

# JAVA KD+-Tree Summary Report

Java KD+-Tree program is converted from C++ ST2I program.

Compared to C++program:

## **Strong points:**

- Can run in both Linux and Windows system directly.
- Easy to build index different datasets without changing original code.
- Can use to do the topK search and range search based on the simple parameter files

## **Weak points:**

- C++'s search speed is faster 3~5 times than Java version's speed. And C++'s IO output speed is faster 7~10 times than Java version.
- C++'s transferring Integer from original float will keep precision, however Java's transferring will lose precision since its transferring based on multiply 10000. And int Java, the float value cannot be greater than 210000+ because it will greater than Integer's max value.
- Java's default heap size is 2GB. When building index file for large datasets, please make sure java's heap size is re-set (like "-Xms8g") to make java's program can utilize enough memory.
- Java's first-time query is very slow and should be the max execute time showing in the report since CPU will take a while to cache the join parameter plan and execute the query. After the first-time query, it will be faster.