Using CNN for Sentiment Attitudes Extraction from Analytical Texts

N.L. Rusnachenko [1] N.V. Loukachevitch [2]

Bauman Moscow State Technical University, Moscow, Russia
kolyarus@yandex.ru

Lomonosov Moscow State University, Moscow, Russia
louk_nat@mail.ru

25 April, 2018

Introduction

Microbloging posts (Twitter)

- Limited and short in length;
- Mostly user reviews $\Rightarrow$ considered a single object for analysis.

Analytical articles

- Large amount of named entities (*NE*):

  Ukraine, Russia, Russian Federation, ...

- Large amount of attitudes between *NE*:
  - Might take several sentences
- Has complicated structure:

  Donald Trump accused China and Russia of "playing devaluation of currencies"

Related

Text Analysis Conference (TAC), Knowledge Base Population (KBP) track[1]:

- Query-based sentiment retrieval task for e_H – entity holder;
- For queried e_H, find all cases when e_H holds sentiment (pos/neg) relation (from/towards) other entity (target)

MPQA 3.0 [DW15]:

- Sentiment attitudes towards entities and events
- Sentence based annotation

[1] https://tac.nist.gov/2014/KBP/Sentiment/index.html

Dataset

- **RuSentRel**[2][LR18] consisted of analytical articles from Internet-portal `inosmi.ru`;
- Text attitudes – manual annotation, sentiment towards *named entities NE* as triplets $\langle Object, Subject, Label \rangle$, where:
  - Object – *NE* or “author”[3]
  - Subject – *NE*
  - Label $\in$ {pos, neg}
- Named entities – automatic, recognizer based on CRF methods [ML16];
- List *S* of synonymous *NE* – manually implemented.

[2] `https://github.com/nicolay-r/RuSentRel`

[3] In this paper considering named entities only

Dataset Statistics

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

Average per doc.	Training collection	Test collection
sentences	74.5	137
text attitudes (pos.)	6.23	14.7
text attitudes (neg.)	9.33	15.6
NE	**194**	**300**
NE (unique)	33.3	59.9

Table 1 : Statistics of RuSentRel corpus

Sentiment Attitude Extraction

- Introducing **context attitude** – a pair of its named entities $\langle NE_1, NE_2 \rangle$ within a context:

  … US intends to impose sanctions against Russia …

- Consider task as follows: given a **context attitude**, we predict its sentiment label: positive, negative, or **neutral**.
- The act of *extraction* – to select only those of them which were predicted as sentiment (non neutral).

– How to complete a set of context attitudes?
– How to predict labels?

Context attitudes equality

Two context attitudes $a_1 = \langle NE_1, NE_2 \rangle$ and $a_2 = \langle NE_3, NE_4 \rangle$ are *equal up to synonyms* $a_1 \simeq a_2$ when both ends related to the same synonym group $S(\cdot)$:

$$S(NE_1) = S(NE_3) \textit{ and } S(NE_2) = S(NE_4) \tag{1}$$

Context attitude set

- Consider context attitudes extraction within a single sentence

Avg. per doc.	Training collection	Test collection
unique positive	6.23	14.7
unique negative	9.33	15.6
unique neutral	**120**	**276**

Table 2 : RuSentRel text attitudes

Total	Training set	Test set
positive	571	—
negative	735	—
neutral	**6584**	**8024**

Table 3 : Context attitudes amount

Attitudes classification

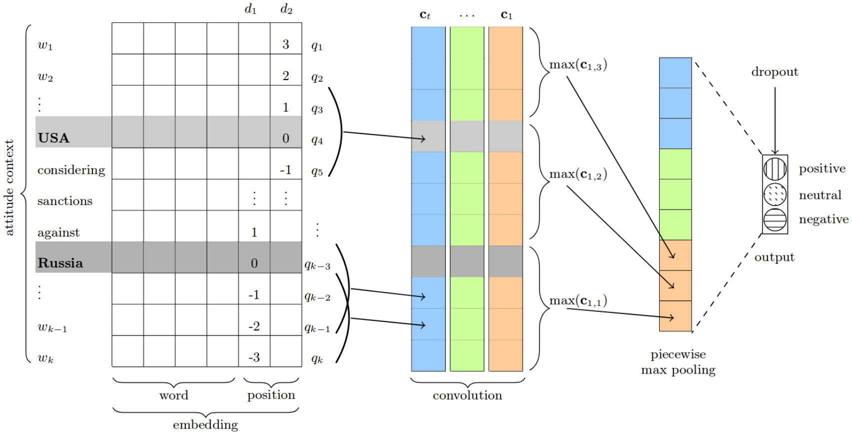


Figure 1 : PCNN[4] [ZLCZ15]

[4] https://github.com/nicolay-r/sentiment-pcnn

Attitudes embedding

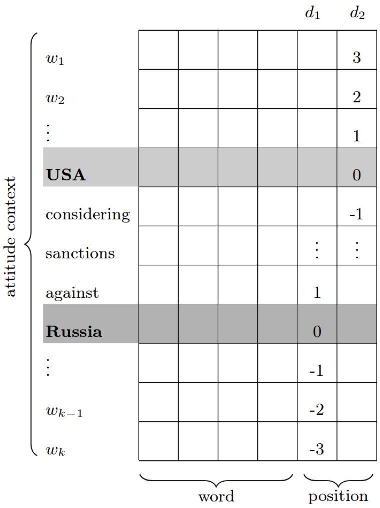


Figure 2 : Attitudes embedding

Convolution

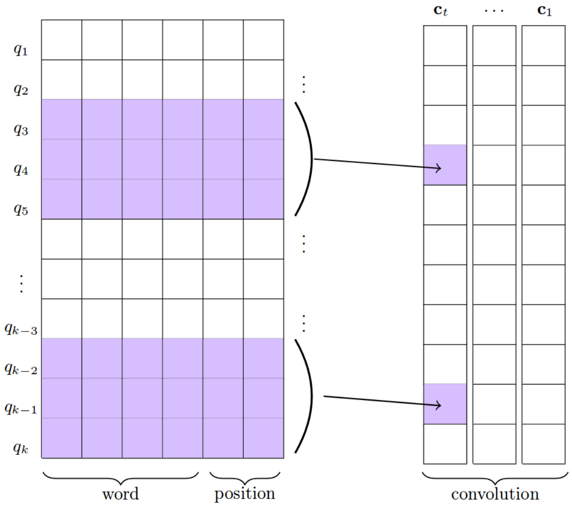


Figure 3 : Convolution

$$w = 3$$

$$m = 4 + 2 = 6$$

$$\mathbf{w} \in \mathbb{R}^{w \cdot m}$$

$$W = \{\mathbf{w_1} \ldots \mathbf{w_t}\}$$

$$c_j = \mathbf{w} q_{j-w+1:j}$$

$$\mathbf{c} = \{c_1, \ldots, c_k\}$$

$$C = \{\mathbf{c}_1, \ldots, \mathbf{c}_t\}$$

Original vs. Piecewise max pooling

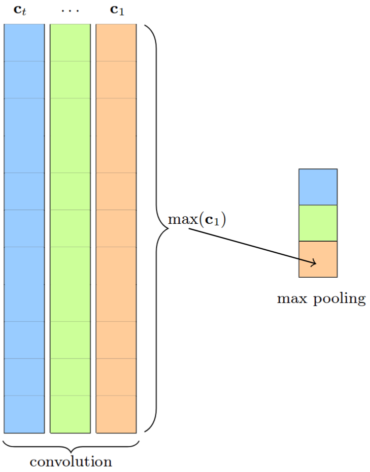


Figure 4 : Original

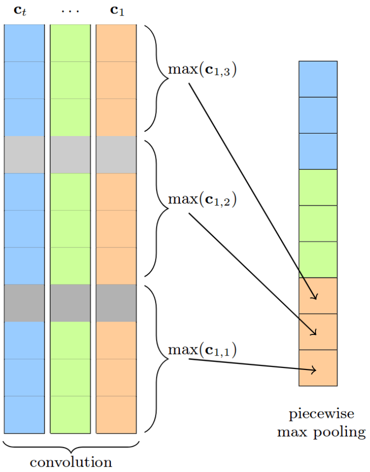


Figure 5 : Piecewise

Output

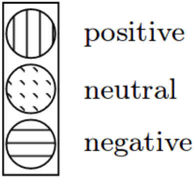


Figure 6 : Output

1. *tanh* maxpool $\mathbf{p}$ activation:

$$d = tanh(\mathbf{p}),$$

2. **Output**[a]:

$$o = W_1 d + b$$

W_1 – hidden layer
b – bias

[a] dropout during training

Training

- **Input** is a sequence of pairs:

$$\{\langle \mathit{embedding}, \mathit{label} \rangle\}$$

- What to train:

$$\{W, W_1, b\}$$

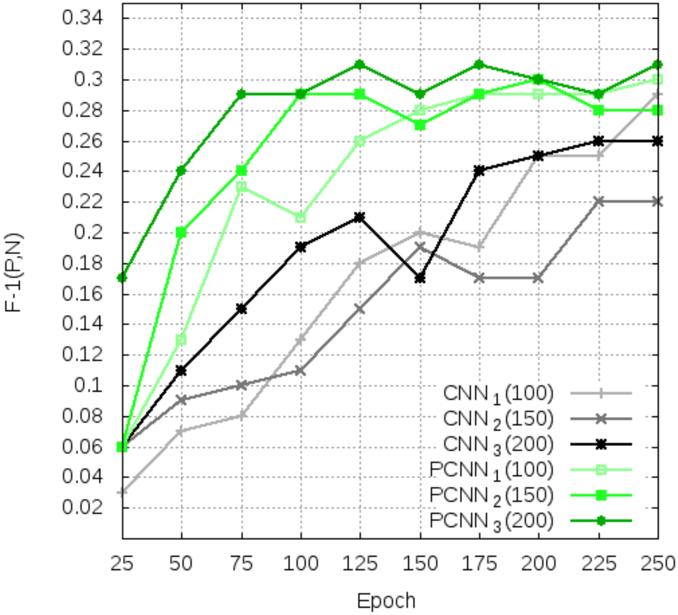
- How to train:
  - Passing batches, *size* $= 50$
  - Error function: *cross-entropy loss*
  - Optimizer: *Adadelta*, $\rho = 0.95, \ \epsilon = 10^{-6}$
  - Use dropout, $\rho = 0.5$

Experiments

Table 4 : Results for sentiment attitudes extraction from RuSentRel corpus

method	*precision*	*recall*	$F_1(P, N)$
Baseline neg	0.03	0.39	0.05
Baseline pos	0.02	0.40	0.04
Baseline distr	0.05	0.23	0.08
KNN	0.18	0.06	0.09
SVM (GRID)	0.09	**0.36**	0.15
Random forest (GRID)	0.41	0.21	0.27
CNN	0.41	0.23	**0.31**
PCNN	**0.42**	0.23	**0.31**
Expert agreement	0.62	0.49	0.55

CNN vs. PCNN



F-1(P,N)
Epoch
0.34
0.32
0.3
0.28
0.26
0.24
0.22
0.2
0.18
0.16
0.14
0.12
0.1
0.08
0.06
0.04
0.02
25
50
75
100
125
150
175
200
225
250
CNN 1 (100)
CNN 2 (150)
CNN 3 (200)
PCNN 1 (100)
PCNN 2 (150)
PCNN 3 (200)

Conclusion

- Proposed CNN-based models significantly outperforms baselines and performs better than NLP-based approaches (Table 4, experiments);
- Best result $F_1(P, N) = 0.31$ is quite low $\Rightarrow$ task still remains significantly complicated;
- Piecewice max pooling prevents from rapid feature reducing $\Rightarrow$ model trains faster (see Figure 16).
- Increasing amount of convolution filters (trainable features) allows model to train faster;

References I

- L. Deng and J. Wiebe, *Mpqa 3.0: An entity/event-level sentiment corpus*, Proceedings of the 2015 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (2015), 1323–1328.
- N. Loukachevitch and N. Rusnachenko, *Extracting sentiment attitudes from analytical texts*, Proceedings of International Conference of Computational Linguistics and Intellectual Technologies Dialog-2018 (2018).
- A. Mozharova, V. and V. Loukachevitch, N., *Combining knowledge and crf-based approach to named entity recognition in russian*, International Conference on Analysis of Images, Social Networks and Texts (2016), 185–195.

References II

Daojian Zeng, Kang Liu, Yubo Chen, and Jun Zhao, *Distant supervision for relation extraction via piecewise convolutional neural networks*, Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing, 2015, pp. 1753–1762.