

# Анализ тональности (sentiment analysis)

Павел Браславский

# ПРИМЕРЫ



eaglebk

8★

Любителям 3d action должно понравится, драйвовое кино. Оригинально снято,приятное звуковое сопровождение, картинка держит в напряжении течении всего просмотра. Сюжет неплохой, он банален и прост, но узнать что же будет дальше хочется. Хочется сказать создателям спасибо.

24 апреля 2016 | Поделиться | Комментировать

👍 0    💬 0



Халисия

5★

Думаю что мальчикам ,всё детство проигравших в компьютерные игры этот фильм понравится (ностальгия у них).Конечно в целом много крутых сцен,но я такие фильмы не понимаю,бредовый сюжет,раздражает прыгающая камера (везде в рекламе пишут ,что это супер находка ,что камера была у артиста на голове) от этой ((супер находки)) всё кружилось в моей голове и когда мы вышли из зала в торговый цен... Читать полностью

24 апреля 2016 | Поделиться | Комментировать

👍 0    💬 0



tatiee

4★

Прикольно атмосферный экшн и пара идей неплохих, и все, что хорошее, от тряски мозгам надоедает и как ни крути мясорубку смысл стремится к нулю при сеансе - к завершению

21 апреля 2016 | Поделиться | Комментировать

👍 0    💬 0



Elusive

8★

Поначалу меня укачивало, потом втянулась. Много экшена, чем дальше, тем больше хардкора (очень подходящее название), и без перерывов на отдых. Местами похоже на counter strike)) сюжет интересно закручен, не даст отвлечься.

21 апреля 2016 | Поделиться | Комментировать

👍 0    💬 0

Дубинин Андрей Автор 2 отзывов

5

отличная модель Опыт использования: менее месяца

**Достоинства:**

Навеска и тормоза Шимано достойного уровня, хорошие педали в базе (широкие, удобные и качественные), отличный накат, удобное сиденье даже для "необстреленной" задницы новичка, очень достойная резина Континенталь- это непосредственно то, что отличает в целом от конкурентов в данном ценовом диапазоне. Пожалуй, лучшее соотношение цена-комплектация, да еще и от GT.

**Недостатки:**

тросы натирают вилку до металла уже через 50км пробега, стяжка в определенных местах жгутиком между собой решает. Больше не выявил.

**Комментарий:**

Писать в недостатки каличную вилку SUNTUR считаю нечестно, в эти деньги это невозможно, пожалуй, единственное, что захочется из навески улучшить в этом велосипеде. Очень доволен.

28 августа 2015, Воронеж



Усачев Дмитрий

5

отличная модель Опыт использования: менее месяца

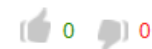
**Достоинства:**

Отличный велосипед. Превосходный накат, как будто везде стоят подгипники от авиатурбин :)). + видимо характеристики резины Continental. Достойная навеска, удобный широкий руль. Быстро и легко рагоняются. Очень информативные тормоза.

**Недостатки:** На хорошем асфальте резина довольно сильно гудит

**Комментарий:** Всем рекомендую !

10 сентября 2015



## Review highlights

### Location



Excellent location close to the beach.

*"Excellent location of the hotel" "Perfect location" "Beautiful location"*

14 reviews

---

### Service



Friendly service, and excellent management.

*"Friendly hotel owner" "Warm welcome" "Excellent service"*

22 reviews

---

### Beach



Beautiful beach.

*"Plenty of beaches"*

9 reviews

---

### Breakfast



Tasty breakfast.

*"Home made Breakfast"*

13 reviews

---

### Food

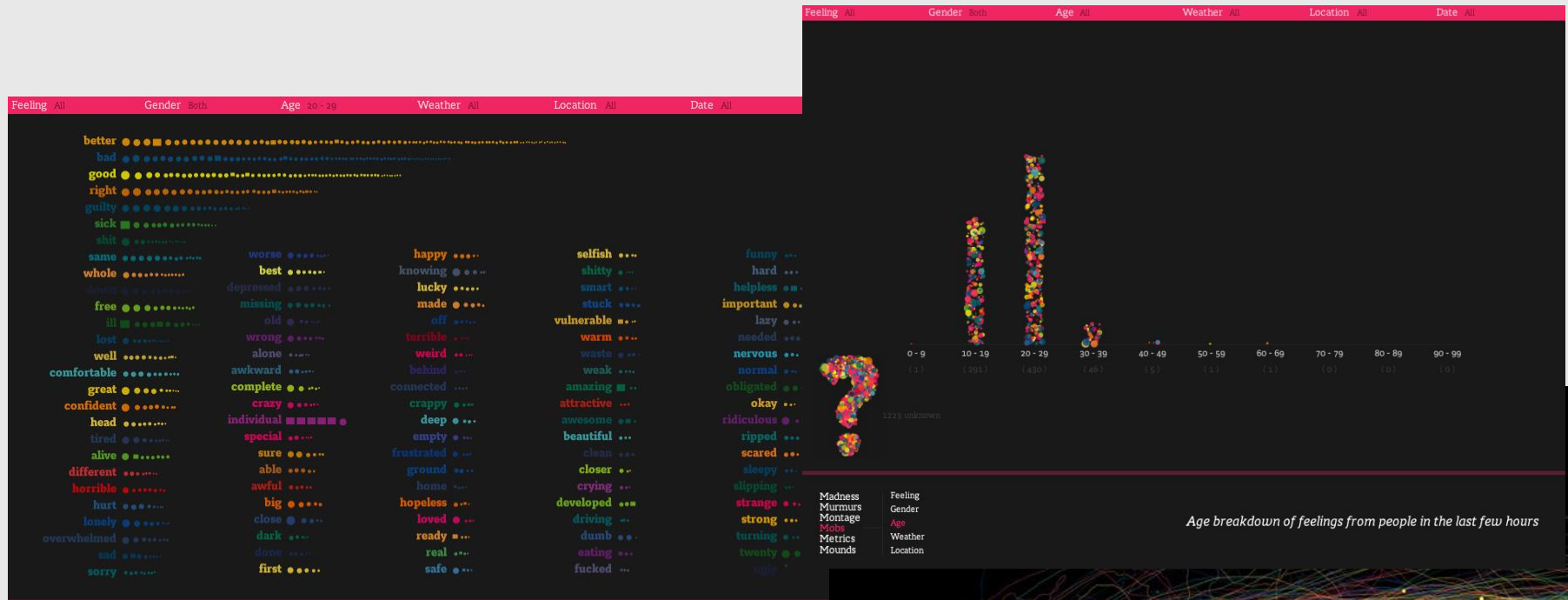


Delicious food.

*"Delicious pastries" "Fresh seafood" "Great food"*

14 reviews

# we feel fine

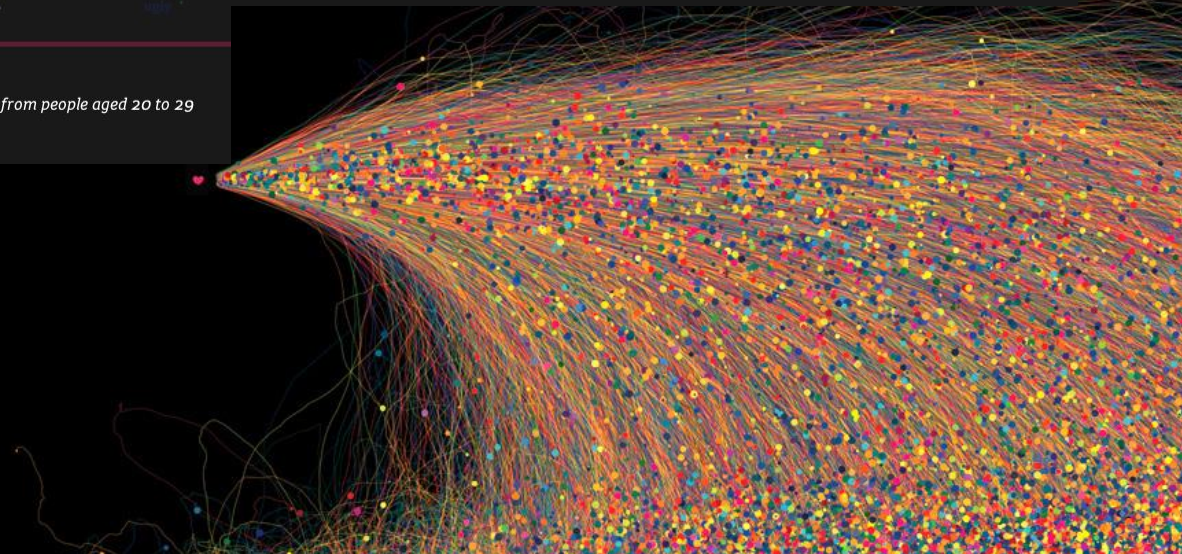


Age breakdown of feelings from people in the last few hours

Feeling breakdown of feelings from people aged 20 to 29

- Madness
- Murmurs
- Montage
- Mobs
- Metrics
- Mounds

- Feeling
- Gender
- Age
- Weather
- Location



Тональность документа:

позитив 0

негатив 0.4

Певица Юлия Чичерина в воскресенье, 18 апреля, даст живой концерт программе «Соль» на РЕН ТВ и расскажет о том, как попала под уголовное преследование на Украине. Об этом сообщается в релизе пресс-службы телеканала. Украинские силовики объявили исполнительницу в уголовный розыск в январе 2015 года. «Это произошло сразу, как только я приехала на Новый год в Луганск – сыграть благотворительный концерт. Новость меня сильно удивила», – приводит канал слова певицы. Во время программы певица передаст своему звукооператору Андрею Прилепскому медаль от Минобороны за концерт для российских военных в Сирии, которую он не смог получить раньше. Чичерина выступила перед военнослужащими на базе Хмеймим в Сирии в январе 2016 года. Кроме нее из группы присутствовал только гитарист. Их поддерживал ансамбль Центрального военного округа и коллектив самодеятельности одного из кораблей группировки ВМФ, действующей в Средиземном море. В 2015 году Чичерина несколько раз дала концерт в Луганске. На Украине Чичерину обвиняют в «посягательстве на территориальную целостность и неприкосновенность» государства. Музыкальная передача «Соль» выходит на РЕН ТВ по воскресеньям в полночь. В программе российские музыканты дают живые концерты, беседуют с ведущим Захаром Прилепиным о жизни, культуре и общественно-политические вопросы, делятся воспоминаниями. В конце марта героем телепередачи стал Юрий Лоза. Он критически высказался о некоторых западных музыкантах. В частности, он сказал, что The Rolling Stones «гитару не настроили за всю жизнь ни разу», а лидер группы Мик Джаггер «ни в одну ноту не попал». Реплики Лозы вызвали бурную дискуссию в сети, которая продолжалась несколько недель.

● Позитив

● Нейтрально

● Негатив

# Sentiment140

Tweet 368

Like 158

+1 78

microsoft surface

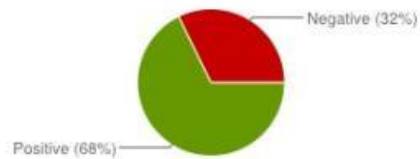
English

Search

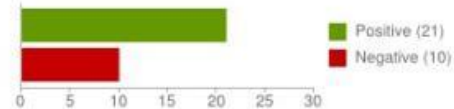
[Save this search](#)

## Sentiment analysis for microsoft surface

Sentiment by Percent



Sentiment by Count



## Tweets about: microsoft surface

[Trillian\\_01](#): They should customise it for the Apple iPad. RT [@Ebenwolfe](#) " **Micros**  
Posted 54 seconds ago

[mactalk](#): "[@techAU](#): Australian **Microsoft Surface** pricing - <http://t.co/Zn3P6MFm> "  
Posted 55 seconds ago

[1862magpie](#): I think..... I want one of these **Microsoft Surface** tablets <https://t.co/L>  
Posted 1 minute ago

[DanDrastic](#): RT [@MyNameisRazzle2](#): "I can't wait to buy the new **Microsoft Surfa**  
Posted 4 minutes ago

[HowardSol](#): RT [@davezatz](#): All this talk of **Microsoft Surface** reminds me how stel  
Posted 4 minutes ago



# Задачи

- определить тональность текста/предложения/твита
- +/0/−; шкала
- объект/субъект
- субъективность/  
тональность (отношение)/  
эмоции (разочарование/грусть/...)/  
личностные характеристики/...

# Приложения

- агрегация отзывов
- мониторинг СМИ/социальных сетей
- поддержка пользователей
- предсказание результатов выборов/биржевых котировок/...

# БАЗОВЫЙ МЕТОД

# Классификация

- классифицируем тексты на 2 класса
- признаки – слова
- отдельно обрабатываем отрицания  
*не очень хороший* → *NOT\_очень NOT\_хороший*
- сохраняем капитализацию  
*отличный фильм vs. фильм просто БОМБА*
- сглаживаем частоту  $\log(tf)$  или даже 0/1
- :-) :-( – хорошие признаки

# Сложные случаи

Когда появилось это пиво, я была в **восторге!!!** Мне очень хотелось его попробовать, но к сожалению, у меня это не получилось. Причина была очень **радостная!!!** В животике у меня находился маленький человечек и **травить** его этой **гадостью** было бы подло! Так вот! После родов и кормления грудью ребенка, я созрела на пиво! Бегом побежала за Клинским "Мохито"! Открыла, глотнула и ..... все!!!! Вкус абсолютно не понятный!!! **То-ли мыла, то-ли шампунь, то-ли зубная паста**, то-ли еще какая-то **гадость!** Короче, пить его **невозможно!** Ну, возможно, это пиво на любителя. Было когда-то у Клинского "Чайна Таун". Тож, на любителя! Многие мои знакомые пили и **восхищались!**

Мы с мужем иногда **любим** побаловать себя пивком к ужину. Почти всегда выбираем одну, давно **проверенную марку**, но тут решили попробовать широко рекламируемое Клинское Мохито. Поэкспериментировали, **ага!** **Хорошо**, что взяли только по одной бутылке.

<http://irecommend.ru/content/shampun-0>

При просмотре этого фильма, я рассчитывала на **интересную захватывающую** фантастику! Но после 20-ти минут мне жутко **хотелось спать)))**  
<...>Просто можно было сделать все более **динамично, интересно**, чтобы от экрана **не оторваться...** А было очень **скучно** и **примитивно(((**

<http://films.imhonet.ru/element/195636/>

# СЛОВАРИ

# Собрать по словарям

## отвратный

Прилагательное, тип склонения по классификации А. Зализняка — 1\*а.

Корень: **-отврат-**; суффикс: **-н**; окончание: **-ый** [Тихонов, 1996].

Приставка: **от-**; корень: **-врат-**; суффикс: **-н**; окончание: **-ый** [Кузнецова, Ефремова, 1986].

## Произношение [править]

МФА: [etˈvratnɨj]

## Семантические свойства [править]

### Значение [править]

- разг.** очень плохой, отвратительный ♦ Итак, если бы можно было какое телесное истребить или в ничто превратить, по учению проповедника новой метафизики, то начала насекомых, растений, минералов и самых **отвратных** материй после уничтожения сих тел не престаи бы существовать; но бестелесное существо принадлежит божеству, а потому из сего выходит весьма странное заключение, что божество сопрягается с такими ничтожными и недостойными его телами. П. С. Батулин, «Исследование книги О заблуждениях и истине», 1790 г. (цитата из Н

### Синонимы [править]

- отвратительный, мерзкий, поганый, **препоганый**

### Антонимы [править]

- великолепный, чудесный

### Гиперонимы [править]

- плохой

## супер II [\[ править \]](#)

### Морфологические и синтаксические свойства [\[ править \]](#)

#### сú-пер

Предикатив, также междометие; неизменяемое.

Корень: --.

### Произношение [\[ править \]](#)

МФА: [ˈsupʲɪr]

### Семантические свойства [\[ править \]](#)

#### Значение [\[ править \]](#)

1. в *знач. сказуемого*, *сленг* оценочная характеристика чего-либо как очень хорошего, превосходного, прекрасного; *отлично, великолепно* ◆ — Жена из меня хреновая, но вдовой буду **супер!** *М. А. Вишневецкая, «Вышел месяц из тумана», 1997 г. (цитата из Национального корпуса русского языка, см. Список литературы)* ◆ — Как жизнь? — **Супер**... В бывшей жизни было всего два супера — суперобложка и суперфосфат. *Михаил Мишин, «Элитарный эксклюзив», 1998 г. (цитата из Национального корпуса русского языка, см. Список литературы)* ◆ Не сказать, чтобы это было чем-то **супер**, скорее раритет для ценителей. *Николай Солдатенков, «Битломанам не дают расслабиться», 2001 г. // «Аргументы и факты» (цитата из Национального корпуса русского языка, см. Список литературы)* ◆ Это же просто **супер**, подумала я. То есть это теперь я уверена, что **супер**. *Юлия Пешкова, «Свадебный лист. Кто последний?», 2002 г. // «Домовой» (цитата из Национального корпуса русского языка, см. Список литературы)* ◆ У тебя все получится! Оля, ты просто **супер!** *Ольга Утешева, «Там, на неведомых дорожках», 2002 г. // «Домовой» (цитата из Национального корпуса русского языка, см. Список литературы)* ◆ — Вот теперь **супер!** Класс! — довольно крякнул Вовка, причём так громко, что разбудил котёнка Классика, дремавшего на диване. *Марина Дружинина, «Открытка», 2003 г. // «Мурзилка» (цитата из Национального корпуса русского языка, см. Список литературы)*



## отстой I [\[ править \]](#)

### Морфологические и синтаксические свойства [\[ править \]](#)

#### от-стой

Существительное, неодушевлённое, мужской род, 2-е склонение (тип склонения ба по классификации А. А. Зализняка).

Приставка: **от-**; корень: **-стой-**. [\[Тихонов, 1996\]](#)

#### Произношение [\[ править \]](#)

МФА: ед. ч. [ɐtˈsɔj], мн. ч. [ɐtˈsɔj]

падеж	ед. ч.	мн. ч.
Им.	отстой	отстои
Р.	отстоя	отстоёв
Д.	отстою	отсто́ям
В.	отстой	отстои
Тв.	отстоём	отсто́ями
Пр.	отстоё	отсто́ях

### Семантические свойства [\[ править \]](#)

#### Значение [\[ править \]](#)

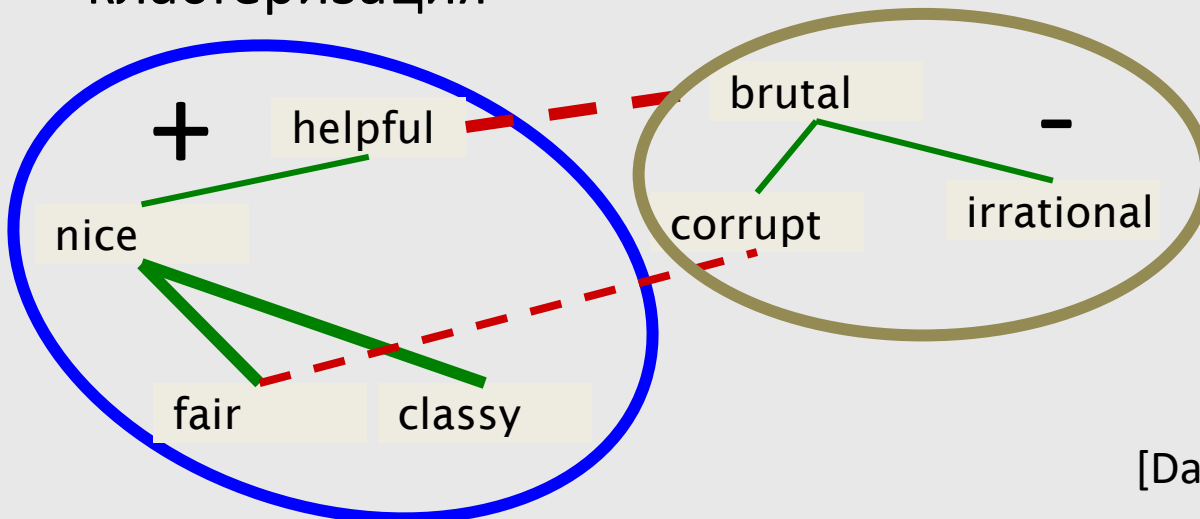
- действие по значению гл. **отстаивать**, **отстоять**; **осаждение** находящихся в жидкости примесей или составных частей ♦ **Не указан пример употребления (см. рекомендации)**.
- результат такого действия; отстоявшийся **слой**, **осадок**, выделившаяся в результате отстаивания часть какой-либо жидкости ♦ **Отработанные масла и отстой топлива из топливных баков необходимо сливать только в специальную тару.** «Типовая инструкция № 23 по предупреждению пожаров и предотвращению ожогов на автомобильном транспорте ТОО Р 200-23-95», 1996 г. (цитата из Национального корпуса русского языка, см. Список литературы)
- перен.**, **сленг** нечто плохое, низкого качества ♦ — Сказала, что «Евгений Онегин» **отстой**, — поджала губы Крестя, — между прочим, русичка предложила нам высказать собственное мнение! Д. Донцова, «Микстура от косоглазия», 2003 г. (цитата из Национального корпуса русского языка, см. Список литературы) ♦ Сколько я знала заколдованных принцев — все они были полный **отстой**. Д. А. Емец, «Таня Гроттер и колодец Посейдона», 2004 г. (цитата из Национального корпуса русского языка, см. Список литературы)
- трансл.** действие по значению гл. **отстаиваться**; временное, обычно продолжительное бездействие транспортных средств; **стоянка** транспорта между рейсами ♦ С Ангары ему откликались и вторили протяжными тоскливыми гудками суда, уходящие на зимний **отстой**. В. Г. Распутин, «Изба», 1999 г. (цитата из Национального корпуса русского языка, см. Список литературы) ♦ Стихший Пал Палыч малым ребёнком озирался среди отстоя казавшихся исполинскими вагонов, внутри которых отсутствовала жизнь, думая, наверное, что очутился в своём личном аду. О. О. Павлов, «Карагандинские девятины, или Повесть последних дней», 2001 г. (цитата из Национального корпуса русского языка, см. Список литературы)

# Расширение словаря

Гипотеза:

- одинаковая полярность
  - сильный и надёжный, ужасный и подлый
- противоположная
  - красивый, но жестокий
- ~1300 исходных прилагательных → поиск по шаблону
- кластеризация

Hatzivassiloglou & McKeown.  
Predicting the Semantic Orientation  
of Adjectives. ACL'1997.



# Словарь фраз

1. Извлечь *фразы* определенной структуры из ОТЗЫВОВ
2. Оценить полярность фраз на основе совместной встречаемости с «хороший/плохой» в вебе
3. Тональность отзыва = средняя полярность входящих в него фраз
4. учет особенностей предметной области

Turney (2002): Thumbs Up or Thumbs Down? Semantic Orientation Applied to Unsupervised Classification of Reviews

# Двусловия с прилагательными

W1	W2	W3 (не извлекается)
JJ	NN or NNS	anything
RB, RBR, RBS	JJ	Not NN nor NNS
JJ	JJ	Not NN or NNS
NN or NNS	JJ	Nor NN nor NNS
RB, RBR, or RBS	VB, VBD, VBN, VBG	anything

# Поточечная взаимная информация

$$\text{PMI}(word_1, word_2) = \log_2 \frac{\text{hits}(word_1 \text{ NEAR } word_2)}{\text{hits}(word_1)\text{hits}(word_2)}$$

$$\text{Polarity}(phrase) = \text{PMI}(phrase, "excellent") - \text{PMI}(phrase, "poor")$$

$$= \log_2 \left( \frac{\text{hits}(phrase \text{ NEAR } "excellent")\text{hits}("poor")}{\text{hits}(phrase \text{ NEAR } "poor")\text{hits}("excellent")} \right)$$

# Пример использования

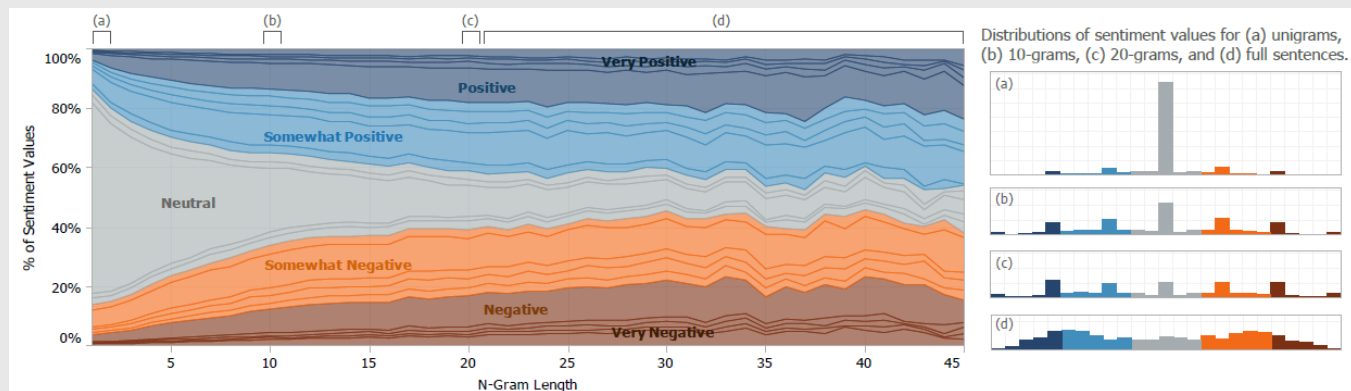
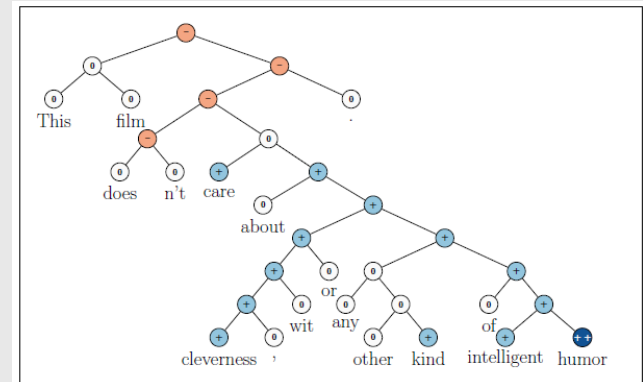
Phrase	POS tags	Polarity
online service	JJ NN	2.8
online experience	JJ NN	2.3
direct deposit	JJ NN	1.3
local branch	JJ NN	0.42
...		
low fees	JJ NNS	0.33
true service	JJ NN	-0.73
other bank	JJ NN	-0.85
inconveniently located	JJ NN	-1.5
<i>Average</i>		0.32

Phrase	POS tags	Polarity
direct deposits	JJ NNS	5.8
online web	JJ NN	1.9
very handy	RB JJ	1.4
...		
virtual monopoly	JJ NN	-2.0
lesser evil	RBR JJ	-2.3
other problems	JJ NNS	-2.8
low funds	JJ NNS	-6.8
unethical practices	JJ NNS	-8.5
<i>Average</i>		-1.2

# Нейронные сети

[Socher et al., 2013]

- разметка 12К предложений на уровне составляющих
- векторы для (слова + синтаксическая структура)



# CNN + pre-trained w2v

Model	MR	SST-1	SST-2	Subj	TREC	CR	MPQA
CNN-rand	76.1	45.0	82.7	89.6	91.2	79.8	83.4
CNN-static	81.0	45.5	86.8	93.0	92.8	84.7	<b>89.6</b>
CNN-non-static	<b>81.5</b>	48.0	87.2	93.4	93.6	84.3	89.5
CNN-multichannel	81.1	47.4	<b>88.1</b>	93.2	92.2	<b>85.0</b>	89.4
RAE (Socher et al., 2011)	77.7	43.2	82.4	—	—	—	86.4
MV-RNN (Socher et al., 2012)	79.0	44.4	82.9	—	—	—	—
RNTN (Socher et al., 2013)	—	45.7	85.4	—	—	—	—
DCNN (Kalchbrenner et al., 2014)	—	48.5	86.8	—	93.0	—	—
Paragraph-Vec (Le and Mikolov, 2014)	—	<b>48.7</b>	87.8	—	—	—	—
CCAЕ (Hermann and Blunsom, 2013)	77.8	—	—	—	—	—	87.2
Sent-Parser (Dong et al., 2014)	79.5	—	—	—	—	—	86.3
NBSVM (Wang and Manning, 2012)	79.4	—	—	93.2	—	81.8	86.3
MNB (Wang and Manning, 2012)	79.0	—	—	<b>93.6</b>	—	80.0	86.3
G-Dropout (Wang and Manning, 2013)	79.0	—	—	93.4	—	82.1	86.1
F-Dropout (Wang and Manning, 2013)	79.1	—	—	<b>93.6</b>	—	81.9	86.3
Tree-CRF (Nakagawa et al., 2010)	77.3	—	—	—	—	81.4	86.1
CRF-PR (Yang and Cardie, 2014)	—	—	—	—	—	82.7	—
SVM <sub>S</sub> (Silva et al., 2011)	—	—	—	—	<b>95.0</b>	—	—



**АСПЕКТ/АТРИБУТ**

# Атрибуты

- Автомобили:

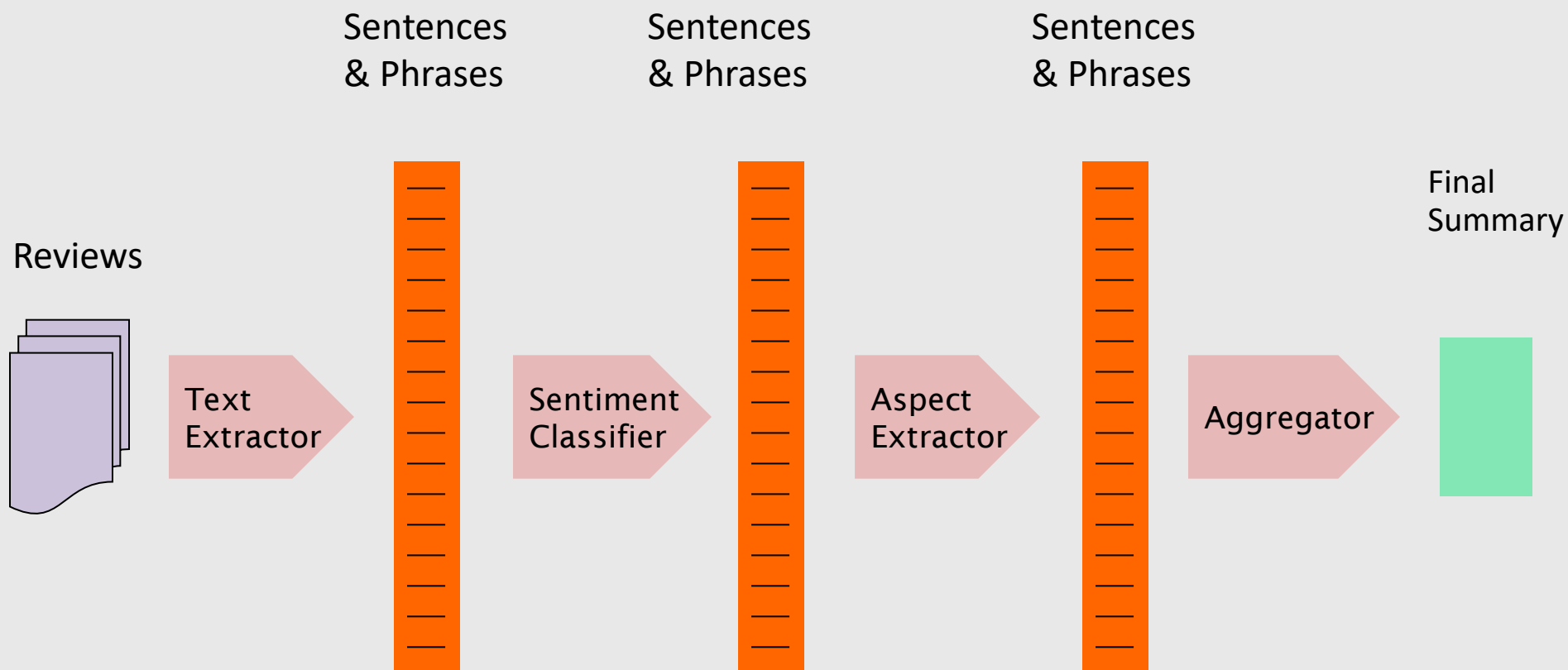
безопасность, комфорт, надежность, внешний вид, цена, дорожные качества

- Рестораны:

кухня, интерьер, сервис, цена

можно составить ручную, можно автоматически извлечь из текста на основании предположения о структуре, близости к тонально окрашенным словам

# Тональный анализ атрибутов



S. Blair-Goldensohn, K. Hannan, R. McDonald, T. Neylon, G. Reis, and J. Reynar.  
Building a Sentiment Summarizer for Local Service Reviews. WWW'2008 Workshop

# **ДРУГИЕ НЕТЕМАТИЧЕСКИЕ ХАРАКТЕРИСТИКИ**

# Другие нетематические аспекты

- жанр документа/авторский стиль
- пол автора
- ирония/сарказм
- раздражение
- депрессия/болезнь Альцгеймера
- личностные качества/черты характера



LHC(b) Experiment...

CMS Experiment...

ATLAS Experiment...

ALICE Experiment...

## Welcome to the Large Hadron Collider

### Tunnelling to the beginning of time

The LHC (Large Hadron Collider) is an international project, in which the UK has a leading role. This site includes the latest news from the project, accessible explanations of how the LHC works, how it is funded, who works there and what benefits it brings us. You can access a wide range of resources for the public, journalists and teachers and students, there are also many links to other sources of information.



### The LHC is asking some Big Questions about the universe we live in

#### How did our universe come to be the way it is?

The Universe started with a Big Bang – but we don't fully understand how or why it developed the way it did. The LHC will let us see how matter behaved a tiny fraction of a second after the Big Bang. Researchers have some ideas of what to expect – but also expect the unexpected!

Read more of [The LHC Big Questions](#).

[Latest News from CERN](#)

### the LHC

the LHC, How does the work, people at CERN,

### Particle Detectives

LHC simulator, check 5 minute presentations, 'Big Question' page...

### Physicists

Communicating about the LHC, and Demonstrations, and Presentations...

### Resources

Links, photos, publications, ...

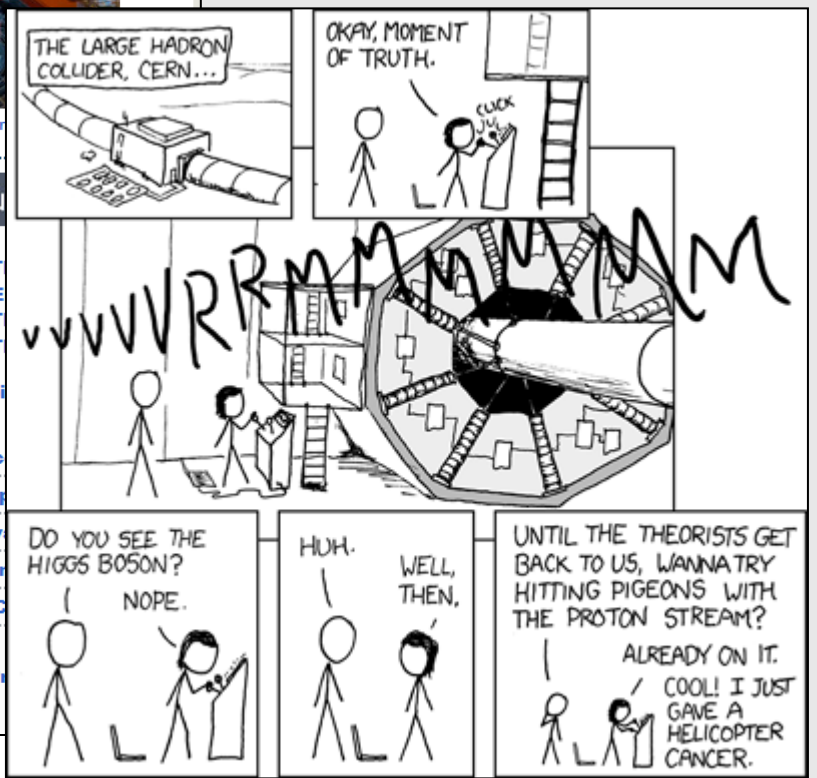
### Contact us

STFC's Science in and Media teams...

Qu

- ST
- CE
- ST
- ST
- Si

- Acc
- Help
- Priv
- Terr
- LHC



## Abstract

We revisit the issue of the limit on the scale of Left-Right symmetry breaking. We focus on the minimal  $SU(2)_L \times SU(2)_R \times U(1)_{B-L}$  gauge theory with the seesaw mechanism and discuss the two possibilities of defining Left-Right symmetry as parity or charge conjugation. In the commonly adopted case of parity, we perform a complete numerical study of the quark mass matrices and the associated left and right mixing matrices without any assumptions usually made in the literature about the ratio of vacuum expectation values. We find that the usual lower limit on the mass of the right-handed gauge boson from the  $K$  mass difference,  $M_{W_R} > 2.5 \text{ TeV}$ , is subject to a possible small reduction due to the difference between right and left Cabibbo angles. In the case of charge conjugation the limit on  $M_{W_R}$  is somewhat more robust. However, the more severe bounds from CP-violating observables are absent in this case. In fact, the free phases can also resolve the present mild discrepancy between the Standard Model and CP-violation in the  $B$ -sector. Thus, even in the minimal case, both charged and neutral gauge bosons may be accessible at the **Large Hadron Collider** with spectacular signatures of lepton number violation.

# Жанры vs. Авторский стиль



Кто написал «Тихий Дон»?  
Кто написал «Роман с кокаином»?

# Стилистические признаки

- длина слов/предложений
- редкие слова
- служебные слова
- пунктуация
- последовательности/частоты POS
- буквенные n-граммы



# Пол автора: признаки

- **Формальность**

$$F = 0.5 * [(freq.noun + freq.adj + freq.prep + freq.art) - (freq.pron + freq.verb + freq.adv + freq.int) + 100]$$

- **Стилистические признаки**

- **последовательности POS**

- слова/суффиксы

- *hmm, lol, smiley*



Анастасия Крылова

@MC\_dolmatinets\_



Follow

Сегодня так классно прошёл день моих каникул ... Весело посидела на консультации по физики  
[#сарказм](#)



Diwoon

@Diwoon\_



Follow

Как я рад, что наступила зима [#ирония ift.tt/2fm8MGj](#)

[View translation](#)



# Юмор/ирония/сарказм

## Humour and irony: one-liners: some pattern

- Jesus *saves*, and at today's prices, that's a miracle! [ambiguity]
- Love is *blind*, but marriage is a real *eye-opener*. [antonymy]
- *Drugs* may lead to nowhere, but at least it's a scenic route. [human weakness]
- Become a *computer programmer* and never see the world again. [common topic / community]
- My software never has bugs; it just develops random features. [??]
- *Sex* is one of the nine reasons for reincarnation; the other eight are unimportant. [language]
- I've got the *body of a god* ...unfortunately is *Buddha*. [irony]

# Юмор/ирония/сарказм

## Irony and humour: some features

- N-grams: frequent sequences of words
- Descriptors: tuned up sequences of words
- POS n-grams: POS templates
- Polarity: polarity of words
- Affectiveness: emotional content (WordNet Affect)
- Pleasantness: degree of pleasure (Whissel's dictionary)
- Funniness: relationship between humor and irony (humour domains and lexical ambiguity)

Tested on Amazon viral effect corpus: (Reyes and Rosso, 2013)

# РЕСУРСЫ



# SentiStrength

Enter text: прекрасные меланхоличные стихи

Detect Sentiment in Russian

Enter text: прекрасные меланхоличные стихи

Detect Sentiment in Russian2

The text 'прекрасные меланхоличные стихи'  
has positive strength **3** and negative strength **-2**

*Approximate classification rationale:* прекрасные[3] меланхоличные[-2] стихи  
[sentence: 3,-2] [result: max + and - of any sentence][overall result = 1 as  
pos>-neg] (Russian)

The text 'прекрасные меланхоличные стихи'  
has positive strength **1** and negative strength **-1**

*Approximate classification rationale:* прекрасные меланхоличные стихи [sentence:  
1,-1] [result: max + and - of any sentence][trinary result = 0 as pos=1 neg=-1]  
(Russian2)

# SemanticAnalyzer

до 200 запросов в день – бесплатно

запрос:

```
{  
  "text": "Иванова уволили, хотя он хороший сотрудник",  
  "object_keywords": "сотрудник",  
  "output_format": "json",  
  "include_strength": 1  
}
```

ответ:

```
{  
  "sentiment": "POSITIVE",  
  "strength": "1.0",  
  "synonyms": "[сотрудник]"  
}
```

<https://market.mashape.com/dmitrykey/russiansentimentanalyzer>

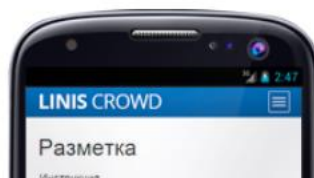
БЕСПОДОБНЫЙ 0.963  
НЕВНЯТНЫЙ 0.953  
ОТЛИЧНЕЙШИЙ 0.935  
ОБАЛДЕННЫЙ 0.933  
БЕЗУМНО 0.924  
НЕПОНЯТНО 0.921  
НЕПРИЯТНО 0.920  
ОТВРАТНЫЙ 0.920  
НЕЖНЫЙ 0.916  
ПОСРЕДСТВЕННЫЙ 0.912  
ШИКАРНЫЙ 0.908  
НЕЛОГИЧНЫЙ 0.906  
КЛЕВЫЙ 0.898  
СРЕДНЕНЬКИЙ 0.895  
НЕПОНЯТНЫЙ 0.895  
СЛАБЕНЬКИЙ 0.892  
НЕЕСТЕСТВЕННЫЙ 0.890  
НЕПЛОХО 0.888  
СУПЕРСКИЙ 0.886  
НЕПЛОХОЙ 0.885  
ТУПО 0.882  
ПРЕВОСХОДНЫЙ 0.881  
НЕОДНОЗНАЧНЫЙ 0.879  
ПРИКОЛЬНЫЙ 0.878  
НЕОБЫЧНЫЙ 0.878  
КРАСОЧНЫЙ 0.876

Словарь тонально окрашенных слов,  
без полярности

Chetviorkin I., Loukachevitch N.  
Extraction of Russian Sentiment  
Lexicon for Product Meta-Domain.  
COLING 2012.

<http://www.cir.ru/SentiLexicon/ProductSentiRus.txt>





Общедоступный тональный словарь и краудсорсинговая платформа для его создания.



Здесь Вы можете принять участие в разметке слов и текстов и помочь созданию общедоступного словаря для автоматической оценки тональности интернет-контента.



Если Вы исследователь, этот онлайн-ресурс позволит Вам привлекать широкие «толпы» добровольцев к разметке слов и текстов для разных задач.



Размеченный тональный словарь, коллекция текстов с тональной разметкой и другие результаты доступны для всех по окончании проектов.

Скачать словарь

Всего 9702 слов в базе

Всего 29106 текстов в базе

Статистика оценок

Кол-во оценок	Кол-во слов
0	7127
1	2207
2	330
3	32
4	2
5	1
8	1
13	1
51	1
Итого	9702

Всего 2821 текста с оценками

Всего 26285 текстов без оценок

Всего 3048 оценок слов

Всего 3048 оценок текстов

Алексеева С.В., Кольцова Е.Ю., Кольцов С.Н. Linis-crowd.org: лексический ресурс для анализа тональности социально-политических текстов на русском языке. IMS-2015  
<http://openbooks.ifmo.ru/ru/file/2203/2203.pdf>

<http://linis-crowd.org/>

# Пример

"закомплексованный"; "-1"; "0"; "-1";  
"законно"; "0"; "0"; "0";  
"законодатель"; "-0,1666666666666667"; "0,372677996249965"; "0";  
"законодательный"; "0"; "0"; "0";  
"законодательство"; "0"; "0"; "0";  
"закоренелый"; "-0,2222222222222222"; "1,13311544746506"; "0";  
"закрадываться"; "-0,2"; "0,748331477354788"; "0";  
"закреплять"; "0"; "0"; "0";  
"закрытие"; "0"; "0"; "0";  
"закупаться"; "0"; "0"; "0";  
"закупка"; "0"; "0"; "0";  
"закуривать"; "0,6666666666666667"; "0,942809041582063"; "1";  
"заложник"; "-1,3333333333333333"; "0,471404520791032"; "-1";  
"заляпывать"; "-0,3333333333333333"; "0,942809041582063"; "0";  
"замалчивать"; "-1,1111111111111111"; "0,737027731190089"; "-1";

# Данные РОМИП/Диалог

- 2011
  - 15К – Отзывы о фильмах (imhonet.ru)
  - 24К – Отзывы о книгах (imhonet.ru)
  - 10К – Отзывы о фотокамерах (Яндекс.Маркет)
- 2012
  - 4К – Новости
  - 0.8К – Блоги – книги (с объектами)
- 2014
  - 0.8К – отзывы о ресторанах/автомобилях (аспекты)
  - 8К – твиты о банках и телекоме
- 2015
  - расширение коллекции твитов
- 2016
  - твиты о банках и телекоме

<http://romip.ru/ru/collections/>

<http://www.dialog-21.ru/evaluation/>

		Neu- tral	Posi- tive	Nega- tive	Total number of tweets
Tele- com	Training collection	4,870	1,354	2,550	8,643
	Gold standard test collection	1,016	226	1,054	2,247
Banks	Training collection	6,977	704	1,734	9,392
	Gold standard test collection	2,240	312	722	3,313

```
<sentence id="1854">
  <speech type="direct">
    &quot;Десятилетний период - это слишком много&quot;;, -
    сказал он, комментируя вопрос журналистов.
  </speech>
  <evaluation>
    -
  </evaluation>
  <url>
    http://www.finmarket.ru/z/nws/news.asp?id=2870916
  </url>
</sentence>
```

```
<sentence id="225">
  <speech type="indirect">
    Главный тренер Манчестер Сити - Роберто Манчини заявил, что
    команда постарается дополнить состав тремя-четырьмя игроками
    в ближайшее время.
  </speech>
  <evaluation>
    +
  </evaluation>
  <url>
    http://neeko.ru/siti-postaraetsya-dopolnit-sostav.html
  </url>
</sentence>
```

```
<query>
```

```
    Canon PowerShot SX20 IS
```

```
</query>
```

```
<text>
```

```
    Не так давно продал свой зеркальный Nikon D40 KIT и купил
    вместо него Canon PowerShot SX20 IS . Обменом ничуть не расстроен -
    теперь получил возможность снимать фото с 20-кратным зумом и
    неплохое HD-видео 720p со стереозвуком, плюс некоторые прикольные
    фишки типа lifeview с поворотным экраном и прочие вкусности. Так же
    есть полностью ручные режимы. Снято с 5 метров с максимальным
    приближением Тоже снято издалека Совсем сблизил
```

```
</text>
```

```
<score-2>
```

```
<score-3>
```

```
<score-5>
```

```
    4
```

```
</score-5>
```

```
<type main="+">
```

```
<real_object>
```

```
    Canon PowerShot SX20 IS
```

```
</real_object>
```

```

<table name="bank_train">
  <column name="id">2999</column>
  <column name="twitid">496654553144890000</column>
  <column name="date">1407246647</column>
  <column name="text">Уже не первый раз слышу негатив про райфазен. Они нормальный были, а
потом часть ушли в ВТБ 24 и они испортились. Ужасно конечно</column>
  <column name="sberbank">NULL</column>
  <column name="vtb">--</column>
  <column name="gazprom">NULL</column>
  <column name="alfabank">NULL</column>
  <column name="bankmoskvy">NULL</column>
  <column name="raiffeisen">-1</column>
  <column name="uralsib">NULL</column>
  <column name="rshb">NULL</column>
</table>

```

```

  <object>ВАЗ 2172 Лада Приора хэтчбек</object>
</meta>
<text>Купил данную машину в 2012 году, выбор пал именно на нее, потому как средства на
покупку были весьма ограничены, многие отговаривали от покупки, утверждая что лучше &quot;
старенькая иномарка&quot; чем &quot;наша приора&quot; но я решил, что лучше взять новое,
чем нарваться на неизвестно откуда взявшуюся &quot;старую иномарку&quot;. Пригнав ее с
салона, был приятно удивлен ходовыми характеристиками, едет плавно, нет резких кренов на
кочках. Гидроусилитель тоже порадовал, руль можно крутить буквально одним пальцем.
Приличная динамика, из отечественных машин вероятно одна из самых &quot;шустрых&quot;. В
общем очень доволен, пока что кроме плановых СТО никаких ремонтов не производил. Общее
впечатление : Машина неплохая, есть недостатки но за такую цену довольно таки неплохо.
</text>
<aspects>
  <aspect category="Whole" from="13" mark="Rel" sentiment="neutral" term="машину" to
="19" type="explicit"/>
  <aspect category="Whole" from="177" mark="Cmpr" sentiment="positive" term=
"иномарка" to="185" type="explicit"/>
  <aspect category="Whole" from="197" mark="Rel" sentiment="negative" term="приора"
to="203" type="explicit"/>
  <aspect category="Whole" from="293" mark="Cmpr" sentiment="negative" term=
"иномарку" to="301" type="explicit"/>
  <aspect category="Driveability" from="345" mark="Rel" sentiment="positive" term=
"ходовыми характеристиками" to="370" type="explicit"/>
  <aspect category="Driveability" from="372" mark="Rel" sentiment="positive" term=
"едет плавно" to="383" type="fct"/>

```

# Pre-trained w2v features

Classifier	Domain			
	Banks		Telecommunication companies	
	Macro (score/rank)	Micro (score/rank)	Macro (score/rank)	Micro (score/rank)
CNN	0.4832 / 21	0.5253 / 21	0.4704 / 41	0.6060 / 36
Two-layer GRU, reversed sequences	0.5517 / 1	0.5881 / 1	0.5594 / 1	0.6569 / 21
Ensemble classifier	0.5352 / 2	0.5749 / 2	0.5403 / 9	0.6525 / 23
Best solution not from our team	0.5252 / 3	0.5653 / 3	0.5493 / 2	0.6822 / 1

[Arkhipenko et al., 2016]