# Document similarity

## Defining a (StringSet) class

* It will store a set of CPP strings
* Create a constructor that takes a file name and load the words from it
  + Ignoring the punctuation marks
  + Turning the text to lower case
* Create a constructor that takes a string and take words from it
* Write member function to add a string to the set
* Write member function to remove a string from the set
* Write member function to clear the entire set
* Write member function to return the number of strings in the set
* Write member function to output all strings in the set
* Overload the + operator to return the union of two StringSet objects
* Overload the \* operator to return the intersection of two StringSet objects.
* Write a member function that computes the similarity between the current StringSet and an input parameter of type StringSet
  + Similarity = = {the size of the set of common words / sqrt of (size of Q \* size of D)}
* Write a program to test all these functions