Оценочные фреймы RuSentiFrames для анализа тональности: создание, тестирование и применение*

Н.В. Лукашевич, Н.Л. Русначенко

В статье описывается лексикон RuSentiFrames для русского языка, содержащий оценочные фреймы для предикатных слов и выражений. Оценочные фреймы включают несколько типов информации о тональности и позитивных или негативных эффектах, что дает дополнительные возможности для разметки текстов в задачах таргетированной тональности. Созданные фреймы были использованы в задаче извлечения тональности отношений между упомянутыми сущностями на основе большой новостной коллекции текстов на русском языке, в результате чего была создана автоматически размеченная коллекция текстов RuAttitudes. Эксперименты показали, что использование RuAttitudes для дообучения моделей типа BERT приводят к повышению качества извлечения тональности отношений.

Ключевые слова: анализ тональности, коннотации, языковые модели.

**Работа выполнена при финансовой поддержке РФФИ (проекты №№16-29-09606 и 20-07-01059).*

Введение

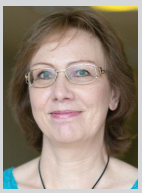
В настоящее время активно развиваются подходы в так называемом таргетированном анализе тональности текстов, при котором модель должна определять отношение к конкретным сущностям, их аспектам (свойствам) или темам. Таргетированный анализ тональностей особенно важен для обработки потока новостей. Предполагается, что новостные тексты содержат разнообразную информацию о позициях по разным вопросам, высказываемых государственными органами, компаниями, мнения частных лиц, положительное или отрицательное отношение упомянутых субъектов друг к другу [1–3].

Анализ тональности сообщений СМИ имеет несколько особенностей, усложняющих анализ, в частности:

— в текстах часто упоминается большое количество сущностей, по отношению к большинству из которых тональность является нейтральной;
— тональность может быть прямо высказана автором (или упоминаемыми субъектами) или может быть выведена из описываемых действий сущностей по отношению друг к другу (имплицитная тональность);
— в значимой доле предложений может встречаться несколько субъектов тональности, которые выражают свое мнение, и несколько объектов тональности, при этом тональности субъектов по отношению к объектам могут быть различными;
— высказанные оценки нужно отличать от негативных/позитивных событий, не оказывающих влияние на выражение и восприятие тональности (наводнение, землетрясение) и негативных/позитивных событий, которые могут имплицитно выражать тональность, например «X уволил Y», «рост экономики в X».

Все это затрудняет определение таргетированной тональности – как авторской, так и между упомянутыми сущностями – и требует создания специализированных моделей и лексических ресурсов.

В данной статье описан словарь оценочной лексики на основе фреймов RuSentiFrames. RuSentiFrames содержит слова-предикаты, которые ссылаются на некоторую ситуацию с несколькими участниками и имеющими оценочные коннотации. Лексикон RuSentiFrames позволяет описать разную тональность



ЛУКАШЕВИЧ Наталья Валентиновна
Московский государственный университет им. М.В. Ломоносова

РУСНАЧЕНКО Николай Леонидович
Московский государственный технический университет им. Н.Э. Баумана

автора к разным участникам ситуации, а также выделить позитивные и негативные эффекты для каждого участника, что дает возможность более детального семантического анализа текста. Для улучшения качества словаря производился лингвистический анализ примеров в корпусах, проводился опрос носителей русского языка с помощью краудсорсинга, делался специализированный дистрибутивный анализ. Также в статье представлен подход по применению лексикона RuSentiFrames для автоматической разметки тональности отношений между упоминаемыми сущностями в большой коллекции текстов. Показано, что использование такого корпуса для предобучения модели BERT позволяет улучшить качество извлечения тональности отношений.

Близкие работы

Структурированные словари для анализа тональности

Большинство словарей оценочной лексики имеют вид простых списков с оценками тональности, которые не могут отразить сложности отношений между участниками, а также учесть различия между позитивными/негативными тональностями и фактами с негативными/позитивными последствиями. Например, в известном словаре MPQA [4] английских оценочных слов указывается, что слово refuse (отказать (1) кто, (2) кому, (3) в чем) имеет негативную тональность. Также описан и русский глагол «отказать» в словаре оценочных слов RuSentiLex [5]. В упрощенных подходах анализа тональности, основанных на таких словарях, предполагается, что сущности, встречающиеся рядом с отрицательным словом, получат отрицательные оценки тональности по отношению к ним.

Однако на самом деле совокупность оценок, связанных со словом «отказать», значительно сложнее. В ситуации, описываемой этим глаголом, участвуют три участника: «кто отказал», «кому отказал», «в чем отказал». Глагол «отказать» сообщает, что:

- имеется отрицательное отношение первого участника ситуации ко второму и третьему участникам;
- для второго участника ситуации имеются явные негативные последствия отказа;
- для первого участника ситуация относительно нейтральна, и никаких негативных последствий из употребления данного глагола не следует;
- при этом позиция автора к ситуации в целом и участникам не высказана.

Таким образом, с данным глаголом связан ряд оценочных ассоциаций, так называемых коннотаций. Для их описания нужно создавать специализированные структурированные словари оценочной лексики, которые позволяют для такого слова-предиката описывать набор участников ситуации и ассоциированные с ними тональности.

В работе [6] рассматриваются фреймы коннотаций для глаголов (connotation frames), в которых описываются предположения об оценках сущностей X (подлежащее) и Y (прямое дополнение), когда они употребляются с глаголом V (XVY): как автор относится к X и как к Y, как X и Y относятся друг к другу, улучшается или ухудшается состояние X и Y после действия V. В работе рассматривается 900 наиболее частотных переходных глаголов английского языка.

В работах [7, 8] авторы рассматривают оценочные фреймы для немецких глаголов. Каждый фрейм состоит из совокупности ролей, ассоциированных с глаголами, полярности, а также (положительных или отрицательных) эффектов, связанных с ролями. Рассматриваются только позиции субъекта и объекта глагола. Также описывается так называемая сигнатура глагола. Сигнатура указывает на фактуальность ролей в зависимости от различных факторов (таких как отрицание, наклонение и т. д.). Например, если глагол «verhindern» (затруднять) не отрицается в предложении, то его объект не находится в фактическом состоянии и его возможная полярность не должна учитываться.

В [6] рассматриваются события, которые положительно или отрицательно влияют на сущности (goodFor/badFor). Например, «снижение X» плохо для X, но «создать X» хорошо для X. В этой статье рассматривается вывод тональности, когда тональность выражается по отношению к плохим или хорошим событиям.

Словари для анализа тональности на русском языке

Опубликовано несколько русскоязычных словарей оценочных слов для анализа тональности. В работе [10] описан автоматически порожденный русскоязычный словарь оценочной лексики в сфере продуктов и услуг

 DOI: 10.22204/2410-4639-2023-119-120-03-04-77-91 № 3-4 (119-120) июль–декабрь 2023 г.

(ProductSentiRus). ProductSentiRus получен путем применения модели машинного обучения с учителем к коллекциям отзывов пользователей в нескольких предметных областях. Он представлен в виде списка из 5 000 слов, упорядоченных по убыванию вероятности их оценочности, без разметки негативной или позитивной тональностей.

Словарь оценочных слов и выражений RuSentiLex [5] представляет собой список слов и выражений, имеющих несколько атрибутов. Единицы лексикона RuSentiLex классифицируются по четырем категориям тональности (положительные, отрицательные, нейтральные или положительные/отрицательные) и трем источникам тональности (мнение, эмоция или факт). Многозначные слова в лексиконе, имеющие разную направленность тональности в разных значениях, связаны с соответствующими понятиями русского тезауруса RuTes [11], что может помочь устранить неоднозначность тональности в определенных областях или контекстах.

Словарь LINIS Crowd был создан методом краудсорсинга [12]. Лексикон предназначен для выявления тональностей в пользовательском контенте (блоги, социальные сети), связанном с социальными и политическими вопросами. Каждое слово оценивалось не менее чем тремя добровольцами в контексте трех разных текстов и оценивалось по шкале от –2 (отрицательный) до +2 (положительный).

Несколько мультиязычных оценочных словарей включают русский язык. Словарь Чена – Скиены (Chen – Skiena, 2 876 слов) [13] был сгенерирован для 136 языков путем распространения тональностей от заданных исходных слов, включая русский. Авторы работы [14] породили русский вариант лексикона EmoLex на основе автоматического перевода английского словаря, размеченного путем краудсорсинга (4 412 русских слов). Более подробный список и анализ словарей оценочной лексики для русского языка представлен в работе [15].

Словарь оценочных фреймов RuSentiFrames

Структура фреймов

Словарь RuSentiFrames содержит совокупность оценочных фреймов, которые связаны со словами-предикатами. Предикаты – это слова или выражения, которые описывает ситуацию с некоторыми участниками. В данной работе оценочный фрейм представляет собой набор положительных или отрицательных ассоциаций (коннотаций), связанных с предикатным словом или выражением. Типы коннотаций, которые передаются в оценочных фреймах, следующие:

— отношение автора текста к участникам ситуации;
— положительные или отрицательные отношения между участниками;
— положительное или отрицательное воздействие на участников;
— положительные или отрицательные эмоциональные состояния участников, связанные с описываемой ситуацией.

Чтобы описать участников ситуации во фрейме, нужно обозначить семантические роли, специфичные для предикатов [16, 17]. Для этого используется подход ресурса PropBank [18], в котором смысловые аргументы отдельных глаголов нумеруются начиная с нуля. Для конкретного глагола Arg0 обычно представляет собой аргумент, демонстрирующий черты прототипического агента [19], а Arg1 обозначает прототипический объект ситуации или тему.

Все утверждения во фрейме снабжены оценкой уверенности, которая в настоящее время имеет два значения: 1, если предполагается, что это утверждение верно почти всегда, или 0.7, если утверждение рассматривается как верное по умолчанию. Утверждения о нейтральных тональностях, эффектах и состояниях участников ситуации во фрейм не включаются.

Рассмотрим оценочный фрейм для глагола «осудить», представленный на *рисунке 1*.

В разделе “roles” представлено, что описываемая фреймом ситуация имеет четырех участников. Для каждой роли описано, связана ли данная роль с одушевленными субъектами (SMB), неодушевленными объектами (SMTH) или участник может быть любым (ANY).

Этот фрейм описывает, что «тот, кто осуждает» A0 отрицательно относится к осуждаемому A1 из-за A2 и знает, что наказание A3 также является чем-то негативным. A1 отрицательно относится к A0 (к осуждающему) и к наказанию A3. Последствия для A1 и состояние A1 отрицательное. Эффекты обозначаются знаками «–» или «+», чтобы подчеркнуть, что эта информация отличается от тональностей. Для глагола-примера мы не можем предположить позицию автора по отношению к описываемой ситуации и ее

Фрейм: осудить

```
«roles»: {
        [«A0»: «тот, кто осуждает», “SMB”],
        [«A1»: «тот, кого осуждают», “SMB”],
        [«A2»: «мотив, то, за что осуждают», “SMTH”],
        [«A3»: «то, на что осуждают», “SMTH”]},
“polarity”: [
        [“A0”,” A1”,”neg”,1.0], [“A0”,” A2”,”neg”,1.0],
        [“A0”,” A3”,”neg”,1.0], [“A1”,” A0”,”neg”,1.0],
        [“A1”,” A3”,”neg”,1.0]],
“effect”: [
        [“A1”,”-”,1.0]],
 “state”: [
        [“A1”,”neg”,1.0]]}
```

Рис. 1. *Пример описания фрейма «осудить» в коллекции RuSentiFrames.*

участникам, поэтому соответствующие утверждения во фрейме отсутствуют.

Если слова или выражения являются близкими по смыслу и имеют одинаковые роли и связанные с ними коннотации, то они относятся к одному и тому же фрейму.

Таблица 1. *Количественные характеристики RuSentiFrames*

Тип лексической единицы	Количество
Глаголы	3 239
Существительные	986
Словосочетания	2 553
Другие	12
Уникальные единицы	6 788
Всего	7 034

Таблица 2. *Фрагмент алфавитного словника RuSentiFrames*

Лексическая единица	Фрейм
апеллировать	ЖАЛОВАТЬСЯ
апелляция	ЖАЛОВАТЬСЯ
аплодировать	ПООЩРИТЬ, ПОХВАЛИТЬ
арест	АРЕСТОВАТЬ
арест по подозрению	АРЕСТОВАТЬ
арестовать	АРЕСТОВАТЬ
арестовать по подозрению	АРЕСТОВАТЬ
арестовывать	АРЕСТОВАТЬ
арестовывать по подозрению	АРЕСТОВАТЬ
ассистировать	ПОМОЧЬ
беднеть	УХУДШИТЬСЯ
бедствовать	НУЖДАТЬСЯ, БЕДСТВОВАТЬ
без внимания	ИГНОРИРОВАТЬ

Лексический состав фреймов RuSentiFrames

С конкретным фреймом могут быть связаны следующие типы языковых выражений:

— отдельные слова, в основном глаголы и существительные: *отказать, отказ, осудить, осуждение, благодарить*;
— идиомы: *вешать лапшу на уши, взять за горло*;
— коллокации: *нанести вред, нанести обиду, нанести поражение*;
— глаголы или существительные с предлогами в постпозиции, что дает возможность снизить неоднозначность исходного слова, например: *выступать против, завязывать с*;
— свободные словосочетания, синонимичные единицам фрейма.

Например, оценочный фрейм под названием «запретить» связан с 53 словами и словосочетаниями, включая такие выражения, как: *налагать запрет, наложение запрета, закрывать доступ, закрытие доступа, прекращение доступа, прекратить доступ, налагать вето* и др.

Табл. 1 содержит количественные данные по составу RuSentiFrames. *Табл. 2* содержит фрагмент алфавитного словника к фреймам.

В настоящее время лексикон RuSentiFrames содержит 311 фреймов, с которыми связано 7 034 слова и выражения.

Тестирование RuSentiFrames

Для тестирования RuSentiFrames использовались разные подходы. В частности, размечались наборы конкретные примеры для слов-предикатов, в результате можно было вывести средние оценки тональностей и сравнить с тональностями, описанными в соответствующем фрейме. Для оценки согласия экспертов по описанию фреймов были выполнены параллельные описания фреймов двумя экспертами-лингвистами. При оценке согласия оценивалось совпадение списка коннотаций и их тональность. Совпадение оценивалось как *F*-мера

DOI: 10.22204/2410-4639-2023-119-120-03-04-77-91 № 3-4 (119-120) июль–декабрь 2023 г.

между двумя разметками, ее значение составило 0.76 [20].

Одним из интересных экспериментов, связанных с проверкой описаний RuSentiFrames, была верификация фреймов на основе опросов носителей языка (краудсорсинга). Эксперимент был реализован в краудсорсинговой системе Яндекс.Толока [21]. Рассмотрим данное тестирование более подробно.

Для эксперимента были отобраны слова, соответствующие одному из следующих критериев:

— слова похожи по значению, но имеют разные смысловые оттенки, в частности, некоторые различия в отношении авторов к участникам, и, следовательно, эти слова относятся к разным фреймам. Например, коннотации глагола *укокошить* отличаются от коннотаций глагола *убить* тем, что, помимо негативного отношения автора к основному участнику, предположительно имеется отрицательное отношение автора ко второму участнику (кто был убит);
— слова близки по смыслу, но имеют различную тональность или интенсивность тональности отношений между участниками ситуации. Например, в предикате *смеяться* тот, кто смеется, будет предположительно отрицательно относиться к объекту ситуации. Но в глаголе *насмехаться* это негативное отношение должно быть более интенсивным;
— слова-синонимы, относящиеся к одному фрейму, но разные по стилю или выражению. В эту группу также входят случаи, когда в одном фрейме указано слово и синонимичная коллокация: *надеяться – возлагать надежду*, *доверять – оказать доверие*;
— слова-антонимы. В этом случае хотелось понять, как будут модифицироваться коннотации для слов-антонимов: *запретить – разрешать*, *улучшить – ухудшить*, *нарушать – соблюдать*.

Для проведения краудсорсингового эксперимента было отобрано 89 предикатов в соответствии с вышеупомянутыми принципами. Затем были порождены специальные предложения с этими предикатами и вопросы о тональностях, связанных с этими предикатами. Предложения были построены искусственно с использованием нейтральных имен, как Иванов или Петров, чтобы не вызывать у аннотаторов особых ассоциаций. Например, предложение для фразы «воздать по заслугам» было такое: *Иванов воздал Петрову по заслугам*. Затем были заданы следующие вопросы:

— *Как Иванов относится к Петрову?*
— *Как Петров относится к Иванову?*
— *Как автор относится к Иванову?*
— *Как автор относится к Петрову?*

Респондентам было предложено выбрать один вариант ответа из ранжированных вариантов: от очень отрицательного (–2) до очень положительного (+2).

В результате тестирования была отмечена корреляция оценок экспертов, которые создавали ресурс RuSentiFrames, и ответов носителей русского языка. Было найдено несколько неточностей в описаниях RuSentiFrames. Проблемой эксперимента оказалось некоторое различие шкал оценок, в которых опрашивались люди в эксперименте и в описаниях RuSentiFrames.

Использование RuSentiFrames в задаче извлечения тональности отношений

Лексикон RuSentiFrames был применен для автоматической разметки коллекции текстов для задачи извлечения оценочных отношений из текстов, то есть для определения того, как упомянутые в тексте сущности относятся друг к другу. Такой корпус был затем использован для предобучения нейросетевых моделей, что привело к улучшению качества извлечения отношений. Далее в нашей статье будут рассмотрены процедура автоматической разметки и результаты использования размеченного корпуса для предобучения модели BERT [22].

Автоматическая разметка новостной коллекции с помощью RuSentiFrames

Основными этапами автоматической разметки новостей для извлечения тональности отношений между сущностями были следующие [23].

На первом этапе проводился анализ заголовков статей. Предполагалось, что наиболее кратко и явно отношения между сущностями (людьми, организациями, странами) выражены в заголовках статей. Поэтому были извлечены те статьи, где в заголовке были по крайней мере две именованные сущности и хотя бы один предикат из RuSentiFrames.

На втором этапе вычислялась тональность отношений между сущностями в заголовке. Если все предикаты в заголовке указывают на позитивную тональность между сущностями (например, *одобрить*), то предполагается, что отношение между сущностями позитивное. Если есть хотя бы один предикат с отрицательной тональностью (*осудить*), то предполагается, что отношения между сущностями негативные. Также учитывались отрицательные частицы типа *не*, которые меняют тональность на противоположную.

На третьем этапе извлеченная из заголовка тональность переносится на предложения внутри новости, которые обычно более сложно построены, то есть если та же пара сущностей встречается в одном предложении внутри текста, то считаем, что в этом предложении выражена та же тональность, что и в заголовке. Таким образом, заголовки и предложения из новостей, в которых выявлена тональность отношений между сущностями на основе RuSentiFrames, становятся автоматически размеченными данными для извлечения оценочных отношений.

Наконец, для улучшения качества автоматически размеченной коллекции были рассчитаны тональности отношений между сущностями в среднем по коллекции, отобраны пары сущностей с явно негативными и позитивными отношениями, которые вместе встречались выше заданного порога (доверенные пары). Далее в размеченную коллекцию для машинного обучения включаются только те предложения, тональность отношений сущностей в которых совпала со средними оценками по коллекции. Таким образом, средняя оценка тональности отношений между сущностями по коллекции используется как дополнительный фильтрующий фактор, увеличивающий точность разметки.

Дополнительно учитывается, что сущности типа Location, не совпадающие с названиями стран, обычно в новостях не входят ни в какие оценочные отношения с другими сущностями. Таким образом, это дает возможность образовать нейтральный класс отношений для предобучения классификатора отношений на три класса: позитивные, негативные, нейтральные.

Для разметки был взят новостной корпус, собранный в 2017 г. из разных новостных источников. В новостных статьях были выделены именованные сущности с помощью инструмента от DeepPavlov (http://docs.deeppavlov.ai/en/master/features/models/NER.html). В результате обработки коллекции новостей из 8 млн новостных статей было извлечено более 250 тыс. предложений, содержащих оценочные отношения между сущностями; полученный корпус был назван RuAttitudes (корпус может быть получен по ссылке https://github.com/nicolay-r/RuAttitudes). *Табл. 3* содержит примеры доверенных пар с высокой уверенностью предсказанной тональности.

Постановки задачи и модели

Качество извлечения отношений между сущностями тестировалось на размеченной коллекции RuSentRel (https://github.com/nicolay-r/RuSentRel) [1], которая содержит статьи по международным отношениям с портала новостного ресурса ИноСМИ. В коллекции новостные тексты размечены триплетами по тональности (например, США, Россия, neg) на уровне документа.

Таблица 3. *Примеры сущностей с высокой вероятностью тональности отношений (доверенные пары)*

Субъект тональности	Объект тональности	Доля вхождений пары с определенной тональностью
Канада	Украина	0.90
Пентагон	Украина	0.90
Украина	МВФ (Международный валютный фонд)	0.80
Трамп	ИГИЛ	-0.79
Россия	ИГИЛ	–0.79
Турция	ИГИЛ	–0.83
Азербайджан	Армения	–0.93
Карабах	Азербайджан	–0.94

DOI: 10.22204/2410-4639-2023-119-120-03-04-77-91 № 3-4 (119-120) июль–декабрь 2023 г.

Извлечение оценочных отношений рассматривается как классификационная задача, организованная в формате следующих экспериментов:

1. *Двуклассовая задача*, в которой все модели должны определить класс тональности заведомо оценочного отношения.
2. *Трехклассовая задача*, в которой каждая модель может классифицировать поданный на вход «контекст с отношением» как положительный, негативный либо нейтральный классы.

Оценка результатов в обоих случаях проводится по метрике *F1(P,N)*, то есть производится макроусреднение *F1* меры позитивного и негативного классов. Нейтральный класс не учитывается при вычислении макромеры, поскольку в текстах большинство отношений между сущностями нейтральные, но задачей извлечения тональности отношений все-таки является нахождение оценочных отношений.

Для конкретной модели процесс обучения (и соответствующей оценки) выполняется в следующих режимах:

1. *Обучение с учителем* – извлечение отношений основано на обучающих данных коллекции RuSentRel;
2. *Применение опосредованного обучения (от англ. distant supervision)* – процесс обучения моделей предполагает дополнительное использование автоматически размеченной коллекции RuAttitudes.

Форматы применения коллекции RuAttitudes были следующими: *предобучение и дообучение* – модели изначально обучались с использованием опосредованного обучения RuAttitudes, после которого следовало дообучение (от англ. fine-tuning) предобученной модели на основе контекстов коллекции RuSentRel.

Для оценки результатов использовались два типа экспериментов:

— кросс-валидационное – в каждом разбиении объединяется коллекция RuAttitudes с каждым обучающим блоком коллекции RuSentRel;
— фиксированное – TRAIN/TEST-разбиение – обучающий набор представляет собой комбинацию RuAttitudes с частью TRAIN.

Измерение средних значений точности (метрика Accuracy) проводилось каждые 5 эпох. Процесс обучения прекращался при превышении лимита в 200 эпох. Для избежания проблем переобучения моделей используется механизм dropout.

Для проведения экспериментов были рассмотрены языковые модели с архитектурой BERT [22]. Далее рассмотрим основные особенности представления входной информации в моделях такой архитектуры. Входная информация моделей BERT представляет собой последовательность токенов (атомарных частей слов, являющихся частью словаря модели), опционально разделенную специальным символом [SEP] на две независимые части: TextA и TextB. Такое разделение является особенностью установления связи между словами и предложениями текста путем проведения дополнительных задач на этапе предобучения моделей: 1) предсказания маскированных токенов (от англ. Masking Language Modeling) и 2) определение, является ли TextB продолжением TextA (Natural Language Inference (NLI)).

Для проведения классификации входная последовательность модели предусматривает специальный токен [CLS] (токен класса), зарезервированный перед началом основной входной последовательности. Применение языковых моделей BERT в задачах классификации выполняется с введением классификационного слоя, который отвечает за сопоставление двух последовательностей (TextA и TextB) множеству выходных классов задачи [24].

Используются два способа формирования последовательностей [24]:

— вопросно-ответный (**QA**) – подразумевает составление вопроса и записи последнего в последовательность TextB для последовательности TextA;
— Natural Language Inference (**NLI**) – подразумевает указание ожидаемой информации, которая должна быть выведена из TextA.

Таким образом, для проведения экспериментов выбраны следующие форматы представления частей входных последовательностей [24]:

1. «**без $\mathbf{Text_B}$**» – использование последовательности без разделения (TextA).
2. «**$\mathbf{Text_B}$QA**» – дополнение TextA вопросом в TextB.
3. «**$\mathbf{Text_B}$NLI**» – дополнение TextA выводом отношения по контексту в TextB.

Для проведения этапа предобучения и применения RuAttitudes среди перечня перечисленных выше форматов было выбрано представление «$Text_B$NLI», которое далее обозначается как **$\mathbf{NLI_P}$**.

Эксперименты

Предварительное обучение языковых моделей составляет 5 эпох. На этапе предобучения используется скорость обучения (learning rate) $2 \cdot 10^{-5}$ с коэффициентом прогрева (от англ. warm-up), равным 1, то есть в течение первой эпохи. На этапе дообучения коэффициент прогрева существенно меньше, в связи с чем внесение небольших изменений в предобученную модель позволяет сохранить особенности изначального состояния. Длина контекста для языковых моделей несколько отличается, поскольку к предварительному разбору применяется токенизация. В качестве алгоритма разбиения слов на составные части используется реализация по умолчанию. Длина результирующей последовательности ограничена 128 токенами, что позволило покрыть ≈95% примеров без проведения усечений длин контекстов.

В эксперименте рассмотрено применение следующих языковых моделей:

— **mBERT** предобучена на текстовых данных 104 языков и поддерживает регистр букв в представлении входных последовательностей [22]. Модель mBERT доступна и распространена только в размере (формате) BASE (https://huggingface.co/bert-base-multilingual-cased);

— **RuBERT** [25] является дообученной версией mBERT на русскоязычных новостных данных и статьях энциклопедии «Википедия»;

— **SentRuBERT** (https://huggingface.co/DeepPavlov/rubert-base-cased-sentence) представляет собой версию RuBERT, дообученную на базе переведенных с помощью сервиса Google Translate на русский язык текстов корпуса SNLI [26] и русскоязычной части текстов коллекции XNLI [27].

Входными параметрами модели являются контексты с упомянутыми в них парами именованных сущностей. *Табл. 4* приводит статистику числа извлеченных контекстов документов коллекции RuSentRel для разбиений фиксированного TRAIN/TEST-формата.

В случае двуклассовой классификации контексты нейтрального класса (NEU) коллекций RuAttitudes не рассматриваются. Аналогичная ситуация и для коллекции RuSentRel. Дополнительно применяется балансировка данных по числу контекстов классов тональности методом дублирования (oversampling), существующих для достижения объема, равного числу контекстов наибольшего класса. В случае объединенного обучения на коллекциях RuSentRel и RuAttitudes такая балансировка применяется после объединения извлеченных контекстов обеих коллекций.

В *табл. 5* приведены результаты языковых моделей [28] (официальная таблица результатов: https://github.com/nicolay-r/RuSentRel-Leaderboard). Влияние опосредованного обучения оказывает прирост в ~2–4% в двуклассовом эксперименте и ~10–13% в случае трех классов. Преимущество использования русскоязычно ориентированных моделей перед mBERT наблюдается в экспериментах с тремя классами. Так, при использовании RuBERT прирост качества составил от 10% по $F1_{cv}$ и 6–9% по метрике $F1_t$ (*табл. 6*). Применение SentRuBERT улучшает показатели модели RuBERT на 6% в случае, когда в опосредованном обучении используется RuAttitudes (*табл. 6*). Наилучший результат достигается при использовании NLI-формата представления входных данных.

Заключение

В статье описывается структурированный словарь оценочной лексики русского языка RuSentiFrames. Единицей словаря является оценочный фрейм, который перечисляет набор семантических ролей, связанных со словом, и описывает их положительные или отрицательные ассоциации (коннотации). Коннотации могут описывать отношение автора текста к участникам ситуации, положитель-

Таблица 4. *Число контекстов, извлеченных на этапе подготовки данных из коллекций RuSentRel и RuAttitudes. Максимально допустимое число термов в контексте ограничено значением 50*

Коллекция	POS	NEG	NEU
RuSentRel (TRAIN)	511	727	6 530
RuSentRel (TEST)	–	–	8 004
RuAttitudes	123 281	161 275	15 429
RuSentRel (TRAIN) + RuAttitudes	123 832	162 002	21 959

DOI: 10.22204/2410-4639-2023-119-120-03-04-77-91 № 3-4 (119-120) июль–декабрь 2023 г.

Таблица 5. *Результаты применения опосредованного обучения для предобученных состояний модели BERT: mBERT, RuBERT, SentRuBERT; вторая колонка – применение RuAttitudes в обучении с форматом представления с выводом отношения по контексту в TextB («NLI_p») (наилучший результат для каждого состояния модели BERT выделен жирным шрифтом, второй наилучший результат отмечен подчеркиванием)*

Модель	RuAttitudes	Результаты обучения моделей			
		2 класса		3 класса	
		$F1_{c_v}$	$F1_t$	$F1_{c_v}$	$F1_t$
mBERT (**NLI_p** + «**без $TEXT_B$**»)	☑	68.9	67.7	30.5	31.1
mBERT («**без $TEXT_B$**») –	–	67.0	**68.9**	26.9	30.0
mBERT (**NLI_p** + **$Text_BQA$**)	☑	**69.6**	65.2	30.1	35.5
mBERT (**$Text_BQA$**)	–	66.5	65.4	28.6	33.8
mBERT (**NLI_p** + **$Text_BNLI$**)	☑	69.4	68.2	**33.6**	36.0
mBERT (**$Text_BNLI$**)	–	67.8	58.4	29.2	**37.0**
RuBERT (**NLI_p** + «**без $TEXT_B$**»)	☑	70.0	**69.8**	35.6	35.4
RuBERT («**без $TEXT_B$**»)	–	67.8	66.2	**36.8**	37.6
RuBERT (**NLI_p** + **$Text_BQA$**)	☑	69.6	68.2	34.8	37.0
RuBERT (**$Text_BQA$**)	–	69.5	66.2	32.0	35.3
RuBERT (**NLI_p** + **$Text_BNLI$**)	☑	**71.0**	68.6	**36.8**	**39.9**
RuBERT (**$Text_BNLI$**)	–	68.9	66.4	29.4	39.6
SeniRuBERT (**NLI_p** + «**без $TEXT_B$**»)	☑	70.0	69.8	37.9	39.8
SentRuBERT («**без $TEXT_B$**»)	–	69.3	65.6	34.0	35.2
SentRuBERT (**NLI_p** + **$Text_BQA$**)	☑	69.6	64.2	38.4	**41.9**
SentRuBERT (**$Text_BQA$**)	–	**70.2**	67.1	34.3	38.9
SentRuBERT **(NLI_p** + **$Text_BNLI$**)	☑	**70.2**	67.7	**39.0**	38.0
SentRuBERT (**$Text_BNLI$**)	–	69.8	67.6	33.4	32.7
Среднее $\Delta F1$	☑	+1.8%	+3.7%	+13.5%	+10.0%
Среднее	–	68.5	65.7	31.6	35.6

Таблица 6. *Сравнение улучшения качества работы моделей RuBERT и mBERT в результате дообучения моделей; средние значения отмечены зеленым – если прирост качества более 5%, серым – если прирост менее 5%*

Модель	Прирост качества работы модели, %			
	2 класса		3 класса	
	$\Delta F1_{cn}$	$\Delta F1_t$	$\Delta F1_{cn}$	$\Delta F1_t$
RuBERT/mBERT («**без $TEXT_B$**»)	1.60	3.10	16.72	13.83
RuBERT/mBERT (**$Text_BQA$**)	0.0	4.60	15.61	4.23
RuBERT/mBERT (**$Text_BNLI$**)	2.31	0.59	9.52	10.83
Среднее	1.30	2.76	13.95	9.63

Таблица 7. *Сравнение улучшения качества работы моделей SentRuBERT и RuBERT в результате дообучения моделей; средние значения отмечены зеленым – если прирост качества более 5%, серым – если прирост менее 5% и красным – прирост отсутствует*

Модель	Прирост качества работы модели, %			
	2 класса		3 класса	
	$\Delta F1_{cn}$	$\Delta F1_t$	$\Delta F1_{cn}$	$\Delta F1_t$
SentRuBERT/RuBERT («**без $TEXT_B$**»)	-1.41	1.75	2.99	-0.25
SentRuBERT/RuBERT (**$Text_BQA$**)	0.0	-5.87	10.34	13.24
SentRuBERT/RuBERT (**$Text_BNLI$**)	-1.13	-1.31	5.98	-4.76
Среднее	-0.85	-1.81	6.44	2.74

ные или отрицательные отношения между участниками, положительное или отрицательное воздействие на участников в результате ситуации, а также положительные или отрицательные эмоциональные состояния участников, связанные с описываемой ситуацией. В статье также описан краудсорсинговый эксперимент для тестирования фреймов RuSentiFrames. В настоящее время лексикон RuSentiFrames содержит 311 фреймов, с которыми связано 7 034 слова и выражения.

Словарь RuSentiFrames использовался для автоматической разметки текстовой коллекции новостей для извлечения отношений между упоминаемыми сущностями (корпус RuAttitudes). Включение этого корпуса в процесс обучения моделей типа BERT позволило улучшить извлечение тональности отношений.

Литература

1. ***N. Loukachevitch, N. Rusnachenko***
В Proc. *Lecture Notes in Computer Science. TSD 2018. Text, Speech and Dialogue, (Czech Republic, Brno, 11–14 September, 2018), Vol. 11107*, Eds P. Sojka, A. Horák, I. Kopeček et al., DE, Cham, Springer-Verlag, 2018, 11107, pp. 448–458.
DOI: 10.1007/978-3-030-00794-2_4.
2. ***A. Golubev, N. Rusnachenko, N. Loukachevitch***
В *Proc. International Conference "Dialogue-2023" (Online Conference, 14–16 June, 2023),* Iss. 22, pp. 130–141.
DOI: 10.28995/2075-7182-2023-22-130-141.
3. ***F. Hamborg, K. Donnay, B. Gipp***
В *Diversity, Divergence, Dialogue. iConference 2021, Ser. Lecture Notes in Computer Science, Vol. 12646,* Eds K. Toeppe, H. Yan, S.K.W. Chu, DE, Cham, Springer-Verlag, 2021, pp. 156–166.
DOI: 10.1007/978-3-030-71305-8_12.
4. ***T. Wilson, J. Wiebe, P. Hoffmann***
В *Proc. Human Language Technology Conference and Conference on Empirical Methods in Natural Language Processing (Canada, Vancouver, 6–8 October, 2005),* USA, PA, East Stroudsburg, ACL, 2005, pp. 347–354.
5. ***N. Loukachevitch, A. Levchik***
В *Proc. Tenth International Conference on Language Resources and Evaluation (LREC'16) (Slovenia, Portorož, 23–28 May 2016)*, FR, Paris, ELRA, 2016, pp. 1171–1176.
6. ***H. Rashkin, S. Singh, Y. Choi***
В *Proc. Association for Computational Linguistics Conference ACL-2016 (DE, Berlin, 7–12 August, 2016), Vol. 1,* USA, PA, East Stroudsburg, ACL, 2016, pp. 311–322.
DOI: 10.18653/v1/P16-1030.
7. ***M. Klenner, M. Amsler, N. Hollenstein***
В *Proc. KONVENS-2014*, DE, Hildesheim, 2014, s.n. 2014, pp. 106–115.
DOI: 10.5167/uzh-99628.
8. ***M. Klenner, M. Amsler***
В *Proc. Tenth International Conference on Language Resources and Evaluation (LREC'16) (Slovenia, Portorož, 23–28 May 2016)*, FR, Paris, ELRA, 2016, pp. 2888–2891.
9. ***L. Deng, J. Wiebe***
В *Proc. 14th Conference of the European Chapter of the Association for Computational Linguistics (Sweden, Gothenburg, 26–30 April, 2014)*, USA, PA, East Stroudsburg, ACL, 2014, pp. 377–385.
DOI: 10.3115/v1/E14-1040.
10. ***I. Chetviorkin, N. Loukachevitch***
В *Proc. COLING-2012 (India, Mumbai, 8–15 December, 2012),* IN, Mumbai, The COLING 2012 Organizing Committee Publ., 2012, pp. 593–610.
11. ***N. Loukachevitch, B. Dobrov***
В *Proc. Seventh Global Wordnet Conference (Estonia, Tartu, 25–29 January, 2014),* Eds H. Oray, C. Fellbaum, P. Vossen, EE, Tartu, University of Tartu Press, 2014, pp. 154–162.
12. ***O. Koltsova, S. Alexeeva, S. Kolcov***
В *Computational Linguistics and Intellectual Technologies: Proc. International Conference "Dialogue-2016" (RF, Moscow, 1–4 July, 2016)*, RF, Moscow, RGGU, pp. 277–287.
13. ***Y. Chen, S. Skiena***
В *Proc. 52nd Annual Meeting of the Association for Computational Linguistics, (USA, MD, Baltimore, 23–25 June, 2014)*, USA, PA, East Stroudsburg, ACL, pp. 383–389. DOI: 10.3115/v1/P14-2063.
14. ***S. Mohammad, P. Turney***
Computational Intelligence, 2013, **29**(3), 436.
DOI: 10.1111/j.1467-8640.2012.00460.x.
15. ***E. Kotelnikov, T. Peskisheva, A. Kotelnikova***
В *Proc. Conference on Artificial Intelligence and Natural Language. AINL 2018. Communications in Computer and Information Science (Russia, Snt. Petersburg, 17–19 October, 2018), Vol. 930*, DE, Cham, Springer-Verlag, 2018, pp. 139–151. DOI: 10.1007/978-3-030-01204-5_14.
16. ***R. Jackendoff***
Semantic Structures, USA, MA, Cambridge, The MIT Press, 1992, 338 pp.
17. ***C.J. Fillmore, C.F. Baker***
В *Proc. NAACL* 2001 Workshop: Wordnet and Other Lexical Resources: Applications, Extensions and Customizations (USA, PA, Pittsburgh, 3-4 June, 2001), USA, PA, Pittsburgh, ACL, 2001, 193 pp.
18. ***M. Palmer, D. Gildea, P. Kingsbury***
Computational linguistics, 2005, **31**(1), 71.
DOI: 10.1162/0891201053630264.
19. ***D. Dowty***
Language, 1991, **67**(3), 547. DOI: 10.2307/415037.
20. ***N. Loukachevitch, N. Rusnachenko***
В *Computational Linguistics and Intellectual Technologies: Proc. International Conference "Dialogue-2020" (RF, Moscow, 17–20 June, 2020)*, RF, Moscow, RGGU Publ., 2020, pp. 541–552.
21. ***I. Matueva, N. Loukachevitch***
В *Recent Trends in Analysis of Images, Social Networks and Texts: 9th International Conference. AIST 2020 (RF, Moscow, Skolkovo, 15–16 October, 2020)*, DE, Cham, Springer-Verlag, 2021, pp. 52–64.
DOI: 10.1007/978-3-030-71214-3.
22. ***J. Devlin, M. Chang, K. Lee, K. Toutanova***
B Proc. 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (USA, MN, Minneapolis, 2–7 June, 2019), USA, PA, East Stroudsburg, ACL, 2019, **1**, pp. 4171–4186.
DOI: 10.18653/v1/N19-1423.

23. ***N. Rusnachenko, N. Loukachevitch, E. Tutubalina***
В *Proc. International Conference on Recent Advances in Natural Language Processing (RANLP 2019) (Bulgaria, Varna, 2–4 September, 2019)*, BG, Varna, INCOMA Ltd., 2019, pp. 1022–1030.
DOI: 10.26615/978-954-452-056-4_118.

24. ***C. Sun, L. Huang, X. Qiu***
В *Proc. 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (USA, MN, Minneapolis, 2–7 June, 2019)*, *Vol. 1*, USA, PA, East Stroudsburg, ACL, 2019, pp. 380–385.
DOI: 10.18653/v1/N19-1035.

25. ***Y. Kuratov, M. Arkhipov***
В *Computational Linguistics and Intellectual Technologies: Proc. International Conference "Dialogue-2019" (RF, Moscow, 29 May – 1 June, 2019)*, RF, Moscow, RGGU, 2019, pp. 333–339.

26. ***S.R. Bowman, G. Angeli, C. Potts, C. D. Manning***
В *Proc. 2015 Conference on Empirical Methods in Natural Language Processing (PT, Lisbon, 17–21 September, 2015)*, USA, NY, New-York, Currain Associates, 2015, pp. 632–642.
DOI: 10.18653/v1/D15-1075.

27. ***A. Conneau, K. Khandelwal, N. Goyal, V. Chaudhary, G. Wenzek, F. Guzmán, E. Grave, M. Ott, L. Zettlemoyer, V. Stoyanov***
В *Proc. 58th Annual Meeting of the Association for Computational Linguistics (online conference, 5–10 July, 2020)*, USA, PA, East Stroudsburg, ACL, 2020, pp. 8440–8451.
DOI: 10.18653/v1/2020.acl-main.747.

28. ***Н.Л. Русначенко***
Труды ИСП РАН, 2021, **33**(3), 199.
DOI: 10.15514/ISPRAS-2021-33(3)-14.

English

Sentiment Frames for Russian: Creation, Testing, Use*

Natalia V. Loukachevitch
Lomonosov Moscow State University
1/4 Leninskie Gory,
Moscow, 119899, Russia
louk_nat@mail.ru

Nikolay L. Rusnachenko
Bauman Moscow State Technical University
5-1 Baumanskaya 2nd Str.,
Moscow, 105005, Russia
kolayrus@yandex.ru

Abstract

In this paper we describe the lexicon RuSentiFrames for Russian, where predicate words and expressions are collected and linked to so-called sentiment frames conveying several types of presupposed information on attitudes and effects. We applied the created frames in the task of extracting attitudes from a large mass-media news collection of texts in Russian. Experiments have shown that the use of RuAttitudes for fine-tuning of BERT-like models leads to an increase in the quality of extracting the sentiment of relations.

Keywords: sentiment analysis, connotation, language models.

**The work was financially supported by RFBR (projects 16-29-09606 and 20-07-01059).*

Images & Tables

Фрейм: осудить

«roles»: {
[«A0»: «тот, кто осуждает», "SMB"],
[«A1»: «тот, кого осуждают», "SMB"],
[«A2»: «мотив, то, за что осуждают», "SMTH"],
[«A3»: «то, на что осуждают», "SMTH"]},
"polarity": [
["A0"," A1","neg",1.0], ["A0"," A2","neg",1.0],
["A0"," A3","neg",1.0], ["A1"," A0","neg",1.0],
["A1"," A3","neg",1.0]],
"effect": [
["A1","-",1.0]],
"state": [
["A1","neg",1.0]]}

Fig. 1. *An example of a description of the frame "condemn" in the RuSentiFrames collection.*

Table 1. *Quantitative characteristics of RuSentiFrames*

Type of lexical item	Count
Verbs	3 239
Nouns	986
Phrases	2 553
Other	12
Unique entries	6 788
Total	7 034

Table 2. *Fragment of the alphabetical dictionary RuSentiFrames*

Lexical item	Frame Value
апеллировать (to appeal)	ЖАЛОВАТЬСЯ (to complain)
апелляция (appeal)	ЖАЛОВАТЬСЯ (to complain)
аплодировать (to applaud)	ПООЩРИТЬ, ПОХВАЛИТЬ (encourage, praise)
арест (arrest)	АРЕСТОВАТЬ (to arrest)
арест по подозрению (arrest on suspicion)	АРЕСТОВАТЬ (to arrest)
арестовать (to arrest someone)	АРЕСТОВАТЬ (to arrest)
арестовать по подозрению (to arrest someone on suspicion)	АРЕСТОВАТЬ (to arrest)
арестовывать (to arrest somebody)	АРЕСТОВАТЬ (to arrest)
арестовывать по подозрению (to arrest somebody on suspicion)	АРЕСТОВАТЬ (to arrest)
ассистировать (to assist)	ПОМОЧЬ (to help)
беднеть (to become poor)	УХУДШИТЬСЯ (to get worse)
бедствовать (to be poor)	НУЖДАТЬСЯ, БЕДСТВОВАТЬ (need, to be poor)
без внимания (without attention)	ИГНОРИРОВАТЬ (to ignore)

Table 3. *Examples of entities with a high probability of relationship sentiment (trusted pairs)*

Sentiment Subject	Sentiment Object	Proportion of pair occurrence of the defined sentiment
Canada	Ukraine	0.90
The Pentagon	Ukraine	0.90
Ukraine	IMF (International Monetary Fund)	0.80
Trump	IGIL	-0.79
Russia	IGIL	–0.79
Turkey	IGIL	–0.83
Azerbaijan	Армения	–0.93
Karabakh	Azerbaijan	–0.94

DOI: 10.22204/2410-4639-2023-119-120-03-04-77-91 № 3-4 (119-120) июль–декабрь 2023 г.

Table 4. *The number of contexts extracted at the data preparation stage from the RuSentRel and RuAttitudes collections; the maximum allowed number of terms in a context is limited to 50*

Collection	POS	NEG	NEU
RuSentRel (TRAIN)	511	727	6 530
RuSentRel (TEST)	–	–	8 004
RuAttitudes	123 281	161 275	15 429
RuSentRel (TRAIN) + RuAttitudes	123 832	162 002	21 959

Table 5. *Results of using mediated learning for pretrained states of the BERT model: mBERT, RuBERT, SentRuBERT; second column – application of RuAttitudes in training with a presentation format with inference of attitude by context in TextB ("NLIp") (the best result for each state of the BERT model is in bold, the second best result is underlined)*

Model	RuAttitudes	Model training results			
		2 classes		3 classes	
		$F1_{c_v}$	$F1_t$	$F1_{c_v}$	$F1_t$
mBERT (**NLI_p** + «**без $TEXT_B$**»)	☑	68.9	67.7	<u>30.5</u>	31.1
mBERT («**без $TEXT_B$**»)	–	67.0	**68.9**	26.9	30.0
mBERT (**NLI_p** + **$Text_BQA$**)	☑	**69.6**	65.2	30.1	35.5
mBERT (**$Text_BQA$**)	–	66.5	65.4	28.6	33.8
mBERT (**NLI_p** + **$Text_BNLI$**)	☑	<u>69.4</u>	<u>68.2</u>	**33.6**	<u>36.0</u>
mBERT (**$Text_BNLI$**)	–	67.8	58.4	29.2	**37.0**
RuBERT (**NLI_p** + «**без $TEXT_B$**»)	☑	<u>70.0</u>	**69.8**	<u>35.6</u>	35.4
RuBERT («**без $TEXT_B$**»)	–	67.8	66.2	**36.8**	37.6
RuBERT (**NLI_p** + **$Text_BQA$**)	☑	69.6	68.2	34.8	37.0
RuBERT (**$Text_BQA$**)	–	69.5	66.2	32.0	35.3
RuBERT (**NLI_p** + **$Text_BNLI$**)	☑	**71.0**	<u>68.6</u>	**36.8**	**39.9**
RuBERT (**$Text_BNLI$**)	–	68.9	66.4	29.4	<u>39.6</u>
SeniRuBERT (**NLI_p** + «**без $TEXT_B$**»)	☑	70.0	69.8	37.9	39.8
SentRuBERT («**без $TEXT_B$**»)	–	69.3	65.6	34.0	35.2
SentRuBERT (**NLI_p** + **$Text_BQA$**)	☑	69.6	64.2	38.4	**41.9**
SentRuBERT (**$Text_BQA$**)	–	**70.2**	67.1	34.3	38.9
SentRuBERT **(NLI_p** + **$Text_BNLI$**)	☑	**70.2**	<u>67.7</u>	**39.0**	38.0
SentRuBERT (**$Text_BNLI$**)	–	69.8	67.6	33.4	32.7
Mean $\Delta F1$	☑	+1.8%	+3.7%	+13.5%	+10.0%
Mean	–	68.5	65.7	31.6	35.6

Table 6. *The performance quality of the RuBERT model in comparison with the mBERT performance as a result of models fine-tuning; the average values are highlighted in green (if the quality increase is more 5%) or gray (if the increase less than 5%)*

Model	Increase in model performance quality, %			
	2 classes		3 classes	
	$\Delta F1_{c_v}$	$\Delta F1_t$	$\Delta F1_{c_v}$	$\Delta F1_t$
RuBERT/mBERT («**без $TEXT_B$**»)	1.60	3.10	16.72	13.83
RuBERT/mBERT (**$Text_BQA$**)	0.0	4.60	15.61	4.23
RuBERT/mBERT **($Text_BNLI$**)	2.31	0.59	9.52	10.83
Mean	1.30	2.76	13.95	9.63

Table 7. The performance quality of the SentRuBERT model in comparison with RuBERT performance as a result of additional training of models; average values are marked in green (if the quality increase more than 5%), gray (if the increase less than 5%) and red (if there is no increase)

Model	Increase in model performance quality, %			
	2 classes		3 classes	
	$\Delta F1_{c_v}$	$\Delta F1_t$	$\Delta F1_{c_v}$	$\Delta F1_t$
SentRuBERT/RuBERT («**без $TEXT_B$**»)	-1.41	1.75	2.99	-0.25
SentRuBERT/RuBERT (**$Text_B$QA**)	0.0	-5.87	10.34	13.24
SentRuBERT/RuBERT (**$Text_B$NLI**)	-1.13	-1.31	5.98	-4.76
Mean	-0.85	-1.81	6.44	2.74

References

1. ***N. Loukachevitch, N. Rusnachenko***
In Proc. *Lecture Notes in Computer Science. TSD 2018. Text, Speech and Dialogue, (Czech Republic, Brno, 11–14 September, 2018), Vol. 11107*, Eds P. Sojka, A. Horák, I. Kopeček et al., DE, Cham, Springer-Verlag, 2018, 11107, pp. 448–458.
DOI: 10.1007/978-3-030-00794-2_4.
2. ***A. Golubev, N. Rusnachenko, N. Loukachevitch***
In *Proc. International Conference "Dialogue-2023" (Online Conference, 14–16 June, 2023),* Iss. 22, pp. 130–141.
DOI: 10.28995/2075-7182-2023-22-130-141.
3. ***F. Hamborg, K. Donnay, B. Gipp***
In *Diversity, Divergence, Dialogue. iConference 2021, Ser. Lecture Notes in Computer Science, Vol. 12646,* Eds K. Toeppe, H. Yan, S.K.W. Chu, DE, Cham, Springer-Verlag, 2021, pp. 156–166.
DOI: 10.1007/978-3-030-71305-8_12.
4. ***T. Wilson, J. Wiebe, P. Hoffmann***
In *Proc. Human Language Technology Conference and Conference on Empirical Methods in Natural Language Processing (Canada, Vancouver, 6–8 October, 2005),* USA, PA, East Stroudsburg, ACL, 2005, pp. 347–354.
5. ***N. Loukachevitch, A. Levchik***
In *Proc. Tenth International Conference on Language Resources and Evaluation (LREC'16) (Slovenia, Portorož, 23–28 May 2016)*, FR, Paris, ELRA, 2016, pp. 1171–1176.
6. ***H. Rashkin, S. Singh, Y. Choi***
In *Proc. Association for Computational Linguistics Conference ACL-2016 (DE, Berlin, 7–12 August, 2016), Vol. 1,* USA, PA, East Stroudsburg, ACL, 2016, pp. 311–322.
DOI: 10.18653/v1/P16-1030.
7. ***M. Klenner, M. Amsler, N. Hollenstein***
In *Proc. KONVENS-2014*, DE, Hildesheim, 2014, s.n. 2014, pp. 106–115.
DOI: 10.5167/uzh-99628.
8. ***M. Klenner, M. Amsler***
In *Proc. Tenth International Conference on Language Resources and Evaluation (LREC'16) (Slovenia, Portorož, 23–28 May 2016)*, FR, Paris, ELRA, 2016, pp. 2888–2891.
9. ***L. Deng, J. Wiebe***
In *Proc. 14th Conference of the European Chapter of the Association for Computational Linguistics (Sweden, Gothenburg, 26–30 April, 2014)*, USA, PA, East Stroudsburg, ACL, 2014, pp. 377–385.
DOI: 10.3115/v1/E14-1040.
10. ***I. Chetviorkin, N. Loukachevitch***
In *Proc. COLING-2012 (India, Mumbai, 8–15 December, 2012)*, IN, Mumbai, The COLING 2012 Organizing Committee Publ., 2012, pp. 593–610.
11. ***N. Loukachevitch, B. Dobrov***
In *Proc. Seventh Global Wordnet Conference (Estonia, Tartu, 25–29 January, 2014),* Eds H. Oray, C. Fellbaum, P. Vossen, EE, Tartu, University of Tartu Press, 2014, pp. 154–162.
12. ***O. Koltsova, S. Alexeeva, S. Kolcov***
In *Computational Linguistics and Intellectual Technologies: Proc. International Conference "Dialogue-2016" (RF, Moscow, 1–4 July, 2016)*, RF, Moscow, RGGU, pp. 277–287.
13. ***Y. Chen, S. Skiena***
In *Proc. 52nd Annual Meeting of the Association for Computational Linguistics, (USA, MD, Baltimore, 23–25 June, 2014)*, USA, PA, East Stroudsburg, ACL, pp. 383–389.
DOI: 10.3115/v1/P14-2063.
14. ***S. Mohammad, P. Turney***
Computational Intelligence, 2013, **29**(3), 436.
DOI: 10.1111/j.1467-8640.2012.00460.x.
15. ***E. Kotelnikov, T. Peskisheva, A. Kotelnikova***
In *Proc. Conference on Artificial Intelligence and Natural Language. AINL 2018. Communications in Computer and Information Science (Russia, Snt. Petersburg, 17–19 October, 2018), Vol. 930*, DE, Cham, Springer-Verlag, 2018, pp. 139–151.
DOI: 10.1007/978-3-030-01204-5_14.
16. ***R. Jackendoff***
Semantic Structures, USA, MA, Cambridge, The MIT Press, 1992, 338 pp.
17. ***C.J. Fillmore, C.F. Baker***
In *Proc. NAACL* 2001 Workshop: Wordnet and Other Lexical Resources: Applications, Extensions and Customizations (USA, PA, Pittsburgh, 3-4 June, 2001), USA, PA, Pittsburgh, ACL, 2001, 193 pp.
18. ***M. Palmer, D. Gildea, P. Kingsbury***
Computational linguistics, 2005, **31**(1), 71.
DOI: 10.1162/0891201053630264.
19. ***D. Dowty***
Language, 1991, **67**(3), 547. DOI: 10.2307/415037.
20. ***N. Loukachevitch, N. Rusnachenko***
In *Computational Linguistics and Intellectual Technologies: Proc. International Conference "Dialogue-2020" (RF, Moscow, 17–20 June, 2020)*, RF, Moscow, RGGU Publ., 2020, pp. 541–552.
21. ***I. Matueva, N. Loukachevitch***
In *Recent Trends in Analysis of Images, Social Networks and Texts: 9th International Conference. AIST 2020 (RF, Moscow, Skolkovo, 15–16 October, 2020)*, DE, Cham, Springer-Verlag, 2021, pp. 52–64.
DOI: 10.1007/978-3-030-71214-3.
22. ***J. Devlin, M. Chang, K. Lee, K. Toutanova***
In *Proc. 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language*

Technologies (USA, MN, Minneapolis, 2–7 June, 2019), USA, PA, East Stroudsburg, ACL, 2019, **1**, pp. 4171–4186.
DOI: 10.18653/v1/N19-1423.

23. ***N. Rusnachenko, N. Loukachevitch, E. Tutubalina***
In *Proc. International Conference on Recent Advances in Natural Language Processing (RANLP 2019) (Bulgaria, Varna, 2–4 September, 2019),* BG, Varna, INCOMA Ltd., 2019, pp. 1022–1030.
DOI: 10.26615/978-954-452-056-4_118.

24. ***C. Sun, L. Huang, X. Qiu***
In *Proc. 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (USA, MN, Minneapolis, 2–7 June, 2019), Vol. 1,* USA, PA, East Stroudsburg, ACL, 2019, pp. 380–385.
DOI: 10.18653/v1/N19-1035.

25. ***Y. Kuratov, M. Arkhipov***
In *Computational Linguistics and Intellectual Technologies: Proc. International Conference "Dialogue-2019" (RF, Moscow, 29 May – 1 June, 2019)*, RF, Moscow, RGGU, 2019, pp. 333–339.

26. ***S.R. Bowman, G. Angeli, C. Potts, C. D. Manning***
In *Proc. 2015 Conference on Empirical Methods in Natural Language Processing (PT, Lisbon, 17–21 September, 2015)*, USA, NY, New-York, Currain Associates, 2015, pp. 632–642.
DOI: 10.18653/v1/D15-1075.

27. ***A. Conneau, K. Khandelwal, N. Goyal, V. Chaudhary, G. Wenzek, F. Guzmán, E. Grave, M. Ott, L. Zettlemoyer, V. Stoyanov***
In *Proc. 58th Annual Meeting of the Association for Computational Linguistics* (*online conference, 5–10 July, 2020)*, USA, PA, East Stroudsburg, ACL, 2020, pp. 8440–8451.
DOI: 10.18653/v1/2020.acl-main.747.

28. ***N.L. Rusnachenko***
Proc. ISP RAS [Trudy ISP RAN], 2021, **33**(3), 199 (in Russian).
DOI: 10.15514/ISPRAS-2021-33(3)-14.