completed.
14
15 Loading and preprocessing data (LSH) with threshold=0.5, num_perm=256...
16 Generating MinHash signatures (LSH)...
17 Building LSH index...
18 Finding nearest neighbors...
19 Nearest neighbor search completed.
20
21 Loading and preprocessing data (LSH) with threshold=0.5, num_perm=512...
22 Generating MinHash signatures (LSH)...
23 Building LSH index...
24 Finding nearest neighbors...
25 Nearest neighbor search completed.
26
27 Final Results Table:
28
29                Type  BuildTime   QueryTime    TotalTime  fraction of the true K most similar documents that are reported by
                                                           LSH method as well
                                                           Parameters
30 Brute-Force-Jaccard   0.00  9403.713814  9403.713814   100%
31       LSH-Jaccard   3927.45   4.870000  3932.320000    6% Perm=128, threshold=0.5
32       LSH-Jaccard   4340.56   6.490000  4347.050000    5% Perm=256, threshold=0.5
33       LSH-Jaccard   5343.29   9.770000  5353.050000    5% Perm=512, threshold=0.5
34
```

```
35
36 Process finished with exit code 0
37
```