Approximately searching dataset for 2 nearest neighbors of the give test file:

+-----+------------+------------------+

|docID| file\_path| distCol|

+-----+------------+------------------+

| 1|sof\_doc2.txt|0.8715596330275229|

| 0|sof\_doc1.txt|0.8727272727272728|

+-----+------------+------------------+

so sof\_doc2 is the number 1 recommendation, and sof\_doc1 is the second.