

## О ПРОБЛЕМЕ ОПРЕДЕЛЕНИЯ ЭМОТИВНОЙ ОЦЕНКИ ТЕКСТА

Куличок Н.С.

Белорусский государственный университет информатики и радиоэлектроники  
г. Минск, Республика Беларусь

Парамонов А.И. – канд. техн. наук, доцент

В данной работе рассматриваются вопросы анализа эмотивности цифровых текстов на естественном языке. Представлены основные понятия, обозначены проблемы и описаны известные пути их решения.

Постоянный рост объемов текстовой информации в цифровом виде приводит к возникновению потребности в новых инструментах для ее анализа. На сегодня достаточно эффективно решаются задачи классификации текстов и автоматического синтаксического анализа. Известны подходы по семантическому анализу текстов. На современном этапе развития автоматической обработки текстов актуальной и востребованной становится задача эмотивной оценки текстов. Анализ эмотивности является достаточно сложной задачей [1], что связано с трудностями при выделении нужной эмоциональной лексики в текстах, а также с определением самого эмотивного пространства, количества и состава его измерений. Ежедневно множество данных генерируется пользователями соцсетей, новостными порталами, различными блогами. Анализ эмотивности также требуется для мониторинговых, аналитических и сигнальных систем, для систем документооборота и рекламных платформ, таргетированных по тематике веб-страниц и многих других. Область применения анализа эмотивности текстов обширна.

Эмотивность определяется как свойство языка выражать психологические (эмоциональные) состояния и переживания человека. В речи она реализуется с помощью специальных языковых средств — эмотивов, которые представлены на каждом уровне любого естественного языка и образуют его эмотивный код (ЭК). Каждый текст имеет свой ЭК. ЭК языка — семиотическая система корреляций между психическими состояниями (эмоциями) коммуникантов и их концептуализацией. Существует разные характеристики эмотивности: эмотивный фон, эмотивная тональность, эмотивная окраска, эмотивная направленность, эмотивная модальность, эмотивные интенции текста [2].

Наибольшего успеха исследования эмотивности текста достигли при анализе его тональности. Анализ тональности - класс методов контент-анализа в компьютерной лингвистике, предназначенный для автоматизированного выявления в текстах эмоционально окрашенной лексики и эмоциональной оценки авторов (мнений) по отношению к объектам, речь о которых идет в тексте [3]. Анализ тональности играет важную роль в принятии решений и в системе рекомендаций. Например, перед покупкой продуктов людей интересуют отзывы других покупателей о данном продукте, при прочтении новостей – мнение других людей об описываемых событиях. Однако современные объемы информации не позволяют самостоятельно оценивать все тексты целиком. Анализ тональности упрощает эту задачу, поскольку описывает полярность текста, так что пользователь может напрямую узнать, является ли данный текст положительным или отрицательным, не читая его целиком.

Для определения настроения в анализе тональности используются три термина: объект, о котором дается мнение, особенности этого объекта, а также владелец мнения об объекте [4]. Анализ тональности выполняет задачу классификации в 3 этапа: на уровне документа, уровне предложения и уровне характеристик. Уровень классификации документа используется там, где задача состоит в том, чтобы найти общую полярность темы независимо от того, кто придерживается мнения. Классификация по уровням предложений предполагает, что каждое предложение придерживается единого мнения. На уровне характеристик выполняется анализ различных характеристик объекта. Анализ тональности включает в себя предварительную обработку данных, выбор характеристик и классификацию, а затем определение полярности данных. Предварительная обработка данных включает в себя токенизацию, удаление стоп-слов, выделение корней, лемматизацию и т. д.

Методы анализа тональности делятся на 2 большие группы: основанные на машинном обучении и основанные на правилах и словарях.

Решение проблемы анализа тональности текста позволит эффективнее обрабатывать цифровые данные в автоматизированном режиме с целью формирования эмоционального контекста материалов или выявления несоответствующих общему фону тестов.

### Список использованных источников:

1. Пазельская А., Соловьев А. Метод определения эмоций в текстах на русском языке // *The international conference on computational linguistics and intellectual technologies "Dialogue 2011"*: конференция. — Москва, 2011. — С. 510–522.
2. Ленько Галина Николаевна Уровни анализа текстовой эмотивности (на примере текстов художественного стиля) // *Вестник ЛГУ им. А.С. Пушкина*. 2014. №2. — 10 с.
3. Гербик, А. И. Анализ тональности текста / А. И. Гербик, П. Е. Дорошкевич, А. И. Свито // *Компьютерные системы и сети: материалы 53-й СНТК*. — Минск: БГУИР, 2017. — С. 167–168.
4. Алгоритмы анализа тональности текста / Н. С. Иванин и другие // *BIG DATA and Advanced Analytics: collection of materials of the third international scientific and practical conference. (Minsk, Belarus, May 3 – 4, 2017) / editorial board : M. Batura [et al.]*. — Минск : BSUIR, 2017. — С. 150-154.