Computational Linguistics and Intellectual Technologies:
Proceedings of the International Conference “Dialogue 2023”

June 14–16, 2023

The role of Indicators in Argumentative Relation Prediction

Sidorova Elena
A.P. Ershov Institute of
Informatics Systems, Siberian Branch,
Russian Academy of Sciences,
Novosibirsk, Russia
lsidorova@iis.nsk.su

Akhmadeeva Irina
A.P. Ershov Institute of
Informatics Systems, Siberian Branch,
Russian Academy of Sciences,
Novosibirsk, Russia
i.r.akhmadeeva@iis.nsk.su

Kononenko Irina
A.P. Ershov Institute of
Informatics Systems, Siberian Branch,
Russian Academy of Sciences,
Novosibirsk, Russia
irina_k@cn.ru

Chagina Polina
A.P. Ershov Institute of
Informatics Systems, Siberian Branch,
Russian Academy of Sciences,
Novosibirsk, Russia
p.chagina@gmail.com

Abstract

The article presents a comparative study of methods for argumentative relation prediction based on a neural network approach. The distinctive feature of the study is the use of argumentative indicators in the preparation of the training sample. The indicators are generated based on the discourse marker dictionary. The experiments were carried out using an annotated corpus of scientific and popular science texts, including 162 articles available on the ArgNet-Bank Studio web platform. A set of all argumentative relations is described by internal connections of arguments and include the conclusion and the premise. In the first stage of training set construction, fragments of text that included two consecutive sentences were examined. In the second stage, indicators were retrieved from the corpus texts and, for each indicator, statements presumably corresponding to the premise and conclusion of the argument were extracted. In total, 4.2 thousand indicator-based training contexts and 13.6 thousand pairs of sentences were obtained from the corpus with annotation of the presence of an argumentative relation. Based on this training sample, four classifiers were built: without indicators, with marking indicators in sentences using tags, taking into account segmentation of text based on indicators, with segmentation and tags. The results of the experiments on argumentative relation prediction are presented.

Keywords: argument mining; text corpus; argument annotation of text; argumentation indicator; argument scheme; argumentative relation prediction

DOI: 10.28995/2075-7182-2023-22-477-485

Исследование роли индикаторов при извлечении аргументативных отношений

Сидорова Елена
Институт систем информатики
им. А.П. Ершова СО РАН,
Новосибирск, Россия
lsidorova@iis.nsk.su

Ахмадеева Ирина
Институт систем информатики
им. А.П. Ершова СО РАН,
Новосибирск, Россия
i.r.akhmadeeva@iis.nsk.su

Кононенко Ирина
Институт систем информатики
им. А.П. Ершова СО РАН,
Новосибирск, Россия
irina_k@cn.ru

Чагина Полина
Институт систем информатики
им. А.П. Ершова СО РАН,
Новосибирск, Россия
p.chagina@gmail.com

Аннотация

В статье проводится сравнительное исследование методов извлечения аргументативных отношений на основе нейросетевого подхода. Особенностью исследования заключается в использовании индикаторов аргументации при подготовке обучающей выборки. Индикаторы сгенерированы на основе словаря дискурсивных маркеров и задаются набором лексико-синтаксических шаблонов. Для экспериментов использовался размеченный корпус научных и научно-популярных текстов, включающий 162 статьи, размещенные на веб-платформе ArgNetBank Studio. Множество всех аргументативных отношений описываются внутренними связями аргументов и включают заключение и посылку. Построение обучающей выборки проходило в два этапа. На первом этапе рассматривались фрагменты текста, включающие два подряд идущих предложения, и отмечалось наличие или отсутствие аргументации. Считалось, что аргументация присутствует, если фрагмент включал заключение и хотя бы одну посылку одного и того же аргумента из разметки. На втором этапе осуществлялся поиск индикаторов и для каждого индикатора извлекались утверждения, предположительно соответствующие посылке и заключению аргумента. Каждый такой набор размечался аналогично по наличию аргументативного отношения в аннотации. Всего на основе корпуса было получено около 4,2 тысяч обучающих контекстов на основе индикаторов и 13,6 тысяч пар предложений с разметкой наличия аргументативной связи. На основе данной обучающей выборки было построено четыре классификатора: без учета индикаторов, с разметкой индикаторов в предложениях с помощью тегов, с учетом сегментации текста на основе индикаторов, с сегментацией и тегами. Приведены результаты экспериментов по извлечению аргументативных отношений.

Ключевые слова: анализ аргументации; корпус текстов; аргументативная разметка текста; индикатор аргументации; схема аргумента; извлечение аргументативных отношений

1 Introduction

Over the past two decades, the study of argumentation involves, in particular, describing the structure of a text in the form of statements connected by relations of support or conflict. Argument Mining is a field of computational linguistics, which has been actively developing during the last decade. Its goal is to automatically extract arguments represented by a sequence of statements ("premises") leading to a certain conclusion ("thesis") from texts. Automating the extraction of arguments from texts became a priority area only a few years ago [8].

The analysis of argumentation presented in a natural language text requires not only the extraction of arguments and argument chains supporting or disproving a thesis (abstract argumentation), but also the exploration of the structure of each argument and its role and relevance to the argument as a whole (structural argumentation). Models or schemes of arguments are used to describe different ways of reasoning [16]. The best known compendium of structured argumentation that has found application in practical systems of argument analysis is that of D. Walton [18]. It contains about 60 argumentation schemas, based on which an ontology of argumentation (AIF-ontology) was constructed in [11].

One of the main conditions for the development of this field is the creation of corpuses of texts with argumentative annotation. The best known resource with argumentation annotation is the AIFdb database, formerly the Araucaria corpus [7], which contains news articles, records of parliamentary and political debates, etc. — a total of 170 corpora of varying size and quality in 14 languages. However, the main research languages are still English and, to a lesser extent, German, and the data themselves have different annotation schemes, making them impossible or very difficult to use combined. There are very few such resources for the Russian language. The annotated corpus of sentences with annotation of the presence of argumentation ("for" or "against") was developed as part of the RuARG-2022 competition [6]. In [2] a web-based resource for the analysis of argumentation in popular science discourse is presented. The annotation model is based on the ontology of argumentation and D. Walton's argumentation schemes [18].

An important linguistic aspect of the study of discourse is the registration of discourse markers - linguistic instruments of structuring discourse, which play a key role in the process of its understanding. Thus, indicators of argumentation simplify the identification and reconstruction of the steps of argumentation that are carried out in an argumentative dialogue or text [5]. The aim of our work is to investigate the role of indicators in detecting argumentative relations and evaluating their effectiveness. The main research tools are annotated text corpora and dictionaries of indicators of argumentative relations.

2 Related works

The solution of Argument Mining (AM) task involves solving the following subtasks, which can be formulated as classification problems:

1. Detection of text fragments containing argumentation (Argument Detection);
2. Classification of statements according to the used argumentation scheme (Argument Component Classification);
3. Identification of relationships between argument components (Argumentative Relation Prediction);
4. Classification of arguments according to the classes presented in the ontology of argumentation (Argument Classification).

According to the multiple reviews [8, 13-14, 17, 19] it is clear that modern pre-trained Deep Learning models (DL), such as BERT, have shown good results on many AM related tasks and they are currently one of the main tools in the field of AM. The subtask that is called either Edge Prediction or Relation Prediction is considered the most difficult part of Argument Mining. Currently there are not so many papers dedicated to the applying of modern NLP techniques to the Relation Prediction problem. The results demonstrated by modern DL models, however, remain comparable to the results of classical models, such as, for example, SVM. It was shown in [3] that, despite having a superior performance on the Argument Component Classification problem with F1 score = 0.86 against 0.79 as the best of the other models, the BERT-based (BERT-base-uncased) model was inferior in performance on the Argumentative Relation Prediction (ARP) task. Trained on the CDCP corpus it obtained F1 score = 0.15 against 0.34 of the LSTM-based model [9] and 0.27 of the SVM with GloVE embeddings as an input.

Lexical features are applied when teaching classical ML models in Argument Detection and Argument Component Classification tasks [14]. When analyzing argumentation in Russian-language texts, it is necessary to study the composition, structure, and role of both primary and secondary connectors of the Russian language [15] used as indicators of argumentation.

From the analysis of recent works we can conclude that the problem of Argumentative Relation Prediction is far from being solved, and, depending on the data and their annotation, a broad range of modern techniques can be applied: from traditional ML models with various features to DL models, and lexical and syntactic features are an important part of the training of classical models. Also, works on Argument Mining do not pay enough attention to the role of lexical features, such as indicators of argumentation (markers) and n-grams, genre segmentation of the text, and the possibility to apply knowledge of the rhetorical structure of the text. For Russian, this problem is even more relevant due to the small amount of annotated data.

3 Corpus of texts with argumentation markup

For this study we used the annotated corpus of scientific and popular science texts, including 162 articles available on the ArgNetBank Studio web platform (https://geos.iis.nsk.su/arg). Each text was annotated according to the AIF (Argument Interchange Format) standard [4], by constructing an oriented connected graph (see Fig. 1) with two types of vertices: information vertices, which correspond to the statements (rectangular blocks), and relation vertices, which indicate the connections between the statements (oval blocks).

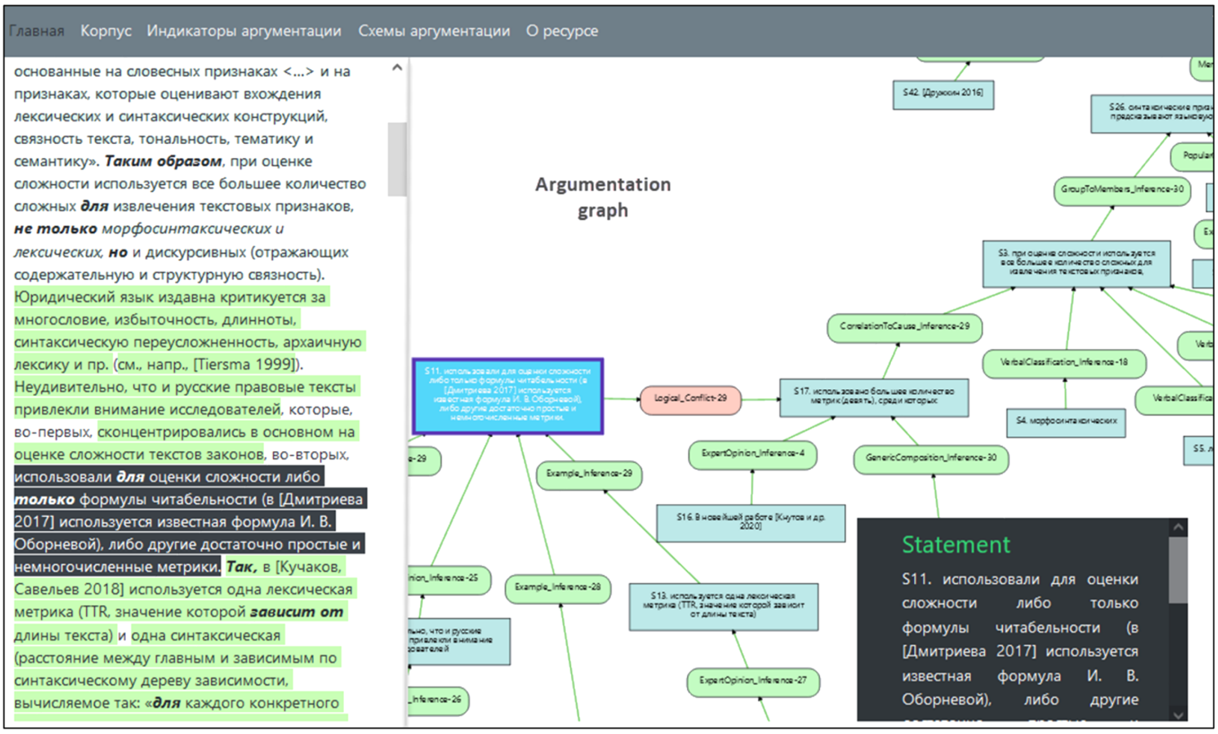

Figure 1: Argumentative text markup on the ArgNetBank Studio platform

A set of all argumentative relations in a graph is described by internal connections of arguments according to the typical models (schemes) of reasoning. Each argument has a conclusion and a premise (or several premises) in the case of support or conflict. The annotated corpora can be downloaded in json format and used in further research. The corpus articles belong to different subject areas and are relatively small (3,500 words on average). A higher-quality, expert-tested sub-corpus was selected from the corpus, from which texts were taken for testing. Training was performed on the entire corpus.

Argumentation indicators play an important role in the research. On the ArgNetBank Studio platform to bring the user's attention to the arguments presented explicitly in the text and to assist in highlighting the boundaries of statements in the text and in choosing the scheme of argumentation, a preliminary linguistic processing of texts is performed, which reveals specific cues in the text expressed as various kinds of verbal cliches, which are potentially indicators of argumentation (Fig. 1). The list of significant words which can act as indicators or anchor words of indicator constructions is heterogeneous and fundamentally incomplete. Therefore, the platform contains additional tools for finding, exploring and creating indicators (Fig. 2).

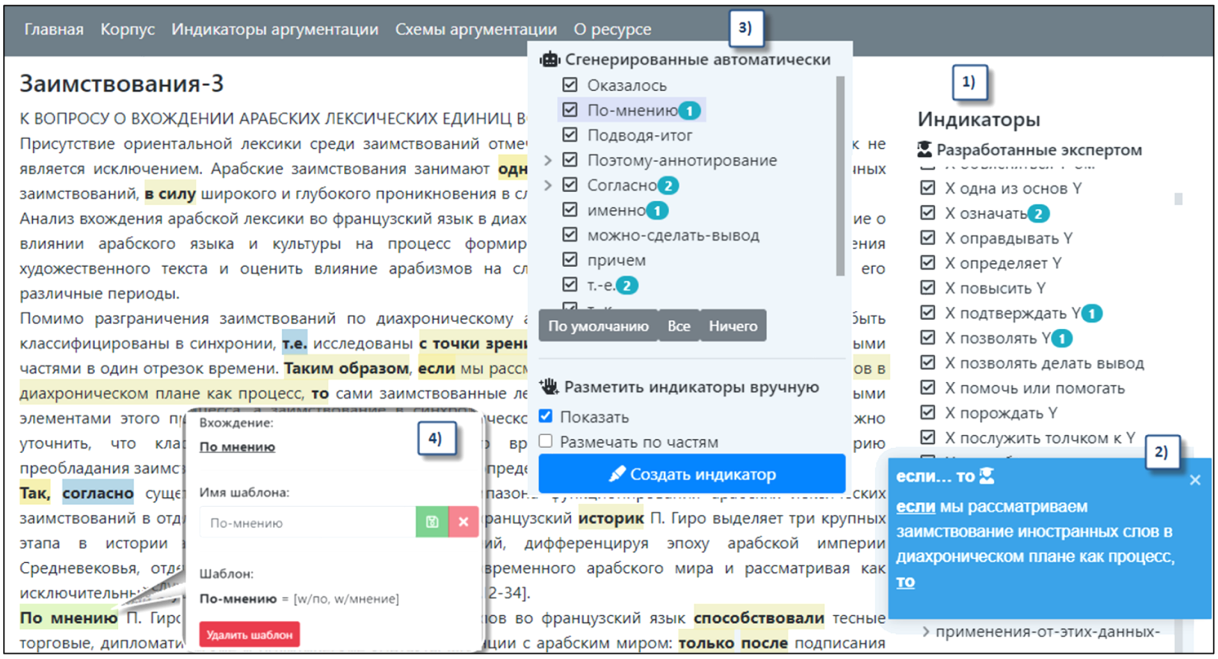

Figure 2: Argumentation indicators annotation on the ArgNetBank Studio platform

Developed tools provide loading dictionaries of lexico-syntactic patterns for search of indicators of argumentation (1), indicators annotation in the text (2), and concordance creation. On the basis of the fragment selected by a user (3), a pattern is automatically generated (4), in which the structure and normalized form of lexical units, punctuation marks and gaps (if the user has highlighted a splitted fragment) are captured.

4 Indicators of argumentation analysis

Indicators of argumentation are words and constructions used in a discourse that indicate the presence of an argument in the text. They help identify the presence of arguments and their components, identify the boundaries of statements in the text, reconstruct the relations between statements, and relate the argument to a certain scheme of reasoning (a form of deduction that expresses the relationship between premises and conclusions).

The indicator can signal different pragmatic aspects of argumentation [12]:

1. the degree of confidence the author has in the statement: *по-видимому ‘seemingly’, уверен ‘sure that’*;
2. the relation of inference between two statements (presence of argumentation): *следует что ‘it follows that, если…то ‘if…then’*;
3. the type of argumentative relation: *поскольку ‘due to’* (support) *vs. хотя ‘although’*(conflict);
4. the role of the statement in the inference: *потому что ‘because of’* (premise*) vs. поэтому ‘that’s why’* (conclusion);
5. the semantic-ontological relation on which the typical scheme of reasoning is based in this case: *по причине ‘by reason of’, X вызывает Y ‘X causes Y’* (causation), *в частности ‘in particular’, например ‘for example’* (hyper-hyponymy), *похожий ‘similar’* (analogy);
6. the structure of argumentation (multiple vs. sequential argumentation) *к тому же ‘besides’, не говоря уже о ‘not to mention’* (multiple argumentation) vs. *в конце концов ‘eventually’* (sequential argumentation).

The original list of discourse markers contained 294 items, from which a list of 143 markers was manually selected. This list was also extended with previously developed indicators for expert opinion reasoning extraction.

Indicators are described in a formal pattern language that allows the use of tokens, arbitrary character sequences, auxiliary patterns, alternatives and gaps.

All markers and their contexts of use were extracted from the corpus of texts in order to study the indicators. For each marker, the context of its use was divided into three statements: the main statement, which included the indicator, and the right and left contexts. For each statement, its role in the structure of the argument was identified. Thus, the data have the following representation:

pattern | main | left | right | main_arg | left_arg | right_arg | same_arg | text | sent_n

where:

- pattern name (*pattern*) - name of the anticipated indicator;
- main statement (*main*) - the sentence containing the marker;
- left context (*left*) - part of the sentence preceding the indicator entry; if the marker is close to the beginning of the sentence, the sentence preceding the main statement (if any) is also taken;
- right context (*right*) - the sentence following the main statement;
- argumentation parameters for the main (*main_arg*), left (*left_arg*) and right (*right_arg*) statements - presence of argumentation and roles in the argument structure, which take values: 0 no argumentation, 1 the premise of the argument, 2 the conclusion of the argument, 3 the premise in one argument and the conclusion in another;
- binary argumentative relation parameter (*same_arg*) - the presence in the main statement and the left context of the premise and conclusion of the same argument (in any order), which indicates the presence of an argumentative relation and implicitly means that the marker is a true indicator of argumentation;
- *text* - reference to the text where the marker was encountered;
- sentence number (*sent_n*) - reference to the sentence in which the marker was encountered

A total of 4,207 patterns and their contexts were obtained from the corpus. Of these, 972 cases contained an argumentative relation. In other words, in only 23% of the cases the marker was a connector between the premise and the conclusion of an argument.

In addition, there were 1,496 cases of simultaneous occurrence of a premise and a conclusion in the same statement, which corresponds to the situation of sequential argumentation, i.e., when the statement is an intermediate (non-leaf) vertex in the argumentation graph.

In terms of identifying the boundaries of argumentative structures, indicators can be divided into the following functional groups:

1. Patterns that break a single sentence, containing a premise and a conclusion, into parts and specify the boundaries of statements.
   Эти варианты различны для разных видов контаминированной речи, ***например*** *воспроизведение английской или русской речи немца не похоже на передачу речи китайца.*
   'These variants are different for different types of contaminated speech, ***for example****, the reproduction of English or Russian speech by a German is not similar to the transmission of speech by a Chinese.'*
2. Patterns that are on the edge of a sentence and signal that the nearest sentence is part of the argument structure.
   Свою семантическую значимость пропозиция обретает только в рамках высказывания. ***Поэтому*** *необходимо обратиться к вербальным способам актуализации пропозиций победы в исследуемых текстах.*
   'A proposition acquires its semantic significance only within the framework of an utterance. ***Therefore****, it is necessary to turn to verbal ways of actualizing the propositions of victory in the texts under study.'*
3. Patterns with a gap that contains either a conclusion or a premise within it.
   Но тот факт, что радионуклид был выявлен на такой обширной территории, говорит о том, что активность в выбросе была весьма высокой.
   'But the fact that the radionuclide was detected over such a vast area suggests that the activity in the release was very high.'

An analysis of the relative positions of premise and conclusion with respect to the marker showed that most indicators (about 90%) allow us to accurately indicate which of the context statements will play the role of premise or conclusion in the case of argumentation detection. Thus, the use of the indicator looks promising, both for improving the quality of argumentative relations extraction, and for postprocessing involving the identification of the roles of statements in the argument structure.

5 Argument relation prediction

The training set construction consisted of two stages. In the first stage, fragments of text that included two consecutive sentences were examined and the presence or absence of argumentation was noted. Argumentation was considered to be present if the fragment included a conclusion and at least one supporting or refuting premise of the same argument from the annotation. In the second step, indicators were retrieved from the corpus texts and, for each indicator, statements presumably corresponding to the premise and conclusion of the argument were extracted. Each such set was annotated similarly by the presence of an argumentative relation in the annotation. In total, 4,207 indicator-based training contexts and 13,655 pairs of sentences were obtained from the corpus with annotation of the presence of an argumentative relation. Thus, the data for the experiments included about 18 thousand examples, of which 2,617 were positive examples and about 15,5 thousand were negative examples.

The ruRoberta (ai-forever/ruRoberta-large) model was used to represent the Russian text, where the two contexts are provided as input separated by the special token [SEP]. We use encoding output for the [CLS] token as the relation representation between two contexts. Then a fully connected neural network consisting of two linear layers with a ReLU activation function and a dropout layer between them is applied to the representation. Finally, a Softmax function was used to obtain the probability distribution of the argumentative relation. We used the following configurations to construct classifiers that predict the presence of argumentative relations.

1. Independent classification (**simple-model**): the classifier is applied to embeddings of sentences obtained by a sliding window of 2 sentences.
2. Independent classification (**simple-indicator-model**): the difference from simple-model is that additionally the argumentation indicators are marked with a special punctuation mark (`*`) similar to the work [1].
3. Classification taking into account the segmentation based on indicators (**context-model**): the classifier is applied to the statements obtained as the left and main indicator contexts; in the absence of an indicator, the partitioning is performed on sentences.
4. Classification with marking indicators (**context-indicator-model**): the difference from context-model is that additionally the argumentation indicators are marked with a special punctuation mark (`*`).

We carried out 5-fold cross validation over our dataset, with the same parameters used for all models in the process: learning rate = 3e-7, batch size = 4, epochs = 5. The results of the experiments are presented in Table 1.

Classifier	Precision	Recall	F1
simple-model(1)	19.87	51.94	28.55
simple-indicator-model(1)	19.85	46.04	27.63
context-model(1)	20.70	**65.90**	31.30
context-indicator-model(1)	**21.32**	65.25	**31.95**
simple-model(2)	41.38	53.13	46.31
simple-indicator-model(2)	41.20	54.14	46.60
context-model(2)	43.47	**66.65**	52.29
context-indicator-model(2)	**44.33**	66.48	**52.86**

Table 1: The results of the experiments

Analysis of the results of experiment (1) reveals problems with the quality of the corpus annotation (a high-quality annotation constitutes only about half of all the annotated data) and the problem of disagreement between different annotators. For the part of the corpus annotated by several experts, the agreement was 0.78 for annotating argumentative statements and 0.55 for annotating argumentative relations. Compared to the results of other studies [8] (for non-experienced annotators k = 0.58, and for experts k = 0.83) the data give worse results, which seems to be related both to the complexity of the annotation scheme and to the studied genre itself.

To solve this problem, the dataset was further processed to remove "badly" annotated texts: texts with abnormally low argumentation coverage were removed. The results of experiment (2) show a stable improvement in the quality of all classifiers.

Overall, the experimental data show that on this corpus, the use of indicators improves the quality of the classifiers performance on all three metrics. And segmentation based on indicators is more effective than simply marking indicators.

6 Conclusion

In this paper we continued our investigation of the role of indicators in argument extraction. While previously we considered only the problem of sentence detection, in this study the focus was on identifying the argumentative connection between two statements. The distinctive features of the applied approach include a) the study of Russian-language texts of scientific and popular science genre, b) the use of a corpus annotated according to one of the most difficult for automatic processing standards of argumentative annotation, c) the construction of one universal classifier instead of a chain of classifiers used consistently to solve the problem [8], d) the integration of the indicator approach with deep learning methods. Additionally, we have taken into account the drawback associated with the exclusion from consideration of text fragments that do not contain indicators.

Thus, further research will be related to the study of the following issues: a) improving the quality of annotation by developing annotation methodology for texts of scientific and popular science genres; b) enriching and refining the vocabulary of argumentation indicators; c) developing independent classifiers that identify whether a marker is an indicator in a given context; d) exploring the role of indicators for classifying argumentation schemes.

Acknowledgements

The work was funded by Russian Science Foundation according to the research project no. 23-21-00325.

References

[1] Alibaeva K., Loukachevitch N. Analyzing COVID-related Stance and Arguments using BERT-based Natural Language Inference // Computational Linguistics and Intellectual Technologies: Proceedings of the International Conference “Dialogue 2022”, 2022. – P. 8–17.

[2] Ilina Daria, Kononenko Irina, Sidorova Elena. On Developing a Web Resource to Study Argumentation in Popular Science Discourse. In Computational Linguistics and Intellectual Technologies // Proceedings of the International Conference “Dialog-2021”. 2021. — P. 318–327.

[3] Chen Ting, BERT Argues: How Attention Informs Argument Mining (2021). Honors Theses. 1589.

[4] Chesñevar C.I., McGinnis J., Modgil S., Rahwan I., Reed C., Simari G., South M., Vreeswijk G., Willmott S. Towards an argument interchange format. — The knowledge engineering review, 21(4), 2006. — P. 293-316.

[5] Eemeren F.H. VAN, Houtlosser P., Snoeck Henkemans F. Argumentative Indicators in Discourse: A Pragma-Dialectical Study. — Dordrecht: Springer. 2007.

[6] Kotelnikov E., Loukachevitch N., Nikishina I., Panchenko A. RuArg-2022: Argument Mining Evaluation // Computational Linguistics and Intellectual Technologies: Proceedings of the International Conference “Dialogue 2022”. 2022. — P. 1–16.

[7] Lawrence John, Chris Reed. AIFdb corpora // In Proceedings of the Fifth International Conference on Computational Models of Argument (COMMA 2014). — Pitlochry, 2014. — P. 465–466.

[8] Lawrence J., Reed C. Argument mining: A survey. — Computational Linguistics, 45(4), 2019. — P. 765-818.

[9] Niculae V., Park J., Cardie C. Argument mining with structured SVMs and RNNs // In: Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics, 2017. — Vol. 1: Long Papers. — P. 985–995.

[10] Pimenov I.S. Compatibility of Arguments from Different Functional Groups in Scientific Texts [Sochetaemost` argumentov raznykh funktcional`nykh grupp v nauchnykh tekstakh] // Philology. Theory & Practice. [Filologicheskie nauki. Voprosy` teorii i praktiki] Tambov: Gramota, 2022, vol. 11. pp. 3672-3680.

[11] Rahwan I, Banihashemi B, Reed C, Walton D, Abdallah S. Representing and classifying arguments on the semantic web. — The Knowledge Engineering Review, 26(4), 2011. — P. 487-511.

[12] Sidorova E.A., Ahmadeeva I.R., Kononenko I.S., Chagina P.M. (2022), Argumentation Extraction Based on Indicator Approach [Izvlechenie argumentatcii na osnove indikatornogo podhoda], Proceedings of the 20-th Russian Conference on Artificial Intelligence RCAI-2022 [Trudy 20-oi natcionalnoi konferentcii po iskusstvennomu intellektu KII-2022], Moscow, pp.219-233.

[13] Schaefer Robin, Stede Manfred. Argument mining on twitter: A survey // it - Information Technology. — Vol. 63(1). — P. 45–58.

[14] Stede Manfred. Automatic argumentation mining and the role of stance and sentiment // Journal of Argumentation in Context, 2020. — Vol. 9(1). — P. 19–41.

[15] Toldova S., Pisarevskaya D., Vasilyeva M., Kobozeva M. The cues for rhetorical relations in Russian: "Cause-Effect" relation in Russian Rhetorical Structure Treebank // Computational Linguistics and Intellectual Technologies. Papers from the Annual International Conference “Dialogue”, 2018. — P. 747–761.

[16] Toulmin S. The Uses of Argument. – Cambridge: Cambridge University Press, 2003.

[17] Vecchi Eva Maria, Falk Neele, Jundi Iman, Lapesa Gabriella. Towards argument mining for social good: A survey // Proceedings of the 59th Annual Meeting of the 9th Annual Meeting of the Association for Computational Linguistics, 2021. — P. 1338–1352.

[18] Walton D., Reed C., Macagno F. Argumentation schemes. — Cambridge University Press, 2008.

[19] Zhang G., Nulty P., Lillis D. A decade of legal argumentation mining: Datasets and approaches // International Conference on Applications of Natural Language to Information Systems. – Springer, Cham, 2022. — P. 240-252.