README

Use the **make** command to compile all the C files.
It executes three commands:

gcc -fopenmp -o clusterSearch clusterSearch.c -O3
gcc -fopenmp -o preprocess preprocess.c -O3
gcc -fopenmp -o linearSearch linearSearch.c -O3

Please use **setenv OMP_NUM_THREADS 32** (This makes the code very fast as most of the code is parallelized)

To run the linear scan approach use this command:

./linearSearch input.txt query.txt permSize

input.txt → input data file
query.txt → query file
permSize → Permutation size (32)

(This has been the best implementation with pruning which on an average takes 3.5 secs / 10 queries)

To run the Clustering approach use these commands:

./preprocess input.txt numCluster permSize
./clusterSearch data.dat query.txt numCluster permSize

input.txt → input data file
query.txt → query file
permSize → Permutation size (32)
numClusters → number of clusters (Takes about 2 minutes for 20 clusters of size 32 permutation)
(Please use the same number of clusters for both commands or else the program will **segfault**)

(This takes on an average takes 6.5 secs / 10 queries)

The codes are tested to work for permutations size up to 256.
(Also please check that no one is executing something heavy on linprog, this will impact execution time.)