Using Convolutional Neural Networks for Sentiment Attitudes Extraction from Analytical Texts

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

1. Bauman Moscow State University, Moscow, Russia kolyarus@yandex.ru
2. Lomonosov Moscow State University, Moscow, Russia louk_nat@mail.ru

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) (Query-based sentiment retrieval task for e_H – entity holder)
MPQA 3.0 (Sentiment attitudes towards entities and events) [1]

Dataset

- **RuSentRel v1.0** [2] consisted of analytical articles from Internet-portal `inosmi.ru`;
- Text attitudes – manual annotation, sentiment towards *named entities* NE as triplets $\langle Object, Subject, Label\ (pos, neg)\rangle$,
- Named entities – automatic, recognizer based on CRF methods [3];
- List of synonymous NE – manually implemented.

Table 1: Statistics of RuSentRel v1.0 corpus

Parameter	Train	Test
Number of documents	44	29
sentences (avg./doc.)	74.5	137
mentioned NE (avg./doc.)	194	300
unique NE (avg./doc.)	33.3	59.9
positive pairs of NE (avg./doc.)	6.23	14.7
negative pairs of NE (avg./doc.)	9.33	15.6
Share of attitudes in a single sentence	76.5%	73%
neutral pairs of NE (avg./doc.)	**120**	**276**

Piecewise CNN architecture

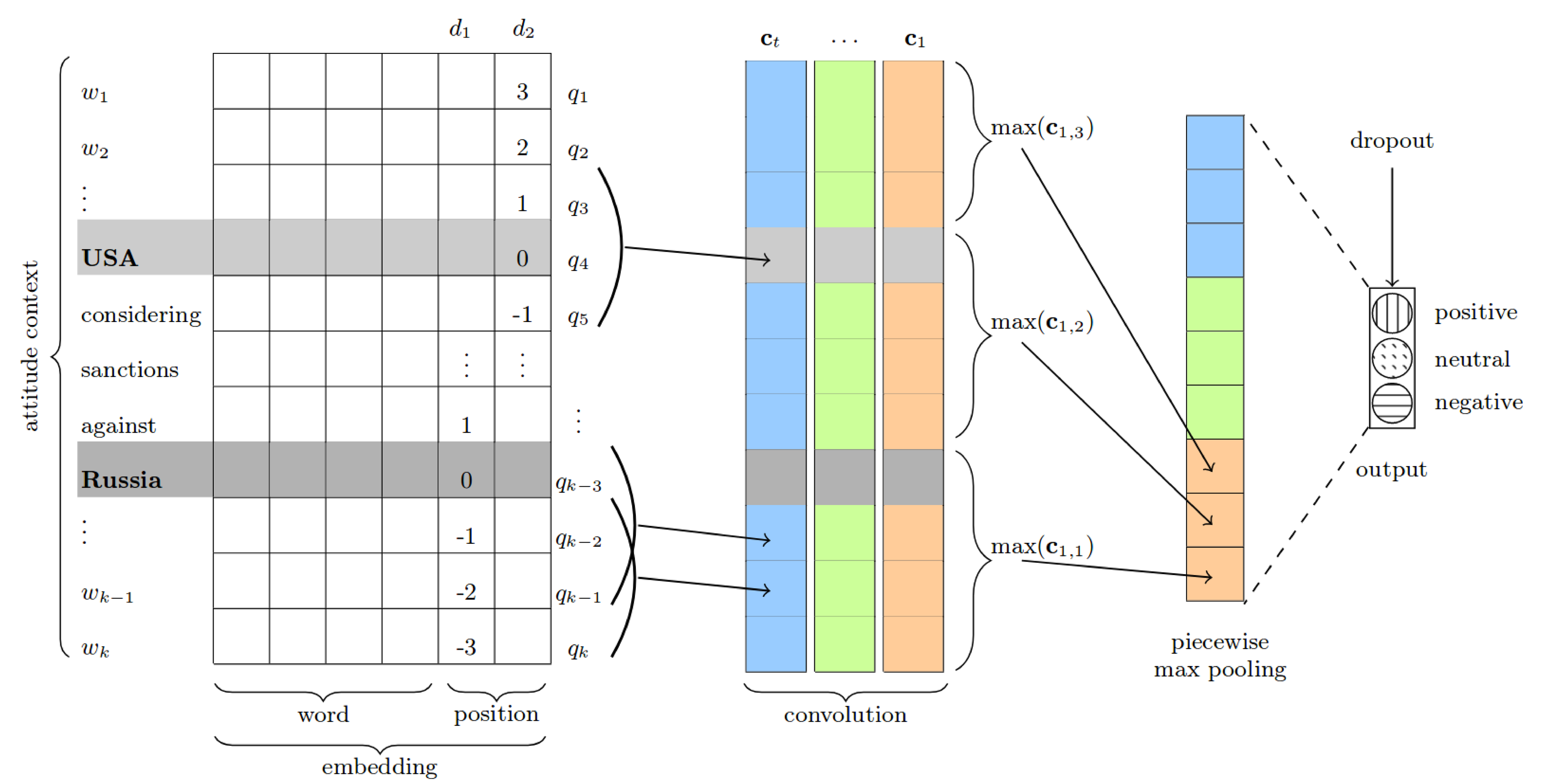

Figure 1: Piecewise CNN architecture [4]

$\{q_1, \ldots, q_k\}$ – word embeddings, $\{\mathbf{c_1}, \ldots, \mathbf{c_n}\}$ – filters. *Piecewise max pooling* – is an opration that applies per each $\mathbf{c_i}$ in pieces: before, after, and between NE's.
Implementation: https://github.com/nicolay-r/sentiment-pcnn

Results

Table 2: Conventional models vs. CNN-based models

Method	P	R	$F_1(P,N)$
KNN	0.18	0.06	0.09
NB Gauss	0.06	0.15	0.11
NB Bernoulli	0.13	0.21	0.16
SVM	0.35	0.15	0.15
SVM (Grid)	0.09	0.36	0.15
Random forest	0.44	0.19	0.27
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

Model Training Comparison

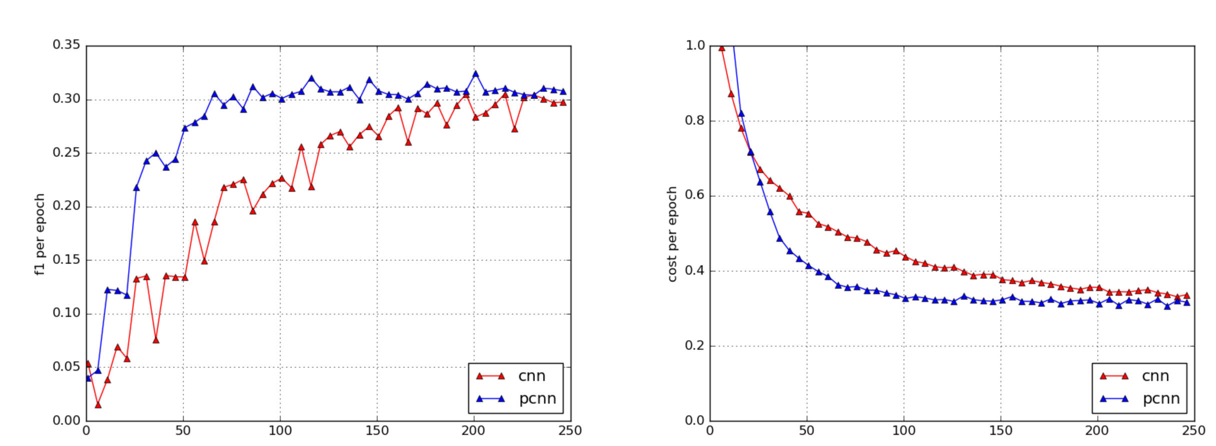

Figure 2: Using default train/test separation of RuSentRel v1.0 collection; $filters = 200$; $window\ size = 3$; *left*: $F_1(P,N)$ results per epoch for test subset; *right*: cost values per epoch; Using piecewise CNN results in training speed, and latter reach better results faster than vanilla CNN

Conclusion

- Proposed CNN-based models significantly outperforms baselines and performs better than NLP-based approaches (Table 2);
- 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 (Figure 2).
- Increasing amount of convolution filters (trainable features) allows model to train faster;

References

[1] Eunsol Choi, Hannah Rashkin, Luke Zettlemoyer, and Yejin Choi. Document-level sentiment inference with social, faction, and discourse context. In *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, volume 1, pages 333–343, 2016.

[2] N. Loukachevitch and N. Rusnachenko. Extracting sentiment attitudes from analytical texts. *In Proceedings of International conference Dialog-2018*, 2018.

[3] 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*, pages 185–195, 2016.

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

Acknowledgements

The work is supported by the Russian Foundation for Basic Research (project 16- 29-09606)