

Извлечение аспектов, тональности и категорий аспектов на основании отзывов пользователей о ресурсах и автомобилях

Иванов В. В. (nomemm@gmail.com),
Тутубалина Е. В. (tutubalinaev@gmail.com),
Минг азов Н. Р. (nicrotek547@gmail.com),
Алимова И. К. (alimovaIlseyar@gmail.com)

Казанский Федеральный Университет, Казань, Россия

Ключевые слова: анализ тональности текстов, SentiRuEval, отзывы пользователей,
извлечение аспектов, категории аспектов

Извлечение аспектов, настроений и категорий аспектов в пользователе Отзывы о ресурсах и автомобилях

Иванов В.В. (nomemm@gmail.com),
Тутубалина Е.В. (tutubalinaev@gmail.com),
Минг азов Н.Р. (nicrotek547@gmail.com),
Алимова И.С. (alimovaIlseyar@gmail.com)

Казанский федеральный университет, Казань, Россия

В данной статье описывается метод решения задач анализа тональности на основе аспектов в предметных областях отзывов о ресурсах и автомобилях. Эти задачи были сформулированы в инициативе «Оценки настроений для России» (SentiRuEval-2015). Во время SentiRuEval-2015 мы сосредоточились на трех подзадачах: извлечение явных терминов аспектов из отзывов пользователей (задача A), классификация настроений на основе аспектов (задача C), а также автоматическая категоризация аспектов (задача D).

В классификации настроений на основе аспектов (задачи C и D) мы предлагаем два контролируемых метода, основанных на модели максимальной энтропии и машинах опорных векторов (SVM), соответственно, которые используют набор характеристик, состоящий из терминов в контексте термина аспекта и особенностей на основе лексикона. В задаче C мы достигли 40% макроусредненной F-меры для автомобилей и 40,05% для отзывов о ресурсах. Мы достигли 65,2% макроусредненной F-меры для автомобилей и 86,5% для отзывов о ресурсах в задаче D. Этот метод занял первое место среди 4 команд в обеих предметных областях. Классификатор SVM основан на характеристиках унитарной и поточечной взаимной информации для оценки, специфичной для категории, и связывания каждого аспекта с соответствующей категорией в предметной области.

Иванов В.В., Тутубалина Е.В., Минг-зов Н.Р., Алимова И.С.

В задаче А мы тщательно оценили эффективность метода, основанного на статистических характеристиках, включенных в модель случайных полей. К сожалению, этот метод не показал значительного улучшения по сравнению с их одним уровнем. Однако его результаты также представлены в статье.

Ключевые слова: анализ настроений на основе аспектов, сентиментальность, отзывы пользователей, извлечение аспектов, категоризация аспектов.

1. Введение

За последнее десятилетие анализ мнений (также называемый анализом настроений) был важной задачей обработки естественного языка (NLP). Поскольку онлайн-обзоры значительно влияют на решения людей о покупках, идентификация настроений имеет ряд применений, включая отслеживание мнений людей о фильмах, книгах, продуктах и т. д.

В данном исследовании мы описываем наш подход к решению задачи по анализу настроений, которая была сформулирована в качестве отдельного трека в инициативе «Оценки настроений для русского языка» (SentiRuEval-2015). Задача SentiRuEval касается анализа настроений пользователей в отзывах о ресторанах и автомобилях. Задача состоит из нескольких подзадач: выделение аспектов (задачи А и В), классификация тональности аспектов (задача С) и определение категорий аспектов и обобщение тональности обзора (задачи D и E). Основная цель задачи SentiRuEval — найти слова и выражения, обозначающие важные аспекты ресторана или автомобиля на основе мнений пользователей, и классифицировать их по классам полярности и категориям аспектов (Loukachevitch et al., 2015).

Было проведено большое количество исследований в области анализа настроений на основе аспектов, которые хорошо описаны в Liu (2012) и Pand and Lee (2008). Традиционные подходы к анализу мнений основаны на извлечении частот встречающихся фраз, содержащих прилагательные, извращенные славяры (Turney, 2002; Popescu and Etzioni, 2007). В современных статьях реализованы вероятностные тематические модели, такие как скрытое распределение Дирихле (LDA) и случайное поле (CRF) для задач многоаспектного анализа (Moghaddam and Ester, 2012; Choi and Cardie, 2010). Анализ настроений в английском языке был тщательно изучен, и существует множество хорошо зарекомендовавших себя методов и универсальных словарей настроений, содержащих несколько тысяч терминов. Однако исследования по анализу настроений в русском языке были менее успешными. В 2011–2013 гг. исследования были сосредоточены на решении задачи анализа настроений в хэштегах анализа настроений ROMIP (Четверкин, Лукачевич, 2013; Котельников, Клековкина, 2012; Блинов и др., 2013; Фролов и др., 2013).

Мы используем модель случайных полей, применяемую к задаче извлечения аспектов. В задаче С для классификации настроений на основе аспектов мы предлагаем метод, основанный на модели максимальной энтропии, которая использует набор характеристик частоты термина в контексте термина аспекта и признаков на основе лексики. Классификатор для определения категорий аспектов основан на модели SVM с набором функций, специфичных для категорий. В задаче С мы достигли 40% макроусредненной F-меры для автомобилей и 40,05% для отзывов о ресторанах. Мы достигли 65,2% F-меры для автомобилей и 86,5% для отзывов о ресторанах в задаче D.

Оставшаяся часть теста организована следующим образом. В Разделе 2 мы представляем соответствующую работу по анализу настроений. В разделе 3 мы описываем предлагаемые подходы. В разделе 4 представлены результаты экспериментов. Наконец, в разделе 5 мы обсудим результаты.

2. Соответствующая работа

В этой статье мы сосредоточимся на обнаружении трех основных элементов обзора: аспектных терминов, настроений по поводу этих аспектов и категорий аспектов. В течение последнего десятилетия было предложено большое количество методов идентификации этих элементов.

Извлечение термина аспекта. Существует несколько широко используемых методов, рассматривающих задачу как задачу классификации (Попеску и др., 2005), как задачу маркировки последовательностей (Якоб, Гуревич, 2010; Кириченко и др., 2014; Чернышев, 2014), как тему моделирование или традиционную задачу классификации (Moghaddam and Ester, 2012; Zhao et al., 2014). Проблема классификации состоит в том, чтобы определить, являются ли существительные и именные словосочетания объектом мнения или нет. Попеску и др. (2005) использовали интаксические паттерны, связанные с тональностью, из лексиконов общего назначения, чтобы идентифицировать высказанные словосочетания. Пория и др. (2014) предложили подход, основанный на правилах, основанный на знаниях и деревьях зависимостей предложений. Эти подходы ограничены из-за более низких результатов при извлечении низкочастотных аспектов или созданных вручную правил зависимостей для ложного извлечения. В работах (Кириченко и др., 2014; Чернышев, 2014) авторы предложили две модификации стандартной схемы моделей маркировки последовательностей.

Полнота термина аспекта. Большинство ранних подходов к классификации аспектов основывались на начальных словах или сгенерированном вручную лексиконе, который содержит сильно положительные или сильно отрицательные слова. Терни (2002) предложил неконтролируемый метод, основанный на оценке тональности каждой фразы, которая рассматривается как взаимная информация между фразой и двумя начальными словами. В последних работах широко применяются методы машинного обучения для решения задачи классификации настроений (Pang et al., 2002; Pang, Lee, 2008; Блинов и др., 2013; Кириченко и др., 2014). Moghaddam and Ester (2012) предложили расширение модели LDA для извлечения аспектов и их оценок настроений путем рассмотрения зависимости между аспектами и их полнотой настроений. Тем не менее, тематические модели обеспечивают более низкую производительность при многоаспектной классификации предложений, чем классификатор SVM в трех разных областях (Lu et al., 2011).

Определение категории аспекта. Автоматическая категоризация эксплицитных аспектов в категории аспектов изучалась как задача суммирования настроений. Moghaddam and Ester (2012) использовали это как часть скрытой проблемы добычи полезных ископаемых. Были проведены работы по группировке аспектных терминов из обзорных текстов для анализа тональности в задании 4 международного семинара по семантической оценке (SemEval-2014). Задача оценивалась с помощью F-меры, и наилучшие результаты были получены с помощью SVM-классификатор с набором признаков и информацией из размеченных обзоров (Pontiki et al., 2014; Кириченко и др., 2014).

Было проведено несколько исследований по анализу настроений на русском языке, связанных с оценочными событиями российских систем анализа настроений (Четворкин и Лукачев, 2013). Фролов и др. (2013) предложили основанный на логике подход с семантическими фильтрами фактов для анализа настроений пользователей в отзывах о книгах. Блинов и др.

Иванов В.В., Тутубалина Е.В., Минг азов Н.Р., Алимова И.С.

(2013) показали преимущества метода машинного обучения по сравнению с лексическим подходом для отзывов пользователей на русском языке и использовали ручные эмоциональные словари.

3. Описание системы

В этом разделе мы опишем наши подходы к трем задачам спектроанализа настроений пользователей о ресторанах и автомобилях. Модель CRF использовалась для автоматического извлечения явных аспектов (задача A). Мы применили подходы машинного обучения для задач C и D, основанные на модели мешка слов и наборе функций на основе лексик, которые описаны в разделах 3.2 и 3.3 соответственно. Для нормализации текста на этапе preprocessing использовался морфосинтаксический анализатор Mystem.

3.1. Извлечение аспекта

Целью извлечения аспектов является выявление извлечения новых явных аспектов продукта (задача A). Поскольку задачу можно рассматривать как частный случай задачи маркировки предложений, мы используем случайные поля (Lafferty et al., 2001).

Эксплицитные аспекты обозначают некоторые части или характеристики описываемого объекта, такие как передний привод (передний привод), руль (руль), динамика (динамика) в обзорах автомобилей; стол (стол), официант (официант), блюдо (блюдо) в отзывах о ресторанах. В следующих примерах мы рассмотрим пользовательские фразы о явных аспектах.

Для обучения CRF мы используем схему Inside-Outside-Begin и пасивно-активный алгоритм краткое описание признаков, используемых для представления текущего маркера w_i , представлено ниже: текущий маркер w_i , текущий маркер w_i в пределах окна (w_{i-2}, \dots, w_{i+2}); теги части речи текущего токена; теги части речи маркера в окне $tag_{i-2}, \dots, tag_{i+2}$; количество ожиданий маркеров в обучающем наборе; наличие токена в созданных вручную доменно-зависимых словарях.

3.2. Классификация настроений на основе аспектов

Задача классификации настроений направлена на прогнозирование полярности (положительной, отрицательной, нейтральной или обеих) каждого аспекта отзывов о продукте. Мы применили классификатор максимальной энтропии с параметрами по умолчанию, основанный на модели мешка слов и наборе функций на основе лексик, описанных в разделе 3.2.2.

Следующие примеры иллюстрируют аспекты (выделенные курсивом) с различными полярностями из обзоров. Некоторые фразы типа «персонал улыбающийся, приветливый». («улыбчивый, приветливый персонал»), «общее впечатление: отличная машина» («общее впечатление: отличная машина») или «просторный салон, удобно сидеть пассажиру сзади» («просторный салон, пассажир мог удобно расположиться за водителем») содержат ильный положительный или отрицательный контекст рядом с термином аспекта. Поэтому такие случаи могли

Извлечение аспектов, настроений и категорий аспектов в отзывах пользователей

правильно классифицировать извлечение биграмм во фразах. Комплексный анализ чувственных фраз, таких как «заказывал бифштекс, нет слов как вкусно» («Я заказал бифштекс, нет слов, чтобы описать, насколько это было вкусно») и «в городе компьютер будут отображаться очень неприятные цифры» («в городе компьютер будет показывать очень неприятные цифры») показывает, что существует дисбаланс между полярными словами вкусно (вкусно), неприятные (неприятно) и терминами аспекта.

Мы используем комбинации термина аспекта и термина контекста, чтобы классифицировать эти случаи. Сложные фразы с обоими оттенками, такие как «отмечаю жесткую жесткость сидений, но привыкаешь, главное сидеть удобно» («Отмечаю некоторую жесткость сидений, но к этому привыкаешь, главное сидеть удобно») или «горячее неплохое», но нагиль было непохоже» («горячие блюда неплохие, но нагиль не похожи») можно узнать по наличию сюзанно (но).

Учитывая контекст термина аспекта, для извлечения признаков генерирую тсбиграммы слов двух типов: (i) биграммы контекста с использованием текста в окне контекста термина аспекта; (ii) биграммы на основе аспектов как комбинация аспектного термина и контекстного слова в контекстном окне. Окно контекста термина аспекта w_i обозначает последовательность (w_{i-4}, \dots, w_{i+4}).

3.2.1. Словарь настроений, созданный вручную Мы

собрали отзывы пользователей с сайта otzovik.com: 7 526 отзывов о ресторанах и 4 952 отзыва об автомобилях. Чтобы сделать корпус более точным, мы включили только положительные отзывы с общим рейтингом 5 в положительный корпус и отрицательные отзывы с общим рейтингом 1 или 2 в отрицательный корпус. Плюсы (Преимущества) и Минусы (Недостатки) — это части отзыва, описывающие веские причины, по которым автору отзыва нравится или не нравится продукт с ответственно. Для каждой области мы отобрали лучшие Наречий, прилагательных, глаголов, окрасящие существительные слова, которые выражают аспекты, глаголы действия наиболее распространенные прилагательные. Созданный вручную словарь состоит примерно из 741 положительного и 362 отрицательных слов в домене ресторанов и включает 1576 положительных и 741 отрицательных слов в домене автомобилей. Мы объединяем два словаря для достижения лучших результатов оценки.

Для функций, основанных на лексике, мы используем следующие оценки: каждое слово в sen значение взвешивается его расстоянием от данного аспекта:

$$оценка(w) = \frac{sc(w)}{e^{|i-j|}}$$

где i, j - позиции термина аспекта и слова, $sc(w)$ - оценка словотональности, равная 1 для положительных слов и -1 для отрицательных слов, извлеченная из словаря тональности.

3.2.2. Классификационные признаки для полярности терминов аспекта

Каждый обзор представлен в виде вектора признаков, для каждого аспекта признаки извлекаются из аспекта и его контекста в предложении. Краткое описание функций, которые мы используем, представлено ниже:

- символьные n -граммы: для отбора признаков расматривались трочные n -граммы символов для $n=2, \dots, 4$ с частотой документа больше двух.

Иванов В.В., Тутубалина Е.В., Мингазов Н.Р., Алимова И.С.

- униграммы на основе лексикона: извлекаются униграммы из лексикона настраиваний для выборок характеристик.
- контекстные n-граммы: униграммы (отдельные слова) и биграммы извлекаются из контекстного окна. Мы извлекаем эти n-граммы для нескольких комбинаций: (i) замена термина аспекта с ловеком аспект; (ii) замена слов настраивания с ловеком полярности pos или neg; (iii) замена тональных слов тегами части речи. • биграммы на основе аспектов: биграммы, сгенерированные как комбинация амотермина аспекта и слова в контекстном

окне. Мы извлекаем эти биграммы для нескольких комбинаций, описанных выше.

- признаки на основе лексикона: признаки рассчитываются следующим образом: максимальная оценка тональности; минимальный балл тональности; сумма оценок тональности слов; сумма оценок положительных слов; сумма оценок отрицательных слов. Слова настраивания отрицанием двигаются от оценок настраивания в сторону противоположной полярности.

В связи с ограниченным размером контекстного окна и сложностью классификации аспекта как отрицательным, так и с положительным отношением к его термину, мы создаем специальное правило для таких случаев: если предложение (предложения) содержит термин аспекта, слово-очетание но, а (но) и классификатор предсказывает нейтральную метку для аспекта, мы помечаем аспект обеими метками.

3.3. Автоматическая категоризация явных аспектов в категоризации аспектов

Цель задачи D состоит в том, чтобы отнести каждый аспект к одной из predetermined категорий. В отзывах о ресторанах присутствуют следующие категории аспектов: еда, обслуживание, интерьер, цена, общее. Для автомобилей аспектными категориями являются управляемость, надежность, безопасность, внешний вид, комфорт, стоимость, общее.

Опишем задачу автоматической категоризации явных аспектов на следующих примерах. Некоторые аспекты, такие как пищевые продукты (например, бифштекс (бифштекс), утка по-пекински (утка по-пекински)) или компоненты автомобиля (например, гидросистема (гидросистема руля), двигатель (двигатель)) классифицируются явным знанием человека-аннотатора.

Категории пищевых продуктов и автокомпонентов – это продукты питания и транспортные средства соответственно. Метка категории некоторых явных аспектов зависит от контекста отзыва пользователя. В примерах «машина с кошельком и отработала полностью» («машина с кошельком и кошельком»), «пробовал откусывать руль машина едет ровно» («поэкспериментировал с ведущим колесом и машина работает ровно»), «машина с вмятинами для фанатов.» («автомобиль предназначен для любителей») и «довольно красивая машина» («довольно красивая машина») категории аспектного термина машина (автомобиль) – стоимость, управляемость, целое, внешний вид соответственно.

Мы рассмотрели задачу как проблему классификации текста и обучили классификатор SVM по словесательной минимальной оптимизации (SMO). Для каждого термина аспекта w_i мы извлекли амотермин аспекта и признаки из окна контекста (w_{i-2}, \dots, w_{i+2}). Лексикона для конкретных категорий основаны на баллах за каждый термин w в тренировочном тесте:

Извлечение аспектов, настроений и категорий аспектов в отзывах пользователей

оценка(ж) = PMI (ж, кошка) - PMI (ж, прочее)

где PMI — точечная взаимная информация cat обозначает контексты всех аспектов в конкретной категории, oth обозначает контексты аспектов в других категориях.

Классификатор SVM основан на модели набора слов и других функциях, описанных ниже:

- словесные n-граммы: термин аспекта и униграммы из контекста термина аспекта извлекаются для выбора признаков.
- Категорийные признаки: для каждой категории отдельно рассматриваются следующие признаки: максимальное количество баллов в разрезе; минимальный балл по контексту; сумма баллов слов в контексте; средний балл слов в контексте;

4. Экспериментальные результаты.

В экспериментальных целях мы использовали обучающую выборку из 200 аннотированных отзывов и тестовую выборку из 200 отзывов для каждого домена, представленную организаторами задания SentiRuEval.

4.1. Результаты производительности

Официальные результаты, полученные нашими подходами на тестовой выборке, представлены в таблицах 1, 2а, 2б и 3. В таблицах представлены официальные исходные результаты и результаты других участников по макросредней F-мере как основному показателю качества в задаче (Лукачевич и др., 2015).

Для задачи А точное совпадение и частичное совпадение использовали для вычисления F1-меры. Таблицы 1а и 1б показывают, что наш метод, основанный на модели CRF, не дал значительного улучшения по сравнению с базовым уровнем.

Для задачи С макросредняя F-мера рассматривается как среднее значение между F-мерой положительного класса, отрицательного класса и F-мерой обоих классов. Из табл. 2а видно, что по макросредней F1-мере наш классификатор не окупается по сравнению с подходом с run_id 4_1, основанным на модели Gradient Boosting Classifier. Наш подход имеет 0,13 % и 0,06 % улучшения в макросреднем показателе F1 по сравнению с подходом с run_id 3_1, занимая второе место в ресторанах и банках соответственно. Наши тиражи не могли быть оценены из-за технических проблем с отправкой.

В таблице 3 показаны официальные исходные результаты и результаты метода, занявшего второе место по макросредней F-мере в задаче D. Этот метод занял первое место среди 4 команд в обеих предметных областях. Наилучший подход обеспечивает повышение макрпоказателя F1 на 0,06 % и 0,09 % по сравнению с базовым уровнем в доменах ресторанов и автомобилей с соответственно.

Иванов В.В., Тутубалина Е.В., Минг азов Н.Р., Алимova И.С.

Таблиц а 1а. Показатели производительнос ти при извлечении явных ас пектов
в домене рес торанов (задача А)

	Точ ное с овпадение				Час тичное с овпадение		
	Мак ро Р	Мак ро R	Мак ро F	Мак ро R	Мак ро F		
Наш метод	0,3515	0,5331	0,5331	0,6507	0,4399	0,5109	0,5506 0,6901 0,6070 0,6886
Подх од, занявший первое ме с то	0,7916	0,7284					
Официальный базовый уровень	0,5570	0,6903	0,6084	0,6580	0,6960	0,6651	

Таблиц а 1б. Показатели производительнос ти при извлечении
явные ас пекты в облас ти автомобилей (задача А)

	Точ ное с овпадение				Час тичное с овпадение		
	Мак ро Р	Мак ро R	Мак ро F	Мак ро R	Мак ро F		
Наш метод	0,6411	0,5363	0,5749	0,7264	0,6117	0,6498	0,6619 0,6560 0,6513 0,7917
Подх од, занявший первое ме с то	0,7272	0,7482					
Официальный базовый уровень	0,5747	0,6287	0,5941	0,7449	0,6720	0,6966	

Таблиц а 2а. Показатели эффективнос ти в клас сификац ии
задач в домене рес торанов (задача С)

Run_id	Micro P	Micro R	Micro F	Мак ро Р	Мак ро R	Мак ро F		
Официальный базовый уровень	0,7104	0,7104	0,7104	0,3209	0,2506	0,2671	0,6194	0,6194 0,6194 0,2517
1_1	0,2454	0,2379	0,6194	0,6194	0,6194	0,2517	0,2454	0,2379
1_2								
3_1	0,6696	0,6696	0,6696	0,3223	0,2430	0,2696	0,8249	0,8249 0,8249 0,5872
4_1	0,5569	0,5545	0,7671	0,7671	0,7671	0,4582	0,3729	0,4081
Наш подх од								

Таблиц а 2б. Показатели эффективнос ти в клас сификац ии
задач в облас ти автомобилей (задача С)

Run_id	Micro P	Micro R	Micro F	Мак ро Р	Мак ро R	Мак ро F		
Официальный базовый уровень	0,6192	0,6192	0,6192	0,2949	0,2685	0,2648	0,6471	0,6471 0,6471 0,3399
1_1	0,3194	0,3293	0,6531	0,6531	0,6531	0,3563	0,3297	0,3422
1_2								
3_1	0,5589	0,5589	0,5589	0,3016	0,2621	0,2794	0,7428	0,7428 0,7428 0,5725
4_1	0,5667	0,5684	0,6252	0,6252	0,6252	0,3507	0,3262	0,3345
1_3								
Наш подх од	0,7110	0,7111	0,7111	0,4481	0,3761	0,4001		

Извлечение аспектов, настроений и категорий аспектов в отзывах пользователей

Таблица 3. Показатели эффективности при категоризации аспектов в обеих предметных областях (задача D)

	Рестораны				Детские заведения		
	Макро P	Макро R	Макро F	Макро P	Макро R	Макро F	
Наш подход	0,8960	0,8414	0,8653	0,6854	0,6355	0,6521	0,8627 0,7963 0,8110
Второй результат	0,7146	0,5750	0,6077	0,8742	0,7737	0,7996	0,6672 0,5190 0,5636
Официальный базовый уровень							

4.2. Эксперименты по аблации

Мы провели эксперименты по аблации, чтобы изучить преимущества функций, которые мы используем для модели CRF и методов машинного обучения. Таблицы 4а, 4б и 5 показывают эксперименты по аблации для задач А и С в наборе для тестирования, удаляя по одной каждой отдельной категории признаков из полного набора. Анализ ошибок в таблицах 4а и 4б показывают, что признаки набора двух предыдущих и двух следующих токенов снижают наши результаты в задаче А в области ресторанов. Наиболее эффективные функции для задачи С основаны на биграммах на основе аспектов, которые включают комбинации термина аспекта и других слов из контекстного окна.

Таблица 4а. Результаты экспериментов по аблации при извлечении аспектов о ресторанах (задача А)

	Точное совпадение			Частичное совпадение		
	p	p	F1	p	p	F1
все функции	0,3515	0,5331	0,5331	0,6507	0,4399	0,5109 0,3382 0,4971
без словарей	0,3961	0,3850	0,6921	0,4821		
без частот без	0,6503	0,4322	0,5068	0,7313	0,4755	0,5612
всех токенов в пределах (wi-2,...,wi)	0,6105	0,4065	0,4751	0,7118	0,4667	0,5471
без всех токенов внутри (wi,...,wi+2)	0,6471	0,4375	0,5104	0,7272	0,4865	0,5681
без токенов, содержащих все функции в пределах (wi-1,...,wi+1)	0,7311	0,4801	0,5644	0,6476	0,4416	0,5120

Иванов В.В., Тутубалина Е.В., Мингазов Н.Р., Алимova И.С.

Таблица 4б. Результаты экспериментов по абляции при извлечении аспектов автомобилей (задача А)

	Точное совпадение			Частичное совпадение		
	п	р	F1	п	р	F1
все функции	0,6411	0,5363	0,5749	0,7264	0,6117	0,6498
без словарей без	0,7303	0,6191	0,6556	0,6380	0,5364	0,5742
частот без всех						
токенов в пределах (wi-2,...,wi)	0,6281	0,5217	0,5609	0,7341	0,6077	0,6498
без всех токенов внутри (wi,...,wi+2)	0,6144	0,5328	0,5624	0,7022	0,6197	0,6453
без токенов, которые содержали все функции внутри (wi-1,...,wi+1)	0,6414	0,5356	0,5742	0,7264	0,6091	0,6472

Таблица 5. Результаты экспериментов по абляции при классификации настраиваемых параметров (задача С)

	Рестораны				Персонализация		
	макрос	Р макрос	Р макрос	F макрос	Р макрос	Р макрос	F
Все функции	0,4582	0,3729	0,4081	0,4481	0,3761	0,4001	
без персонажа n-граммы	0,4479	0,3659	0,4000	0,4480	0,3750	0,3994	
без лексикона униграммы	0,4259	0,3651	0,3921	0,4213	0,3669	0,3869	
без аспектно-	0,4261	0,3396	0,3728	0,4380	0,3746	0,3951	
ориентированных биграмм без биграмм	0,4255	0,3586	0,3906	0,4370	0,3717	0,3941	
n-граммы без оценок на основе лексикона	0,4629	0,3681	0,4050	0,4374	0,3747	0,3959	

Таблица 6. Результаты экспериментов по удалению признаков при категоризации аспектов (задача D)

Комбинация функций	Рестораны			Персонализация		
	п	р	Ф	п	р	Ф
n-граммы с лова n-	0,7650	0,7193	0,7388	0,6554	0,6060	0,6219
граммы с лова + единичная с овокупная оценка n-	0,6800	0,6296	0,6461			
грамм с лова + оценка, специфичные для предметной области	0,8960	0,8414	0,8653	0,6854	0,6355	0,6521

Эксперименты для задачи D представлены в таблице 6. С помощью этих экспериментов по удалению признаков мы показываем, что наиболее важными признаками являются признаки, специфичные для предметной области, которые основаны на точечной взаимной информации для категории и включают четыре различных расчета баллов в контексте термина аспекта.

5. Вывод

В этой статье мы описали контролируемые методы анализа тональности отзывов пользователей о ресторанах и автомобилях. При извлечении явных аспектов (задача A) мы предложили метод, основанный на синтаксических и статистических признаках, включенных в модель условных случайных полей. Метод не показал какого-либо значительного улучшения по сравнению с официальными исходным уровнем. При извлечении настроений по отношению к явным аспектам (задача C) наш метод был основан на модели максимальной энтропии на наборе признаков, основанных на лексиконе, и двух типах признаков частоты терминов: контекстных n -граммах и биграмах на основе аспектов. Мы продемонстрировали, что при использовании этих признаков эффективность классификации повышается по сравнению с базовыми макроусредненными F-мерами с 0,267 до 0,408 для ресторанов и с 0,265 до 0,4 для автомобилей. При категоризации явных аспектов в категории аспектов (задача D) мы предложили классификатор SVM, основанный на признаках униграмм и точечной взаимной информации для расчета оценки, специфичной для категории. Мы достигли 65,2% макроусредненной F-меры для автомобилей и 86,5% для отзывов о ресторанах в задаче D. Этот метод занял первое место среди 4 команд в обеих предметных областях. Для дальнейшей работы мы планируем предоставлять анализ ошибок описанных методов.

Благодарности

Работа выполнена за счет субсидии Правительства РФ на поддержку Программы повышения конкурентоспособности Казанского федерального университета и при поддержке РФФИ (проект РФФИ 13-07-00773).

Рекомендации

1. Блинов П., Клековкина М., Котельников Е., Пестов О. (2013), Исследование лексического подхода и методов обучения для анализа тональности, Вычислительная лингвистика и интеллектуальные технологии, Vol. 2(12), с. 48–58.
2. Чернышевич М. (2014), IHS R&D Беларусь: междоменное извлечение характеристик продукта с использованием условных случайных полей, SemEval 2014, с. 309–313.
3. Четверкин И., Лукашевич Н. (2013), Оценка систем анализа настроений. на русском языке, ACL 2013, с. 14.
4. Чой Ю. Карди К. (2010), Иерархическое последовательное обучение для извлечения мнений и их атрибутов, Материалы конференции ACL 2010, краткие статьи, с. 269–274.
5. Якоб Н., Гуревич И. (2010), Извлечение целевых мнений в одно- и междоменных условиях с условными случайными полями, Материалы конференции 2010 г. по эмпирическим методам обработки естественного языка, с. 1035–1045.

Иванов В.В., Тутубалина Е.В., Минг азов Н.Р., Алимова И.С.

6. Фролов А. В., Поляков П. Ю, Плешков В. В. (2013), Использование семантических фильтров в приложениях к sentiment-анализу рецензий на книги. Режим доступа: www.dialog-21.ru/digests/dialog2013/materials/pdf/FrolovAV.pdf 7. Кириченко С., Чжу С., Черри С., Мокс аммад С.М (2014), NRC-Канада-2014: Выявление аспектов и настроений в отзывах клиентов, SemEval 2014, с.тр. 437–442.
8. Лафферти Дж., МакКаллум А., Перейра Ф.К. (2001), Условные случайные поля вероятностные модели для сегментации и маркировки данных последовательности, Материалы 18-й Международной конференции по машинному обучению 2001 г. (ICML 2001), с.тр. 282–289. .
9. Лю Б. (2012), Анализ настроений и анализ мнений, Обобщающие лекции по технологиям человеческого языка, том 5(1), с.тр. 1–167.
10. Лукашевич Н., Блинов П., Котельников Е., Рубцова Ю, Иванов В., Тутубалина Е. (2015), SentiRuEval: тестирование систем объектно-ориентированного анализа настроений на русском языке, Материалы международной конференции «Диалог -2015», с.тр. 3–9.
11. Лу Б., Отт М, Карди К., Цю Б.К. (2011), Многоаспектный анализ настроений с использованием тематических моделей, Семинары по интеллектуальному анализу данных (ICDMW), 11-я международная конференция IEEE, 2011 г. , с.тр. 81–88.
12. Моргенштерн С., Эстер М (2012), Разработка моделей LDA для анализа мнений на основе аспектов, Материалы 21-й международной конференции ACM по управлению информацией и знаниями, с.тр. 803–812.
13. Панг Б., Ли Л., Вайтманган С. (2002), Большой палец вверх?: классификация настроений с использованием методов машинного обучения Труды конференции ACL-02 по эмпирическим методам обработки естественного языка, том. 10, с.тр. 79–86.
14. Панг Б., Ли Л. (2008), Сбор мнений и анализ настроений, Основы и тенденции информационного поиска т. 1, с. 2(1–2), с.тр. 1–135.
15. Понтики М., Папагеоргиу Х., Галанис Д., Андруцопулос И., Павлопулос Дж., Манандхар С. (2014), Семевал-2014, задача 4: Анализ настроений на основе аспектов, Материалы 8-го Международного семинара по семантической оценке (SemEval 2014), с.тр. 27–35.
16. Попеску А. М., Этциони О. (2007), Извлечение характеристик продукта и мнений из обзоров, Обработка естественного языка и анализ текста, с.тр. 9–28.
17. Poria S., Cambria E., Ku LW, Gui C., Gelbukh A. (2014), Основанный на правилах подход к извлечению аспектов из обзоров продуктов, SocialNLP 2014, с.тр. 28–37.
18. Терни П.Д. (2002), Большой палец вверх или большой палец вниз?: семантическая ориентация применяемая неконтролируемой классификации обзоров, Труды 40-го ежегодного собрания ассоциации компьютерной лингвистики, с.тр. 417–424.
19. Чжао Ю, Цинь Б., Лю Т. (2014), Классификация аспектов продукта с использованием двух эффективных отношений аспектов для сбора мнений, китайской компьютерной лингвистики и обработки естественного языка на основе естественно аннотированных больших данных , с.тр. 120–130. .