ADVANCED PROGRAMMING
2015-16
Programming Assignment 3

The assignment involves a programming activity and includes an optional *problem*. The submission must include all required files for compiling and executing the proposed solution. It is allowed to develop the solution into teams, however the proposed solution should be discussed individually.

Introduction:

In this programming assignment, you will design and implement an application that analyzes and manipulates texts. The goal is to familiarize yourself with Python constructs. In class programming we designed a Python script that counts the frequency of words out of a text. In this assignment we will continue with our exploration of Python mechanism for text analysis.

The problem set (mandatory)

1. Design and implement a Python script which computes the distance among two texts. You can exploit the ***tf*idf*** distance.

2. Extend the Python script to determine the distinctive words within a text.

The problem set (optional)

1. Design and implement a Python script which computes a basic recommendation system. The system from five books that you enjoyed yourself checks whether a sixth book is similar to the other five.
2. Evaluate the implementation through suitable test cases.

Grading

The breakdown for the programming assignment is as follows:
 • Problem set (Mandatory): 60%
 • Problem set (Optional): 30%
 • Other factors (code quality): 10%

Submission

The AP homework should be submitted by e-mail by **10-01-2016 (midnight)** to gian-luigi.ferrari@unipi.it with the subject prefix **[AP-HW3]**.

The submission must include all required files for executing the proposed solution.