Advances in Sentiment Analysis of the Large Mass-Media Documents

Nicolay Rusnachenko

nicolay-r.github.io

Newcastle University United Kingdom



Models Data and Finetunning Advances Frameworks and Toolset

Sentiment Analysis

Text classification Targeted sentiment analysis Aspect Level Sentiment Analysis Attitude Definition Sentiment Attitude Extraction Examples

I LOVEEE dogs @beautygirJ5 I love you <3 I enjoyed the food. The game yesterday was intense! @LOLTrish hey long time no see! You put smiles on my face. Today was a good day. I love this notebook!



Positive



Negative

@bigdennis4 nobody asked you! This week is not going as I had hoped life has been like hell... Don't force a joke if it ain't funny I'm learning R programming. So many homeworks !!! Ugh. Can't sleep. Its 1:30am. My Nokia 1110 died.

Models Data and Finetunning Advances Frameworks and Toolset lext classification Targeted sentiment analysis Aspect Level Sentiment Analysis Attitude Definition Sentiment Attitude Extraction Examples

Detalization of Analysis



Models Data and Finetunning Advances Frameworks and Toolset Text classification Targeted sentiment analysis Aspect Level Sentiment Analysis Attitude Definition Sentiment Attitude Extraction Examples

Text classification

The first attempt to propose the task^[1]:

 $\langle d
angle
ightarrow c$

d – documentc – related class positive, negative

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

 $\langle d
angle
ightarrow \textit{positive}$

^[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.

Sentiment Analysis Task Evolution Data and Finetunning Advances

Frameworks and Toolset

Targeted sentiment analysis

Targeted sentiment analysis

Considering entity as an input parameter^[2]:

 $\langle d, \mathbf{e}_i \rangle \rightarrow c$

 e_i – object, or entity

"The picture quality of this camerae at night time is amazing, especially with tripode"

 $\langle d, camera \rangle \rightarrow positive \quad \langle d, tripod \rangle \rightarrow ?$

^[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.

Text classification Targeted sentiment analysis Aspect Level Sentiment Analysis Attitude Definition Sentiment Attitude Extraction Examples

Aspect Based Sentiment Analysis

Focusing on two core tasks^[3]:

- Aspect extraction;
- Aspect sentiment analysis:

 $\langle d, e_j, \frac{a_k}{a_k}
angle o c$

 a_k – aspect, object characteristics

"The picture quality of this camera_e is amazing"^[3]

 $\langle d, \mathit{camera}, \mathit{picture quality}
angle o \mathit{positive}$

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

Data and Finetunning Advances Frameworks and Toolset Text classification Targeted sentiment analysis Aspect Level Sentiment Analysis Attitude Definition Sentiment Attitude Extraction Examples

Attitude Definition

Opinions between mentioned named entities (e_j, e_m) :

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

 a_k – aspect

 $\begin{array}{l} e_m - \text{subject} \\ e_j - \text{object} \\ h_t - \text{author} \\ t_l - \text{time} \\ c - \text{sentiment class (pos, neg)} \end{array}$ " ... Moscow_e dissatisfied with the Warsaw's_e decision ... " $\langle e_m, e_j \rangle \rightarrow \text{neg}$

Text classification Targeted sentiment analysis Aspect Level Sentiment Analysis Attitude Definition Sentiment Attitude Extraction Examples

Sentiment Attitude Extraction

Input:

- **Q** Collection of analytical articles $\langle D_i, E_i \rangle$ (in Russian)
 - Each article includes: document D_i, list of mentioned named entities E_i
- Por synonymous mentions: given a collection of synonyms:

 $Russia_e$, RF_e , Russian Federation_e

Task: For each D_i complete the list of sentiment attitudes (pairs $\langle e_i, e_j, l_{i,j} \rangle$)^[4], with label $l_{i,j} \in \{\text{pos, neg}\}$

^[4] 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.

Models Data and Finetunning Advances Frameworks and Toolset Text classification Targeted sentiment analysis Aspect Level Sentiment Analysis Attitude Definition Sentiment Attitude Extraction Examples

Document-Level Attitude Representation



Models Data and Finetunning Advances Frameworks and Toolset Text classification Targeted sentiment analysis Aspect Level Sentiment Analysis Attitude Definition Sentiment Attitude Extraction **Examples**

Task aspects and problems

- Large amount of named entities (NE);
- 2 Text structure complexities:

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

 $(Trump_{subj}, China_{obj}) \rightarrow negative$ $(Trump_{subj}, Russia_{obj}) \rightarrow negative$

Rule-Based Conventional Classifiers Neural Networks and Embeddings Attention Language Models

Models

Rule-Based Conventional Classifiers Neural Networks and Embeddings Attention Language Models

Approach in Large Document Sentiment Analysis

Contexts as the main idea¹

 Retrieval of attitudes – pos and neg labeling among a set *neutrally* labeled contexts

Prediction:

- Structured output: Text Classification
- Non-structured output: Text Generation

¹ Assumption: a relatively short distance between entities in the text

Rule-Based Conventional Classifiers Neural Networks and Embeddings Attention Language Models

Rule-Based Annotation

Patterns for classification:

• Emoticons^[1], matching words or phrases.

Any algorithm which allows you to perform this annotation.

PROS: fast², minimal amount of RAM to launchCONS: data dynamics

² Relatively in comparison with further methods

Rule-Based Conventional Classifiers Neural Networks and Embeddings Attention Language Models

Conventional Classifiers

- Documents as vectors
- NB, SVM, Random Forest, kNN.
- We can adopt different **kernels** (for the non-linear transformations)
- Every word has a scalar value: Bag-Of-Words



PROS: all text as vector, update. CONS: no connection between words, vectors sparsity

Rule-Based Conventional Classifiers Neural Networks and Embeddings Attention Language Models

Neural Networks (NN) (I)

Words as vectors, or *embeddings*:

• One-hot vector model

 $[0\cdots 0,1,0\cdots 0]$

Classification: $o = W \cdot s + b$

Views of input:

- Windowed (Convolutional NN)
- Sequential (Recurrent NN)
- PROS: non-linear transformations
- CONS: How to establish

connection?



Rule-Based Conventional Classifiers Neural Networks and Embeddings Attention Language Models

Embeddings

Raw documents could be a source of words in contexts $^{\left[5\right] }$

PROS: attempt of domain/general knowledge sharing for AI models, replacement of BoWCONS: time and resources for training on large data

[[]e] Tomas Mikolov et al. "Efficient estimation of word representations in vector space". In: *arXiv preprint arXiv:1301.3781* (2013).



Rule-Based Conventional Classifiers Neural Networks and Embeddings Attention Language Models

Neural Networks with Embeddings



Figure: CNN, Convolution

CONS: limit of window, forgetting information, limit of input in words/tokens

Rule-Based Conventional Classifiers Neural Networks and Embeddings Attention Language Models

Attention mechanism for Machine Translation (MT)

Mechanism for assessing weights of input information, originally for MT^[6]





PROS: widely distributed in other NLP domains, including sentiment analysis

[6] Dzmitry Bahdanau, Kyunghyun Cho, and Yoshua Bengio. "Neural machine translation by jointly learning to align and translate". In: *arXiv preprint arXiv:1409.0473* (2014).

Rule-Based Conventional Classifiers Neural Networks and Embeddings Attention Language Models

Attention for Text Classification^[7]



(a) Context encoder architecture



(b) Quantification of h_j with respect to parameter w [17]; w represents a hidden vector which modifies during model training process

[7] 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. url: https://doi.org/10.1007/ 978-3-030-51310-8_15.

Rule-Based Conventional Classifiers Neural Networks and Embeddings Attention Language Models

Self-Attention

Proposed for the Machine Translation problem^[8]

PROS: Affect on other NLP tasks with different conception of models training, knowledge about language **CONS:** Computation cost $O(N^2)$, where N is an input sequence length

[8] Ashish Vaswani et al. "Attention is all you need". In: Advances in neural information processing systems 30 (2017).

BERT for text classification^[9]

- Pre-training on large amount of data gives us a deep generalized understanding of the language, or language model.
- Text classification: FC-layer application towards the averaged embedded vectors
- Variations: RoBERTa, DistilBERT

Rule-Based Conventional Classifiers Neural Networks and Embeddings Attention Language Models



PROS: Backbone with general knowledge CONS: Input limitation of 512 tokens

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

Rule-Based Conventional Classifiers Neural Networks and Embeddings Attention Language Models

Decoder and Attention Sparsity^[11]

- Generative based: GPT, T5, Longformer, LongT5, BigBIRD
- Text classification: classification layer
- Serialized input/output^[10]

| (a) Full n^2 attention | (b) Sliding window attention | (c) Dilated sliding window | (d) Global+sliding window |
|--------------------------|------------------------------|----------------------------|---------------------------|

PROS: options to train long_input with 4K, 8K, 16K [10] Gaku Morio et al. "Hitachi at SemEval-2022 Task 10: Comparing Graph- and Seq2Seq-based Models Highlights Difficulty in Structured Sentiment Analysis". In: Proceedings of the 16th International Workshop on Semantic Evaluation (SemEval-2022). Association for Computational Linguistics, 2022, pp. 1349–1359. [11] Iz Beltagy, Matthew E Peters, and Arman Cohan. "Longformer: The long-document transformer". In: arXiv preprint arXiv:2004.05150 (2020). 22/54

Supervised Learning Distant Supervision Prompting for fine

Data and Finetunning Advances

Supervised Learning Distant Supervision Prompting for fine

Supervised Learning

By default for AI methods, we consider a training based on manually annotated data by experts



Supervised Learning Distant Supervision Prompting for fine

Supervised Learning Experiments

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

 $(Trump_{subj}, China_{obj}) \rightarrow negative$ $(Trump_{subj}, Russia_{obj}) \rightarrow negative$

Supervised Learning Distant Supervision Prompting for fine

Supervised Learning Experiments

RuSentRel³: articles about Russia's international relations

| Documents | 73 |
|--------------------------------|-------|
| Sentences per document | 105.8 |
| Entities per document | 247 |
| pos and neg pairs per document | 11.47 |

³ https://github.com/nicolay-r/RuSentRel/tree/v1.1

Supervised Learning Distant Supervision Prompting for fine

RuSentRel^[13] Supervised Learning Results, 3-fold cv

| Model | $F_1(P,N)$ |
|-------------------------|------------|
| SentRuBERT | 33.4 |
| AttPCNN _{ends} | 29.9 |
| PCNN | 29.6 |
| Experts agreement | 55.0 |

For MPQA-3.0, $F_1 = 36.0^{[12]}$

^[12] Eunsol Choi et al. "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). 2016, pp. 333–343.
[13] Nicolay Rusnachenko. "Language Models Application in Sentiment Attitude Extraction Task". Russian. In: Proceedings of the Institute for System Programming of the RAS (Proceedings of ISP RAS), vol.33. 3. 2021, pp. 199–222.

Supervised Learning Distant Supervision Prompting for fine

Distant Supervision

Using external knowledge with rule-based or AI pre-trained methods to perform annotation.



PROS: Quick data annotation for further fine-tunning CONS: Noisy labeling

Supervised Learning Distant Supervision Prompting for fine

Distant Supervision Experiments

- **O** News collection: Russian articles from mass-media sources (8.8M);
- Knowledge Base RuSentiFrames⁴: describes sentiment association, conveyed by *predicate* in a form of a verb on noun (311 frames)
 - roles: A0 (agent), A1 (theme);
 - dimensions: authors attitude towards the participants mentioned in text; polarity – score between participants;

| Frame (bragging) | Description | | | |
|------------------|------------------------------|--|--|--|
| entries | bragging, boasting | | | |
| roles | A0: those who bragging | | | |
| | A1: the object of bragging | | | |
| polarity | A0 \rightarrow A1, pos | | | |
| | author \rightarrow AO, neg | | | |

⁴ https://github.com/nicolay-r/RuSentiFrames

Supervised Learning Distant Supervision Prompting for fine

Distant Supervision Application II

Main assumption: news title has a simple structure.

 $\dots \quad \mathsf{Subject}_e \quad \dots \quad \{\mathit{frame}_{\mathtt{A0}\to\mathtt{A1}}\}_k \quad \dots \quad \mathsf{Object}_e \quad \dots$

Distant supervision performed in two steps^[14]:

- Collect the list A of the most-sentiment attitudes (subject \rightarrow object) from news titles using frame A0 \rightarrow A1 polarity across all news titles
- **②** Filter news titles and sentences, which contains at least one pair with A0→A1 score as in A

^[14] 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.

Supervised Learning Distant Supervision Prompting for fine

Frame Title

| Title | | | | | | | | | | |
|--|---|----------------|--------------------|--|--|--|--|--|--|--|
| Tillerson _e : | n _e : USA _e won't remove <i>sanctions_{neg}</i> from Russia _e before the | | | | | | | | | |
| of Crimea _e | | | | | | | | | | |
| | \downarrow USA \rightarrow Russia _{neg} , USA \rightarrow Crimea _{neg} | | | | | | | | | |
| | Most sentiment attitudes | | | | | | | | | |
| Query | Sea | rch results | | | | | | | | |
| USA->Russi | .a _{neg} pai | r found, score | s match; pos: 32%, | | | | | | | |
| | neg: 68% | | | | | | | | | |
| $\mathtt{USA}{ ightarrow}\mathtt{Crime}$ | USA→Crimea _{neg} pair not found | | | | | | | | | |
| \downarrow USA \rightarrow Russia _{neg} | | | | | | | | | | |

| 4 | | | | | | | | | | | |
|--|----------|------|---|--|--|--|--|--|--|--|--|
| Sentence | | | | | | | | | | | |
| Secretary of State USA_e Rex Tillerson _e , speaking in Brussels _e at a meeting | | | | | | | | | | | |
| Foreign _e | heads of | NATO | affiliates stated that the sanctions from Russia | | | | | | | | |
| will only be removed after the return of Crimea _e , according to CNN _e . | | | | | | | | | | | |
| | | 31/5 | | | | | | | | | |

Supervised Learning Distant Supervision Prompting for fine

Datasets

RuAttitudes – automatically marked up collection of texts using the Distant Supervision approach over a large amount of mass-media short news per 2017 year.

| Documents | 134442 |
|------------------------|--------|
| Attitudes per document | 2.26 |

Supervised Learning Distant Supervision Prompting for fine

RuSentRel^[13] distant-supervision results, 3-fold cv

| Model | $F_1(P,N)$ |
|----------------------------|------------|
| SentRuBERT (pretrain + ft) | 37.9 |
| AttPCNN _{ends} | 32.2 |
| SentRuBERT | 33.4 |
| AttPCNN _{ends} | 29.9 |
| PCNN | 29.6 |
| Experts agreement | 55.0 |

Supervised Learning Distant Supervision Prompting for fine

Prompts, prompts, prompts!

Provide additional information that mimicking the expected class or region of text to consider.

- Predefined template: QA, NLI
- Sequence of words mimicking the class^[15]
- With abstract tokens serializing a particular task^[16]

^[15] Taylor Shin et al. "Autoprompt: Eliciting knowledge from language models with automatically generated prompts". In: *arXiv preprint arXiv:2010.15980* (2020).
[16] Xiang Lisa Li and Percy Liang. "Prefix-tuning: Optimizing continuous prompts for generation". In: *arXiv preprint arXiv:2101.00190* (2021).

Supervised Learning Distant Supervision Prompting for fine

BERT with prompts^[17]

Input sequences:

- TextA: Input context terms
- TextB: (Optional), as prompt:

$$\underline{\underline{E}}_{subj}$$
 towards $\underline{\underline{E}}_{obj}$ in « $\underline{\underline{E}}_{subj}$... $\underline{\underline{E}}_{obj}$ » is NEG

Context labeling: FC-layer application towards the averaged embedded vectors



[17] Chi Sun, Luyao Huang, and Xipeng Qiu. "Utilizing BERT for aspect-based sentiment analysis via constructing auxiliary sentence". In: *arXiv preprint arXiv:1903.09588* (2019).

Supervised Learning Distant Supervision Prompting for fine

RuSentRel^[13] distant-supervision results, 3-fold cv

| Model | $F_1(P, N)$ |
|--|-------------|
| SentRuBERT (pretrain + ft) + NLI_{prompt} | 39.0 |
| SentRuBERT (pretrain + ft) | 37.9 |
| AttPCNN _{ends} | 32.2 |
| SentRuBERT | 33.4 |
| AttPCNN _{ends} | 29.9 |
| PCNN | 29.6 |
| Experts agreement | 55.0 |



Official RuSentRel leaderboard

Supervised Learning Distant Supervision Prompting for fine

ChatGPT

Illustrates state-of-the art results in zero-short learning!^[18] We use the following prompt template (NLI format)⁵:

What's the attitude of the sentence "[S]" from "[X]" to the target "[Y]". positive or negative.

Experiments with RuSentRel: classification of sentiment labels

- We receive 1% of unknown answers
- 27% of Neutral classes
- F1 = 68.46 (Complarable with prior results in 2 scale format)

^[18] Bowen Zhang, Daijun Ding, and Liwen Jing. "How would Stance Detection Techniques Evolve after the Launch of ChatGPT?". In: *arXiv preprint arXiv:2212.14548* (2022). 5

Supervised Learning Distant Supervision Prompting for fine

Unknown Response⁶

Input

What's the attitude of the sentence: "She has in Syria two strongest allies Israel and Iran ." from "Israel" to the target "Syria". positive or negative.

ChatGPT

It is not possible to determine the attitude of the sentence towards Syria from the perspective of Israel as the sentence does not express any sentiment or opinion about the relationship between Israel and Syria. It simply states that Israel is one of two allies of Syria

⁶ whether we should consider "allies" as sentiment?

Supervised Learning Distant Supervision Prompting for fine

Unknown Response⁷

Input

What's the attitude of the sentence: "Russia waging a hybrid war against Europe without the military component that is on Ukraine ", - believes klimkin ." from "klimkin" to the target "Ukraine". positive or negative

ChatGPT

It is difficult to determine the specific attitude of the speaker, "klimkin," towards Ukraine based on the provided sentence alone. The statement is primarily focused on describing Russias actions towards Europe, and mentions Ukraine in relation to that. Without more context it is not possible to know the attitude of Klimkin towards Ukraine.

7 Unknown position of the speaker

Supervised Learning Distant Supervision Prompting for fine

Neutral Examples with Explanation

Input

What's the attitude of the sentence: "Iran and Syria are in the same group for the 2018 World Cup qualifying round to be held in Russia : thus , the new mentor of the Syrian team hossam-al-saed to compete for a ticket to Russia with a Portuguese specialist carlos-queiros leading the team iran , the publication" from "iran" to the target "Syria". positive or negative.

ChatGPT

It is neutral, as it is simply stating a fact about the groups for the 2018 World Cup qualifying round and the coaches for the teams of Iran and Syria. It does not express a positive or negative attitude towards Syria.

Supervised Learning Distant Supervision Prompting for fine

Neutral Examples with Explanation⁸

Input

What's the attitude of the sentence:"Belarus included in Customs-Union together with Russia , Kazakhstan , armenia and kyrgyzstan ." from "Belarus" to the target "Russia". positive or negative.

ChatGPT

The attitude of the sentence towards Russia is neutral. It simply states that Belarus is included in a Customs-Union with Russia, Kazakhstan, Armenia, and Kyrgyzstan, without expressing any positive or negative sentiment.

⁸ Different way it treats the expression, however we convey the presence of positive attitude due to the «union»

Supervised Learning Distant Supervision Prompting for fine

ChatGPT alternatives

[TBA] Google Bard⁹ based on LaMDA^[19]

• Transformer pretrained on large amout of chat dialogs

⁹ https://blog.google/technology/ai/bard-google-ai-search-updates/ [19] Romal Thoppilan et al. "Lamda: Language models for dialog applications". In: *arXiv* preprint arXiv:2201.08239 (2022).

Frameworks and Toolset

AREkit - Text Opinion Sampler



AREkit – Document level Attitude and Relation Extraction toolkit for sampling mass-media news into datasets for your ML-model training and evaluation



github.com/nicolay-r/AREkit

AREnets - attention based neural networks



AREnets* – Tensorflow-based framework which lists attentive implementation of the CNN, RNN-based, applicable for Relation Extraction classification tasks as well as API for custom model implementation



github.com/nicolay-r/AREnets

Awesome Sentiment Attitude Extraction



A curated list of works devoted to sentiment attitude extraction of large mass-media documents



github.com/nicolay-r/awesome-...

Conclusion

- The informative analysis of the large Mass-media texts is a granular analysis:
 - $\bullet~$ Text Classification \to Targeted Sentiment Analysis \to Aspect-based Analysis \to Attitude Extraction
- Attitude extraction is considered as text classification problem of small text parts¹⁰
- The latest advances is self-attention which lead us to transformers that can memorize information from massive aomount of the pretrained texts

¹⁰ Generative transformers with the largest input of 16K tokens.

Conclusion

- Rule-based
- Linear classifiers + features
- Neural Networks + embedding + attention + features
- Language Models
- Language Models + prompts
- Zero-Shot Language Models + prompts

The crucial part of optimizations are **prompts**^[20] ... early in a form of features and later closer to output clarification

[20] Shuofei Qiao et al. "Reasoning with Language Model Prompting: A Survey". In: *arXiv preprint arXiv:2212.09597* (2022).

Thank you for attention!



https://nicolay-r.github.io

ARElight – Inference Attitudes

The complete example is available here¹¹. For SentRuBERT¹²:



11 https://raw.githubusercontent.com/nicolay-r/ARElight/main/data/ texts-inosmi-rus/e1.txt 12 http://172.17.0.2/examples/demo/wui_bert.py

SentRuBERT Attention weights analysis

SentRuBERT (head 2, layers from left-to-right: 2, 4, 8, 11)¹³



13 ... playing such a-game, #S finally lost the-trust of #O and countries #E. [SEP] #S to #O in context: "#S has finally lost the-trust of #O" [SEP]

SentRuBERT Attention weights analysis (II)

+ 4 epochs on RuAttitudes¹⁴



14 ... playing such a-game, #S finally lost the-trust of #O and countries #E. [SEP] #S to #O in context: "#S has finally_{frame} lost_{frame} the-trust_{frame} of #O" [SEP]

SentRuBERT Attention weights analysis (III)

 15 + 4 epochs for finetunning on RuSentRel



15 ... playing such a-game, #S finally lost the-trust of #O and countries #E. [SEP] #S to #O in context: "#S has finally_{frame} lost_{frame} the-trust_{frame} of #O" [SEP]

Sampling

In the case of conventional neural networks (frames, features):

| id | doc_id | labe | text | a | s_ind | t_ind | sent_ind | entity_values | entity_types | frames | frame connots uint | syn_subjs | syn_objs entities | pos_tags |
|--------|--------|------|---------|---|-------|-------|----------|--------------------------|---------------|--------|--------------------|-----------|-------------------|--|
| _0i_0o | | D | 0 <[.] | > <[.]> <[.]> <[.]> при это subject неоднократн | > 5 | 19 | 1 | москва,нато,россии | GPE, ORG, GPE | | | 5 | 19 5, 19, 25 | 15, 15, 15, 11, 13, 15, 2, 14, 15, 13, |
| o1_i0_ | | D | 0 <[.]: | > <[.]> <[.]> при это subject неоднократн | > 5 | 25 | 1 | москва,нато,россии | GPE,ORG,GPE | | | 5 | 25 5,19,25 | 15, 15, 15, 11, 13, 15, 2, 14, 15, 13, |
| o2_i0_ | 1 | D | 0 <[.]: | > <[.]> *** лри это object неоднократно | 19 | 5 | 1 | моства,нато,поссии | GDE,ORG,GPE | | | 19 | 5 5,19,25 | 15,15,15,11,13,15,2,14,15,13 |
| o3_i0_ | 1 | 0 | 0 <[.]: | > <[.]> <[> мов. эт. не эднократно подч | 19 | 25 | 1 | I MOVEL HIS 'D, NUCL HIS | O E and and | Er | amoc | 19 | 25 5,19,25 | 15,15,15,11,13,15,2,14,15,13, |
| o4_i0_ | 1 | 0 | 0 obje | ect намер н чести с нк чи против subject | 4 | 0 | (| на.рс эси н | G E PERSON | | antes | 4 | 0.0,4 | 15,12,14,11,15 |
| o5_i0_ | 1 | 0 | 0 <[.]> | > <[.]> <[.]> при это object неоднократно | 25 | 5 | 1 | москва,нато,россии | GPE, ORG, GPE | | | 25 | 5 5,19,25 | 15, 15, 15, 11, 13, 15, 2, 14, 15, 13, |
| 06_i0_ | | D | 0 <[.] | > <[.]> <[.]> при это е неоднократно подч | 25 | 19 | 1 | москва, нато, россии | GPE, ORG, GPE | | | 25 | 19 5, 19, 25 | 15,15,15,11,13,15,2,14,15,13, |
| o7 i0 | | D | 0 subi | lect намерен ввести-санкции против object | 0 | 4 | (| сша россии | GPE.PERSON | 2 | 2 | 0 | 40.4 | 15.12.14.11.15 |

In case of BERT-based language models (TextA, TextB):

| id. | doc | d label | kt a | te | t_b | | s ind | t_ind_sent_in | entity_values | entity_types | entities |
|--------|-----|---------|-----------------------------------|--|--------------------|--|------------------------|---------------|---------------------------------|--------------------------------|-------------------|
| 00_10 | | 0 | марта президент #E #S провел п | ереговоры с лидерами стран ^е #S | к #О в контексте : | << #S провел переговоры с лидерами стран #O >> | 4 | 10 | 3 сша,джо байден,евросоюза, в | GPE, PERSON, ORG, GPE, GPE, OR | 3,4,10,12,21,24 |
| 00_i1_ | | 0 | марта президент #E #S провел п | ереговоры с лидерами стран ^е #S | к ИО в контексте : | << #S провел переговоры с лидерами стран #О в #Е выз# | P 4 | 24 | 3 сша, джо байден, евросоюза, в | GPE, PERSON, ORG, GPE, GPE, OR | 3,4,10,12,21,24 |
| 02_10 | | 0 | марта президент #О #S провел п | ереговоры с лидерами стран #S | к #О в контексте : | << #S #0 >> | 4 | 3 | 3 сша,джо байден,евросоюза, в | GPE, PERSON, ORG, GPE, GPE, OR | 3,4,10,12,21,24 |
| o2 i1 | | 0 | марта президент #О #S провел п | ереговоры с лидерами стран# #S | к #О в контексте : | : << #S провел переговоры с лидерами стран #E в #E вызи | • 4 | 21 | 3 сша, джо байден, евросоюза, в | GPE, PERSON, ORG, GPE, GPE, OR | 3,4,10,12,21,24 |
| o4_i0_ | | 0 | 5 крайне зависим от #О в плане по | ставок нефти и газа. ИS | к ИО в контексте : | << #S крайне зависим от #O>> | 0 | 4 | 4 европейский союз, россии | ORG,GPE | 0,4 |
| o4_i1_ | | 0 | скольку #Е является важным узл | ом транспортировки россий? #5 | к ИО в контексте : | << #S ee конфликт c #O >> | 5 | 13 | 7 украина, европу, россией | GPE,GPE,GPE | 1,9,13 |
| o6_i0_ | | 0 | марта президент #Е #О провел п | ереговоры с лидерами страния | к ИО в контексте : | << #S провел переговоры с лидерами стран #O >> | 10 | 4 | З сша, джо байден, евросоюза, в | GPE, PERSON, ORG, GPE, GPE, OR | 3,4,10,12,21,24 |
| 06_i1_ | | 0 | марта президент #Е #О провел п | ереговоры с лидерами стран##S | к ИО в контексте : | : << #S провел переговоры с лидерами стран #О в #Е выз# | 24 | 4 | З сша,джо байден,евросоюза, в | GPE, PERSON, ORG, GPE, GPE, OR | 3,4,10,12,21,24 |
| 08_10_ | | 0 | марта президент #О ИЕ провел п | ереговоры с лидерами стран##S | к ИО в контексте : | < #S #E провел переговоры с лидерами стран #O >> | 10 | 3 | 3 сша,джо байден,евросоюза,0 | GPE, PERSON, ORG, GPE, GPE, OR | 3,4,10,12,21,24 |
| 08_i1_ | | 0 | марта президент #О #Е провел п | ереговоры с лидерами стран##5 | к ИО в контексте : | : << #S в #E вызвав внимание рынка и предположения о те | P 10 | 21 | 3 сша,джо байден,евросоюза, в | GPE, PERSON, ORG, GPE, GPE, OR | \$3,4,10,12,21,24 |
| 08_12_ | | 0 | марта президент #О #Е провел п | ереговоры с лидерами стран##S | к #О в контексте : | : << #S #E провел переговоры с лидерами стран #О в #E # | 24 | 3 | 3 сша,джо байден,евросоюза, в | GPE, PERSON, ORG, GPE, GPE, OR | 3,4,10,12,21,24 |
| 08_i3_ | | 0 | марта президент #О #Е провел п | ереговоры с лидерами стран #S | к ИО в контексте : | : << #S удалось уговорить #O >> | 24 | 21 | З сша, джо байден, евросоюза, в | GPE, PERSON, ORG, GPE, GPE, OR | 3,4,10,12,21,24 |
| 012_10 | | 0 | сле нача ийско-украин-ког | о конф мкта страны #S одна##S | к #О в контексте : | одна за другой вводят с отношении #O >> | 5 | 12 | 0 запада россии | U C.GPE | 5,12 |
| o13_i0 | _ | 0 | настоящее в ни нь ф ч не ни | / #S / А продолжается, и р* #S | к #О в контексте : | < 15 A+2 -> | 5 | 7 | 6 россией, украинов | G EX PE | 5,7 |
| 013_11 | | 0 | CKORLEY #O D. WICH D' MAN YOU | ом 7 ортировки россий# #5 | к ИО в контексте : | « К знаняет свар ным уз/ зм'т анспортировки российся | 13 | 1 | 7 украина, европу, р Влией | G VE (PE GPI | 1,9,13 |
| o15_i0 | | 0 | О крайно захисии ин иб и плане на | станок ненити и газа. ИЗ | к #О в контексте : | < #S spasified 32million OT #Course | 4 | 0 | 4 европейский союз,россии | ORG, GPE | 0,4 |
| 015_11 | | 0 | скольку #Е является важным узл | ом транспортировки российся #5 | к ИО в контексте : | < #S ee kondynakt c #O >> | 13 | 9 | 7 украина, европу, россией | GPE,GPE,GPE | 1,9,13 |
| 017_i0 | _ | 0 | сле начала российско-украинског | о конфликта страны #О одн##S | к ИО в контексте : | : << #S одна за другой вводят в отношении #O >> | 12 | 5 | 0 запада,россии | LOC, GPE | 5,12 |
| o18_i0 | _ | 0 | марта президент #S #O провел п | ереговоры с лидерами стран##5 | к ИО в контексте : | : << #\$ #0 >> | 3 | 1 4 | 3 сша,джо байден,евросоюза, в | GPE, PERSON, ORG, GPE, GPE, OR | 3,4,10,12,21,24 |
| 018_11 | | 0 | марта президент #S #O провел п | ереговоры с лидерами стран##5 | к ИО в контексте : | : << #S провел переговоры с лидерами стран #E в #E выз# | 21 | 4 | 3 сша,джо байден,евросоюза, в | GPE, PERSON, ORG, GPE, GPE, OR | \$3,4,10,12,21,24 |
| o20_i0 | | 0 | марта президент #S #E провел п | ереговоры с лидерами стран ^е #S | к #О в контексте : | : << #S #E провел переговоры с лидерами стран #O >> | 2 | 10 | 3 сша,джо байден,евросоюза, в | GPE, PERSON, ORG, GPE, GPE, OR | 3,4,10,12,21,24 |
| 020_11 | _ | 0 | марта президент #S #E провел п | ереговоры с лидерами стран ^е #S | к ИО в контексте : | : << #S #E провел переговоры с лидерами стран #O в #E # | • 3 | 24 | 3 сша, джо байден, евросоюза, в | GPE, PERSON, ORG, GPE, GPE, OR | 3,4,10,12,21,24 |
| 020_12 | | 0 | марта президент #S #E провел п | ереговоры с лидерами страния | к #О в контексте : | : << #S в #Е вызвав внимание рынка и предположения о те | 21 | 10 | 3 сша, джо байден, евросоюза, в | GPE, PERSON, ORG, GPE, GPE, OR | 3,4,10,12,21,24 |
| o20_i3 | | 0 | марта президент #S #E провел п | ереговоры с лидерами стран ^е #S | к #О в контексте : | : << #S удалось уговорить #О >> | 21 | 24 | 3 сша, джо байден, евросоюза, в | GPE, PERSON, ORG, GPE, GPE, OR | 3,4,10,12,21,24 |
| 024_10 | _ | 0 | настоящее время конфликт межд | / #О и #S продолжается, и р• #S | к ИО в контексте : | << #S и #О >> | 7 | 5 | 6 россией, украиной | GPE,GPE | 5,7 |
| 024_i1 | | 0 | скольку #S является важным узл | ом транспортировки россий/ #5 | к ИО в контексте : | : << #S является важным узлом транспортировки российся | • 3 | 13 | 7 украина, европу, россией | GPE,GPE,GPE | 1,9,13 |