



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 №7

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

Домашняя задача 1

- Система рубрикации должна классифицировать поток документов по двум рубрикам.
- Эксперт отнес к первой рубрике 75 документов, ко второй рубрике – 50 документов.
- Система отнесла:
 - - к первой рубрике 100 документов, из них 50 правильно.
 - - ко второй рубрике 40 документов, из них 30 правильно.
- Найти макро-характеристики качества классификации (точность, полноту, F-меру) - и микро-характеристики (точность, полноту, F-меру).

Домашняя задача 2

- Даны документы и их классы C1 и C2
- $D1=(X1, X2, X3)$ C1
- $D2=(X1, X2, X4)$ C1
- $D3=(X4, X5, X6)$ C2
- Определить класс документа на основе метода наивного Байеса
- $D4 (X1, X4, X5)$

2. Решение