

Изотова Е.Д. Интеллектуальное извлечение эмоционального отклика потребителей лекарственных средств из отзывов на естественном языке // Учен. зап. Казан. ун-та. Сер. Естеств. науки. – 2014. – Т. 156, кн. 4. – С. 52–59.

УДК 519.767.6+615.036.2

ИНТЕЛЛЕКТУАЛЬНОЕ ИЗВЛЕЧЕНИЕ ЭМОЦИОНАЛЬНОГО ОТКЛИКА ПОТРЕБИТЕЛЕЙ ЛЕКАРСТВЕННЫХ СРЕДСТВ ИЗ ОТЗЫВОВ НА ЕСТЕСТВЕННОМ ЯЗЫКЕ

Е.Д. Изотова

Аннотация

В статье рассмотрены подходы алгоритма машинного обучения с учителем с целью извлечения эмоциональных откликов пользователей о лекарственных средствах (ЛС) из отзывов на естественной языке. Произведена оценка сбалансированности классов отзывов (положительных/отрицательных/нейтральных) в сети Интернет среди потребителей ЛС. Подобрано оптимальное соотношение классов в обучающей выборке для обучения классификатора. Составлен комплекс правил подготовки текста отзыва для построения вектора: перевод всего текста в нижний регистр, исключение стоп-слов (местоимений, частиц, местоимений), обработка отрицания в предложении. Составлены словари униграмм: частотный словарь и словарь по методу взаимной информации. Приведены результаты работы классификатора по каждой комбинации изучаемых параметров системы классификации эмоциональных откликов пользователей.

Ключевые слова: лекарственные средства, машинное обучение, обучение с учителем, нейронные сети, тональность текста, «сумка слов», словарь униграмм, метод взаимной информации, естественный язык, русский язык.

Литература

1. *Зырянов С.К., Белоусов Ю.Б.* Организация и развитие службы фармаконадзора // Фарматека. – 2005. – № 16. – URL: <http://www.pharmateca.ru/ru/archive/article/6317>, свободный.
2. *Кукес В.Г.* Клиническая фармакология. – М.: ГЭОТАР-Медиа, 2009. – 1052 с.
3. *Fletcher A.P.* Spontaneous adverse drug reaction reporting vs event monitoring: a comparison // J. R. Soc. Med. – 1991. – V. 84, No 6. – P. 341–344.
4. *Астахова А.В., Брайцева Е.В., Лепехин Е.В.* Контроль безопасности лекарственных средств // Фармация. – 2000. – № 4. – С. 92–96.
5. *Фитилев С.Б.* Служба безопасности лекарств в Российской Федерации // Новая аптека. – 1998. – № 2. – С. 13–19.
6. *Мельчук И.А.* Опыт теории лингвистических моделей «смысл ⇔ текст»: семантика, синтаксис. – М.: Языки рус. культуры, 1999. – 211 с.

7. *Pang B., Lee L., Vaithyanathan S.* Thumbs up sentiment classification using machine learning techniques // Proc. Conf. on Empirical Methods in Natural Language Processing (EMNLP). – 2002. – V. 5. – P. 79–86. – doi: 10.3115/1118693.1118704.
8. *Wiebe J., Riloff E.* Creating subjective and objective sentence classifiers from unannotated texts // Proc. 6th Int. Conf. on Computational Linguistics and Intelligent Text Processing. – 2005. – P. 486–497. – doi: 10.1007/978-3-540-30586-6_53.
9. *Steven M.* On understanding and classifying web queries: Ph.D. Thesis. – Chicago, IL, USA: Illinois Institute of Technology, 2006. – 96 p.
10. *Руднева А.Б.* ГИА. Русский язык. Универсальный справочник. – М.: Яуза-Пресс, 2014. – 208 с.
11. *Kennedy A., Inkpen D.* Sentiment classification of movie reviews using contextual valence shifters // Computational Intelligence. – 2006. – V. 22, No 2. – P. 110–125.

Поступила в редакцию
28.08.14

Изотова Екатерина Дмитриевна – ассистент кафедры биохимии и биотехнологии,
Казанский (Приволжский) федеральный университет, г. Казань, Россия.
E-mail: *izotova.e.d@gmail.com*