

ООП

Создайте класс, который рассчитывает TF-IDF (про TF-IDF можно прочитать здесь: <https://ru.wikipedia.org/wiki/TF-IDF>).

Некоторые замечания и пожелания:

1. Для расчёта IDF используйте любой корпус текстов;
2. Класс должен содержать публичный метод, который в качестве аргумента принимает текст пользователя, а возвращает - TF-IDF для каждого слова в тексте (например, список кортежей (*слово, tf-idf*));
3. При создании объекта класса значения IDF должны инициализироваться следующим образом:
 - a. Если файла со значениями IDF нет (т.е. вы ещё их не считали), то вычисляем эти значения, присваиваем их переменной класса и сохраняем в файл;
 - b. Если файл со значениями IDF уже есть (т.е. вы их уже вычислили), то в переменная класса содержит данные из этого файла. Это нужно для того, чтобы при каждом создании экземпляра класса, не нужно было вычислять эти значения, а можно было бы просто загрузить их из файла.
4. Продумайте, как будет выглядеть этот класс:
 - a. Какие переменные он содержит (возможно, какие-то из них приватные)?
 - b. Какие аргументы принимает конструктор?
 - c. Какие методы содержит класс: какие из них приватные, какие - публичные?
 - d. Может быть, этот класс является наследником другого класса?