

Transfer Learning for Improving Results on Russian Sentiment Datasets

Anton Golubev

Bauman Moscow State Technical University

Natalia Loukachevitch

Lomonosov Moscow State University

Datasets

- General Sentiment
 - NewsQuotes – ROMIP 2012
- Entity-oriented sentiment
 - Telecom 2015
 - Banks 2015
 - Telecom 2016
 - Banks 2016

Previous results

(Golubev, Lokachevitch, 2020)

Model	Accuracy	F_1 macro	F_1^{+-} macro	F_1^{+-} micro
Telecom 2015	–	–	48.80	53.60
SVM	62.86	58.29	50.27	54.70
CNN	60.80	57.52	49.92	53.23
LSTM	64.46	58.94	52.10	56.03
BiLSTM	65.54	59.35	53.01	56.83
BERT-single	72.48	67.04	58.43	62.53
BERT-pair-QA	74.00	67.83	58.15	62.92
BERT-pair-NLI	74.66	68.24	59.17	64.13
BERT-single (C)	76.55	69.12	61.34	66.23
BERT-pair-QA (C)	76.63	68.54	63.47	67.51
BERT-pair-NLI (C)	76.40	68.83	63.14	67.45
Manual	–	–	70.30	70.90

BERT-Based Models (RuBERT conversational)

Additional sentences

- pair-NLI: *"The sentiment polarity of MASK is..."*
- pair-QA: *"What do you think about MASK?"*

Masking entities for entity-oriented task

- *"Sberbank is a safe place where you can keep your savings."*
- *"MASK is a safe place where you can keep your savings."*

Masking entities for general task

- *"56% of Rambler Group was sold to Sberbank."*
- *"MASK = 56% of Rambler Group was sold to Sberbank."*

BERT-based models

Sentence-single model

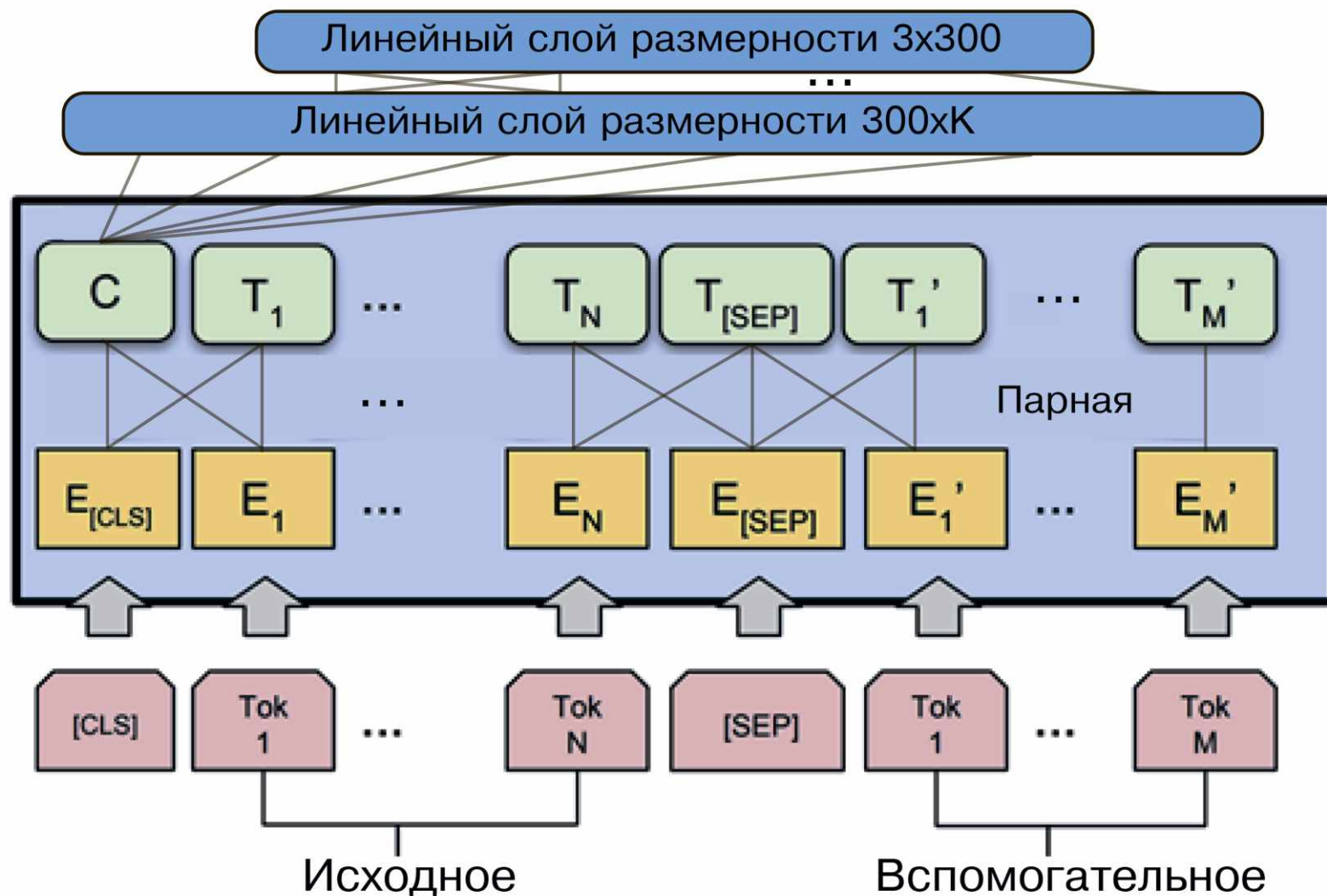
- initial sentence as an input;
- classification layer over the whole sentence on the top;

Sentence-pair model

- *[CLS]*–initial sentence–*[SEP]*–additional sentence;
- classification layer over the final representation of *[CLS]*;

Soft-max function on the last layer

Architecture



Automatically annotated data

- Large Russian news corpus
- Sentiment vocabulary: RuSentiLex
- Automatically labeled data
 - General collection
 - Thematic collection

Automatically labeled general collection

- Extraction of sentences with monosemous positive or negative nouns naming persons
 - Initial words are masked
- 822 negative words, 108 negative words
 - Champion, hero
 - Outsider, swindler, liar
- Example of sentences
 - A MASK is one who, on a **gratuitous** basis, **helps the development** of science and art, provides them with material **assistance** from their own funds
 - Such **irresponsibility** --- **non-payments** --- hits not only the MASK himself, but also throughout the house in which he lives"

Automatically labeled thematic collection

- Extraction of sentences containing target words (banks and communication companies)
- Presence of positive or negative words
 - Distance between entity and sentiment word not more than 4 words
- Sentences not included
 - With negation
 - With quotation
- Neutral sentences
 - Organizations according to named entity recognition module without known sentiment words

Training on created data

- **additional general and neutral thematic data** only and studying dependence of the results on sentiment class distribution;
- additional general and neutral thematic data **mixed** with the benchmark training set;
- **generated data** (the data of previous steps are extended with sentiment-oriented thematic examples) mixed with the benchmark training set;
- **two-step approach**: independent sequential training on additional dataset at the first step and on the benchmark training set at the second step;
- **three-step approach**: independent sequential training in three stages using: automatic general data and, the thematic examples.

Results of two-step training

Dataset	Model	Accuracy	F_1 macro	F_1 micro	F_1 macro
ROMIP-2013	BERT-single	79.95	71.16	85.39	85.61
	BERT-pair-QA	80.21	71.29	85.72	85.93
	BERT-pair-NLI	80.56	71.68	86.14	86.19
	Current SOTA	80.28	70.62	85.52	85.68
SRB-2015 Banks	BERT-single	86.06	79.11	64.87	66.73
	BERT-pair-QA	86.34	79.58	65.29	67.02
	BERT-pair-NLI	87.62	80.72	68.44	71.39
	Current SOTA	86.88	79.51	67.44	70.09
SRF-2015 Telecom	BERT-single	77.11	69.76	61.89	66.95
	BERT-pair-QA	78.14	70.03	64.53	68.29
	BERT-pair-NLI	77.96	69.68	64.52	68.21
	Current SOTA	76.63	68.54	65.47	67.51
SRB-2016 Banks	BERT-single	81.94	74.08	67.24	70.68
	BERT-pair-QA	84.36	77.43	72.32	74.06
	BERT-pair-NLI	84.19	75.63	68.52	70.89
	Current SOTA	82.28	74.06	69.53	71.76
SRF-2016 Telecom	BERT-single	75.82	69.78	65.04	74.22
	BERT-pair-QA	77.25	69.71	67.35	76.22
	BERT-pair-NLI	77.59	69.84	68.11	75.93
	Current SOTA	–	70.68	66.40	76.71

Three-step approach: best achieved results

Dataset	Model	Accuracy	F_1 macro	F_1^+ macro	F_1^+ micro
ROMIP-2013	BERT-single	80.27	71.78	85.82	86.07
	BERT-pair-QA	80.78	72.09	86.14	86.42
	BERT-pair-NLI	82.33	72.69	86.77	87.04
	Current SOTA	80.28	70.62	85.52	85.68
SRE-2015 Banks	BERT-single	87.65	80.79	65.74	67.46
	BERT-pair-QA	87.92	81.12	66.47	68.55
	BERT-pair-NLI	88.14	81.63	68.76	72.28
	Current SOTA	86.88	79.51	67.44	70.09
SRE-2015 Telecom	BERT-single	77.85	70.42	62.29	67.38
	BERT-pair-QA	79.21	70.94	65.68	69.11
	BERT-pair-NLI	79.12	71.16	65.71	70.65
	Current SOTA	76.63	68.54	63.47	67.51
	Manual [12]	–	–	70.30	70.90
SRE-2016 Banks	BERT-single	83.21	75.31	68.45	71.69
	BERT-pair-QA	85.59	78.93	74.05	75.12
	BERT-pair-NLI	85.43	76.85	70.23	72.07
	Current SOTA	82.28	74.06	69.53	71.76
SRE-2016 Telecom	BERT-single	76.79	70.64	66.16	75.27
	BERT-pair-QA	78.42	70.54	68.65	77.45
	BERT-pair-NLI	78.62	71.18	69.36	76.85
	Current SOTA	–	70.68	66.40	76.71

Still difficult examples

- Sberbank of Russia -- 170 years on the queue market
- While we are waiting for a Sberbank employee, I could have gone to lunch 3 times
- Positive to Beeline is not found
 - **MTS** does not work! Forever out of reach. The connection is constantly interrupted. We transfer the whole family to **Beeline**.

Conclusion

- We presented a method for automatic generation of annotated sample from a Russian news corpus using distant supervision technique.
- We compared different options of combining additional data with benchmark train samples and improved current state-of-the-art results by more than 3\% using BERT models together with the transfer learning approach.
- The best variant was three-step approach of sequential training on general, thematic and benchmark train samples with intermediate freezing of the model weights.
- On one of benchmarks, the BERT-NLI model treating a sentiment classification problem as a natural language inference task, reached human level according to one of the metrics.