

Methods of Sentiment Detection Towards Aspect of Economic and Social Development in Russian Sentences

A. Y. Poletaev¹, I. V. Paramonov¹, E. I. Boychuk¹

DOI: [10.18255/1818-1015-2024-4-362-383](https://doi.org/10.18255/1818-1015-2024-4-362-383)

¹P.G. Demidov Yaroslavl State University, Yaroslavl, Russia

MSC2020: 68T50

Research article

Full text in Russian

Received October 21, 2024

Revised November 7, 2024

Accepted November 13, 2024

The article is devoted to the task of the sentiment detection towards an aspect of economic and social development in Russian sentences. The aspect, the attitude to which is determined, can be either explicitly mentioned or implied. The authors investigated possibilities of using neural network classifiers and proposed an algorithm for determining the sentiment towards an aspect based on semantic rules implemented with the use of constituency trees. The sentiment towards an aspect is determined in two stages. At the first stage, aspect terms (explicitly mentioned events or phenomena associated with the aspect) are found in the sentence. At the second stage, the sentiment towards an aspect is calculated as the sentiment towards the aspect term that is most closely associated with the aspect. The paper proposes several methods for searching the aspect terms. The performance was assessed on a corpus of 468 sentences extracted from election campaign materials. The best result for neural network classifiers was obtained using the BERT-SPC neural network pretrained on the task of identifying the sentiment towards an explicitly mentioned aspect, the macro F-score was 0.74. The best result for the semantic rule-based algorithm was obtained using the method of aspect term searching based on semantic similarity, the macro-F-score was 0.63. When combining BERT-SPC and the rule-based algorithm into an ensemble, the macro-F-score was 0.79, which is the best result obtained in this work.

Keywords: sentiment analysis; sentiment detection; sentiment towards an aspect; implicit aspect; semantic rules; publicism; constituency tree

INFORMATION ABOUT THE AUTHORS

Poletaev, Anatoliy Y. (corresponding author)	ORCID iD: 0000-0003-0116-4739 . E-mail: anatoliy-poletaev@mail.ru Senior lecturer
Paramonov, Ilya V.	ORCID iD: 0000-0003-3984-8423 . E-mail: ilya.paramonov@fruct.org PhD, Associate professor
Boychuk, Elena I.	ORCID iD: 0000-0001-6600-2971 . E-mail: elena-boychouk@rambler.ru Dr. Sc., Associate professor

Funding: Russian Science Foundation (Project no. 23-21-00495).

For citation: A. Y. Poletaev, I. V. Paramonov, and E. I. Boychuk, "Methods of sentiment detection towards aspect of economic and social development in Russian sentences", *Modeling and Analysis of Information Systems*, vol. 31, no. 4, pp. 362–383, 2024. DOI: [10.18255/1818-1015-2024-4-362-383](https://doi.org/10.18255/1818-1015-2024-4-362-383).

Методы определения тональности по отношению к аспектам социально-экономического развития в предложениях на русском языке

А. Ю. Полетаев¹, И. В. Парамонов¹, Е. И. Бойчук¹

DOI: [10.18255/1818-1015-2024-4-362-383](https://doi.org/10.18255/1818-1015-2024-4-362-383)

¹Ярославский государственный университет им. П.Г. Демидова, Ярославль, Россия

УДК 004.912

Научная статья

Полный текст на русском языке

Получена 21 октября 2024 г.

После доработки 7 ноября 2024 г.

Принята к публикации 13 ноября 2024 г.

Статья посвящена задаче определения тональности по отношению к аспектам социально-экономического развития в предложениях на русском языке. Аспект, отношение к которому определяется, может как упоминаться явно, так и подразумеваться. Авторами были исследованы возможности применения нейросетевых классификаторов, а также предложен алгоритм определения тональности по отношению к аспекту, основанный на семантических правилах, реализованных с использованием деревьев синтаксических единиц. Тональность по отношению к аспекту определяется в два этапа. На первом этапе в предложении отыскиваются аспектные термины — явно упоминаемые события или явления, связанные с аспектом. На втором этапе тональность по отношению к аспекту определяется как тональность по отношению к аспектному термину, который теснее всего связан с аспектом. В работе предлагается несколько методов поиска аспектных терминов. Качество оценивалось на корпусе из 468 предложений, извлечённых из материалов предвыборной агитации. Лучший результат для нейросетевых классификаторов был получен с использованием нейронной сети BERT-SPC, предобученной на задаче определения тональности по отношению к явно упоминаемому аспекту, макро-F-мера составила 0.74. Лучший результат для алгоритма, основанного на семантических правилах, был получен при использовании метода поиска аспектных терминов на основе семантической схожести, макро-F-мера составила 0.63. При объединении BERT-SPC и алгоритма, основанного на правилах, в ансамбль была получена макро-F-мера, равная 0.79, что является лучшим результатом, полученным в рамках работы.

Ключевые слова: анализ тональности; определение тональности; тональность по отношению к аспекту; неявно упоминаемые аспекты; семантические правила; публицистический стиль; дерево синтаксических единиц

ИНФОРМАЦИЯ ОБ АВТОРАХ

Полетаев, Анатолий Юрьевич (автор для корреспонденции)	ORCID iD: 0000-0003-0116-4739 . E-mail: anatoliy-poletaev@mail.ru Старший преподаватель
Парамонов, Илья Вячеславович	ORCID iD: 0000-0003-3984-8423 . E-mail: ilya.paramonov@fruct.org Канд. физ.-мат. наук, доцент
Бойчук, Елена Игоревна	ORCID iD: 0000-0001-6600-2971 . E-mail: elena-boychouk@rambler.ru Доктор филол. наук, доцент

Финансирование: Российский научный фонд (проект № 23-21-00495).

Для цитирования: А. Ю. Poletaev, I. V. Paramonov, and E. I. Boychuk, “Methods of sentiment detection towards aspect of economic and social development in Russian sentences”, *Modeling and Analysis of Information Systems*, vol. 31, no. 4, pp. 362–383, 2024. DOI: [10.18255/1818-1015-2024-4-362-383](https://doi.org/10.18255/1818-1015-2024-4-362-383).

Введение

Анализ тональности — одно из направлений компьютерной лингвистики, целью которого является определение выраженного в тексте авторского отношения [1], обычно положительного, отрицательного или нейтрального. Определяться может отношение автора как к теме текста в целом (общая тональность), так и к отдельным аспектам темы (тональность по отношению к аспекту) [2].

Автор может выражать отношение к аспекту, даже если не упоминает его явно. Рассмотрим предложение «Горожане тратят слишком много времени на очереди в больницах, врачам нужно брать пример с качества работы банковской системы». В нём автор выражает отрицательную тональность по отношению к качеству здравоохранения (аспекту социально-экономического развития), при этом не используя термин «здравоохранение», и положительную тональность по отношению к банковской системе (также являющейся аспектом социально-экономического развития), используя термин «банковская система». В таком случае аспект «банковская система» называется явно упоминаемым (*explicit aspect*), а аспект «здравоохранение» — неявно упоминаемым (*implicit aspect*) [3].

Задача определения тональности по отношению к упоминаемому в предложении аспекту заключается в определении, относится ли автор предложения к некоторому аспекту, о котором идёт речь в предложении, положительно, отрицательно или нейтрально. Отличие от задачи определения тональности по отношению к явно упоминаемому аспекту состоит в том, что неизвестно, сообщает ли автор о своём отношении к аспекту с использованием самого названия аспекта, или же он делает это как-то иначе. Например, в обоих предложениях «Хочу сказать, что проехать в наш посёлок стало невозможно» и «Хочу затронуть тему состояния дорожной сети и сказать, что проехать в наш посёлок стало невозможно» упоминается аспект состояния дорожной сети, по отношению к которому выражается отрицательная тональность, но только во втором случае он упоминается явно.

Необходимо отметить, что определение неявно упоминаемых в предложении аспектов (*implicit aspect detection*) представляет собой отдельную задачу, которая обычно решается независимо от задачи определения тональности по отношению к аспекту [4] и в данной статье не рассматривается. Задаче определения неявно упоминаемых аспектов в публицистических предложениях на русском языке посвящена одна из предыдущих работ авторов [5].

Поскольку публицистические тексты часто выражают настроения в обществе, определение тональности, выражаемой в них по отношению к аспектам социально-экономического развития, позволяет оценивать отношение общества к происходящим событиям и определять, какие проблемы являются наиболее значимыми [6]. Однако из-за того, что в настоящий момент для русского языка задача определения тональности по отношению к неявно упоминаемому аспекту практически не исследована, при проведении прикладных исследований приходится использовать менее точные методы, например, анализировать общую тональности текста, посвящённого интересующему аспекту [7].

Данная работа посвящена методам решения задачи определения тональности по отношению к аспектам социально-экономического развития в публицистических предложениях на русском языке. Цель работы — разработка методов определения тональности по отношению к аспектам. Задачи работы включают как исследование возможностей определения тональности с использованием существующих архитектур нейронных сетей, так и разработку нового алгоритма определения тональности, основанного на семантических правилах, а также построение ансамблевого классификатора, объединяющего эти методы.

Работа организована следующим образом. В разделе 1 приведён обзор существующих методов, которые могут быть использованы для решения поставленной задачи. Раздел 2 описывает набор данных, использованный для экспериментов. В разделе 3 приведены результаты экспериментов с нейросетевыми классификаторами. В разделе 4 описан алгоритм определения тональности, осно-

ванный на семантических правилах. Эксперименты с построением ансамбля нейросетевого классификатора и алгоритма, основанного на правилах, приведены в разделе 5. В заключении обсуждаются полученные в ходе экспериментов результаты и подводятся итоги работы.

1. Обзор существующих методов

Для английского языка опубликовано несколько работ, посвящённых задаче определения тональности по отношению к аспекту [4, 8]. Существуют как методы определения тональности с помощью семантических правил, использующих знания экспертов-лингвистов, так и методы, использующие различные алгоритмы машинного обучения.

В работе [9] для определения тональности (положительной или отрицательной) по отношению к неявно упоминаемым аспектам политической жизни общества используются семантические правила. На первом этапе эксперт отмечает в тексте термины, относящиеся к интересующему аспекту (аспектные термины). Например, для аспекта «правительство» аспектными терминами являются как само слово «правительство», так и «суды» или «режим». На втором этапе с помощью тонального словаря и набора из 15 семантических правил определяется, сообщается ли о термине положительная или отрицательная информация. При применении предложенного метода к набору из 300 текстов с Reddit доля правильных ответов составила 0.92. Необходимо отметить, что описанный в данной работе метод не позволяет автоматически определять тональность по отношению к неявно упоминаемому аспекту, так как он требует, чтобы эксперт вручную отметил в тексте термины, связанные с аспектом.

Работа [10] посвящена определению тональности (также положительной или отрицательной) по отношению к аспектам, неявно упоминаемым в заголовках новостей и коротких текстах о финансовых рынках. Для определения в предложении аспектных терминов использовались три метода: задание экспертом ключевых слов для каждого аспекта; использование метрики семантической схожести (вычисляемой с помощью BERT) для определения связанных с каждым аспектом слов предложения; гибридный метод, использующий метрику схожести, вычисляемую по отношению ко всем ключевым словам, и агрегирующий результаты с помощью Bi-LSTM. Для определения тональности по отношению к аспекту также были использованы три метода: нейронная сеть Bi-LSTM, применяемая к вектору-эмбедингу предложения (эмбединги были обучены на Wikipedia); классификатор на основе BERT, применяемый к фрагментам предложения, синтаксически связанным со словами, связанным с аспектным термином; гибридная модель из BERT и Bi-LSTM. Оценка качества методов производилась на наборе FiQA 2018, состоящем из 1057 записей. Наилучший результат с F-мерой 0.92 был получен при использовании гибридных моделей из BERT и Bi-LSTM как для определения аспектных терминов, так и для определения тональности. При использовании остальных методов качество было существенно ниже, вплоть до F-меры, равной 0.74.

Метод, использующий как семантические правила, так и машинное обучение, описывается в работе [11], посвящённой задаче определения отношения к демонетизации (положительного, отрицательного или нейтрального). Применяемый в этой работе метод может также применяться для определения тональности по отношению к произвольным аспектам социально-экономического развития. На первом этапе с помощью словаря в предложении отыскиваются слова, выражающие мнения. На втором этапе с помощью анализа синтаксической структуры предложения определяется, относятся ли эти мнения к какому-либо явно упоминаемому в предложении объекту. Если мнение относится к какому-либо объекту, то этот объект считается явно упоминаемым аспектом; если нет, то считается, что мнение относится к неявно упоминаемому аспекту. Тональность мнений о явно упоминаемых аспектах определялась с помощью алгоритма VADER [12], а о неявно упоминаемых — с помощью полносвязной нейронной сети, получающей на вход эмбединги слов Word2Vec. Для агрегации мнений использовалась рекуррентная нейронная сеть. При применении

метода на трёх наборах твитов размером от 6000 записей авторами были получены значения F-меры от 0.85 до 0.88.

Определение тональности по отношению к неявно упоминаемым аспектам в несколько этапов является достаточно распространённым подходом [8]. При этом на первом этапе в тексте обнаруживаются термины, относящиеся к целевому аспекту, а на втором этапе производится непосредственно определение тональности по отношению к этим терминам.

Для русского языка большая часть методов аспектно-ориентированного анализа тональности разработана для анализа отзывов на товары и услуги, тем не менее некоторые из них могут использоваться для создания методов определения тональности по отношению к аспектам социально-политического развития. Так, для поиска в предложении аспектных терминов, выполнявшегося вручную в [9], может быть использован метод, предложенный в работе [13] для поиска аспектных терминов в текстах отзывов. Он основан на семантическом сходстве различных слов предложения и названия самого оцениваемого товара. Для вычисления метрики сходства использовались эмбединги gensim, обученные на наборе из 40 тысяч отзывов. Для принятия решения о том, является ли слово аспектным термином, использовался алгоритм на основе деревьев решений и кластерного анализа. При валидации на данных из корпуса SentiRuEval-2015 F-мера составила 0.72 при точности 0.81 и полноте 0.65.

Кроме того, для русского языка описано несколько методов определения тональности по отношению к явно упоминаемым аспектам. В предыдущей работе авторов [14] было показано, что среди классификаторов, основанных на машинном обучении, с этой задачей лучше всего справляется BERT-SPC: макро-F-мера на корпусе CABSAR составила 0.71 (при трёхклассовой классификации). Также в работе был предложен алгоритм, основанный на семантических правилах. Его качество на CABSAR оказалось немного ниже, чем у BERT-SPC: макро-F-мера составила 0.70. Значительное улучшение качества было достигнуто при объединении в ансамбль BERT-SPC и алгоритма, основанного на семантических правилах: макро-F-мера ансамбля достигла 0.81.

Таким образом, для определения тональности по отношению к аспектам могут успешно применяться как методы машинного обучения, так и алгоритмы, основанные на семантических правилах. При определении тональности по отношению к неявно упоминаемым аспектам часто требуется решение вспомогательной задачи поиска в предложении аспектных терминов.

2. Разметка корпуса текстов

Для проведения экспериментов по определению тональности по отношению к аспекту был собран и размечен корпус публицистических предложений. В качестве источника были выбраны материалы предвыборной агитации, поскольку в ней встречается достаточно много тональных предложений и оценочных высказываний [15].

Для составления корпуса было использовано 10 случайно выбранных агитационных статей в газетах и журналах региональных выборов, проходивших в России в 2005–2015 годах. Статьи были разделены на предложения, слишком короткие предложения (менее 4 слов) были удалены. Чтобы обеспечить беспристрастность разметки, предложения были анонимизированы: имена и фамилии, географические названия, названия политических партий и компаний были заменены на вымышленные. Анонимизация проводилась экспертом-лингвистом вручную, так, чтобы тональность предложений не искажалась.

Для разметки было выбрано 10 аспектов социально-экономического развития:

- газификация;
- качество здравоохранения;
- качество и стоимость услуг ЖКХ;
- качество управления;
- проведение выборов;

Table 1. Classes ratio in the corpus**Таблица 1.** Соотношение классов в корпусе

Аспект	Положит.	Нейтр.	Отриц.	Итого
газификация	12	6	8	26
качество здравоохранения	8	8	12	28
качество и стоимость услуг ЖКХ	9	16	48	73
качество управления	46	15	35	96
проведение выборов	11	36	28	75
развитие промышленности	16	21	21	58
развитие сельского хозяйства	11	14	7	32
состояние дорожной сети	5	10	15	30
состояние жилья	3	5	11	19
строительство АЭС	22	5	4	31
Итого	143	136	189	468

- развитие промышленности;
- развитие сельского хозяйства;
- состояние дорожной сети;
- состояние жилья;
- строительство атомной электростанции.

Разметка проводилась 9 волонтерами. Перед ними была поставлена задача определить для каждого предложения:

- аспекты социально-экономического развития из списка выше, о которых идёт речь в предложении;
- тональность (которая может быть положительной, отрицательной или нейтральной) по отношению к каждому из аспектов.

Предложения были случайно распределены между разметчиками так, чтобы каждое предложение было оценено хотя бы 5 разметчиками. В случае, если мнения всех разметчиков об упоминании аспекта и отношении к нему совпадали, такая запись сразу же включалась в корпус. В противном случае пример рассматривался экспертом-лингвистом, который принимал решение о том, включать ли запись в корпус, и с какой меткой тональности.

В результате был получен корпус из 468 записей, каждая из которых состоит из текста предложения, аспекта социально-экономического развития и метки тональности. Всего в корпус попали 377 уникальных предложений. В 83 % из них упоминается только один аспект, в 12 % предложений – два аспекта, в 5 % предложений – три или четыре аспекта. При этом для 43 % предложений, в которых упоминается два или более аспекта, тональность по отношению к ним не совпадает, а в 10 % таких предложений к одному из аспектов выражается положительная тональность, а к другому – отрицательная. Явное упоминание аспектов встречается всего в 32 % случаев, т.е. в 68 % случаев встречаются только их неявные упоминания.

Соотношение аспектов и классов тональности в полученном корпусе приведено в таблице 1.

Доли предложений, входящих в положительный, нейтральный и отрицательный классы, составляют 31%, 29% и 40% соответственно, что свидетельствует о слабом дисбалансе классов. В то же время, для отдельных аспектов присутствует достаточно сильный дисбаланс классов: например, 71% упоминаний строительства атомной электростанции положительны, а 65% упоминаний качества и стоимости услуг ЖКХ отрицательны. При этом практически для всех аспектов преобладают именно тональные упоминания, только для аспектов «проведение выборов» и «развитие сельского хозяйства» преобладающим классом являются нейтральные упоминания.

Table 2. BERT-SPC sentiment detection performance

Тональность	Точность	Полнота	F-мера	Число предложений
положительная	0.70	0.65	0.67	143
нейтральная	0.69	0.67	0.68	136
отрицательная	0.78	0.84	0.81	189
макроусреднение	0.72	0.72	0.72	468

Таблица 2. Качество определения тональности при использовании BERT-SPC

Доля правильных ответов — 0.73

Table 3. BERT-SPC sentiment detection confusion matrix

		Предсказанная тональность			Итого
		положительная	нейтральная	отрицательная	
Факт. тон.	положительная	93	26	24	143
	нейтральная	25	91	20	136
	отрицательная	15	15	159	189

Таблица 3. Матрица ошибок при использовании BERT-SPC

3. Нейросетевые классификаторы

В данном разделе описываются эксперименты по применению нейронных сетей для определения тональности по отношению к аспектам социально-экономического развития.

Эксперименты проводились с классификатором BERT-SPC [16] с предобученными весами *rubert-base-cased-sentence*¹, поскольку они лучше всего показали себя при решении сходной задачи — определении тональности по отношению к явно упоминаемому аспекту [14]. Для реализации нейронных сетей была использована библиотека PyTorch [17].

3.1. Классификатор, обучаемый только на корпусе предвыборной агитации

Первый эксперимент проводился с обучением BERT-SPC только на корпусе предвыборной агитации. Для оценки качества использовалась кросс-валидация с разделением на 5 подвыборки, сбалансированных по классам тональности. Из-за высокой вычислительной сложности обучения нейросетевого классификатора со всеми возможными комбинациями гиперпараметров вначале была выбрана лучшая комбинация значений гиперпараметров из 10 случайно составленных, а затем был произведён перебор значений некоторых гиперпараметров: коэффициент дропаута (dropout rate) от 0 до 0.75 с шагом 0.05; размер батча (batch size), равный 1, 4, 8, 12, 16; скорость обучения (learning rate) от 0.001 до 0.05 с шагом 0.001; коэффициент затухания весов (weight decay rate) от 0.0001 до 0.005 с шагом 0.0001.

Наивысшая макро-F-мера была достигнута при 512 нейронах в скрытом слое и следующих значениях гиперпараметров: коэффициент дропаута (dropout rate) — 0.25, размер батча (batch size) — 16, оптимизатор SGD, скорость обучения (learning rate) — 0.005, коэффициент затухания весов (weight decay rate) — 0.001, максимум эпох обучения — 30, остановка обучения при отсутствии повышения качества на валидационной выборке за 5 эпох. Полученные метрики качества и матрица ошибок приведены в таблицах 2 и 3 соответственно.

Для сравнительно небольшого объёма обучающих данных было получено достаточно высокое качество. Оно практически идентично полученному в [14] для сходной задачи с помощью модели, обученной на существенно большем (на порядок) объёме данных. Тем не менее нужно отметить достаточно низкую полноту определения предложений с положительной тональностью — больше

¹<https://huggingface.co/DeepPavlov/rubert-base-cased-sentence>

Table 4. Sentiment detection performance of BERT-SPC trained in two stages

Тональность	Точность	Полнота	F-мера	Число предложений
положительная	0.72	0.73	0.73	143
нейтральная	0.67	0.70	0.69	136
отрицательная	0.82	0.78	0.80	189
макроусреднение	0.74	0.74	0.74	468

Таблица 4. Качество определения тональности при использовании BERT-SPC, обученного в два этапа

Доля правильных ответов – 0.74

Table 5. Sentiment detection confusion matrix of BERT-SPC trained in two stages

		Предсказанная тональность			Итого
		положительная	нейтральная	отрицательная	
Факт. тон.	положительная	105	22	16	143
	нейтральная	24	95	17	136
	отрицательная	17	24	148	189

Таблица 5. Матрица ошибок при использовании BERT-SPC, обученного в два этапа

трети из них ошибочно определяются как нейтральные или отрицательные. В то же время, предложения с отрицательной тональностью определяются и достаточно точно, и достаточно полно.

3.2. Классификатор, дообучаемый на корпусе предвыборной агитации

В ходе второго эксперимента классификатор BERT-SPC обучался в два этапа. На первом этапе модель обучалась задаче определения тональности по отношению к явно упоминаемому аспекту, для чего использовалась обучающая часть корпуса CABSAR без записей из Twitter [18] и параметры, с которыми был получен лучший результат в [14]. На втором этапе полученная модель дообучалась на корпусе предвыборной агитации, с использованием того же разделения для кросс-валидации и той же методики выбора гиперпараметров, что и при первом эксперименте. Наивысшая макро-F-мера по результатам второго этапа была достигнута со следующими значениями гиперпараметров: коэффициент дропаута (dropout rate) – 0.1, размер батча (batch size) – 16, оптимизатор SGD, скорость обучения (learning rate) – 0.005, коэффициент затухания весов (weight decay rate) – 0.001, максимум эпох обучения – 15, остановка обучения при отсутствии повышения качества на валидационной выборке за 5 эпох. Полученные метрики качества и матрица ошибок приведены в таблицах 4 и 5 соответственно.

По сравнению с первым экспериментом качество определения тональности возросло за счёт предложений с положительной тональностью. Сильнее всего увеличилась полнота обнаружения положительной тональности – на 0.08. Для отрицательной тональности возросла точность, но снизилась полнота: 11 предложений, ранее верно определявшиеся как отрицательные, классификатор стал относить к нейтральным или положительным.

Для каждого аспекта по отдельности была рассчитана доля правильных ответов. Результаты, упорядоченные по доле правильных ответов, приведены в таблице 6.

Лучше всего определяется тональность по отношению к аспекту «качество здравоохранения», также результат выше среднего получен для аспектов «состояние дорожной сети», «качество и стоимость услуг ЖКХ», «строительство атомной электростанции» и «качество управления». В то же время, значительно ниже среднего оказался результат для аспектов «проведение выборов» и «газификация». Таким образом, результаты достаточно сильно отличаются для некоторых аспектов

Table 6. Accuracy and number of sentences by aspect**Таблица 6.** Доля правильных ответов и число предложений для аспектов

Аспект	Доля правильных ответов	Число предложений
качество здравоохранения	0.89	28
состояние дорожной сети	0.80	30
качество и стоимость услуг ЖКХ	0.79	73
строительство АЭС	0.77	31
качество управления	0.77	96
состояние жилья	0.74	19
развитие промышленности	0.72	58
развитие сельского хозяйства	0.72	32
проведение выборов	0.65	75
газификация	0.58	26

со схожим вкладом в социально-экономическое развитие, в частности, для газификации и строительства атомной электростанции, а также для качества управления и проведения выборов.

Для доли правильных ответов и числа предложений были рассчитаны коэффициенты ранговой корреляции ρ_s Спирмена и τ_b Кендалла, их значения составили -0.05 и -0.11 соответственно. Проверка по критерию Стьюдента показала незначимость связи (p ложного отклонения истинной гипотезы о существовании связи 0.88 и 0.65 соответственно), что даёт основания считать, что доля правильных ответов для аспекта не зависит от числа предложений. Следовательно, нейросетевой классификатор может достаточно хорошо обучиться определять тональность даже по отношению к аспектам, предложений с которыми в обучающем наборе немного.

3.3. Дополнительные эксперименты

Для лучшего изучения задачи определения тональности по отношению к аспекту были проведены два дополнительных эксперимента с нейросетевыми классификаторами.

В ходе первого эксперимента изучалась возможность использования классификатора, обученного задаче определения тональности по отношению к явно упоминаемому аспекту, для определения тональности по отношению к аспекту, без дообучения. Поскольку модели архитектуры BERT в некоторых случаях могут показывать хороший результат и без дообучения [19], а задачи определения тональности по отношению к аспекту и к явно упоминаемому аспекту близки, было решено проверить, насколько существенным окажется снижение качества на другой задаче по сравнению со снижением качества на другом корпусе для той же задачи. Для этого качество классификатора BERT-SPC, обученного на обучающей части CABSAR без записей из Twitter, оценивалось на следующих корпусах:

- тестовой части CABSAR без записей из Twitter, макро-F-мера 0.71 (явно упоминаемые аспекты);
- тестовой части RuSentNE-2023 [20], макро-F-мера 0.53 (явно упоминаемые аспекты);
- корпусе политической агитации, макро-F-мера 0.33 (и явно, и неявно упоминаемые аспекты).

Таким образом, при изменении задачи качество снизилось достаточно существенно — на 0.38 , тогда как для той же задачи, но на другом корпусе снижение составило всего 0.18 . Следовательно, без дообучения использовать нейросетевой классификатор, обученный определению тональности по отношению к явно упоминаемому аспекту, нельзя.

Во втором эксперименте изучалось качество определения тональности с помощью нейросетевых классификаторов IAN [21], использующего эмбединги ELMo. Эксперимент проводился, поскольку при небольших объёмах обучающей выборки более простые архитектуры нейронных сетей, такие как IAN, часто показывают лучший результат, чем более сложные, такие как BERT-SPC. Наивысшая

полученная макро-F-мера составила 0.70, т. е. для задачи определения тональности по отношению к аспекту IAN показывает качество немного ниже, чем BERT-SPC. Такое же соотношение между метриками качества этих двух классификаторов было получено в предыдущей работе авторов при существенно большем объёме обучающей выборки для задачи определения тональности по отношению к явно упоминаемому аспекту [14].

4. Алгоритм, основанный на семантических правилах

Несмотря на то, что с помощью нейросетевых классификаторов было получено достаточно высокое качество определения тональности, в рамках данной работы также предлагается алгоритм определения тональности по отношению к аспекту, основанный на семантических правилах. Такие алгоритмы могут объединяться с нейросетевыми классификаторами в ансамбль, качество которого будет выше, чем обоих методов по отдельности, что показано в предыдущей работе авторов по смежной теме [14].

Алгоритм использует описанный в [14] метод определения тональности по отношению к явно упоминаемым аспектам с помощью набора семантических правил, применяющихся к дереву синтаксических единиц. Результат его работы – метка положительной, отрицательной или нейтральной тональности. Обозначим результат для предложения S и явно упоминаемого аспекта t как $E(S, t)$.

Для определения общей тональности предложения используется рекурсивный алгоритм определения тональности, описанный в работе [22]. Обозначим результат его работы для предложения S как $G(S)$.

4.1. Общая схема алгоритма

Для определения тональности по отношению к аспекту алгоритм использует аспектные термины – называемые в предложении события и явления, связанные с целевым аспектом социально-экономического развития. Один аспектный термин может быть связан с аспектом более или менее тесно, чем другой. Например, в предложении «Несмотря на то, что управляющая компания делает всё возможное, проблем с отоплением всё больше и больше» есть два аспектных термина, связанных с аспектом «качество и стоимость услуг ЖКХ»: «управляющая компания» и «отопление». При этом термин «отопление» связан с аспектом теснее, чем «управляющая компания», поскольку отрицательная тональность по отношению к отоплению важнее для определения тональности по отношению к качеству и стоимости услуг ЖКХ, чем положительная тональность по отношению к управляющей компании.

Определение тональности выполняется в два этапа. На первом этапе в S ищутся аспектные термины, из которых формируется список $T_a = (t_1, \dots, t_n)$, упорядоченный по убыванию тесноты связи с аспектом. Если аспектных терминов обнаружено не было, то результатом работы алгоритма является общая тональность предложения $G(S)$.

На втором этапе для каждого аспектного термина $t_i \in T_a$, $i = 1, \dots, n$, последовательно определяется тональность по отношению к термину $e_i = E(S, t_i)$. При нахождении первого значения e_i , являющегося положительным или отрицательным, это значение и становится результатом работы алгоритма. Если для всех $t \in T_a$ тональность оказалась нейтральной, то результат работы алгоритма – общая тональность предложения $G(S)$.

Поскольку по мнению эксперта-лингвиста для аспекта «качество управления» аспектным термином может являться практически любое событие или явление, для определения тональности по отношению к этому аспекту всегда используется общая тональность предложения.

Было разработано три метода поиска в предложении аспектных терминов: на основе перечня терминов, на основе семантической схожести и гибридный.

4.2. Метод поиска аспектных терминов на основе перечня терминов

В данном методе для поиска аспектных терминов используется составленный экспертом-лингвистом для каждого из аспектов социально-экономического развития перечень потенциальных аспектных терминов:

- газификация: газификация, газ;
- качество здравоохранения: здравоохранение, медицинское учреждение, больница, поликлиника, медобслуживание;
- качество и стоимость услуг ЖКХ: ЖКХ, тариф, отопление, вывоз мусора, управляющая компания;
- развитие промышленности: промышленность, производство, предприятие, АПК, железнодорожный узел;
- развитие сельского хозяйства: сельское хозяйство, сельхозпредприятие, АПК;
- состояние дорожной сети: дорога, улица;
- состояние жилья: жильё, дом, жилой фонд;
- строительство атомной электростанции: строительство АЭС, АЭС, атомная станция.

При использовании данного метода в T_a сначала включаются все упоминания первого потенциального аспектного термина из перечня, затем второго и так далее.

Для аспекта «проведение выборов» метод неприменим, так как потенциальных аспектных терминов слишком много.

4.3. Метод поиска аспектных терминов на основе семантической схожести

При использовании этого метода аспектные термины ищутся на основе семантической схожести. Формирование списка аспектных терминов в предложении происходит в три этапа.

На первом этапе составляется список кандидатов в аспектные термины. В него включаются все подлежащие, дополнения и главные части составных именных сказуемых из предложения S , а также словосочетания, в которых они являются главными словами. Для поиска слов и словосочетаний используется дерево синтаксических единиц, которое строится с помощью алгоритма, описанного в работе [23].

На втором этапе для каждого кандидата рассчитывается метрика семантической схожести со строкой-названием аспекта, рассчитываемая как косинусное расстояние между векторами эмбедингов кандидата и аспекта [24], рассчитываемые с помощью Fasttext, предобученного в рамках проекта RusVectōrēs [25] (модель `geowac_tokens_none_fasttextskipgram_300_5_2020`).

На третьем этапе формируется итоговый список T_a , в которые попадают кандидаты со значением метрики схожести, большей 0.5, упорядоченные по убыванию её значения.

4.4. Гибридный метод поиска аспектных терминов

Данный метод формирования списка аспектных терминов представляет собой гибрид двух методов, описанных выше: аспектные термины выделяются на основе семантического сходства с потенциальными аспектными терминами из перечня.

Введём обозначения: T_p — список потенциальных аспектных терминов, $P(a)$ — список аспектных терминов, найденных в предложении S на основе семантического сходства с a . Итоговый T_a формируется в результате конкатенации всех $P(t_p)$, составленных для каждого $t_p \in T_p$ вне зависимости от того, упоминается ли сам t_p в предложении.

4.5. Дополнительные семантические правила

В описанный в [14] набор семантических правил, используемых алгоритмом определения тональности по отношению к аспекту, было добавлено семь правил, отражающих знания лингвистов о существующих в русском языке способах выражения тональности. В примерах этого подраздела

жирным шрифтом выделяется аспектный термин, а курсивом — часть предложения, приводящая к применению правила.

Следующие три правила были добавлены для обработки связи между синтаксической группой, содержащей аспектный термин, и другими синтаксическими группами, тональность которых также относится к аспекту:

1. Если в первую часть сложного предложения входит аспектный термин, а во вторую — местоимение в том же роде и числе, то тональностью по отношению к аспекту считается тональность синтаксической группы второй части сложного предложения: «У нас было градообразующее **предприятие**, теперь *оно*, к сожалению, утратило своё значение» («оно, к сожалению, утратило своё значение», тональность отрицательная).
2. Если упоминается подготовка, то тональностью по отношению к аспекту считается тональность синтаксической группы, содержащей информацию о подготовке: «Начало **строительства АЭС** позволило бы решить ряд вопросов округа, но в настоящее время *подготовительные работы* идут ни шатко ни валко» («подготовительные работы идут ни шатко ни валко», тональность отрицательная).
3. Если автор выносит оценку общему понятию и сообщает, что аспектный термин входит в это понятие, то тональностью по отношению к аспекту считается тональность синтаксической группы, сказанное о котором относится и к аспекту: «Экономика района находится не в лучшем состоянии, *это касается и промышленности, и сельского хозяйства*» («Экономика района находится не в лучшем состоянии», тональность отрицательная).

Также были добавлены четыре правила для обработки различных закономерностей выражения тональности по отношению к аспектам социально-экономического развития:

1. Если автор сообщает, что было начато использование аспектного термина, а также о том, что ситуация в целом улучшилась, отношение автора к аспекту считается положительным, даже если использован соединительный союз: «В 2003 году начался пуск природного **газа**, и сегодня жители не знают забот» (аспект — газификация). Это правило применяется для синтаксических групп сложносочинённых предложений и однородных подлежащих, в первую часть которых (в первое простое предложение или в группу первого однородного подлежащего) входит аспектный термин в роли подлежащего или прямого дополнения, а также один из глаголов «начинаться», «начаться», «начать», «начинать» или образованное от него слово. В таком случае, если тональность первой части синтаксической группы нейтральна, а второй — положительна, тональность по отношению к аспекту также считается положительной.
2. Если автор сообщает, что что-либо негативное для аспектного термина не существует, то отношение к аспекту считается положительным: «У нас *нет* «тёмных лошадок», которых **избиратели** бы не знали» (аспект — проведение выборов). Это правило применяется, если синтаксическая группа с отрицательной тональностью по отношению к аспекту («тёмных лошадок, которых избиратели бы не знали») выступает в роли подлежащего в предложении, синтаксическая группа сказуемого в котором имеет семантику отсутствия: либо включает глагол «быть» и отрицание, либо состоит из одного слова «нет».
3. Если автор сообщает, что аспектный термин требует улучшения, то отношение к аспекту считается отрицательным: «*Нужно*, чтобы **управляющая компания** начала работать нормально» (аспект — качество и стоимость услуг ЖКХ). Это правило применяется к предложениям или к синтаксическим группам сказуемых, содержащих аспектный термин в качестве подлежащего или дополнения (как прямого, так и непрямого). В случае, если группа сказуемого содержит одно из слов «необходимо», «нужно», «нет», а группа подлежащего имеет положительную тональность, тональность по отношению к аспекту считается отрицательной.

Table 7. Sentiment detection performance of the algorithm with different methods of aspect terms searching**Таблица 7.** Качество работы алгоритма с различными методами поиска аспектных терминов

Тональность	Метод поиска аспектных терминов									Число предл.
	на основе перечня			на основе семантической схожести			гибридный			
	Точн.	Полн.	F-мера	Точн.	Полн.	F-мера	Точн.	Полн.	F-мера	
положительная	0.68	0.57	0.62	0.61	0.62	0.62	0.61	0.63	0.62	143
нейтральная	0.45	0.62	0.52	0.51	0.49	0.50	0.51	0.47	0.49	136
отрицательная	0.75	0.63	0.69	0.76	0.77	0.76	0.75	0.77	0.76	189
Макроусреднение	0.63	0.61	0.61	0.63	0.63	0.63	0.62	0.62	0.62	468
Доля правильных ответов	0.61			0.64			0.64			

4. Если автор сообщает, что будет обращать особое внимание на аспектный термин, то отношение к аспекту считается положительным: «*Моя задача как депутата — максимально ускорить **газификацию***» (аспект — газификация). Это правило применяется к предложениям и к синтаксическим группам сказуемых, включающим в себя местоимение первого лица. Если в синтаксическую группу входит аспектный термин, а также хотя бы одно из слов «первоочередной», «приоритет», «задача», «нужно», «за», и синтаксическая группы нейтральна, тональность по отношению к аспекту считается нейтральной.

4.6. Оценка качества

Для оценки качества работы алгоритма он был реализован на языке Python с использованием библиотек PyMorphy2 [26] и Stanza [27], а также алгоритма построения деревьев синтаксических единиц, описанного в работе [23].

Метрики качества определения тональности и матрицы ошибок для алгоритма при использовании каждого из трёх методов поиска аспектных терминов приведены в таблицах 7 и 8. Качество определения тональности существенно ниже, чем при использовании нейросетевого классификатора: лучшая макро-F-мера на 0.11 ниже, чем у BERT-SPC. Такая разница вызвана менее качественным определением положительной и отрицательной тональности: F-меры для положительного и отрицательного классов ниже, чем у BERT-SPC, на 0.11 и 0.19 соответственно. При этом отрицательная тональность определяется практически так же хорошо, как и с помощью BERT-SPC: полнота ниже всего на 0.01.

При использовании метода поиска на основе семантической схожести и гибридного метода качество в среднем выше на 1–3 %, чем при использовании метода на основе перечня аспектных терминов, за счёт более полного обнаружения тональных предложений. Особенно сильно отличается полнота определения отрицательной тональности: при использовании метода на основе перечня терминов она составила 0.63, а при использовании двух других методов — 0.77. В то же время при использовании метода на основе перечня терминов значительно выше точность определения положительной тональности: 0.68 против 0.61 при использовании остальных методов. При использовании гибридного метода алгоритм относит к тональным на 6 предложений больше, чем при использовании метода на основе семантической схожести, при этом только 2 из этих 6 предложений действительно являются тональными. Таким образом, при использовании метода на основе семантической схожести больше предложений определяются как положительные или отрицательные, что повышает полноту, но также может привести к снижению точности.

Для оценки влияния добавленных семантических правил на качество определения тональности была рассчитана макро-F-мера для алгоритма без добавленных правил, только с описанными в ра-

Table 8. Sentiment detection confusion matrices of the algorithm with different methods of aspect terms searching**Таблица 8.** Матрицы ошибок алгоритма с различными методами поиска аспектных терминов

		Метод поиска аспектных терминов								
		на основе перечня			на основе семантической схожести			гибридный		
		Предсказанная тональность								
		полож.	нейтр.	отриц.	полож.	нейтр.	отриц.	полож.	нейтр.	отриц.
Факт. тон.	полож.	82	46	15	89	36	18	90	34	19
	нейтр.	27	85	24	41	67	28	43	64	29
	отриц.	12	58	119	15	29	145	15	28	146

Table 9. Accuracy by aspect with different methods of aspect term searching**Таблица 9.** Доля правильных ответов для аспектов при использовании различных методов поиска аспектных терминов

Аспект	Метод поиска аспектных терминов			BERT-SPC
	на основе перечня	на основе семантической схожести	гибридный	
качество здравоохранения	0.61	0.68	0.71	0.89
состояние дорожной сети	0.60	0.67	0.67	0.80
качество и стоимость услуг ЖКХ	0.59	0.67	0.67	0.79
строительство АЭС	0.61	0.74	0.74	0.77
состояние жилья	0.68	0.79	0.79	0.74
развитие промышленности	0.59	0.64	0.60	0.72
развитие сельского хозяйства	0.72	0.59	0.56	0.72
проведение выборов	—	0.63	—	0.65
газификация	0.58	0.58	0.62	0.58

Для аспекта «качество управления» доля правильных ответов алгоритма составила 0.59, доля правильных ответов BERT-SPC — 0.77

боте, посвящённой определению тональности по отношению к явно упоминаемому аспекту [14]. При использовании для поиска аспектных терминов метода на основе перечня она составила 0.57 (изменение на 0.04), при использовании метода на основе семантической схожести — 0.61 (изменение на 0.02), при использовании гибридного метода — 0.60 (изменение на 0.02). Таким образом, добавление новых правил привело к повышению качества, но и без них алгоритм показал бы близкий результат.

Чтобы проанализировать качество определения тональности по отношению к различным аспектам, была рассчитана доля правильных ответов для каждого аспекта при использовании каждого из методов. Полученные значения приведены в таблице 9, также справочно приведены доли правильных ответов BERT-SPC, обученного в два этапа.

Доли правильных ответов для различных аспектов варьируются достаточно сильно при использовании метода поиска аспектных терминов на основе семантической схожести и гибридного метода. При этом аспекты, для которых тональность определяется более качественно и менее качественно, в основном оказались такими же, что и для BERT-SPC. Исключением оказался аспект

«качество управления», для которого только 59 % ответов оказались правильными, тогда как BERT-SPC даёт правильный ответ в 77 % случаев.

При использовании метода поиска аспектных терминов на основе перечня качество для подавляющего большинства аспектов приблизительно одинаковое — доля правильных ответов находится в диапазоне от 0.58 до 0.61. Исключение представляют только аспекты «развитие сельского хозяйства» и «состояние жилья», для которых качество существенно выше — доля правильных ответов составляет 0.72 и 0.68 соответственно. Нужно отметить, что для аспекта «развитие сельского хозяйства» доля правильных ответов при использовании этого метода совпадает с полученной с помощью BERT-SPC — 0.79.

При использовании метода на основе семантической схожести качество для различных аспектов варьируется гораздо сильнее: от 0.58 для аспекта «газификация» до 0.79 для аспекта «состояние жилья». Несмотря на то, что среднее качество при использовании данного метода выше, чем при использовании метода на основе перечня аспектных терминов, для аспекта «развитие сельского хозяйства» он наоборот приводит к существенно худшему результату — доля правильных ответов ниже на 0.13. Также при использовании метода на основе семантической схожести доля правильных ответов для аспекта «состояние жилья» на 0.05 превышает полученную с помощью BERT-SPC.

При использовании гибридного метода доля правильных ответов та же, что и при использовании метода на основе семантической схожести, за исключением аспектов «качество здравоохранения» (доля правильных ответов повысилась на 0.03), «газификация» (повысилась на 0.04), «развитие промышленности» (снизилась на 0.04) и «развитие сельского хозяйства» (снизилась на 0.03). При использовании данного метода получается достичь более высокой доли правильных ответов для аспекта «газификация», что превышает как результат, полученный при использовании остальных методов, так и результат, полученный с помощью BERT-SPC.

4.7. Анализ ошибок

Для предложенного алгоритма определения тональности на основе семантических правил был проведён выборочный анализ ошибок. Для этого для каждого метода поиска аспектных терминов было случайно отобрано 50 предложений, по 10 для каждого из аспектов, тональность которых была определена неверно. Для аспекта «состояние жилья» при использовании для поиска аспектных терминов семантической схожести и гибридного метода таких предложений оказалось всего 4, поэтому было добавлено по дополнительному предложению для аспекта «состояние дорожной сети».

Для каждого предложения определялась непосредственная причина, по которой тональность была определена неверно. Всего было выделено 7 групп таких причин.

1. Неверный морфологический или синтаксический разбор. Ошибки этой группы возникают из-за несовершенства морфологического анализатора, построителя деревьев синтаксических связей или алгоритма построения дерева синтаксических единиц по дереву синтаксических связей.
2. Неверное определение тональности одиночного слова или устойчивого словосочетания. Ошибки этой группы возникают из-за несовершенства тонального словаря.
3. Неверный поиск аспектных терминов. Ошибки этой группы возникают из-за того, что аспектный термин не был найден, или, наоборот, слово, не являющееся аспектным термином, было определено как аспектный термин. Например, в нейтральном предложении «Дворы это ответственность местных властей, а областной депутат должен контролировать их деятельность и реагировать на жалобы избирателей» (аспект — «состояние жилья») при использовании гибридного метода не был обнаружен термин «дворы», из-за чего тональность была ошибочно определена как отрицательная из-за слова «жалобы».

4. Средство выражения тональности не обрабатывается никаким из семантических правил. Ошибки этой группы возникают из-за того, что набора правил, используемых алгоритмом, не хватает для того, чтобы обнаружить, что автор предложения выразил положительную или отрицательную тональность по отношению к аспекту. Например, в предложении «При переходе на новую системы оплаты услуг ЖКХ плата горожан за отопление существенно выросла» (аспект — «качество и стоимость услуг ЖКХ») отрицательная тональность выражается с помощью сообщения о росте нагрузки на горожан, и, чтобы верно определить тональность, требуется дополнительный алгоритм определения того, что рост платы — это негативный факт, в отличие, например, от роста зарплат или пенсий.
5. Применённое семантическое правило неверно обрабатывает средство. Ошибки этой группы возникают из-за того, правило было применено там, где на самом деле его применять не требовалось. Например, в предложении «По-человечески я прекрасно понимаю эту позицию: работы нет, дети уезжают, дорог нет, и люди считают, что властям нет дела до их проблем» (аспект — качество управления) применяется одно из правил рекурсивного выведения тональности, описанных в работе [22]: если автор сообщает, что он понимает или принимает что-либо, то тональность считается положительной. Тем не менее именно в этом предложении автор, несмотря на понимание, выносит отрицательную оценку деятельности власти и общему состоянию дел в обществе.
6. Средство выражения тональности не может быть обработано с помощью алгоритма. Ошибки этой группы возникают из-за того, что для выражения тональности используется более сложное языковое средство, которое невозможно обработать с помощью синтаксических деревьев и правил, реализованных с их помощью. Например, в предложении «Регион эти деньги найдёт, моя задача — привести средства в наш город» (аспект — «качество управления») положительная тональность выражается с помощью порядка слов, показывающего уверенность в качестве управления регионом и в росте благосостояния округа.
7. Ответ алгоритма может быть признан верным. В эту группу отнесены предложения, смысл которых может быть интерпретирован по-разному, и возможна такая интерпретация, при которой тональность совпадает с результатом работы алгоритма, но не совпадает с разметкой. Например, тональность предложения «А работа в комитете по бюджету позволяет изыскивать средства для решения проблем здравоохранения, поскольку этот комитет обеспечивает финансами все остальные направления» по отношению к аспекту «качество здравоохранения» была определена разметчиками как отрицательная, а алгоритмом — как положительная. Возможны две интерпретации предложения: и то, что в здравоохранении много проблем, и на их решение приходится именно изыскивать средства; и то, что на решение проблем здравоохранения будут найдены средства, и ситуация улучшится. Другой пример — предложение «Естественно, начало газификации и активизация промышленных предприятий увеличит нагрузку на наши дороги, которые уже и дорогами можно назвать с большой натяжкой». Тональность по отношению к аспектам «газификация» и «активизация промышленных предприятий» была определена разметчиками как положительная, поскольку автор сообщает о будущем улучшении ситуации, а алгоритмом — как отрицательная, поскольку и газификация, и промышленность увеличат нагрузку на дороги, с которыми ситуация уже плохая, поэтому отрицательный эффект для социально-экономического развития важнее, чем положительный.

Соотношение групп ошибок при использовании каждого из методов поиска аспектных терминов приведено в таблице 10. Неверный морфологический или синтаксический разбор и несовершенство тонального словаря достаточно слабо повлияли на качество работы алгоритма — ими было вызвано всего 6 % и 8 % ошибок соответственно. Неверный поиск аспектных терминов пред-

Table 10. Groups of errors with different aspect term searching methods**Таблица 10.** Соотношение групп ошибок при использовании различных методов поиска аспектных терминов

Группа ошибок	Метод поиска аспектных терминов			В среднем
	на основе перечня	на основе семантической схожести	гибридный	
неверный морфологический или синтаксический разбор	0.04	0.06	0.06	0.06
неверное определение тональности одиночного слова или устойчивого словосочетания	0.08	0.10	0.06	0.08
неверный поиск аспектных терминов	0.20	0.22	0.16	0.19
средство выражения тональности не обрабатывается никаким из семантических правил	0.26	0.20	0.20	0.22
применённое семантическое правило неверно обрабатывает средство	0.10	0.08	0.16	0.11
средство выражения тональности не может быть обработано с помощью алгоритма	0.22	0.18	0.18	0.19
ответ алгоритма может быть признан верным	0.10	0.16	0.18	0.15

ставляет более серьёзную проблему: им вызвано около 20 % ошибок, причём при использовании гибридного метода поиска доля таких ошибок ниже, чем при использовании других методов. Из-за несовершенства набора семантических правил — отсутствия требуемых и неверного применения существующих — было допущено в сумме до трети всех ошибок. При этом ошибки из-за неверного применения существующих правил допускались значительно реже, чем из-за отсутствия требуемого правила. Также почти в каждом пятом предложении средство выражения тональности в принципе не может быть обработано с помощью дерева синтаксических единиц. Наконец, в среднем в 15 % случаев результат работы алгоритма верен при определённой интерпретации предложения, несмотря на несовпадение с разметкой.

5. Ансамблевый классификатор

Поскольку в работе, посвящённой определению тональности по отношению к явно упоминаемым аспектам [14], было показано, что ансамбль, объединяющий основанный на правилах алгоритм и нейросетевой классификатор, позволяет получить более высокое качество, чем при использовании методов по отдельности, было принято решение построить ансамблевый классификатор из BERT-SPC, обученного в два этапа, и алгоритма, основанного на семантических правилах, из раздела 4.

Для построения ансамбля был использован вспомогательный классификатор на основе логистической регрессии, предсказывающий по весам последнего слоя BERT-SPC, является ли результат BERT-SPC верным. Выходом ансамблевого классификатора является результат BERT-SPC, если вспомогательный классификатор определил, что этот результат верный, или результат работы алгоритма, основанного на семантических правилах, в противном случае.

Для оценки качества определения тональности по отношению к аспекту с помощью ансамблевого классификатора было создано три варианта ансамбля, по одному для каждого из методов поиска

Table 11. Sentiment detection performance of the ensemble classifier with different methods of aspect terms searching**Таблица 11.** Качество работы ансамблевого классификатора с различными методами поиска аспектных терминов

Тональность	Метод поиска аспектных терминов									Число предл.
	на основе перечня			на основе семантической схожести			гибридный			
	Точн.	Полн.	F-мера	Точн.	Полн.	F-мера	Точн.	Полн.	F-мера	
положительная	0.74	0.78	0.76	0.72	0.78	0.75	0.72	0.78	0.75	143
нейтральная	0.77	0.72	0.75	0.77	0.68	0.72	0.77	0.68	0.72	136
отрицательная	0.85	0.86	0.86	0.84	0.84	0.85	0.84	0.86	0.85	189
Макроусреднение	0.79	0.79	0.79	0.78	0.77	0.77	0.78	0.77	0.77	468
Доля правильных ответов	0.79			0.78			0.78			

Table 12. Sentiment detection confusion matrices of the ensemble classifier with different methods of aspect terms searching**Таблица 12.** Матрицы ошибок ансамблевого классификатора с различными методами поиска аспектных терминов

		Метод поиска аспектных терминов								
		на основе перечня			на основе семантической схожести			гибридный		
		Предсказанная тональность								
		полож.	нейтр.	отриц.	полож.	нейтр.	отриц.	полож.	нейтр.	отриц.
Факт. тон.	полож.	111	19	13	111	18	14	111	18	14
	нейтр.	23	98	15	27	93	16	27	92	17
	отриц.	16	10	163	16	10	163	16	10	163

аспектных терминов. Чтобы точнее оценить применимость для определения тональности ансамблевого классификатора, объединяющего нейросетевую модель и алгоритм, основанный на семантических правилах, для аспекта «проведение выборов» во всех вариантах ансамбля использовался метод поиска аспектных терминов на основе семантической схожести, несмотря на то, что это может затруднить сравнение применимости методов поиска аспектных терминов при создании ансамбля.

Качество работы ансамбля оценивалось с помощью кросс-валидации. BERT-SPC обучался с теми же гиперпараметрами, при использовании которых ранее был получен наилучший результат. Итоговая F-мера вспомогательного классификатора на основе логистической регрессии составила 0.68, что показывает его неплохую объясняющую силу. Метрики качества всех вариантов ансамблевого классификатора и матрицы ошибок приведены в таблицах 11 и 12 соответственно.

При использовании метода поиска аспектных терминов на основе перечня качество возрастает достаточно существенно по сравнению с BERT-SPC — макро-F-мера выше на 0.05. Улучшение достигнуто за счёт того, что меньшее число положительных или отрицательных предложений ошибочно относятся к нейтральным: полнота обнаружения положительной и отрицательной тональности возросла на 0.05 и 0.08 соответственно. Также снизилось число ошибочного определения предложений с положительной тональностью как отрицательных и наоборот: если BERT-SPC допустил 16 и 17 таких ошибок соответственно, то ансамблевый классификатор — 13 и 16 соответственно. При использовании метода поиска аспектных терминов на основе семантической схожести качество возрастает не так сильно: макро-F-мера на 0.02 ниже, чем при использовании метода на основе перечня. Это связано с тем, что слишком большое число нейтральных предложений ошибочно

Table 13. Accuracy of ensemble classifier by aspect**Таблица 13.** Доля правильных ответов ансамблевого классификатора для аспектов

Аспект	BERT-SPC	Алгоритм, основанный на правилах	Ансамблевый классификатор
качество здравоохранения	0.89	0.61	0.89
состояние дорожной сети	0.80	0.60	0.83
качество и стоимость услуг ЖКХ	0.79	0.59	0.81
строительство атомной электростанции	0.77	0.61	0.87
качество управления	0.77	0.59	0.81
состояние жилья	0.74	0.68	0.79
развитие промышленности	0.72	0.59	0.78
развитие сельского хозяйства	0.72	0.72	0.84
проведение выборов	0.65	0.63	0.72
газификация	0.58	0.58	0.62

относятся к положительным или отрицательным: полнота определения нейтральной тональности на 0.04 ниже, чем при использовании метода на основе перечня. Качество определения тональности ансамблевым классификатором при использовании гибридного метода поиска аспектных терминов практически не отличается от качества при использовании метода на основе семантической схожести.

Доля правильных ответов ансамблевого классификатора для каждого из аспектов при использовании метода поиска аспектных терминов на основе перечня приведена в таблице 13 вместе с долей правильных ответов BERT-SPC и алгоритма, основанного на семантических правилах. Для аспекта «качество здравоохранения» качество не изменилось по сравнению с BERT-SPC, для всех остальных аспектов оно возросло. Сильнее всего доля правильных ответов возросла для аспектов «строительство атомной электростанции» (на 0.10) и «развитие сельского хозяйства» (на 0.12). В результате доля правильных ответов превысила 0.7 для всех аспектов, кроме «газификации», для которой она возросла незначительно — на 0.04 до 0.62.

Заключение

В статье рассмотрены разработка и апробация методов определения тональности по отношению к аспектам социально-экономического развития в публицистических предложениях на русском языке. Апробация проводилась на специально собранном и размеченном корпусе политической агитации, состоящем из 468 предложений.

При использовании нейросетевого классификатора BERT-SPC была получена макро-F-мера определения тональности 0.74, что является достаточно высоким результатом. Нужно отметить, что этот результат был получен при дообучении классификатора, предобученного на задаче определения тональности по отношению к явно упоминаемым аспектам на существенно большем наборе данных.

В рамках работы был создан новый алгоритм определения тональности по отношению к аспектам, использующий ранее предложенный авторами алгоритм определения тональности по отношению к явно упоминаемым аспектам, а также разработаны методы поиска аспектных терминов: на основе задаваемого экспертом перечня, на основе семантической схожести и гибридный. Лучшая макро-F-мера, равная 0.63, была получена при использовании метода на основе семантической схожести. Этот результат существенно ниже, чем у нейросетевого классификатора, однако для некоторых аспектов методы дали примерно одинаковую долю правильных ответов.

Лучшее качество определения тональности в ходе работы было получено при объединении нейросетевого классификатора и алгоритма, основанного на семантических правилах, в ансамбль. Макро-F-мера составила 0.79, а для отдельных аспектов доля правильных ответов приблизилась к 0.90.

Подводя итоги, нужно отметить, что решалась достаточно сложная и слабо изученная для русского языка задача, тем не менее по результатам работы удалось добиться достаточно высокого качества, позволяющего решать многие практические задачи.

Дальнейшие направления исследований могут включать в себя как эксперименты с дообучением моделей, обученных для схожих задач анализа тональности, для которых доступны открытые корпуса, так и создание новых, более точных алгоритмов, основанных на семантических правилах, которые будут лучше определять тональность по отношению к отдельным аспектам. Отдельным направлением исследований может стать разработка новых методов поиска аспектных терминов. В этой области наиболее перспективным является использование семантической схожести, вычисляемой на основе эмбедингов.

References

- [1] B. Liu, *Sentiment Analysis and Opinion Mining*. Springer, 2022, 167 pp.
- [2] W. Zhang, X. Li, Y. Deng, L. Bing, and W. Lam, “A survey on aspect-based sentiment analysis: Tasks, methods, and challenges”, *IEEE Transactions on Knowledge and Data Engineering*, vol. 35, no. 11, pp. 11 019–11 038, 2022. DOI: [10.1109/TKDE.2022.3230975](https://doi.org/10.1109/TKDE.2022.3230975).
- [3] O. Alqaryouti *et al.*, “Aspect-based sentiment analysis using smart government review data”, *Applied Computing and Informatics*, vol. 20, no. 1/2, pp. 142–161, 2024. DOI: [10.1016/j.aci.2019.11.003](https://doi.org/10.1016/j.aci.2019.11.003).
- [4] A. Nazir, Y. Rao, L. Wu, and L. Sun, “Issues and challenges of aspect-based sentiment analysis: A comprehensive survey”, *IEEE Transactions on Affective Computing*, vol. 13, no. 2, pp. 845–863, 2020. DOI: [10.1109/TAFFC.2020.2970399](https://doi.org/10.1109/TAFFC.2020.2970399).
- [5] A. Poletaev, I. Paramonov, and E. Kolupaev, “Methods of implicit aspect detection in Russian publicism sentences”, *Modeling and Analysis of Information Systems*, vol. 31, no. 3, pp. 226–239, 2024, in Russian. DOI: [10.18255/1818-1015-2024-3-226-239](https://doi.org/10.18255/1818-1015-2024-3-226-239).
- [6] A. D. Kazun, “Construction of social problems in the media and agenda-setting theory: The limits of concepts’ compatibility”, *Monitoring of Public Opinion: Economic and Social Changes*, no. 3 (133), pp. 159–172, 2016, in Russian. DOI: [10.14515/monitoring.2016.3.09](https://doi.org/10.14515/monitoring.2016.3.09).
- [7] A. Guseva, I. Kuznetsov, P. Bochkarev, and D. Smirnov, “Digital shadow of Russian international megaprojects of NPP construction abroad: Assessment of the tone of utterances”, *Modern High Technologies*, no. 12 (1), pp. 26–34, 2022, in Russian. DOI: [10.17513/snt.39432](https://doi.org/10.17513/snt.39432).
- [8] W. Zhang, X. Li, Y. Deng, L. Bing, and W. Lam, “A survey on aspect-based sentiment analysis: Tasks, methods, and challenges”, *IEEE Transactions on Knowledge and Data Engineering*, vol. 35, no. 11, pp. 11 019–11 038, 2023. DOI: [10.1109/TKDE.2022.3230975](https://doi.org/10.1109/TKDE.2022.3230975).
- [9] E. G. Brunova, Y. V. Bidulya, and A. A. Gorbunov, “Aspect-based sentiment analysis of political discourse”, *Tyumen State University Herald. Humanities Research. Humanitates*, vol. 7, no. 3 (27), pp. 6–22, 2021, in Russian. DOI: [10.21684/2411-197X-2021-7-3-6-22](https://doi.org/10.21684/2411-197X-2021-7-3-6-22).
- [10] Muljono, B. Harjo, and R. Abdullah, “Aspect-based sentiment analysis for financial review with implicit aspect and opinion using semantic similarity and hybrid approach”, *International Journal of Intelligent Engineering & Systems*, vol. 17, no. 5, pp. 646–658, 2024. DOI: [10.22266/ijies2024.1031.49](https://doi.org/10.22266/ijies2024.1031.49).

- [11] K. Ananthajothi, K. Karthikayani, and R. Prabha, “Explicit and implicit oriented aspect-based sentiment analysis with optimal feature selection and deep learning for demonetization in India”, *Data & Knowledge Engineering*, vol. 142, p. 102 092, 2022. DOI: [10.1016/j.datak.2022.102092](https://doi.org/10.1016/j.datak.2022.102092).
- [12] C. Hutto and E. Gilbert, “VADER: A parsimonious rule-based model for sentiment analysis of social media text”, in *Proceedings of the International AAI Conference on Web and Social Media*, vol. 8, 2014, pp. 216–225.
- [13] N. Chechneva, “Simple and efficient approach to the aspect extraction from customers’ product reviews”, in *Proceedings of the 26th Conference of Open Innovations Association FRUCT*, 2020, pp. 67–73. DOI: [10.23919/fruct48808.2020.9087546](https://doi.org/10.23919/fruct48808.2020.9087546).
- [14] A. Poletaev, I. Paramonov, and E. Boychuk, “Automatic detection of sentiment towards explicit aspect in Russian publicism sentences using syntactic structure”, in *Proceedings of the 36th Conference of Open Innovations Association FRUCT*, IEEE, 2024, pp. 593–602.
- [15] M. A. Pil’gun, “Rechevye osobennosti politicheskoy kommunikacii”, *Proceedings of Kazan University. Humanities Sciences Series*, vol. 152, no. 2, pp. 236–246, 2010, in Russian.
- [16] Y. Song *et al.*, “Targeted sentiment classification with attentional encoder network”, in *Artificial Neural Networks and Machine Learning – ICANN 2019: Text and Time Series*, Springer International Publishing, 2019, pp. 93–103. DOI: [10.1007/978-3-030-30490-4_9](https://doi.org/10.1007/978-3-030-30490-4_9).
- [17] J. Ansel *et al.*, “PyTorch 2: Faster machine learning through dynamic Python bytecode transformation and graph compilation”, in *ASPLOC’24: Proceedings of the 29th ACM International Conference on Architectural Support for Programming Languages and Operating Systems, Volume 2*, ACM, 2024, pp. 929–947. DOI: [10.1145/3620665.3640366](https://doi.org/10.1145/3620665.3640366).
- [18] A. Naumov *et al.*, “Neural-network method for determining text author’s sentiment to an aspect specified by the named entity”, in *CEUR Workshop Proceedings*, vol. 2648, 2020, pp. 134–143.
- [19] Y. Wang, L. Wu, J. Li, X. Liang, and M. Zhang, “Are the BERT family zero-shot learners? A study on their potential and limitations”, *Artificial Intelligence*, vol. 322, p. 103 953, 2023. DOI: [10.1016/j.artint.2023.103953](https://doi.org/10.1016/j.artint.2023.103953).
- [20] A. Golubev, N. Rusnachenko, and N. Loukachevitch, “RuSentNE-2023: Evaluating entity-oriented sentiment analysis on Russian news texts”, in *Computational Linguistics and Intellectual Technologies: Papers from the Annual conference “Dialogue” (2023)*, vol. 22, 2023, pp. 130–141. DOI: [10.28995/2075-7182-2023-22-130-141](https://doi.org/10.28995/2075-7182-2023-22-130-141).
- [21] D. Ma *et al.*, “Interactive attention networks for aspect-level sentiment classification”, in *Proceedings of the 26th International Joint Conference on Artificial Intelligence*, AAAI Press, 2017, pp. 4068–4074. DOI: [10.5555/3171837.3171854](https://doi.org/10.5555/3171837.3171854).
- [22] A. Y. Poletaev, I. V. Paramonov, and E. I. Boychuk, “Semantic rule-based sentiment detection algorithm for Russian publicism sentences”, *Modeling and Analysis of Information Systems*, vol. 30, no. 4, pp. 394–417, 2023, in Russian. DOI: [10.18255/1818-1015-2023-4-394-417](https://doi.org/10.18255/1818-1015-2023-4-394-417).
- [23] A. Y. Poletaev, I. V. Paramonov, and E. I. Boychuk, “Algorithm of constituency tree from dependency tree construction for a Russian-language sentence”, *Informatics and Automation*, vol. 22, no. 6, pp. 1323–1353, 2023, in Russian. DOI: [10.15622/ia.22.6.3](https://doi.org/10.15622/ia.22.6.3).
- [24] D. Chandrasekaran and V. Mago, “Evolution of semantic similarity – a survey”, *ACM Computing Surveys (CSUR)*, vol. 54, no. 2, pp. 1–37, 2021. DOI: [10.1145/3440755](https://doi.org/10.1145/3440755).

- [25] A. Kutuzov and E. Kuzmenko, “WebVectors: A toolkit for building web interfaces for vector semantic models”, in *Analysis of Images, Social Networks and Texts: 5th International Conference, AIST 2016, Yekaterinburg, Russia, April 7–9, 2016, Revised Selected Papers*. Cham: Springer International Publishing, 2017, pp. 155–161, ISBN: 978-3-319-52920-2. DOI: [10.1007/978-3-319-52920-2_15](https://doi.org/10.1007/978-3-319-52920-2_15).
- [26] M. Korobov, “Morphological analyzer and generator for Russian and Ukrainian languages”, English, in *Analysis of Images, Social Networks and Texts*, ser. Communications in Computer and Information Science, vol. 542, Springer International Publishing, 2015, pp. 320–332, ISBN: 978-3-319-26122-5. DOI: [10.1007/978-3-319-26123-2_31](https://doi.org/10.1007/978-3-319-26123-2_31).
- [27] P. Qi *et al.*, “Stanza: A Python natural language processing toolkit for many human languages”, in *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics: System Demonstrations*, Association for Computational Linguistics, 2020, pp. 101–108. DOI: [10.18653/v1/2020.acl-demos.14](https://doi.org/10.18653/v1/2020.acl-demos.14).