**Semeval-2007 task 14**

Английский

Ссылка: <https://web.eecs.umich.edu/~mihalcea/affectivetext/>

Статья с описанием: [strapparava2007semeval.pdf](https://docs.yandex.ru/docs/view?url=ya-disk%3A%2F%2F%2Fdisk%2FЛитература%2Ffighting%20post-truth%2Femotions%20detection%2Fstrapparava2007semeval.pdf&name=strapparava2007semeval.pdf&uid=541722536&nosw=1)

Постановка задачи: Дан заголовок новостной статьи, по нему нужно определить эмоциональный окрас заголовка (1 число от -100 до 100) и эмоциональный профиль – вектор интенсивностей эмоций читателя (6 чисел от 0 до 100).

Разметка: Разметкой занималось 6 человек с помощью веб-интерфейса. На странице появляется 1 заголовок, 6 слайдеров для эмоций (злость, страх, радость, грусть, удивление), из которых составляется эмоциональный профиль, и ещё 1 слайдер для эмоционального окраса. Эмоциональный профиль состоит из 6 целых чисел от 0 до 100, 0 = эмоция не представлена в заголовке, 100 = максимальная эмоциональная загрузка. Эмоциональный окрас представлен только одним целым числом от -100 до 100, 0 = нейтральный заголовок, 100 = максимально позитивный заголовок, -100 = максимально негативный. Разметчики указывали эмоции, опираясь на слова/фразы с эмоциональным окрасом и общее впечатление от заголовка, а также им советовали следовать их “first intuition” и использовать полный диапазон слайдеров.

Файлы: Папки trial и test – тренировочная и тестовая выборка соответственно. XML-файл содержит заголовки новостных статей из “The New York Times”, “CNN”, “BBC”, “Google news”. В тренировочной выборке: 250, в тестовой: 1000. Файл emotions.gold содержит эмоциональные профили. В файле valence.gold для каждой статьи содержится эмоциональный окрас.

Лучший результат: Bi-LSTM + Attention

Статья: [anoop22readers.pdf](https://docs.yandex.ru/docs/view?url=ya-disk%3A%2F%2F%2Fdisk%2FЛитература%2Ffighting%20post-truth%2Femotions%20detection%2Fanoop22readers.pdf&name=anoop22readers.pdf&uid=541722536&nosw=1)

**REN-10k и RENh-4k (2022)**

Английский

Пока не доступны, но описаны в статье: [anoop22readers.pdf](https://docs.yandex.ru/docs/view?url=ya-disk%3A%2F%2F%2Fdisk%2FЛитература%2Ffighting%20post-truth%2Femotions%20detection%2Fanoop22readers.pdf&name=anoop22readers.pdf&uid=541722536&nosw=1)

Постановка задачи: Дан новостной документ(заголовок + abstract новостной статьи), по нему нужно определить эмоциональный профиль – нормированный вектор интенсивностей эмоций читателя.

Разметка: С помощью Rappler Mood Meter разметчики голосовали за различные эмоции, таким образом получали долю голосов за каждую эмоцию. Выбирались только популярные по голосам статьи.

RENh-4k: содержит 4000 новостных документа (заголовок + abstract) и эмоциональный профиль из 5 эмоций: страх, злость, счастье, вдохновение, грусть. Статьи собирались по темам: здоровье, социальные проблемы и др.

REN-10k: то же самое только 10000 статей, а также статьи содержат больше текста и больше эмоций: страх, веселье, злость, раздражение, безразличие, счастье, вдохновение, грусть.

Лучший результат: Bi-LSTM + Attention

Статья: [anoop22readers.pdf](https://docs.yandex.ru/docs/view?url=ya-disk%3A%2F%2F%2Fdisk%2FЛитература%2Ffighting%20post-truth%2Femotions%20detection%2Fanoop22readers.pdf&name=anoop22readers.pdf&uid=541722536&nosw=1)

**GoodNewsEveryone (2020)**

Английский

Статья с полным описанием: [bostan20goodnewseveryone.pdf](https://docs.yandex.ru/docs/view?url=ya-disk%3A%2F%2F%2Fdisk%2FЛитература%2Ffighting%20post-truth%2Femotions%20detection%2Fbostan20goodnewseveryone.pdf&name=bostan20goodnewseveryone.pdf&uid=541722536&nosw=1)

Постановка задачи: Дан заголовок + abstract новостной статьи, по нему нужно определить эмоцию читателя, внутри текста, выделить в тексте участки, указывающие на того, кто испытывает эмоции, причину, цель эмоции.

Разметка: Датасет с 5000 новостными статьями. Для каждой статьи дан заголовок и ссылка на статью, для них размечены эмоции со стороны писателя и читателя, их интенсивность.

В первой фазе разметки текста собирались статьи. 3 разметчика отвечали на 2 вопроса: какая эмоция доминантная в заголовке (ответ – выбор из 16 вариантов) и вызовет ли заголовок какие-либо эмоции в читателе (да или нет). Статьи автоматически ранжируются по согласованности разметчиков в ответе на 1 вопрос у тех, кто одинаково ответил на 2 вопрос. 5000 лучших по рейтингу заголовков переходят во 2 фазу разметки.

Во второй фазе для каждой статьи участвовало 5 разметчиков (резидентов США), они указывали эмоцию проявляемую в тексте, кто её испытал, событие, которое стало причиной эмоции, по отношению к кому направлена эмоция(если есть), эмоции читателя.

Файлы: Опросники находятся в файлах task1.pdf, task2.pdf. Данные находится в файле gne-release-v1.0.jsonl.

В статье также приводилась baseline модель (Bi-LSTM-CRF) для поиска причины эмоции, кто её испытал, цели и ситуации. (F1-score < 0.5)

**Semeval-2019 task 3**

Английский

Ссылка: [соревнование](https://competitions.codalab.org/competitions/19790#results)

Статья с полным описанием: [ankush19semeval.pdf](https://docs.yandex.ru/docs/view?url=ya-disk%3A%2F%2F%2Fdisk%2FЛитература%2Ffighting%20post-truth%2Femotions%20detection%2Fankush19semeval.pdf&name=ankush19semeval.pdf&uid=541722536&nosw=1)

Постановка задачи: дан текст переписки с контекстом, нужно определить список эмоций пользователя, написавшего последнее сообщение.

Разметка: Все диалоги были размечены 7-ю разметчиками, по итогу для каждого диалога выбиралась эмоция за которую больше всего проголосовали.

Файл: csv-файл, столбцы Turn1, Turn2, Turn3 – сообщения пользователя 1, пользователя 2, пользователя 1 соответственно, Label – эмоция в 3 сообщении.

Один из лучших результатов: BERT + HRLCE

Статья: [huang19semeval2019.pdf](https://docs.yandex.ru/docs/view?url=ya-disk%3A%2F%2F%2Fdisk%2FЛитература%2Ffighting%20post-truth%2Femotions%20detection%2Fhuang19semeval2019.pdf&name=huang19semeval2019.pdf&uid=541722536)