Directory Structure:-

1) Indexer.py
2) Search.py
3) FieldQueriesSearch.py
4) SearchPhase2.py
5) Readme.md

Optimizations used:-

1) Binary Search to find posinglist from index files corresponding to given word.

Improvements in time and space complexity:-

1) The time complexity reduced to logn instead of n where n is number of index files.

Index Creation Time and Size:-

1) Time taken:- 600-650seconds
2) Size:- 400.3MB

Format for final Index is as follows:-

Three things have been stored in the index:-

a) DocID to title map
   Format: key: DocID, value: TItle of Wikipedia page.
b) Word Posting List
   Format: key: Word, value: docId: t1i0c0b1r0 : tfidf of word in doc.
c) DocID to words List
   Format: key: DocId, value: words : tfidf