

**ИНФОРМАТИКА И ИНФОРМАЦИОННЫЕ ПРОЦЕССЫ/INFORMATICS AND INFORMATION PROCESSES**DOI: <https://doi.org/10.60797/IRJ.2026.168.15> EDN: МЕНДКУ**СРАВНИТЕЛЬНЫЙ АНАЛИЗ МЕТОДОВ КЛАССИФИКАЦИИ ТЕЛЕФОННЫХ ОБРАЩЕНИЙ:
КЛАССИЧЕСКИЕ АЛГОРИТМЫ И ТРАНСФОРМЕРНЫЕ МОДЕЛИ**

Научная статья

Никулин Д.А.^{1,*}, Офицеров В.П.²¹ ORCID : 0009-0004-2103-0003;² ORCID : 0000-0002-8055-3785;^{1,2} Московский авиационный институт (национальный исследовательский университет), Москва, Российская Федерация

* Корреспондирующий автор (nikulas.nikulin[at]yandex.ru)

Предложена: 16.04.2026; Принята: 22.05.2026; Опубликовано: 17.06.2026

Аннотация

В статье представлен сравнительный анализ методов классификации телефонных обращений клиентов на основе текстовых транскриптов. Рассматриваются классические алгоритмы машинного обучения, использующие статистические методы векторизации текста (Bag-of-Words, TF-IDF, n-граммы), а также современные трансформерные модели семейства BERT. Проведено экспериментальное сопоставление подходов по метрикам качества классификации (Precision, Recall, F1-score), вычислительной сложности и пригодности для внедрения в корпоративные тикет-системы. Особое внимание уделено обработке русскоязычных текстов, содержащих ошибки автоматического распознавания речи. Установлено, что трансформерные модели обеспечивают прирост качества классификации до 13,4% по F1-score при существенном увеличении вычислительных затрат. Предложены рекомендации по практическому применению различных классов моделей.

Ключевые слова: классификация текста, телефонные обращения, машинное обучение, трансформеры, BERT, SVM, тикет-система, NLP.

A COMPARATIVE ANALYSIS OF METHODS FOR CLASSIFYING TELEPHONE ENQUIRIES: CLASSICAL ALGORITHMS AND TRANSFORMER MODELS

Research article

Nikulin D.A.^{1,*}, Ofitserov V.P.²¹ ORCID : 0009-0004-2103-0003;² ORCID : 0000-0002-8055-3785;^{1,2} Moscow Aviation Institute (National Research University), Moscow, Russian Federation

* Corresponding author (nikulas.nikulin[at]yandex.ru)

Suggested: 16.04.2026; Accepted: 22.05.2026; Published: 17.06.2026

Abstract

The article presents a comparative analysis of methods for classifying customer telephone enquiries based on text transcripts. Classic machine learning algorithms that utilise statistical text vectorisation methods (Bag-of-Words, TF-IDF, n-grams) are examined, as well as modern transformer models from the BERT family. An experimental comparison of the approaches was conducted using classification quality metrics (Precision, Recall, F1-score), computational complexity, and suitability for implementation in corporate ticketing systems. Particular attention is paid to the processing of Russian-language texts containing errors from automatic speech recognition. It has been established that transformer models provide an increase in classification quality of up to 13.4% in terms of F1-score while significantly increasing computational costs. Recommendations are suggested for the practical application of various classes of models.

Keywords: text classification, telephone enquiries, machine learning, transformers, BERT, SVM, ticketing system, NLP.

Введение

Автоматизация обработки телефонных обращений является важной задачей современных корпоративных информационных систем. Рост нагрузки на контакт-центры требует применения интеллектуальных методов анализа текстовых данных.

Типовой конвейер обработки обращений включает:

- автоматическое распознавание речи;
- формирование текстового транскрипта;
- классификацию обращения;
- маршрутизацию заявки.

Ключевым этапом является классификация, определяющая дальнейшую обработку обращения.

Цель работы — сравнительный анализ классических и нейросетевых методов классификации.

Обзор литературы

Задача классификации текстов подробно рассмотрена в фундаментальных работах по обработке естественного языка. В работе Jurafsky и Martin [1] представлены базовые методы NLP и модели обработки речи.

Методы статистической обработки текста, включая TF-IDF и векторные модели, рассмотрены в [2]. Применение классических алгоритмов для задач call-центров исследовано в работах [3], [4], [5].

Метод опорных векторов показал высокую эффективность в задачах классификации текста [6].

Современные исследования сосредоточены на трансформерных моделях. Модель BERT предложена в работе Devlin и др. [7], а её улучшенные версии представлены в [8], [9], [10].

Для многоязычных задач применяются модели XLM-R [11]. Русскоязычные трансформеры исследованы в [12].

Таким образом, современное состояние области характеризуется переходом от классических алгоритмов к глубоким нейросетевым архитектурам.

Постановка задачи

Задача формализуется как многоклассовая классификация:

$$f : X \rightarrow Y$$

где X — множество текстов, Y — множество классов.

Рассматриваются классы:

- консультация;
- лицензирование;
- доступ к portalу;
- обучение;
- сотрудничество.

Методы исследования

4.1. Классические модели

Используются методы векторизации:

- Bag-of-Words.
- TF-IDF.
- n-граммы.

Алгоритмы:

- SVM.
- Naive Bayes.
- Logistic Regression.
- Random Forest.
- Decision Tree.

4.2. Трансформерные модели

Рассматриваются модели:

- ruBERT.
- ruRoBERTa.
- ruELECTRA.
- DeBERTa.

Методика эксперимента

Объем датасета — 401 запись.

5.1. Распределение классов

Таблица 1 - Распределение записей по классам

DOI: <https://doi.org/10.60797/IRJ.2026.168.15.1>

Класс обращения	Количество записей, шт
Сотрудничество	104
Лицензирование	94
Обучение	84
Консультация	71
Доступ к portalу	48

5.2. Предобработка

Исходные данные представляли собой аудиозаписи телефонных обращений, для которых были получены текстовые транскрипты с использованием системы автоматического распознавания речи (ASR). Полученные тексты приведены к унифицированному виду: выполнены очистка от служебных символов, нормализация регистра, удаление нерелевантных элементов и анонимизация персональных данных, таких как ФИО, номера телефонов, адреса и другие чувствительные данные.

На этапе подготовки данных использовалось разбиение датасета на обучающую и тестовую выборки в соотношении 80/20. Итоговая оценка качества моделей и построение матриц ошибок выполнялись на основе 5-кратной кросс-валидации по полному подготовленному набору данных.

5.3. Метрики

Для оценки качества обученных моделей использовались метрики:

- Precision.
- Recall.
- F1-мера.
- Матрица ошибок (confusion matrix).

5.3. Параметры обучения трансформерных моделей

В таблице 2 приведены параметры, используемые при обучении трансформерных моделей.

Таблица 2 - Параметры обучения трансформерных моделей

DOI: <https://doi.org/10.60797/IRJ.2026.168.15.2>

Параметр	Значение
Batch size	4
Learning rate	5×10^{-6}
Оптимизатор	AdamW
Max Sequence Length	512
Epochs	до 100
Early Stopping	10 эпох

Подбор параметров носил прикладной характер и был ориентирован на достижение устойчивого качества классификации при ограничениях используемого вычислительного стенда.

Была проведена адаптация следующих моделей:

- mDeBERTa-v3;
- ruBERT-tiny;
- ruRoBERTa-large;
- ruBERT-base-cased;
- xlm-roberta-base;
- ruELECTRA-large;
- ruELECTRA-medium.

5.4. Аппаратные характеристики

Параметры обучения были подобраны в соответствии с техническими характеристиками вычислительной техники, на которой производилось обучение моделей, а также путем ручного подбора. Характеристики тестового стенда приведены в таблице 3.

Таблица 3 - Характеристики вычислительного стенда

DOI: <https://doi.org/10.60797/IRJ.2026.168.15.3>

Параметр	Значение
GPU	RTX 5060 Ti 16 ГБ
CPU	Ryzen 5 5600X
RAM	16 ГБ

Результаты

6.1. Классические методы

Результаты обучения классических алгоритмов представлены в таблице 4.

Таблица 4 - Результаты обучения классических алгоритмов

DOI: <https://doi.org/10.60797/IRJ.2026.168.15.4>

Модель	F1-score (взвешенное)	Precision	Recall	Время обучения, сек	Время предсказания, мс
TF-IDF + NaiveBayes	0,7915	0,7899	0,7955	0,5	0,153

Модель	F1-score (взвешенное)	Precision	Recall	Время обучения, сек	Время предсказания, мс
N-gram(1-3) + SVM	0,7830	0,7833	0,7830	1,4	0,348
TF-IDF + SVM	0,7777	0,7777	0,7781	0,8	0,268
BoW + NaiveBayes	0,7737	0,7780	0,7731	0,2	0,079
TF-IDF + LogReg	0,7702	0,7754	0,7681	4,7	0,262
N-gram(1-3) + LogReg	0,7642	0,7701	0,7631	4,9	0,402
BoW + LogReg	0,6757	0,6802	0,6758	17,6	0,119
BoW + RandomForest	0,6620	0,6932	0,6683	1,3	0,522
BoW + SVM	0,6564	0,6600	0,6559	0,2	0,086
TF-IDF + RandomForest	0,6514	0,6752	0,6559	1,7	0,608
BoW + DecisionTree	0,5944	0,6024	0,5910	0,3	0,086
TF-IDF + DecisionTree	0,5672	0,5666	0,5686	0,7	0,169
N-gram(1-3) + DecisionTree	0,5610	0,5592	0,5636	1,4	0,232

Наилучшие результаты по F1-мере продемонстрировал наивный Байесовский классификатор в сочетании с признаковым представлением текста на основе TF-IDF, достигнув значения F1=0,7915.

Матрица ошибок представлена на рисунке 1.

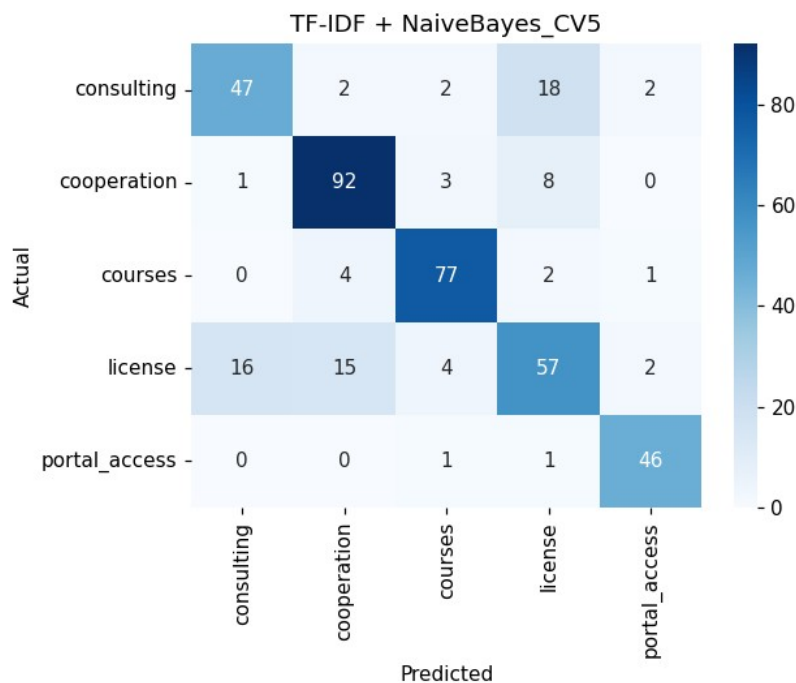


Рисунок 1 - Матрица ошибок модели Naive Bayes с признаковым представлением TF-IDF
DOI: <https://doi.org/10.60797/IRJ.2026.168.15.5>

Гистограмма сравнения классических алгоритмов машинного обучения приведена на рисунке 2.

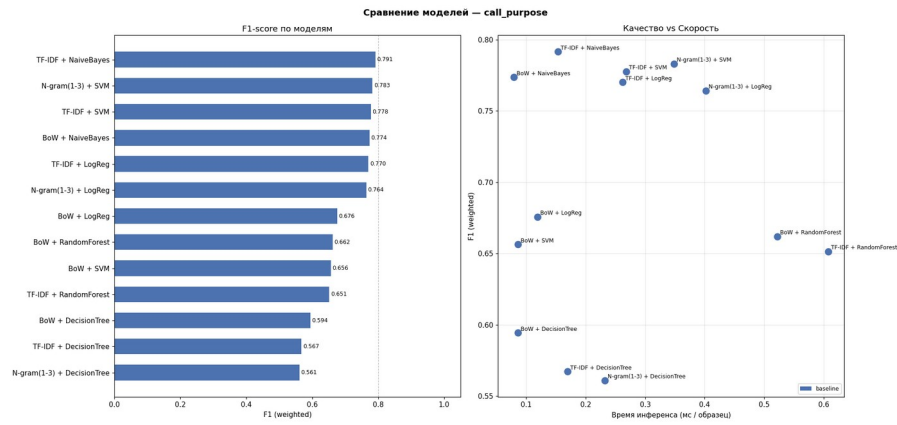


Рисунок 2 - Сравнение классических моделей по F1-score и времени предсказания
DOI: <https://doi.org/10.60797/IRJ.2026.168.15.6>

Вывод: классические алгоритмы машинного обучения демонстрируют приемлемую точность и стабильные результаты в задаче классификации, обладая при этом высокой скоростью работы

6.2. Трансформерные модели

Таблица 5 - Результаты обучения трансформерных моделей

DOI: <https://doi.org/10.60797/IRJ.2026.168.15.7>

Модель	F1-score (взвешенное)	Precision	Recall	Время обучения, с	Время предсказания, мс
XLM-RoBERTa	0,8978	0,8984	0,8978	473,7	3,613
ruRoBERTa-large	0,8881	0,8888	0,8878	1474,0	10,999
RuBERT-tiny	0,8651	0,8665	0,8653	588,7	1,536
RuBERT	0,8605	0,8607	0,8603	604,5	3,574
ruELECTRA-large	0,8558	0,8567	0,8554	1662,3	17,596
mDeBERTa-v3	0,8303	0,8310	0,8304	1786,7	11,510
ruELECTRA-medium	0,7518	0,7503	0,7556	1032,8	4,098

По сравнению с лучшей классической моделью (наивный байесовский классификатор в сочетании с признаковым представлением текста на основе TF-IDF, F1 = 0,7915) модель XLM-RoBERTa повысила значение F1-взвешенного показателя до 0,8978, что соответствует приросту примерно на 13,4%. Матрица ошибок представлена на рисунке 3.

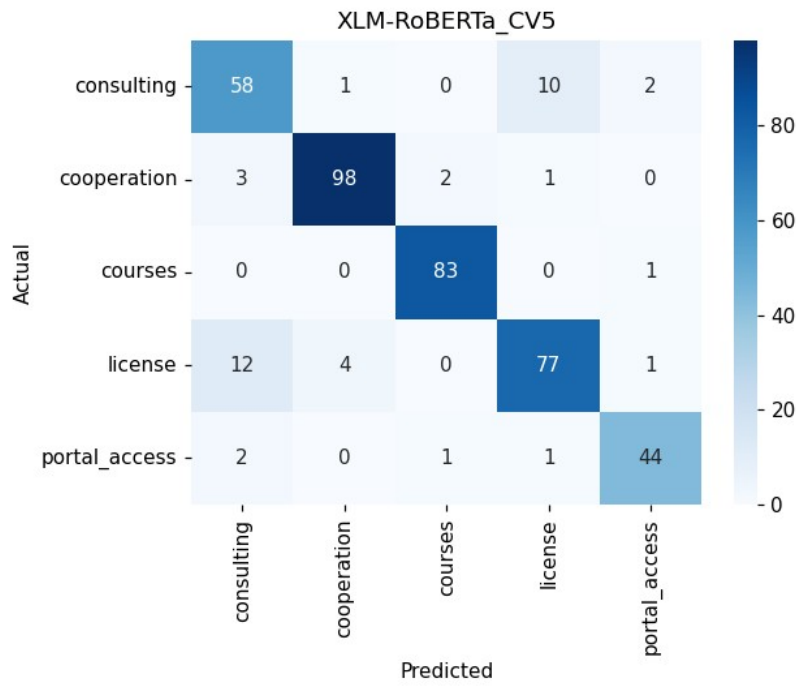


Рисунок 3 - Матрица ошибок для модели XLM-RoBERTa
 DOI: <https://doi.org/10.60797/IRJ.2026.168.15.8>

Гистограмма сравнения моделей по метрике F1 и времени предсказания приведена на рисунке 4.

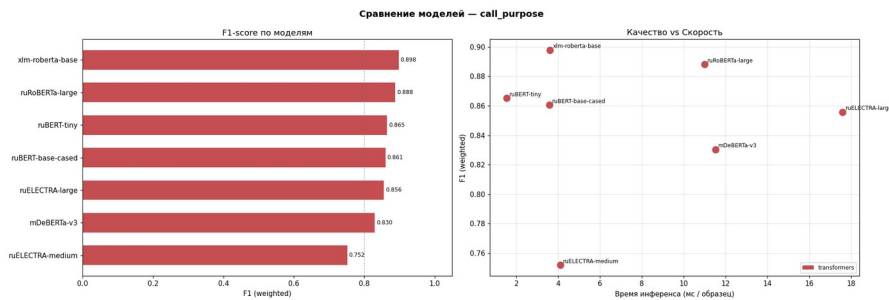


Рисунок 4 - Гистограмма сравнения трансформерных моделей по метрике F1 и времени предсказания
 DOI: <https://doi.org/10.60797/IRJ.2026.168.15.9>

Вывод: использование трансформерных архитектур обеспечивает прирост качества классификации по сравнению с классическими алгоритмами машинного обучения (прирост качества $\approx + 13,4\%$).

Обсуждение

Результаты демонстрируют компромисс между качеством и вычислительной сложностью.

Классические методы обеспечивают высокую скорость работы при умеренной точности.

Трансформерные модели показывают прирост качества ($\sim 13,4\%$), но требуют существенно больших ресурсов. График с фронтом Парето представлен на рисунке 5.

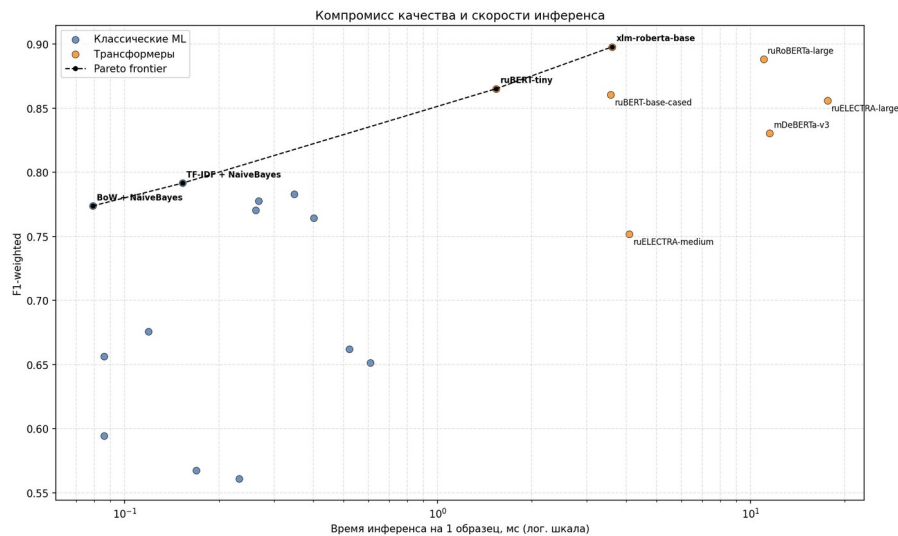


Рисунок 5 - Сравнение моделей по качеству классификации и времени предсказания с выделением границы Парето

DOI: <https://doi.org/10.60797/IRJ.2026.168.15.10>

По результатам сравнительного анализа установлено, что классические методы машинного обучения обеспечивают приемлемое качество классификации при значительно меньших вычислительных затратах.

Трансформерные модели демонстрируют более высокие значения F1-score за счет учета контекстных зависимостей и устойчивости к вариативности формулировок обращений, при этом рост качества сопровождается увеличением вычислительных затрат: если для лучшей классической модели время предсказания составляет 0,153 мс на запись, то для XLM-RoBERTa — 3,613 мс, то есть примерно в 23 раза выше. Полученные данные согласуются с современными исследованиями [9], [10], [11], [12].

Практическая применимость

На основании результатов эксперимента предлагается гибридный подход к построению архитектур: в качестве базового решения (baseline) могут применяться быстрые и вычислительно эффективные алгоритмы, такие как, например, логистическая регрессия, или метод опорных векторов. В качестве основного классификатора целесообразно применять трансформерную модель.

Для систем реального времени, работающих с большим потоком обращений, классические методы могут использоваться как базовый или резервный контур классификации.

Трансформерные модели целесообразно применять в качестве основного механизма интеллектуальной маршрутизации обращений при наличии достаточных вычислительных ресурсов.

Заключение

Проведенный в статье сравнительный анализ классических алгоритмов машинного обучения и современных трансформерных моделей для задачи классификации телефонных обращений в корпоративных тикет-системах подтвердил выдвинутую гипотезу о более высоком качестве классификации, обеспечиваемом трансформерными моделями.

По результатам эксперимента установлено, что классические методы машинного обучения характеризуются меньшей вычислительной сложностью, более высокой скоростью работы и сохраняют практическую актуальность в условиях ограниченных вычислительных ресурсов. Трансформерные модели демонстрируют более высокие значения метрик качества, что делает их предпочтительными для задач, в которых приоритетом является точность классификации.

Ограничения: эксперимент проведен на малом датасете из 401 записи, поэтому дальнейшие исследования целесообразно проводить на более крупных и разнообразных корпусах телефонных обращений.

Перспективы:

- увеличение выборки;
- использование ансамблей;
- учет ошибок ASR.

**Конфликт интересов**

Не указан.

Рецензия

Все статьи проходят рецензирование. Но рецензент или автор статьи предпочли не публиковать рецензию к этой статье в открытом доступе. Рецензия может быть предоставлена компетентным органам по запросу.

Conflict of Interest

None declared.

Review

All articles are peer-reviewed. But the reviewer or the author of the article chose not to publish a review of this article in the public domain. The review can be provided to the competent authorities upon request.

Список литературы на английском языке / References in English

1. Jurafsky D. Speech and Language Processing / D. Jurafsky, J. Martin. — Stanford: Stanford University, 2026. — 600 p. — URL: <https://web.stanford.edu/~jurafsky/slp3/>. (accessed: 15.04.26).
2. Manning C.D. Introduction to Information Retrieval / C.D. Manning, P. Raghavan, H. Schütze. — Cambridge: Cambridge University Press, 2008. — 506 p.
3. Busemann S. Message classification in call centers / S. Busemann, S. Schmeier, R. Arens. // Proceedings of the Sixth Conference on Applied Natural Language Processing; — Seattle: Association for Computational Linguistics, 2000. — P. 158–165.
4. Haffner P. Optimizing SVMs for complex call classification / P. Haffner, G. Tür, J. Wright. // Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing; — Hong Kong: IEEE, 2003. — P. 632–635.
5. Malik S. Classification of Call Transcriptions / S. Malik, M. Idrees, H.M. Danish et al. // VAWKUM Transactions on Computer Sciences. — 2023. — Vol. 11, No. 2. — P. 1–17.
6. Joachims T. Text Categorization with Support Vector Machines / T. Joachims. // Proceedings of the European Conference on Machine Learning; — Chemnitz: Springer, 1998. — P. 137–142. doi: 10.1007/BFb0026683
7. Devlin J. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding / J. Devlin, M.W. Chang, K. Lee et al. // Proceedings of NAACL; — Minneapolis: ACL, 2019. — P. 4171–4186. doi: 10.18653/v1/N19-1423
8. Liu Y. RoBERTa: A Robustly Optimized BERT Pretraining Approach / Y. Liu, M. Ott, N. Goyal [et al.] // arXiv. — 2019. — URL: <https://arxiv.org/abs/1907.11692> (accessed: 15.04.2026).
9. He P. DeBERTa: Decoding-enhanced BERT with Disentangled Attention / P. He, X. Liu, J. Gao [et al.] // arXiv. — 2020. — URL: <https://arxiv.org/abs/2006.03654> (accessed: 15.04.2026).
10. Clark K. ELECTRA: Pre-training Text Encoders as Discriminators Rather than Generators / K. Clark, M.T. Luong, Q.V. Le et al. // Proceedings of ICLR; — Addis Ababa: ICLR, 2020. — P. 1–18.
11. Conneau A. Unsupervised Cross-lingual Representation Learning at Scale / A. Conneau, K. Khandelwal, N. Goyal et al. // Proceedings of ACL; — Online: ACL, 2020. — P. 8440–8451. doi: 10.18653/v1/2020.acl-main.747
12. Zmitrovich D. A Family of Pretrained Transformer Language Models for Russian / D. Zmitrovich, A. Abramov, A. Kalmykov et al. // Computational Linguistics and Intellectual Technologies; — Moscow: RSUH, 2024. — P. 507–524.