

```
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: 9374.22 sec
7
```

```
8
9 Loading and preprocessing data (LSH) with threshold=0.8...
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...
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=0.2...
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	9374.217099	9374.217099	100%
31	LSH-Jaccard		3489.06	3.010000	3492.070000	2% Perm=64, threshold=0.8
32	LSH-Jaccard		3497.85	3.920000	3501.770000	6% Perm=64, threshold=0.5
33	LSH-Jaccard		3522.63	21.560000	3544.200000	0% Perm=64, threshold=0.2

File - apomakrismeno_test

35

36 Process finished with exit code 0

37