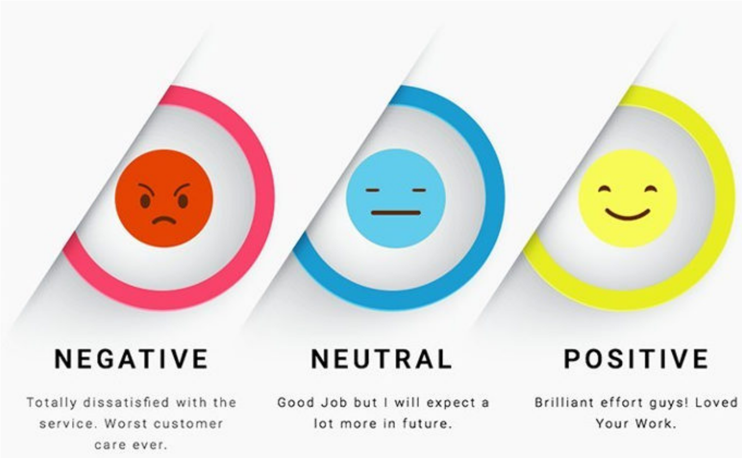
Advances in Sentiment Attitude Extraction task

Nicolay Rusnachenko

Bauman Moscow State Technical University, Moscow, Russia
kolyarus@yandex.ru
nicolay-r.github.io

November 26'th, 2020

Sentiment Analysis



Text classification

The first attempt to propose the task[1]:

$$\langle d \rangle \rightarrow c$$

d – document
c – related class positive, negative

"The picture quality of this camera at night time is amazing"

$$\langle d \rangle \rightarrow positive$$

[1] Peter Turney. "Thumbs Up or Thumbs Down? Semantic Orientation Applied to Unsupervised Classification of Reviews". In: *Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics*. 2002, pp. 417–424.

Targeted sentiment analysis

Considering entity as an input parameter[2]:

$$\langle d, e_j \rangle \to c$$

e_j – object, or entity

"The picture quality of this camera_e
at night time is amazing, especially with tripod_e "

$$\langle d, camera \rangle \to positive \qquad \langle d, tripod \rangle \to ?$$

[2] Long Jiang et al. "Target-dependent twitter sentiment classification". In: *Proceedings of the 49th annual meeting of the association for computational linguistics: human language technologies*. 2011, pp. 151–160.

Aspect Based Sentiment Analysis

Focusing on two core tasks[3]:

1. Aspect extraction;
2. Aspect sentiment analysis:

$$\langle d, e_j, a_k \rangle \rightarrow c$$

a_k – aspect, object characteristics

"The picture quality of this camera$_e$ is amazing ..."[3]

$$\langle d, camera, picture\ quality \rangle \rightarrow positive$$

[3] Bing Liu and Lei Zhang. "A survey of opinion mining and sentiment analysis". In: *Mining text data*. Springer, 2012, pp. 415–463.

Opinion Definition

Defined as follows[3,4]:

$$\langle d, e_j, a_k, h_t, t_l \rangle \rightarrow c$$

h_t – author

t_l – time

[4] Bing Liu et al. "Sentiment analysis and subjectivity.". In: ().

The source of opinion

~~author~~ $\rightarrow$ *object*

Attitude Definition

Opinions between mentioned named entities (e_j, e_m):

$$\langle d, e_j, \color{red}{e_m}, a_k, h_t, t_l \rangle \rightarrow c$$

$\color{red}{e_m}$ – Subject
e_j – Object
(Subject→Object)

Sentiment Attitude Extraction

Example

Text domain: focusing on analytical articles[5];

«As is apparent in Washington_{subj}, there is no place for objectivity on the subject of Russia_{obj}, irrespective of facts and events»

$$(\text{Washington}_{subj}, \mathit{Russia}_{obj}) \rightarrow \text{negative}$$

[5] Natalia Loukachevitch and Nicolay Rusnachenko. "Extracting sentiment attitudes from analytical texts". In: *Proceedings of International Conference on Computational Linguistics and Intellectual Technologies Dialogue-2018 (arXiv:1808.08932)* (2018), pp. 459–468.

Task aspects and problems

1. Large amount of named entities (*NE*):

   Ukraine_e, Russia_e, $\text{Russian Federation}_e$

2. Text structure complexities:

   « Trump_e accused China_e and Russia_e of "playing devaluation of currencies" »

$$(\text{Trump}_{subj}, China_{obj}) \rightarrow \text{negative}$$

$$(\text{Trump}_{subj}, Russia_{obj}) \rightarrow \text{negative}$$

Methods

1. Focusing on Machine learning methods:

$$x_i \rightarrow c_i$$

x_i – input text with opinion
c_i – output label (pos, neg)

2. Methods: will be declared later.

Machine Learning Requirements

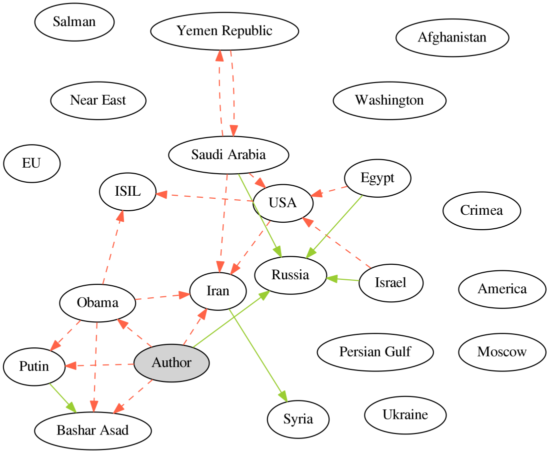
1. Dataset;
2. Document separations: `Train` and `Test`.
   - `Train` – for model training;
   - `Test` – for model evaluation.

Step 1. Manually developed collection

RuSentRel: Contents

- 73 large analytical articles;
- **Text attitudes** – manual annotation, sentiment towards *named entities* (*NE*) as triplets $\langle Subject, Object, Label \rangle$, where:
  - Subject – *NE* or "author"
  - Object – *NE*
  - Label $\in$ {pos, neg}
- **Named Entities** – automatically labeled;
- List *S* of synonymous *NE* – manually implemented.

Document Visualization



Dataset Statistics

- 73 large analytical articles divided into `Train` and `Test` collections (44 in train, 29 in test);

Average per doc.	`Train`	`Test`
Sentences	74.5	137
Attitudes	15	30
Named Entities	**194**	**300**
Named Entities (unique)	33.3	59.9

Attempt #1: How we treat sentiment attitude extraction task

$$\langle d, e_j, e_m, a_k, h_t, t_l \rangle \to c$$

$$e_m \to \text{Subject}$$
$$e_j \to \text{Object}$$
$$a_k \to \text{no aspects}$$
$$h_t \to \text{single author}$$
$$t_l \to \text{current time}$$

Machine Learning Methods

- Conventional Methods – linear and tree-based classifiers (SVM, Random Forest, Gradient Boosting);
- Neural Networks – non linear optimisers.

Question

How to present **input opinion**:

1. *Feature-based* – list of features for document-level opinions;
2. *Context-based* – find sentence with the related attitude participants;

1. Feature-based [Conventional methods]

Participant based:

- The presence in the lists of **countries** or their **capitals**;
- The relative frequency of a NE or the whole synonym group in the document; the order of two named entities;
- Concrete lemmas of named entities **are not used**.

Context based (min, max, and avg values):

- The distance between participants in lemmas;
- Number of commas between the named entities;
- Lexicon-based (vocabulary of entries with preassigned sentiment scores).

2. Context-based [Neural Networks]

1. Introducing **context attitude** – a pair with its named entities (source: Subject, target: Object) in a context

   «Talking about the separation of the Caucasus region$_e$ due to the confrontation between Russia$_{subj}$ and Turkey$_{obj}$ is not necessary, although there is a danger»

2. Additional note: requres convertion from context→document level opinions and vice versa;

Results

Method	Precision	Recall	F-measure
Baseline-neg	0.03	0.39	0.05
Baseline-pos	0.02	0.40	0.04
Baseline-random	0.04	0.22	0.07
SVM	0.09	**0.36**	0.15
Random forest	0.41	0.21	0.27
Gradient boosting	**0.47**	0.21	0.28
Convolutional networks	0.42	0.23	**0.31**
Human labeling agreement	0.62	0.49	0.55

Step 2. Automatically annotated collection

Problems and Motivation

- RuSentRel collection is pretty small;

Lexicons as a Knowledge Base

- Lexicons – vocabulary of pairs $\langle word, label \rangle$;
- We depend on lexicons with more complicated structure, that allows to emphasize the presence of an attitude in context.

RuSentiFrames Lexicon Structure

Describes sentiments and connotations conveyed with a predicate in a verbal or nominal form.

1. Role Designation:
   - A0 is an argument exhibiting features of a Prototypical Agent;
   - A1 is a Theme.
2. Dimentions:
   - the attitude of the author of the text towards mentioned participants;
   - **polarity** – sentiment between participants ;
   - **effects** to participants;
   - mental **states** of participants related to the described situation.

Frame	"Одобрить" (Approve)
roles	A0: who approves A1: what is approved
polarity	A0 → A1, pos, 1.0 A1 → A0, pos, 0.7
effect	A1, pos, 1.0
state	A0, pos, 1.0 A1, pos, 1.0

Table 1: Example description of frame "Одобрить" (Approve) in RuSentiLex lexicon.

Distant Supervision: Lexicons application in data labeling

We apply lexicon to a large news collection, and compose RuAttitudes[6] with the following assumptions:

- News titles usually have a simple structure.

[6] Nicolay Rusnachenko, Natalia Loukachevitch, and Elena Tutubalina. "Distant supervision for sentiment attitude extraction". In: *Proceedings of the International Conference on Recent Advances in Natural Language Processing (RANLP 2019)*. 2019, pp. 1022–1030.

RuAttitudes: Collection of automatically labeled news

We perform news titles annotation in following ways:

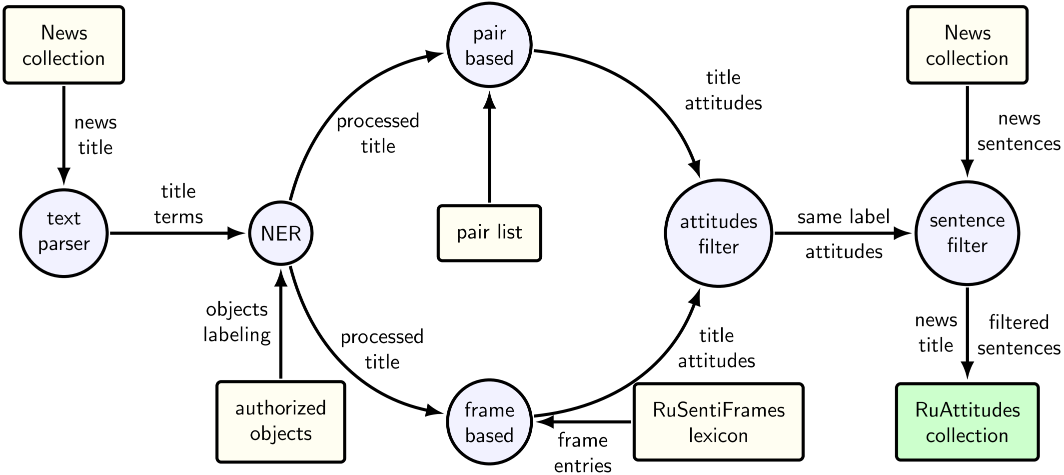
1. **Pair-Based** – attitudes with preassigned labels (using RuSentRel statistics):

$$\langle \mathit{Subject}, \mathit{Object}, \mathit{label} \rangle$$

2. **Frame-Based** – utilizing frame entries from the RuSentiFrames lexicon; matching the following pattern:

$$\ldots \quad \text{Subject}_e \quad \ldots \quad \{\boldsymbol{frame}_{A0 \rightarrow A1}\}_k \quad \ldots \quad \text{Object}_e \quad \ldots$$

News processing workflow



RuAttitudes: News Example

Title
McCain: USA$_e$ continue$_{pos}$ supporting$_{pos}$ Georgia$_e$

↓ USA→Georgia$_{pos}$ (sentence: 5)

« USA$_e$ and in further continue$_{pos}$ support$_{pos}$ freedom, sovereignty and territorial integrity Georgia$_e$ within the internationally recognized borders of the country», – he said.

↓ USA→Georgia$_{pos}$ (sentence: 11)

29'th december prime-minister Kvirikashvili$_e$ reported, that the government of Georgia$_e$ has established first contacts with the new USA$_e$ administration.

Experiments

We additionally adopt the developed collection in training:

- Neural Networks[6];
- Neural Networks with *Attention*[7]: specific module that provide words weighting;

[7] Nicolay Rusnachenko and Natalia Loukachevitch. "Attention-Based Neural Networks for Sentiment Attitude Extraction using Distant Supervision". In: *The 10th International Conference on Web Intelligence, Mining and Semantics (WIMS 2020), June 30-July 3, 2020, Biarritz, France*. 2020. doi: 10.1145/3405962.3405985. url: https://doi.org/10.1145/3405962.3405985.

What is Attention?

Is a Module that provides weighting of terms in context:

leading such a game , $\underline{\mathrm{E}}_{subj}$ will finally $\boldsymbol{lose}_{pos}$ $\boldsymbol{trust\text{-}in}_{pos}$ $\underline{\mathrm{E}}_{obj}$ and country E

Additional source of features during training:

- Which words are related to the class (c) w.r.t. the Object (e_j) and Subject (e_m) (and other terms) of the particular context?

Attempt #2: How we treat sentiment attitude extraction task

$$\langle d, e_j, e_m, a_k, h_t, t_l \rangle \rightarrow c$$

$e_m \rightarrow$ Subject

$e_j \rightarrow$ Object

$a_k \rightarrow$ frames, context words

$h_t \rightarrow$ single author

$t_l \rightarrow$ current time

Results

Proceeding with experiments[7,8]:

Method	Precision	Recall	F-measure
Convolutional networks	0.42	0.23	0.31
Convolutional networks	0.40	0.46	**0.40**
+ Attention	0.42	0.42	**0.41**
Human labeling agreement	0.62	0.49	0.55

[8] Nicolay Rusnachenko and Natalia Loukachevitch. "Studying Attention Models in Sentiment Attitude Extraction Task". In: *Proceedings of the 25th International Conference on Natural Language and Information Systems*. 2020. doi: 10.1007/978-3-030-51310-8_15. url: https://doi.org/10.1007/978-3-030-51310-8_15.

Attention weights analysis

ATT-BLSTM (SL)
leading such a game , E_{subj} will finally ***lose***$_{pos}$ ***trust-in***$_{pos}$ E_{obj} and country E
...
however over the past few months due to combination circumstances E_{subj} gradually ***renew***$_{pos}$ cautions interaction with E_{obj}
...
But E_{subj} consequently emphasizes its ***interest***$_{pos}$ in ***normalizing***$_{pos}$ relationships with E_{obj} (<NUM> february <NUM> year <DOT> took place the visit E at E and its ***conversation***$_{pos}$ with the spiritual leader E and with president E)
ATT-BLSTM (DS)
leading such a game , E_{subj} will finally ***lose***$_{pos}$ ***trust-in***$_{pos}$ E_{obj} and country E
...
however over the past few months due to combination of circumstances E_{subj} gradually ***renew***$_{pos}$ cautious interaction with E_{obj}
...
But E_{subj} consequently emphasizes its ***interest***$_{pos}$ in ***normalizing***$_{pos}$ relationships with E_{obj} (<NUM> february <NUM> year <DOT> took place the visit E at E and its ***conversation***$_{pos}$ with the spiritual leader E and with president E)

$$\ldots \quad \text{Subject}_e \quad \ldots \quad \{\mathit{frame}_{A0 \to A1}\}_k \quad \ldots \quad \text{Object}_e \quad \ldots$$

Conclusion

- Importance of News Titles: in most cases easier to analyse;
- Importance of Lexicons: we may treat them as aspects;
- Conventional methods → Neural Networks → Attentive-Based Neural Networks → ... Language Models[9] (BERT, GPT, etc.);

[9] Jacob Devlin et al. "Bert: Pre-training of deep bidirectional transformers for language understanding". In: *arXiv preprint arXiv:1810.04805* (2018).

Links

http://nicolay-r.github.io