

МЕДВЕДНИКОВА М.А., студентка кафедры ИЗИ ВлГУ

КОСТИНА Н.В., студентка кафедры ИЗИ ВлГУ

К ЗАДАЧЕ АВТОМАТИЗАЦИИ ОЦЕНКИ ПСИХОСЕМАНТИЧЕСКОГО ПРОФИЛЯ ПОЛЬЗОВАТЕЛЯ СОЦИАЛЬНОЙ СЕТИ

Для исследования психосемантического профиля пользователя социальной сети была поставлена задача - создать методику оценки психосемантического профиля пользователя социальной сети и последующей ее автоматизации. В данной работе авторы предлагают использовать для оценки профиля такой параметр, как «эмоциональное шкалирование»

2 источник

В настоящее время актуальным вопросом является распространение пропаганды в Интернете, в частности в социальных сетях. Таким образом, ориентирование информации на потребителя становится важной задачей. Соответственно, возникает необходимость оценки психосемантического профиля узла социальной сети, так как это позволяет выявить особенности пользователя исходя из публикуемых сообщений.

Задача исследования заключается в создании методики оценки психосемантического профиля пользователя социальной сети и последующей ее автоматизации.

Объектом анализа являются микроблоги узла социальной сети.

Предмет исследования – методы автоматизированной оценки и построения психосемантического профиля пользователя социальной сети.

Целью исследования является повышение эффективности предоставления контекстной рекламы, агитационных и прочих материалов, которые будут коррелировать с интересами пользователя.

Результаты таких исследований помогут координировать социальные работы (например, агитационные, продвигаемые бизнес-идеи и т.д.) в социальных сетях.

Предлагается оценивать профиль узла социальной сети, исходя из следующих принципов:

- принцип текстоцентричности;
- принцип объективности.

Первый принцип означает, что данные будут выявляться только на основе публикуемых узлов сообщений узлом, без учета психологических характеристик пользователя. Вводя принцип объективности, авторы стараются минимизировать роль экспертных оценок.

Для создания методики необходимо выявить параметры психосемантического профиля узла социальной сети. В данной статье описана методика проведения эксперимента для оценки параметра «эмоциональное шкалирование».

Авторами был предложен такой статистический параметр, как «эмоциональное шкалирование». Он позволяет оценить эмоциональный окрас текстов. Рассчитывается по следующему алгоритму:

1. Задается некоторое количество шкал, которые будут служить для оценки текста. Для каждой шкалы выбирается одно или несколько слов, наиболее полно ее описывающих. Например, для шкалы «радость» подходящим выбором выглядят слова «радость» или «наслаждение».

2. Для каждого слова текста вычисляется набор индексов PMI-IR, формула которых в общем виде выглядит так:

$$PMI(\text{слово}, \text{шкала}) = \log_2(f(\text{слово}, \text{шкала}) / f(\text{слово}) \wedge f(\text{шкала})), \quad (1)$$

где f – количество результатов в поисковой системе по заданному запросу «шкала» – поисковый запрос по выбранной нами шкале (слова шкалы соединяются логическим оператором «или»), «слово» – поисковый запрос по выбранному нам слову. Для увеличения точности работы из запросов также убираются все результаты отрицанием слов запроса. Полученный индекс отражает то, насколько часто люди употребляют искомое слово и любое из слов шкалы вместе, относительно того, насколько часто они употребляют их по отдельности.

3. Индекс для текста представляет собой усредненное значение индексов слов этого текста (с учетом частоты их употребления в нем).

Эксперимент для выявления ассоциативного ряда проводился следующим образом: испытуемым были предложены шкалы, которые, по мнению авторов, наиболее необходимы. а именно, вдохновение, радость, равнодушие, страх, недовольство, злость, неясность, расстройство. Предлагалось в течение 2х минут указать слова, которыми вы в социальной сети выражаете такое чувство как равнодушие:

Опрашиваемые имели право в качестве ассоциативного ряда вписывать слова, словосочетания и символы,

«Эмоциональное шкалирование» даст представление об эмоциональной структуре текстов, публикуемых пользователем, и, соответственно, об эмоциональном шкалировании текстов, лучше воспринимаемых пользователем. В будущей реализации автоматизированной оценки эмоциональности текста этот процесс будет происходить наоборот. Зная необходимые эмоциональные показатели текста для конкретного пользователя, необходимо будет сгенерировать сообщение заданного содержания с указанными показателями эмоциональности.

Технически формы для ответов эксперимента были созданы в двух вариантах – печатном и электронном. Печатный шаблон подготовлен в MS Word, электронный – сделан средствами Google. Распечатанные анкеты были розданы студентам во время занятий, ссылка на электронный вариант и краткая информация о ней – размещены в социальной сети ВКонтакте.

Анкета также содержала информацию со статистическими данными: возраст, образование, пол, семейное положение, которая тоже может понадобится для дальнейшего анализа.

На данный момент в общей сложности оба типа анкет прошли более 50 человек. Для обработки и статистического анализа данных недостаточно, поэтому сбор информации пока продолжается. Информация из печатных анкет переводится в удобный для последующей обработки электронный вид.

1. Четвериков А.А. (СПбГУ). Дифференцированное измерение эмоциональности текстов с помощью алгоритма PMI-IR (Статья из сборника «Материалы XXXVIII Международной филологической конференции 11–13 марта 2009 г.» Психолингвистика. Часть 1 / Отв. редактор Т. В. Черниговская. – СПб.: Факультет филологии и искусств СПбГУ, 2009. – 53 с.)

2. Костина Н.В., Медведникова М.А. Аналитический обзор методов психосемантики и когнитивистики на предмет их применимости для анализа микроблогов узла социальной сети [Электронный ресурс]: <http://izi.vlsu.ru/НТС/16.pdf> 2012 г.