

Результаты наблюдений

Сухочев Александр и Балакший Андрей

3 мая 2015 г.

Гляньте на эту табличку!

Logistic Regression with 1 0

Accuracy: 0.663265306122

All test count: 98; TP: 7; TN: 58; FP: 10; FN: 23

и

Logistic Regression with tf idf

Accuracy: 0.704081632653

All test count: 98; TP: 1; TN: 68; FP: 0; FN: 29

С Тф идф выдает лучший результат, но посмотрите на ТП! Просто так совпало что у нас больше негативных в тестирующем множестве, поэтому мы не можем сказать что тф идф лучше работает, видно ведь что он хуже определяет Позитивные! Переделывать обучающее множество нельзя, т.е. у нас и при работе программы будет подобное соотношение позитивных/негативных, но спасибо за полезный опыт

1 Таблицы точностей различных машинных обучений и экстракторов.

1.1 Общая таблица

	SE	N(only)-gramm	N-gramm E	SE with not
MNB				
GNB				
SVM 1 0				
SVM tf-idf				
LogReg				

1.2 Более подробно

1.2.1 Standart Extractor

	Correct	Total	Acc	TP	TN	FP	FN
MNB	742	1115	0.6655	61	618	51	322
GNB							
SVM 1 0							
SVM tf-idf							