

**МОСКОВСКИЙ ГОСУДАРСТВЕННЫЙ ТЕХНИЧЕСКИЙ
УНИВЕРСИТЕТ ИМЕНИ Н. Э. БАУМАНА**
Факультет информатики и систем управления
Кафедра теоретической информатики и компьютерных технологий

Лабораторная работа №12
по курсу «Информационный поиск»
«Микро и макроусреднение параметров»

Выполнил:
студент группы ИУ9-21М
Беляев А. В.

Проверила:
Лукашевич Н. В.

Москва 2019

Цель работы

Найти макро-характеристики качества классификации (точность, полноту, F-меру) и микро-характеристики в следующей ситуации:

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.

Ход работы

Составим таблицу:

	Эксперт (реально правильно)	Система выдала	Система выдала правильно	Precision	Recall	F-мера
Класс 1	75	100	50	$50/100 = 0.5$	$50/75 = 0.67$	0.57
Класс 2	50	40	30	$30/40 = 0.75$	$30/50 = 0.6$	0.67
Микро	125	140	80	$80/140 = 0.57$	$80/125 = 0.64$	0.6
Макро	-	-	-	$(0.5+0.75)/2 = 0.625$	$(0.6+0.67)/2 = 0.635$	$(0.57+0.67)/2 = 0.62$

Точность - сколько правильных среди выданных

Полнота - сколько выданных среди правильных

F-мера:

$$F_1 = \frac{1}{\frac{1}{2} \frac{1}{P} + \frac{1}{2} \frac{1}{R}} = \frac{2PR}{P + R}$$