



Министерство науки и высшего образования  
Российской Федерации

Федеральное государственное бюджетное образовательное  
учреждение высшего образования  
«Московский государственный технический университет  
имени Н. Э. Баумана (национальный  
исследовательский университет)»  
(МГТУ им. Н.Э. Баумана)

---

Факультет «Информатика и системы управления»

## ОТЧЕТ по лабораторной работе №4

по дисциплине

«Информационный поиск и извлечение информации из  
текстов»

Студент группы ИУ9-21М

\_\_\_\_\_ С.С. Погосян  
(подпись, дата)

Руководитель

\_\_\_\_\_ Н.В. Лукашевич  
(подпись, дата)

## 1. Постановка задачи

### Задача на дом

- Запрос: отбор кандидатов
- Пользователь отметил релевантными два документа
  - Кандидат отобрать претендент
  - Отбор выбрать претендент
- Объем коллекции – 1 млн. документов
- Df:
  - отбор 70000, кандидат – 70000,
  - Претендент - 30000, отобрать – 50000, выбрать 70000
- Как изменится запрос, если
  - $\alpha=0.7$  (коэффициент учета запроса),
  - $\beta=0.3$  (коэффициент учета релевантных документов),
  - Запрос представляется как вектор частот
  - Документ представляется как нормализованный вектор  $tf.idf$

Полный ход решения доступен по ссылке: [https://github.com/legion15q/sem2/blob/master/num4\\_02/%D0%9E%D1%82%D1%87%D0%B5%D1%82/vpQ0xfVWcw8.jpg](https://github.com/legion15q/sem2/blob/master/num4_02/%D0%9E%D1%82%D1%87%D0%B5%D1%82/vpQ0xfVWcw8.jpg)

## 2. Решение