

Автоматическое определение основного аргументационного утверждения с использованием традиционных моделей машинного обучения

И. Н. Фищева

старший преподаватель кафедры прикладной математики и информатики,
Вятский государственный университет. Россия, г. Киров.
ORCID: 0000-0002-6941-2009. E-mail: fishchevain@gmail.com

Аннотация. В последнее время появляется все больше онлайн-площадок с обсуждениями, спорами и дебатами. В связи с этим актуальность автоматической обработки аргументативных текстов постоянно растет. Определение основного аргументационного утверждения позволяет выявить позицию автора текста. В данной работе исследуется задача бинарной классификации русскоязычных аргументационных предложений на «основное утверждение» и «другие аргументативные дискурсивные единицы» традиционными моделями машинного обучения: случайный лес, наивный байесовский метод, метод опорных векторов, ансамблевый метод AdaBoost, бэггинг (Bagging), реализованные в библиотеке scikit-learn, а также градиентный бустинг из библиотеки XGBoost. Рассматриваются различные виды признаков: дискурсивных маркеров, модальных слов, пунктуационных, морфосинтаксических признаков, позиции предложения в тексте. Лучшие результаты были получены с использованием модели градиентного бустинга.

Ключевые слова: извлечение аргументов, градиентный бустинг, бэггинг, отбор признаков.

Извлечение аргументов направлено на автоматическое обнаружение, классификацию и структурирование аргументации в тексте [10]. Аргументация – это метод убеждения, предназначенный для рационального обоснования какого-либо утверждения при помощи других утверждений, но не способный при этом служить доказательством его истинности [1].

Количество различных онлайн-площадок с обсуждениями, спорами и дебатами постоянно увеличивается, поэтому растет и важность автоматической обработки таких данных. Понимание структуры аргументов позволяет определить не только то, какую позицию занимают стороны, но и почему они придерживаются такого мнения. Это дает ценную информацию в очень разных областях от прогнозирования финансовых рынков до связи с общественностью [7].

Основным текстовым элементом, используемым при анализе аргументации, является аргументативная дискурсивная единица (АДЕ) – фрагмент текста, который имеет единственное аргументационное значение [16]. В качестве АДЕ чаще всего выступают отдельные предложения, но возможны ситуации, когда АДЕ является частью предложения или нескольких предложений.

Анализ аргументов включает три основных этапа [17]:

- 1) сегментирование текста на АДЕ и другие части;
- 2) классификация типа или роли каждой АДЕ;
- 3) выявление и классификация отношений между АДЕ.

Для реализации перечисленных этапов необходимы аннотированные корпуса текста на русском языке. В настоящее время существует достаточно большое количество корпусов с разнообразной аргументативной разметкой [7], однако русскоязычных корпусов немного. В данном исследовании были использованы корпуса в объединенной схеме аргументационной разметки из [5]. Эти корпуса были получены переводом с помощью Google Translate корпусов Argumentative Microtext Corpus (ArgMicro) [11; 13] и Persuasive Essay Corpus (PersEssays) [14]. В работе [6] было показано, что машинный перевод на русский язык англоязычного корпуса ArgMicro позволяет получить качество классификации АДЕ, не уступающее человеческому переводу, поэтому Google Translate был использован в данной работе.

Каждый текст корпуса посвящен некоторой обсуждаемой тематике. В корпусе ArgMicro АДЕ являются частью предложения или целым предложением. В корпусе PersEssays в качестве АДЕ выступают целые предложения.

АДЕ можно классифицировать следующим образом:

- *основное утверждение* – выражает точку зрения автора текста;
- *довод «за»* – поддерживает основное утверждение;
- *довод «против»* – опровергает основное утверждение;
- *нейтральная АДЕ* – не является частью аргументационной структуры.

Рассмотрим три основных этапа извлечения аргументов на примере фрагмента текста из корпуса PersEssays (перевод фрагмента сделан человеком).

Тема. Дети должны расти в большом городе.

Текст. Несомненно, лучше, чтобы дети росли в большом городе. Конечно, необходимо выбрать хороший район. Я верю в это по двум основным причинам: академическим и социальным. Некоторые люди думают, что если ребенок растет в большом городе, то он будет весь день проводить дома за компьютером или играть в видеоигры. Но это не так, если вы живете в районе с другими детьми вашего возраста, как я. Мои друзья и я играем в футбол, катаемся на велосипедах, лазим по деревьям и делаем много других вещей каждый день.

На первом этапе текст разделяется на предложения, так как в данном случае в качестве АДЕ выступает предложение.

На втором этапе каждое предложение получает одну из следующих меток: основное утверждение, довод «за», довод «против», нейтральный узел (Рис. 1). На третьем этапе определяются связи между АДЕ: поддержка или опровержение.

В работе [5] исследовалась классификация доводов «за» и «против». Данная работа является продолжением и исследует задачу классификации основное утверждение / другие АДЕ. Состав рассматриваемых корпусов приведен в таблице 1.



Рис. 1. Пример аргументационной структуры

Таблица 1

Характеристики текстовых корпусов

Корпуса	Тексты	АДЕ		
		Основное утверждение	Остальные АДЕ	Всего
ArgMicro	283	301 (19,5 %)	1,240 (80,5 %)	1,541 (100 %)
PersEssays	399	746 (10,2 %)	6,531 (89,8 %)	7,277 (100 %)
ArgMicro +PersEssays	682	1,047 (11,9 %)	7,771 (88,9 %)	8,818 (100 %)

Задача классификации основного утверждения решалась в различных исследованиях [4; 12; 15]. Однако все эти исследования проводились на англоязычных корпусах.

Эксперименты проводились с целью получения ответов на следующие вопросы:

V1: Какое качество бинарной классификации русскоязычных аргументационных предложений на «основное утверждение» и «другие АДЕ» может быть достигнуто на основе традиционных моделей машинного обучения?

V2: Возможно ли повысить качество классификации за счет расширения обучающего корпуса и как это согласуется с результатами [5]?

V3: Какова значимость различных видов признаков для традиционных классификаторов и как это согласуется с результатами [5]?

Под «традиционными» понимаются такие модели машинного обучения, которые не используют глубокие нейронные сети и в которых формирование эффективного с точки зрения качества классификации набора признаков возлагается на исследователя. Примерами таких традиционных моделей являются метод опорных векторов, наивные методы Байеса и другие.

Для каждой АДЕ была проведена предварительная обработка: токенизация и удаление стоп-слов с использованием пакета библиотек nltk [9], а также лемматизация с использованием программы mystem [8].

Фищева и Котельников [6] показали, что при использовании традиционных методов машинного обучения признаки TF.IDF и word2vec не улучшают качество классификаторов на небольших корпусах. Поэтому в настоящей работе рассматривались только следующие виды признаков:

– лексические признаки – дискурсивные маркеры («поэтому», «я думаю», «в итоге» и другие) и модальные слова («нужно», «может быть», «обязательно» и другие), включая отрицания;

– пунктуационные признаки – запятая, двоеточие, точка с запятой, вопросительный и восклицательный знаки;

– морфосинтаксические признаки – N-граммы на основе частей речи (существительные, местоимения, глаголы, прилагательные и наречия), $N = \{2, 3, 4\}$, грамматические признаки глаголов: время, наклонение, лицо;

– позиция предложения в тексте.

Все указанные виды признаков рассматривались как для текущей АДЕ, так и для предыдущей и следующей АДЕ (если они были). Вектор признаков для АДЕ формировался как конкатенация различных видов признаков (Таблица 2).

Таблица 2

Наборы признаков

Набор признаков (номер набора)	Позиция текущего предложения в тексте	Текущее предложение				Предыдущее предложение				Следующее предложение			
		Дискурсивные маркеры	Модальные слова	Пунктуационные признаки	Морфосинтаксические признаки	Дискурсивные маркеры	Модальные слова	Пунктуационные признаки	Морфосинтаксические признаки	Дискурсивные маркеры	Модальные слова	Пунктуационные признаки	Морфосинтаксические признаки
Все без предыдущего (1)	V	V	V	V	V	—	—	—	—	V	V	V	V
Все без позиции (2)	—	V	V	V	V	V	V	V	V	V	V	V	V

Набор признаков (номер набора)	Позиция текущего предложения в тексте	Текущее предложение				Предыдущее предложение				Следующее предложение			
		Дискурсивные маркеры	Модальные слова	Пунктуационные признаки	Морфосинтаксические признаки	Дискурсивные маркеры	Модальные слова	Пунктуационные признаки	Морфосинтаксические признаки	Дискурсивные маркеры	Модальные слова	Пунктуационные признаки	Морфосинтаксические признаки
Дискурсивные маркеры (3)	—	V	—	—	—	V	—	—	—	V	—	—	—
Лексические признаки (4)	—	V	V	—	—	V	V	—	—	V	V	—	—
Лексические признаки текущего (5)	—	V	V	—	—	—	—	—	—	—	—	—	—

Для обучения классификатора использовались традиционные методы машинного обучения: случайный лес, наивный байесовский метод, метод опорных векторов, ансамблевый метод AdaBoost, бэггинг (Bagging), реализованные в библиотеке *scikit-learn*, а также градиентный бустинг из библиотеки *XGBoost* [18]. Далее приведены результаты для XGBoost и Bagging, показавшие наилучшие результаты.

Исследовались четыре варианта формирования обучающих и тестовых данных:

- обучение на ArgMicro, тестирование на ArgMicro;
- обучение на PersEssays, тестирование на PersEssays;
- обучение на ArgMicro и PersEssays, тестирование на ArgMicro;
- обучение на ArgMicro и PersEssays, тестирование на PersEssays.

В связи с небольшим размером корпусов для каждого из четырех вариантов была реализована пятикратная перекрестная проверка с целью получения несмещенных оценок качества. Разбиения были случайными со стратификацией по классам и сохранялись одними и теми же для всех экспериментов.

В связи с сильной несбалансированностью обоих корпусов в качестве основной метрики качества использовалась макроусредненная F1-мера [3]. Также вычислялись доля правильно классифицированных объектов (accuracy) и макроусредненная точность (precision) и полнота (recall). Полученные результаты для каждого эксперимента усреднялись, а также вычислялось среднее квадратическое отклонение по пяти разбиениям в процедуре перекрестной проверки. В таблице 3 приведены результаты экспериментов.

Таблица 3

Результаты классификации (среднее значение ± среднее квадратическое отклонение)

Обучающие данные	Тестовые данные	Набор признаков	Модель	F1-мера	Точность	Полнота	Доля правильно классифицированных объектов
ArgMicro	ArgMicro	(1)	Bagging	0,7485±0,0316	0,7953±0,0408	0,7230±0,0300	0,8598±0,0195
			XGBoost	0,7453±0,0235	0,7796±0,0254	0,7239±0,0231	0,8553±0,0129
		(2)	Bagging	0,7415±0,0205	0,8408±0,0234	0,7031±0,0209	0,8683±0,0086
			XGBoost	0,7517±0,0381	0,7989±0,0362	0,7249±0,0362	0,8631±0,0176
		(3)	Bagging	0,6211±0,0118	0,6349±0,0120	0,6147±0,0163	0,7787±0,0111
			XGBoost	0,6809±0,0245	0,7075±0,0209	0,6653±0,0248	0,8196±0,0105
		(4)	Bagging	0,6824±0,0232	0,6978±0,0323	0,6727±0,0197	0,8112±0,0204
			XGBoost	0,7058±0,0143	0,7381±0,0244	0,6875±0,0124	0,8332±0,0120
		(5)	Bagging	0,6431±0,0229	0,6378±0,0213	0,6520±0,0273	0,7638±0,0161
			XGBoost	0,6862±0,0180	0,7282±0,0292	0,6664±0,0179	0,8274±0,0125

Обучающие данные	Тестовые данные	Набор признаков	Модель	F1-мера	Точность	Полнота	Доля правильно классифицированных объектов
ArgMicro + PersEssays	(1)	Bagging	0,6977±0,0171	0,8082±0,0170	0,6636±0,0153	0,8514±0,0056	
		XGBoost	0,7130±0,0266	0,7954±0,0222	0,6809±0,0250	0,8527±0,0110	
	(2)	Bagging	0,6288±0,0277	0,8003±0,0490	0,6067±0,0213	0,8345±0,0098	
		XGBoost	0,7202±0,0292	0,8216±0,0441	0,6835±0,0231	0,8592±0,0149	
	(3)	Bagging	0,5688±0,0336	0,6851±0,0500	0,5641±0,0235	0,8105±0,0116	
		XGBoost	0,6621±0,0214	0,7759±0,0370	0,6338±0,0168	0,8378±0,0109	
	(4)	Bagging	0,5834±0,0297	0,7007±0,0188	0,5748±0,0218	0,8138±0,0059	
		XGBoost	0,6867±0,0253	0,8015±0,0410	0,6536±0,0196	0,8475±0,0130	
	(5)	Bagging	0,6564±0,0244	0,701±0,0326	0,6378±0,0210	0,8157±0,0142	
		XGBoost	0,6467±0,0520	0,7867±0,0575	0,6213±0,0401	0,8378±0,0199	
PersEssays	(1)	Bagging	0,7435±0,0168	0,7953±0,0207	0,7123±0,0216	0,9185±0,0050	
		XGBoost	0,7514±0,0094	0,8109±0,0277	0,7171±0,0150	0,9217±0,0050	
	(2)	Bagging	0,7351±0,0274	0,8114±0,0314	0,6955±0,0267	0,9203±0,0070	
		XGBoost	0,7049±0,0153	0,7998±0,0150	0,6637±0,0176	0,9153±0,0025	
	(3)	Bagging	0,7230±0,0093	0,7748±0,0195	0,6920±0,0093	0,9129±0,0047	
		XGBoost	0,7248±0,0097	0,7829±0,0173	0,6912±0,0100	0,9147±0,0036	
	(4)	Bagging	0,7167±0,0195	0,7738±0,0320	0,6844±0,0166	0,9121±0,0076	
		XGBoost	0,7311±0,0144	0,7942±0,0143	0,6956±0,0165	0,9173±0,0032	
	(5)	Bagging	0,6595±0,0177	0,6843±0,0174	0,6429±0,0177	0,8887±0,0048	
		XGBoost	0,6872±0,0146	0,7524±0,0105	0,6552±0,0179	0,9066±0,0021	
ArgMicro + PersEssays	(1)	Bagging	0,7308±0,0210	0,7714±0,0142	0,7041±0,0238	0,9133±0,0042	
		XGBoost	0,7447±0,0200	0,7939±0,0124	0,7142±0,0264	0,9188±0,0037	
	(2)	Bagging	0,7095±0,0160	0,8027±0,0154	0,6683±0,0190	0,9162±0,0024	
		XGBoost	0,6996±0,0203	0,7822±0,0240	0,6620±0,0210	0,9123±0,0047	
	(3)	Bagging	0,7136±0,0171	0,7363±0,0228	0,6967±0,0148	0,9031±0,0071	
		XGBoost	0,7144±0,0149	0,7767±0,0214	0,6801±0,0136	0,9127±0,0049	
	(4)	Bagging	0,7066±0,0196	0,7346±0,0261	0,6869±0,0167	0,9026±0,0082	
		XGBoost	0,7042±0,0192	0,7617±0,0186	0,6728±0,0202	0,9093±0,0046	
	(5)	Bagging	0,6565±0,0171	0,6608±0,0120	0,6535±0,0229	0,8769±0,0032	
		XGBoost	0,6807±0,0141	0,7343±0,0148	0,6524±0,0158	0,9026±0,0035	

Лучший результат для корпуса ArgMicro получен с использованием XGBoost (F1 мера = 0,7517) при обучении только на ArgMicro. Bagging незначительно отстает (F1 мера = 0,7485).

Для корпуса PersEssays лучший результат также показывает XGBoost (F1 мера = 0,7514). Он превосходит Bagging за счет более высокой точности (0,8109 против 0,7953).

Расширение обучающей выборки за счет добавления к ArgMicro корпуса PersEssays в случае тестирования на ArgMicro снижает результаты для XGBoost (на 0,0315) и для Bagging (на 0,0508) за счет увеличения дисбаланса классов (в ArgMicro основных утверждений 19,5 % от общего числа АДЕ, в ArgMicro + PersEssays только 11,9 %).

Добавление ArgMicro к PersEssays так же приводит к ухудшению результатов: для XGBoost F1 мера уменьшается на 0,0067, для Bagging – на 0,0127.

Bagging и XGBoost демонстрируют одинаковую стабильность результатов обучения: дисперсия результатов по разбиениям одинаковая (в среднем по всем экспериментам 0,0212).

Результаты, полученные в ходе данного исследования, согласуются с работой [5] (Таблица 4).

Таблица 4

**Сравнение метрик качества для задачи определения
основного утверждения и задачи классификации доводов «за» / «против»
(среднее значение F1-меры ± среднее квадратическое отклонение)**

Обучающие данные	Тестовые данные	Модель	F1-мера Классификация доводов «за» / «против»	F1-мера Определение основного утверждения
ArgMicro	ArgMicro	XGBoost	0,7921 ± 0,0309	0,7517 ± 0,0381
ArgMicro + PersEssays			0,7678 ± 0,0203	0,7202 ± 0,0292
PersEssays	0,6308 ± 0,0191		0,7514 ± 0,0094	
ArgMicro + PersEssays	0,6510 ± 0,0165		0,7447 ± 0,0200	

В обеих задачах классификации достигнутое качество сопоставимо. При этом в задаче определения основного утверждения F1-мера не сильно зависит от обучающих и тестовых данных. В то же время в задаче классификации доводов наблюдалось существенное колебание F1-меры (0,7678 против 0,6510) при изменении тестовых данных.

Для ответа на вопрос В3 о значимости различных видов признаков проанализируем результаты модели XGBoost в таблице 3, обращая внимание на наборы признаков. Использование лексических признаков недостаточно для получения высокого качества при определении основного утверждения. При тестировании на PersEssays важную роль начинает играть позиция предложения в тексте. Это объясняется особенностью корпуса PersEssays, в котором часто присутствуют два основных утверждения в начале и в конце эссе. В ArgMicro такой особенности нет. Однако при тестировании на ArgMicro F1-мера на наборе признаков (1) не сильно уступает набору (2) (0,7453 против 0,7517 и 0,7130 против 0,7202). Следовательно, признаки предыдущего предложения не слишком полезны, что согласуется с выводами о значимости признаков в [5].

Таким образом, были получены следующие ответы на поставленные вопросы.

В1: Какое качество бинарной классификации русскоязычных аргументационных предложений на «основное утверждение» и «другие АДЕ» может быть достигнуто на основе традиционных моделей машинного обучения? – Лучшие значения качества по макроусредненной F1-мере позволяет достичь модель XGBoost для ArgMicro F1-мера = 0,7517, для PersEssays F1-мера = 0,7514.

В2: Возможно ли повысить качество классификации за счет расширения обучающего корпуса и как это согласуется с результатами [5]? – Нет. В обоих случаях качество снижалось, так как увеличивается дисбаланс обучающих классов. Однако по сравнению с предыдущим исследованием модель XGBoost на тестовом корпусе PersEssays получила более высокие оценки качества (F1-мера 0,7514 и 0,7447 против 0,6308 и 0,6510), тогда как на корпусе ArgMicro качество оказалось несколько ниже (F1-мера 0,7517 и 0,7202 против 0,7921 и 0,7678).

В3: Какова значимость различных видов признаков для традиционных классификаторов и как это согласуется с результатами [5]? – Использование лексических признаков недостаточно для получения хорошего качества при определении основного утверждения. Признаки предыдущих предложений сильное влияние на качество классификатора не оказывают. Данные выводы не противоречат предыдущему исследованию.

Актуальными задачами, которые должны быть решены в дальнейших исследованиях, являются, во-первых, расширение спектра русскоязычных корпусов с аргументативной разметкой; во-вторых, исследование качества анализа аргументации с использованием векторного представления АДЕ на основе глубоких нейросетевых моделей.

Список литературы

1. *Ивлев Ю. В., Новоселов М. М., Бергман А. С.* Аргументация // Гуманитарный портал: Концепты. Центр гуманитарных технологий, 2002–2021. URL: <https://gtmarket.ru/concepts/7226> (дата обращения: 06.11.2021).
2. *Переводчик* // Google. URL: <https://www.nltk.org> (дата обращения: 06.11.2021).
3. *Метрики и оценки: количественная оценка качества прогнозов* // Машинное обучение в Python. URL: <https://scikit-learn.ru/3-3-metrics-and-scoring-quantifying-the-quality-of-predictions/#> (дата обращения: 06.11.2021).
4. *Eckle-Kohler J., Kluge R., Gurevych I.* On the Role of Discourse Markers for Discriminating Claims and Premises in Argumentative Discourse // In Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing, 2015. Pp. 2236–2242.
5. *Fishcheva I. N., Goloviznina V. S., Kotelnikov E. V.* Traditional Machine Learning and Deep Learning Models for Argumentation Mining in Russian Texts // Computational Linguistics and Intellectual Technologies: Papers from the Annual International Conference "Dialogue-2021". Pp. 246–258.
6. *Fishcheva I., Kotelnikov E.* Cross-Lingual Argumentation Mining for Russian Texts // Proceedings of the International Conference on Analysis of Images, Social Networks and Texts (AIST-2019), Springer, 2019. Pp. 134–144.
7. *Lawrence J. and Reed C.* Argument mining: A survey // Computational Linguistics. 2020. Vol. 45 (4). Pp. 765–818.
8. *MyStem* // Яндекс. URL: <https://yandex.ru/dev/mystem/> (дата обращения: 06.11.2021).
9. *NLTK Documentation* // NLTK Project. URL: <https://www.nltk.org> (дата обращения: 06.11.2021).
10. *Palau R. M., Moens M.-F.* Argumentation mining: the detection, classification and structure of arguments in text // Proceedings of the 12th international conference on artificial intelligence and law. ACM, 2009. Pp. 98–107.
11. *Peldszus A., Stede M.* An annotated corpus of argumentative microtexts // Argumentation and Reasoned Action: Proceedings of the 1st European Conference on Argumentation, Lisbon 2015. Vol. 2. London. College Publications, 2015. Pp. 801–816.
12. *Ruggeri F., Lippi M., Torroni P.* Tree-Constrained Graph Neural Networks For Argument Mining // arXiv.org. 2021. URL: <https://arxiv.org/abs/2110.00124> (дата обращения: 06.11.2021).
13. *Skeppstedt M., Peldszus A., Stede M.* More or less controlled elicitation of argumentative text: enlarging a microtext corpus via crowdsourcing // Proceedings of the 5th Workshop in Argumentation Mining, 2018. Pp. 155–163.
14. *Stab C., Gurevych I.* Annotating argument components and relations in persuasive essays // Proceedings of the International Conference on Computational Linguistics, 2014. Pp. 1501–1510.

15. Stance classification of context-dependent claims / Bar-Haim R., Bhattacharya I., Dinuzzo F., Saha A., Slonim N. // In Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics. 2017. Vol. 1, Long Papers. Pp. 251–261.
16. *Stede M., Schneider J.* Argumentation Mining, Synthesis Lectures on Human Language Technologies, San Rafael. Morgan and Claypool Publishers, 2018. 191 p.
17. Tutorial on Argumentation Technology for Artificial Intelligence / Cimiano P., Al-Khatib K., Stein B., Wachsmuth H. // 43rd German Conference on Artificial Intelligence, 2020.
18. XGBoost Documentation // xgboost developers. URL: <https://xgboost.readthedocs.io/> (дата обращения: 06.11.2021).

Automatic determination of the main argumentative statement using traditional machine learning models

I. N. Fishcheva

senior lecturer of the Department of applied mathematics and computer science, Vyatka State University.
Russia, Kirov. ORCID: 0000-0002-6941-2009. E-mail: fishchevain@gmail.com

Abstract. Recently, there are more and more online platforms with discussions, disputes and debates. In this regard, the relevance of automatic processing of argumentative texts is constantly growing. The definition of the main argumentative statement makes it possible to identify the position of the author of the text. This paper investigates the problem of binary classification of Russian-language argumentative sentences into "main statement" and "other argumentative discursive units" by traditional machine learning models: random forest, naive Bayesian method, support vector machine method, AdaBoost ensemble method, bagging implemented in the scikit-learn library, as well as gradient boosting from the XGBoost library. The significance of various types of signs is considered: discursive markers, modal words, punctuation, morphosyntactic signs, the position of the sentence in the text. The best results were obtained using the gradient boosting model.

Keywords: argument extraction, gradient boosting, bagging, feature selection.

References

1. *Ivlev Yu. V., Novoselov M. M., Bergman A. S.* *Argumentaciya* [Argumentation] // *Gumanitarnyj portal: Koncepty – Humanitarian portal: Concepts.* Center for Humanitarian Technologies, 2002–2021. Available at: <https://gtmarket.ru/concepts/7226> (date accessed: 06.11.2021).
2. *Perevodchik* – Translator // Google. Available at: <https://www.nltk.org> (date accessed: 06.11.2021).
3. *Metriki i ocenki: kolichestvennaya ocenka kachestva prognozov* – Metrics and estimates: quantitative assessment of the quality of forecasts // *Mashinnoe obuchenie v Python – Machine learning in Python.* Available at: <https://scikit-learn.ru/3-3-metrics-and-scoring-quantifying-the-quality-of-predictions/#> (date accessed: 06.11.2021).
4. *Eckle-Kohler J., Kluge R., Gurevych I.* On the Role of Discourse Markers for Discriminating Claims and Premises in Argumentative Discourse // In Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing. 2015. Pp. 2236–2242.
5. *Fishcheva I. N., Golviznina V. S., Kotelnikov E. V.* Traditional Machine Learning and Deep Learning Models for Argumentation Mining in Russian Texts // *Computational Linguistics and Intellectual Technologies: Papers from the Annual International Conference "Dialogue-2021".* Pp. 246–258.
6. *Fishcheva I., Kotelnikov E.* Cross-Lingual Argumentation Mining for Russian Texts // *Proceedings of the International Conference on Analysis of Images, Social Networks and Texts (AIST-2019), Springer.* 2019. Pp. 134–144.
7. *Lawrence J. and Reed C.* Argument mining: A survey // *Computational Linguistics.* 2020. Vol. 45(4). Pp. 765–818.
8. MyStem // Yandex. Available at: <https://yandex.ru/dev/mystem/> (date accessed: 06.11.2021).
9. NLTK Documentation // NLTK Project. Available at: <https://www.nltk.org> (date accessed: 06.11.2021).
10. *Palau R. M., Moens M.-F.* Argumentation mining: the detection, classification and structure of arguments in text // *Proceedings of the 12th international conference on artificial intelligence and law.* ACM. 2009. Pp. 98–107.
11. *Peldszus A., Stede M.* An annotated corpus of argumentative microtexts // *Argumentation and Reasoned Action: Proceedings of the 1st European Conference on Argumentation, Lisbon 2015.* Vol. 2. London. College Publications, 2015. Pp. 801–816.
12. *Ruggeri F., Lippi M., Torroni P.* Tree-Constrained Graph Neural Networks For Argument Mining // *arXiv.org.* 2021. Available at: <https://arxiv.org/abs/2110.00124> (date accessed: 06.11.2021).
13. *Skeppstedt M., Peldszus A., Stede M.* More or less controlled elicitation of argumentative text: enlarging a microtext corpus via crowdsourcing // *Proceedings of the 5th Workshop in Argumentation Mining.* 2018. Pp. 155–163.
14. *Stab C., Gurevych I.* Annotating argument components and relations in persuasive essays // *Proceedings of the International Conference on Computational Linguistics.* 2014. Pp. 1501–1510.
15. Stance classification of context-dependent claims / Bar-Haim R., Bhattacharya I., Dinuzzo F., Saha A., Slonim N. // In Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics. 2017. Vol. 1, Long Papers. Pp. 251–261.

16. *Stede M., Schneider J.* Argumentation Mining, Synthesis Lectures on Human Language Technologies, San Rafael. Morgan and Claypool Publishers, 2018. 191 p.

17. Tutorial on Argumentation Technology for Artificial Intelligence / Cimiano P., Al-Khatib K., Stein B., Wachsmuth H. // 43rd German Conference on Artificial Intelligence, 2020.

18. XGBoost Documentation // xgboost developers. Available at: <https://xgboost.readthedocs.io/> (date accessed: 06.11.2021).