

УДК 81'33

АНАЛИЗ ТОНАЛЬНОСТИ ТЕКСТА С ПРИМЕНЕНИЕМ БОЛЬШОЙ ЯЗЫКОВОЙ МОДЕЛИ RUBERT

М. В. Свиридович

*Белорусский государственный университет,
ул. К. Маркса, 31, 220030, г. Минск, Беларусь, sviridovichmv@mail.ru*

Исследуется эффективность применения нейронных сетей на основе архитектуры трансформера для классификации тональности отзывов о кинофильме. В результате проведения эксперимента с использованием двух подходов – предсказания с помощью предобученной модели ruBERT и модификации ruBERT с дальнейшим дообучением – установлено, что небольшая модификация модели позволила повысить общую точность (меткость, ассурасу) классификации при увеличении чувствительности классификации положительных отзывов и снижении чувствительности классификации отрицательных отзывов.

Ключевые слова: анализ тональности текста, большая языковая модель, трансформер, дообучение.

SENTIMENT ANALYSIS USING THE LARGE RUBERT LANGUAGE MODEL

M. V. Sviridovich

*Belarusian State University,
31 Karl Marks St, 220030, Minsk, Belarus, sviridovichmv@mail.ru*

The effectiveness of using neural networks based on the transformer architecture for tonality classification of film reviews is investigated. As a result of the experiment incorporating two approaches – prediction using the pre-trained ruBERT model and modification of ruBERT followed by further training – it was found that a small modification of the model lead to increase in the overall accuracy of the classification, with a simultaneous increase in the recall of positive reviews and decrease in the recall of negative reviews.

Keywords: sentiment analysis; large language model; transformer; additional training.

В компьютерной лингвистике анализ тональности текста представляет собой совокупность методов автоматизированного выявления эмоциональной позиции авторов (мнений) по отношению к объектам высказываний.

К настоящему времени в области анализа тональности наиболее эффективно показали себя нейронные сети, основанные на архитектуре

трансформера, описанного впервые Vaswani et. al. [1]. До появления трансформера большинство языковых моделей основывалось на рекуррентных архитектурах, где учитывалась только предопределенная временная структура связи между элементами последовательности. В трансформере же с помощью механизма self-attention значимость каждого слова определяется относительно всех других слов в контексте, что позволяет учитывать не только временные, но и пространственные связи в тексте. Таким образом, все веса в блоке self-attention вычисляются в результате попарного сравнения линейных трансформаций входных токенов – матриц Key, Query, Value. Это позволяет более эффективно моделировать зависимости между словами в предложении и тексте.

В разработанных вскоре после появления трансформера модификациях был расширен функционал и преодолены ограничения базовой архитектуры. Одной из самых используемых модификаций является BERT (Bidirectional Encoder Representations from Transformers – двунаправленные репрезентации энкодера трансформера). Стандартный трансформер был разработан и обучен для задач sequence-to-sequence (последовательность в последовательность, например для перевода) в рамках одностороннего контекста: предсказывались только токены, следующие за определенным токеном. BERT представляет собой модификацию, основанную на двунаправленности контекста: модель обучена предсказывать токен на любой позиции в тексте. Модификация позволяет еще точнее устанавливать контекстуальные связи между токенами.

Изначально разработанная для генерации токенов, модель BERT была адаптирована и для других задач обработки естественного языка, включая анализ тональности. Целью нашего эксперимента является классификация тональности отзывов (нейтральный, позитивный, негативный) с применением большой языковой модели ruBERT и ее модификации, а также определение эффективности классификации путем сравнения предсказанных значений с истинными классами отзывов, выставленными авторами отзывов. Материалом послужили 209 отзывов зрителей о кинофильме «Мастер и Маргарита» (2023), размещенные на сайте kinopoisk.ru.

В качестве модели была выбрана предобученная разновидность BERT [rubert-base-cased-sentiment](#) из python-библиотеки transformers [2], предобученная на четырех русскоязычных датасетах: корпусе твитов RuTweetCorp; корпусе отзывов покупателей на приобретенные товары [RuReviews](#); корпусе постов в социальных сетях [RuSentiment](#); корпусе отзывов на медицинские учреждения с сайта prodoctorov.ru. Данная модель была напрямую использована для предсказания класса тональности.

Выбранный нами датасет стилистически и содержательно отличается от корпусов, на которых был предобучен базовый ruBERT. С целью его адаптивирования непосредственное под имеющиеся отзывы мы внесли в модель небольшую модификацию: веса предобученного ruBERT были оставлены без изменения («заморожены»), а после последнего слоя добавлено два дополнительных обучаемых слоя [3], что позволило откалибровать модель без ресурсозатратного дообучения BERT. Данная модель была дообучена на 100 отзывах, и ее веса были использованы для предсказания класса сентимента 100 других отзывов. Девять отзывов при этом были использованы для валидации модели.

Описание модели: 1-ый слой размерностью 768, 2-ой 512; слои связаны ReLU функцией активации; для регуляризации применен Dropout (0,1); скорость обучения 0,0001; оптимизатор AdamW; функция потерь – категориальная кросс-энтропия, веса классов сбалансированы исходя из распределения истинных меток; 40 эпох обучения.

Текст отзывов был токенизирован с помощью токенайзера, являющегося частью указанной вариации BERT. Для моделей семейства BERT параметры токенайзера определяются в процессе обучения самой модели.

Результаты классификации с использованием двух вышеописанных моделей представлены в таблицах 1 и 2:

Таблица 1

Распределение истинных классов и классов, предсказанных ruBERT

Истинный класс	Результат модели	Количество	Доля
0	0	20	0.54
	1	7	0.19
	2	10	0.27
1	0	40	0.35
	1	59	0.51
	2	16	0.14
2	0	21	0.37
	1	1	0.02
	2	35	0.61

Таблица 2

**Распределение истинных классов и классов, предсказанных ruBERT+2
обучаемых слоя**

Истинный класс	Результат модели	Количество	Доля
0	0	5	0.33
	1	8	0.53
	2	2	0.13
1	0	5	0.08
	1	50	0.85
	2	4	0.07
2	0	7	0.27
	1	10	0.38
	2	9	0.35

Эффективность классификатора определяется прежде всего показателями precision (точность) и recall (полнота, чувствительность).

Таблица 3

Метрики классификатора ruBERT

Классы, метрики	precision	recall	f1-score	Количество
0	0,25	0,54	0,34	37
1	0,88	0,51	0,65	115
2	0,57	0,61	0,59	57
macro avg	0,57	0,56	0,53	209
weighted avg	0,68	0,55	0,58	209
accuracy	0,55			209

Таблица 4

Метрики классификатора ruBERT+2 обучаемых слоя

Классы, метрики	precision	recall	f1-score	Количество
0	0,29	0,33	0,31	15
1	0,74	0,85	0,79	59
2	0,60	0,35	0,44	26
macro avg	0,51	0,48	0,45	100
weighted avg	0,62	0,60	0,58	100
accuracy	0,64			100

Согласно показателям recall, представленным в таблицах 3 и 4, ruBERT лучше справляется с обнаружением нейтральных и негативных отзывов. Дообученный BERT гораздо лучше базового обнаруживает позитивные отзывы (0,85 против 0,51). Из метрик таблицы 3 следует, что ruBERT также в целом демонстрирует более высокую точность оценок главным образом за счет показателя в 0,88 для положительных отзывов. При этом чувствительность для данного класса достаточно невысокая – 0,51. Дообученный же ruBERT показывает хорошие результаты и точности, и полноты для положительных классов (0,74 и 0,85 соответственно – таблица 4).

В целом дообученный ruBERT проигрывает стандартному в полноте классификаций нейтральных и негативных отзывов, а также в точности определения положительных отзывов. Вместе с тем дообученный ruBERT демонстрирует более высокую меткость (долю правильных классификаций от всех классификаций: 0,64 против 0,55) за счет высоких и сбалансированных показателей для позитивных отзывов и показателя точности для негативных и нейтральных.

Таблица 5

Изменения в результатах классификации

Ухудшение классификации		Улучшение классификации	
Класс	Количество	Класс	Количество
0	6	0	5
1	0	1	17
2	7	2	0

В таблице 5 отражено, как конкретно изменились метки классов при переходе со стандартной модели на дообученную: последняя стала хуже определять негативные отзывы, при этом научившись гораздо лучше определять позитивные. В качестве примера улучшения результата приведем положительный отзыв [4], содержащий элементы восторженной оценки (*«Здорово смотрятся второстепенные персонажи. Некоторые – дикий восторг»*), который базовым ruBERT был классифицирован как нейтральный, а модифицированным – как истинноположительный.

В докладе рассмотрен пример отзыва с непоследовательно выставленной автором нейтральной оценкой [5], классифицированный моделью как отрицательный. Текст отзыва насыщен негативными эмоциональными элементами (*«Абсолютно посредственный, банальный в плане повествования фильм», «Жалко, что так лениво и плоско подошли к грандиозному материалу»* и др.).

Выводы: 1. Несмотря на в целом не очень высокую меткость (accuracy 0,55), ruBERT позволяет в большинстве случаев правильно классифицировать достаточно объемные и стилистически разнообразные тексты. 2. Дообученная с добавлением всего двух полносвязных слоев модель повышает меткость до 0,64 и служит эффективным классификатором положительных отзывов (85%). 3. Оптимально подобранная архитектура для дообучения позволяет более тонко откалибровать модель для конкретного корпуса без необходимости задействования значительных вычислительных ресурсов.

Библиографические ссылки

1. Attention Is All You Need, Vaswani et. al. [Electronic resource]. URL: <https://arxiv.org/pdf/1706.03762.pdf> (date of access: 20.02.2024).
2. RuBERT for Sentiment Analysis [Electronic resource]. URL: <https://huggingface.co/blanchefort/rubert-base-cased-sentiment> (date of access: 20.02.2024).
3. Трансформеры для классификации [Electronic resource]. URL: <https://newtechaudit.ru/klassifikaciya-teksta-s-ispolzovaniem-modelej-transformerov/> (date of access: 20.02.2024).
4. История на все времена [Electronic resource]. URL: <https://www.kinopoisk.ru/user/166693243/comment/3346284/> (date of access: 19.02.2024).
5. Все это мы видели сотни раз [Electronic resource]. URL: <https://www.kinopoisk.ru/user/2100661/comment/3347053/> (date of access: 19.02.2024).