

ИССЛЕДОВАНИЕ МОДЕЛИ АНАЛИЗА НАСТРОЕНИЙ ПОЛЬЗОВАТЕЛЕЙ СЕТИ TWITTER ОТНОСИТЕЛЬНО ТЕРРОРИСТИЧЕСКИХ АТАК.

Демешко А. (Национальный исследовательский университет ИТМО)

Воробьева Алиса Андреевна, кандидат технических наук, факультет безопасности информационных технологий, доцент
(Национальный исследовательский университет ИТМО)

В данной работе предлагается модель автоматической классификации отзывов пользователей в Твиттере после совершения теракта, модель построена с использованием подходов лексики и машинного обучения.

Введение. Твиттер считается одним из известных микроблогов, привлекающих политиков и отдельных лиц для выражения своего мнения по политическим, экономическим и социальным вопросам. Феномен террористических операций является одной из крупнейших проблем безопасности и экономики, стоящих перед миром в последние годы. Комментарии пользователей Твиттера по вопросам терроризма важны для понимания их отношения к террористическим актам. Анализ настроений — это область исследований для понимания и извлечения мнений пользователей.

Основная часть. В работе был использован подход Lexicon для создания помеченного обучающего набора данных, а для построения модели использовался подход машинного обучения. Десятки некоторых слов, связанных с предметной областью, были нейтрализованы, чтобы избежать их негативного эффекта. Функции были выбраны на основе PoS. Применялось мажоритарное голосование между алгоритмами классификации машинного обучения NB, SVM и LR. Производительность алгоритмов классификации измерялась с использованием показателей точности и F1. Полученные результаты сравниваются для определения наилучшего алгоритма классификации для выбора признаков. Результаты показывают, что наша модель достигла точности 94,8% с показателем F1 95,9%.

Выводы. Наша модель достигла высокой производительности с точностью до 94,8% и показателем F1 95,9%. Соответственно, комбинация функций дает лучшие результаты, чем функция одного типа. Это исследование будет полезно для правоохранительных органов и спецслужб в их стремлении разработать эффективные стратегии сдерживания для предотвращения террористических атак в будущем. В дальнейшей работе функции будут

извлекаться на основе N-грамм, а полученные результаты будут сравниваться с функциями на основе PoS.

Демешко А.А.(автор)

Подпись

Воробьева А.А. (научный руководитель)

Подпись