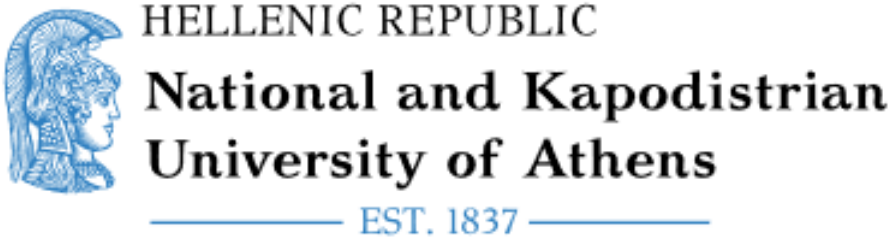
Opinion Mining
2018-2019

Author: Nemanja Nedic
A.M. 1115201400124



Contents

Intro

Compile & Run

Makefile is provided for compilation.

Run: ./recommendation -d input_fule -o out_file

After starting you will be asked but the program to provide (from stdin) missing or necessary info such as the path to lexicon,query file, configuration file, vectorized tweets file.

File Organisation and Documentation

You will find extensive comments regarding each class in header files and more comments on implementation in source files.

ErrorCodes.hpp: error codes used to handle irregular termination
myvector: a std::vector with a std::string id
CosineSimilarity: Class CosineSimilarity
Euclidean: Class Euclidean
Metric: Abstract Class Metric
HashTable: Class HashTable
ReadInput: Functions to parse arguments, config and read data sets
WriteOutput: Functions that write to outputfiles.
utility: Various functions and templates (ie Exhaustive Search, HammingNeighbors, Comparing vectors, Modulo operation, Vector Norms, progress3 bar, ...)

main.cpp: general outline of the program
Cluster: Class Cluster.
ClusterSpace: A set of clusters in the same vector space and implementations of clustering algorithms.
User: User class with his tweets and his crypto values. Here you can find user related functions that assign crypto values, such as RateByNNSimilarity or RateByClusterSimilarity.
Tweet: Tweet class with its score, tokens.