BST vs AVL

This project is a programming assignment in C which aims to build two algorithms (BST and AVL) that will store/search/update/delete information related to textual materials and make performance experiments in order to evaluate your system.

Steps:

1. Read the Input1.txt.
2. Make necessary stop word removal and stemming operations.
3. Store your data into **BST**.
4. Delete the keys in the first sentence of Input1.txt from your BST and re-store your tree.
5. Record execution time in the 3^{rd} and 4^{th} steps.
6. Read the Input1.txt.
7. Make necessary stop word removal and stemming operations.
8. Store your data into **AVL tree**.
9. Delete the keys in the first sentence of Input1.txt from your AVL tree and re-store your tree.
10. Record execution time in the 8^{th} and 9^{th} steps.
11. Compare the results you get in 5^{th} and $10t^{h}$ steps.
12. Do the steps 1^{st}-11^{th} for Input1.txt.

Input1.txt

Text mining studies have gained importance in recent years because of the increasing number of electronic documents like news, social networks, research papers and digital libraries. There is no doubt that this enormous data continues to increase day by day with the contribution of lots of people. Automatically processing, organizing and handling this textual data are a central problem. The key aim of text mining is to allow users to get information from textual materials. Text mining mainly deals with several important applications like information retrieval (IR), classification (i.e., supervised, unsupervised and semi supervised classification), document filtering, summarization, sentiment or opinion classification. Natural Language Processing (NLP), Machine Learning (ML) and Data Mining methods work together to detect patterns from the different types of the documents and classify them in an automatic manner.

Input1.txt (there is no stopwords!!!)

text mining study gain importance recent year increase number electronic document like news social network research papers digital library doubt enormous datum continue increase day day contribution lot people automatically process organize handle textual data central problem key aim text mining allow user get information textual material text mining mainly deal several important application like information retrieval classification supervise unsupervised semi supervise classification document filter summarization sentiment opinion classification natural language processing nlp machine learn datum mining method work together detect pattern different type document classify automatic manner