

## Языковые нейронные сети для сопоставления текстовых описаний товаров

В. А. Зенков<sup>1</sup>, В. С. Панищев<sup>1</sup>, А. Л. Желанов<sup>2</sup>, Д. А. Волков<sup>3</sup>

<sup>1</sup>Юго-Западный государственный университет, г. Курск

<sup>2</sup>Курская академия государственной и муниципальной службы, г. Курск

<sup>3</sup>Российский государственный университет нефти и газа (национальный исследовательский университет) имени И.М. Губкина, г. Москва

**Аннотация:** Статья посвящена применению языковых нейронных сетей для сопоставления текстовых описаний товаров. Проведен анализ методов сравнения текстовых описаний товаров, отмечены преимущества и недостатки каждого метода. Рассмотрен способ сопоставления текстовых описаний товаров, в основу которого положены нейронные сети архитектуры Bert. Проведены эксперименты и тестирования на наборах данных текстовых описаний схожих товаров из разных торговых сетей. Сделаны выводы о качестве сопоставления различных сетей архитектуры Bert.

**Ключевые слова:** нейронные сети, трансформеры, сравнения текстовых описаний, анализ текста, Bert

### Введение

На сегодняшний день крупные представители торговых сетей активно прорабатывают технологии распознавания товаров [1,2]. Подобные решения позволяют с одной стороны повысить эффективность работы товароведов за счет уменьшения доли ручного труда, а с другой стороны предоставить пользователям удобный сервис для получения актуальной информации об интересующем товаре.

Одной из важных задач сопоставления товаров является сопоставление их текстовых описаний. Данная задача заключается в определении, являются ли два или более товара одним и тем же уникальным товаром, несмотря на возможные различия в их текстовых описаниях. Эти различия могут возникать в следствие следующих особенностей текстовых описаний:

- *сокращения и аббревиатуры:* использование сокращений или аббревиатур в названиях товаров, что может затруднить их сопоставление (например, «клубничный» и «клубн.»),

- *языковые различия*: использование разных языков или транслитерации в названиях товаров (например, «бигбон» и «bigbon»),
- *наличие переводов*: производители могут переводить только некоторые части описаний товаров (например, «Черный чай Волшебные ночи» или «Черный чай Magic Nights», «Напиток со вкусом вишни» или «Напиток cherry»),
- *синонимы*: использование разных слов или фраз для описания одного и того же типа товара (например, «молоко» и «молочный продукт»),
- *опечатки и ошибки*: наличие опечаток или ошибок в текстовых описаниях товаров,
- *вариации написания*: разные варианты написания одного и того же слова или фразы (например, «бигбон» и «биг бон»).

При сопоставлении товаров следует учитывать следующие факторы:

- *бренд*: товары должны иметь одинаковый бренд или производителя,
- *вкус*: товары должны иметь одинаковый вкус (клубничный, яблочный и т.д.),
- *тип упаковки*: товары должны иметь одинаковый тип упаковки (если указан в текстовом описании товара),
- *вес*: товары должны иметь одинаковый вес (товары с разной массой имеют разные цены, поэтому их сопоставление нецелесообразно).

Важно не сопоставлять продукты, имеющие одинаковый бренд, но, например, разные вкусы или другие характеристики. Таким образом, в работе рассматривается возможность применения языковых нейронных сетей для сопоставления текстовых описаний сопоставляемых товаров, с учетом перечисленных выше факторов.

---

## **Обзор возможных методов сопоставления текстовых описаний товаров**

На основе обзора литературы были выделены следующие методы сопоставления текстовых описаний товаров: словари синонимов и поиск похожих слов регулярными выражениями, статистический метод TF-IDF (Term Frequency-Inverse Document Frequency), машинное обучение и нейросети (BERT модели). Далее рассмотрены эти методы, а также выделены их положительные и отрицательные стороны.

### **1. Словарь синонимов и поиск похожих слов регулярными выражениями.**

Метод использования словаря синонимов и поиска похожих слов с помощью регулярных выражений может быть применен к проблеме сопоставления товаров, предлагая простой подход к обработке текстовых описаний продуктов. Этот метод направлен на решение некоторых вопросов проблемы сопоставления, в том числе связанных с вариативностью текстовых описаний.

Преимущества использования этого метода в контексте сопоставления товаров:

- простота реализации, что позволяет быстро внедрить базовое решение для сопоставления товаров,
- низкие требования к вычислительным ресурсам, что может быть важно при обработке больших каталогов товаров,
- высокая скорость обработки небольших и средних объемов данных, что обеспечивает быстрое сопоставление товаров в реальном времени.

Однако применение этого метода к проблеме сопоставления товаров имеет ряд существенных ограничений:

- низкая точность при обработке сложных и разнообразных описаний товаров. Это может привести к ошибкам в сопоставлении продуктов
-

с похожими, но разными характеристиками (например, товары одного бренда, но с разными вкусами),

- необходимость регулярного обновления словаря синонимов, что требует постоянного мониторинга и ручной работы, особенно при появлении новых продуктов или изменении существующих,
- ограниченная способность обрабатывать сокращения, аббревиатуры и различные варианты написания (например, «бигбон» и «биг бон»), если они не включены в словарь синонимов,
- сложности с обработкой опечаток и ошибок в текстовых описаниях товаров, так как регулярные выражения обычно требуют точного соответствия,
- ограниченная возможность учета контекста и семантических связей между словами, что может привести к неправильному сопоставлению товаров с похожими, но разными характеристиками.

Таким образом, хотя метод словаря синонимов и регулярных выражений может быть полезен для базового сопоставления товаров, особенно в случаях со стандартизированными описаниями, но, вероятно, не сможет полностью решить все проблемы, возникающие при сопоставлении текстовых описаний товаров.

## **2. Статистический метод TF-IDF (Term Frequency-Inverse Document Frequency)**

Статистический метод оценивает важность терминов в описании товара относительно всего каталога продуктов, что может помочь в решении некоторых проблем сопоставления [3,4].

Преимущества использования TF-IDF в контексте сопоставления товаров:

- способность выделять важные термины в описаниях товаров, потенциально помогая идентифицировать ключевые характеристики, такие как бренд, вкус или тип упаковки,
- улучшенная точность по сравнению с простыми методами на основе ключевых слов, что может помочь в различении схожих, но разных продуктов.

Однако, применение TF-IDF к проблеме сопоставления товаров имеет ряд ограничений:

- ограниченная способность учитывать контекст может затруднить обработку синонимов, сокращений и различных вариантов написания (например, «бигбон» и «биг бон»),
- метод может испытывать сложности при работе с опечатками и ошибками в текстовых описаниях товаров,
- TF-IDF не способен напрямую учитывать семантические связи между словами, что может быть важно при сопоставлении товаров с различными, но связанными описаниями.

Таким образом, метод лучше справляется с задачей сопоставления чем словарь синонимом или регулярные выражения, но также не позволяет решить все проблемы сопоставления текстовых описаний товаров в силу отмеченных выше ограничений.

### **3. Машинное обучение и нейросети (BERT модели)**

Применение методов машинного обучения и нейросетей, в частности модели BERT (Bidirectional Encoder Representations from Transformers) и других трансформер-моделей, представляет собой мощный подход к решению проблемы сопоставления текстовых описаний товаров [5]. Эти методы способны эффективно обрабатывать сложные и разнообразные текстовые описания продуктов, преодолевая многие трудности, связанные с вариативностью текста.

Преимущества использования машинного обучения и нейросетей в контексте сопоставления товаров:

- высокая точность и эффективность в сопоставлении товаров, даже при наличии различий в текстовых описаниях, таких как сокращения, языковые различия, синонимы и вариации написания,
- способность учитывать контекст и семантические особенности текстов, что критически важно для правильного сопоставления товаров с учетом бренда, вкуса, типа упаковки и веса,
- автоматическая адаптация к новым данным и терминологии, что позволяет системе эффективно обрабатывать новые продукты и изменения в существующих описаниях без необходимости постоянного ручного обновления словарей,
- потенциальная способность справляться с опечатками и ошибками в текстовых описаниях товаров благодаря обучению на больших объемах данных.

Однако, применение этого метода также сопряжено с определенными требованиями:

- высокие требования к вычислительным ресурсам, особенно на этапе обучения модели, что может быть проблематично при работе с очень большими каталогами товаров,
- необходимость в большом объеме качественных данных для обучения модели, чтобы она могла эффективно справляться с разнообразием текстовых описаний и правильно учитывать все важные характеристики товаров.

Таким образом, использование машинного обучения и нейросетей, таких как BERT, представляется одним из наиболее перспективных подходов к решению проблемы сопоставления товаров. При правильной реализации и

---

обучении, такой подход может обеспечить высокую точность сопоставления товаров, адаптируясь к изменениям в данных.

### **Бинарная классификация текстовых пар описаний товаров**

Выходом BERT-сети является вектор признаков. Так, если определить степень сходства между различными векторами, то можно сделать вывод о том, относятся ли текстовые описания к одному и тому же товару. Этот подход можно рассматривать как бинарную классификацию на основе обученной языковой BERT модели, где класс 1 – одинаковые товары, класс 0 – различные товары. Процесс работы такого классификатора включает следующие этапы:

1. Получение векторных представлений: каждое текстовое описание пропускается через модель BERT, которая генерирует векторы признаков, отражающие семантическое содержание текстовых описаний.
2. Вычисление сходства: для каждой пары текстовых описаний рассчитывается сходство между их векторными представлениями.
3. Применение порогового значения: если сходство между двумя векторами превышает этот порог, тексты считаются семантически близкими, а товары одинаковыми.

Для определения схожести между этими векторами могут использоваться различные метрики [6], такие, как:

- Евклидово расстояние ( $L_2$  distance):
- Манхэттенская дистанция ( $L_1$  distance):
- Косинусное сходство (от англ. Cosine similarity):
- Сходство Жаккара (от англ. Jaccard similarity):

В общем случае, при работе с векторами признаков, которые обычно являются высокоразмерными и разреженными, косинусное сходство часто

---

оказывается более эффективным. Преимуществом данной метрики является то, что результат нормализован, то есть всегда находится в диапазоне от -1 до 1, что упрощает этап подбора порога классификации.

Косинусное сходство принимает значение от -1 до 1. Значение 1 означает, что векторы сонаправленные, а значение -1 означает, что векторы направлены в противоположные стороны, таким образом, в контексте сравнения текстовых описаний, если косинус угла между их векторами признаков близок к 1, то текстовые описания являются семантически схожими.

Для оценки эффективности сопоставления применяются следующие метрики [7]:

- Average precision ( $AP$ ),
- $F_\beta$ , где  $\beta=1/2$ .

Average precision ( $AP$ ) – показатель эффективности классификатора.  $AP$  является метрикой, которая дает целостное представление о качестве классификатора.  $AP$  дает общую картину качества модели по всему спектру возможных порогов классификации, а то есть отражает равномерность распределения векторов признаков в пространстве признаков.

Формула для вычисления  $AP$ :

$$AP = \int_0^1 p(r) dr,$$

где  $p(r)$  – кривая точность-полнота. То есть  $AP$  является площадью под кривой точность-полнота (от англ. Precision-Recall).

$F_\beta$  мера является средним гармонических точности и полноты взятым с коэффициентом ( $\beta^2 + 1$ ):

$$F_b = (\beta^2 + 1) \frac{Recall * Precision}{Recall + \beta^2 Precision}.$$

При использовании значения  $\beta = 0.5$ , получается удвоенный вклад точности в  $F_\beta$  меру, что является  $F_{1/2}$  мерой. Рассчитав  $F_{1/2}$  для каждого допустимого значения порога классификации, можно получить оптимальное значение этого порога, взяв наибольшее значение  $F_{1/2}$  меры.

Таким образом, для определения качества модели необходимо произвести оценку  $AP$ , а также найти точку, в которой показатель  $F_{1/2}$  меры будет максимальный.

### **Выбор лучшей модели BERT-сети.**

Помимо классической модели BERT от компании Google, существуют различные ее модификации. Рассмотрим некоторые из них и их особенности:

- XLM-RoBERTa: Разработана Facebook AI на основе модели BERT от Google. Это многоязычная модель, обученная на 100 языках с использованием большого объема данных [8].
- DistilBERT: Разработана Hugging Face. Это облегченная версия BERT, которая сохраняет 97% производительности оригинала, но на 40% меньше и на 60% производительнее [9].
- ELECTRA: Совместная разработка Google Research и Stanford University. Использует новый метод предобучения, называемый «replaced token detection» (RTD), где модель должна определить, были ли токены в предложении заменены генератором [10].

Архитектура этих моделей схожа с BERT, различия могут быть в количестве слоев и голов внимания. Однако используются разные методы обучения и/или разные наборы данных. Также может отличаться словарь токенизатора.

Поэтому важно выбрать оптимальную модель BERT для дальнейшего обучения на задаче сопоставления товаров. Так как целью данной работы является получение модели, которая может максимально точно сравнивать

---

различные описания товаров, то необходимо произвести выбор модели из множества других, которая покажет максимальные показатели на тестовой выборке до обучения, также можно учесть момент, чтобы модель была мультязычной, так как данные могут представлять собой различные переводы или транслитерации.

Определение оптимальной модели выполняется на тестовой выборке на основе  $AP$  показателя. Тестовая выборка представляет собой два списка товаров двух разных торговых точек. В первом списке 10000 товаров, во втором 8000. Количество пересечений (одинаковых товаров) между этими списками 5000. Все вычисления производились на ЭВМ со следующими характеристиками: процессор Intel Core i7-13700K, видеокарта NVIDIA GeForce GTX 4060 Ti 16GB, память 64GB DDR3 5200МГц.

Таблица № 1

Пример тестовых данных

Тестовая торговая сеть №1 1	Тестовая торговая сеть №2 2
Крупа рис увелка 800 г для плова по-узбекски	Рис шлифованный для плова по-узбекски Увелка 800г
Гренки хрустец 130 г телятина с аджикой	Сухарики жареные ржаные гренки Хрустец со вкусом телятины с аджикой 130г
Корм для кошек феликс 75 г мясные ломтики говядина	Корм для кошек консервированный Felix Мясные ломтики 75г Говядина
Вино игристое санкт-петербургъ традиционное 0.75 л 10, 5-13% белое брют игристые вина	Вино игристое Санкт-Петербург Традиционное белое брют 10.5-13% 0.75л
Салат листовой шт	Салат листовой
Напиток биоматрикс 0,43л 2,5% кисломолочный авида	Биопродукт кисломолочный Biomatrix 2.5% 430г
Сухарики хрустим багет 45 г лобстер	Сухарики Хрустим Багет со вкусом Лобстер 45г
Сушки старооскольские малютка в сахарной глазури 350г Ваш хлеб	Сушки Старооскольские Малютка в глазури 350г
Хлеб Дарницкий новый 700 г Брянский хлебокомбинат №1	Хлеб Дарницкий новый 700г в упаковке Брянский хлебокомбинат N1

1	2
Нектар сочная долина 0,95л персик яблоко	Нектар Сочная Долина персиково-яблочный с мякотью 0.95л
Батон чайный 300 г тамбовский хлебокомбинат	Батон Чайный 300г Тамбовский хлебокомбинат ОАО
Чай красный Принцесса Ява каркадэ 80г	Чай Принцесса Ява Каркадэ 80г м/у
Печенье Oreо 95г со вкусом шоколада Mondelez	Печенье Oreо 95г с шоколадной начинкой
Конфеты Коркунов ассорти 192г темный и молочный шоколад Одинцовская КФ	Конфеты Коркунов 192г Ассорти Темный и молочный шоколад
Гренки 777 50гр холодец хрен	Гренки Три семерки Холодец с хреном 50г
Квас национальный продукт 1,0л медовый	Квас ржаной Национальный продукт медовый 1л
Гренки Хрустец 130г чеснок	Сухарики жареные ржаные гренки Хрустец со вкусом чеснока 130г
Салат листовой шт	Салат листовой
Напиток биоматрикс 0,43л 2,5% кисломолочный авида	Биопродукт кисломолочный Biomatrix 2.5% 430г
Сухарики хрустим багет 45 г лобстер	Сухарики Хрустим Багет со вкусом Лобстер 45г

Для сравнения были выбраны следующие модели:

- ai-forever/sbert\_large\_nlu\_ru,
- BAAI/bge-m3,
- intfloat/multilingual-e5-large,
- intfloat/multilingual-e5-base,
- intfloat/multilingual-e5-small,
- cointegrated/LaBSE-en-ru,
- ai-forever/ruElectra-large,
- dunzhang/stella\_en\_1.5B\_v5,
- DeepPavlov/rubert-base-cased-sentence.

Все модели находятся на открытых репозиториях HuggingFace [11] и доступны для скачивания, распространяются под MIT лицензией.

Таблица № 2

### Метрика AP для различных моделей

Модель	AP
ai-forever/sbert_large_nlu_ru	0.067
BAAI/bge-m3	0.435
intfloat/multilingual-e5-large	0.439
intfloat/multilingual-e5-base	0.37
intfloat/multilingual-e5-small	0.412
cointegrated/LaBSE-en-ru	0.3
DeepPavlov/rubert-base-cased-sentence	0.031
ai-forever/ruElectra-large	0.044
dunzhang/stella_en_1.5B_v5	0.417

Следуя из данных таблицы, было принято решение взять для дальнейшего обучения intfloat/multilingual-e5-large модель. В действительности, данная модель показывает хорошие результаты на различных бенчмарках, таких как ru-Mteb и Mteb.

model: intfloat/multilingual-e5-large, binary Precision-Recall curve: AP=0.439

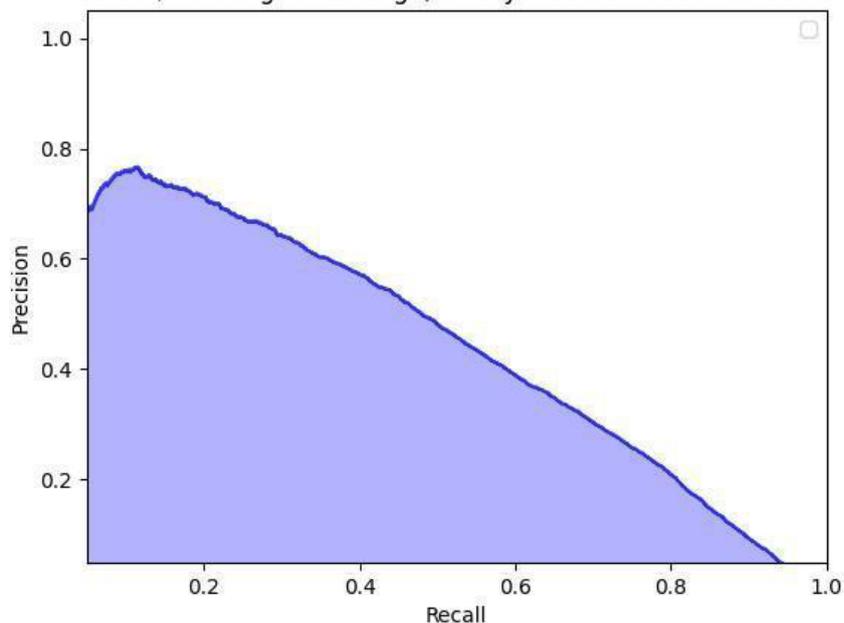


Рис. 1. – Кривая точность-полнота, модель intfloat/multilingual-e5-large

## Обучение модели `intfloat/multilingual-e5-large`

Для обучения модели `intfloat/multilingual-e5-large` использовалась `Cached Multiple Negatives Symmetric Ranking Loss (CMNSRL)` функция потерь, являющаяся модифицированной версией `MultipleNegativesSymmetricRankingLoss (MNSRL)` [12]. Согласно оригинальной статье по обучению модели `intfloat`, при использовании `InfoNCE` в качестве функции потерь использовался низкий температурный коэффициент (от англ. `temperature`) равный 0.01, таким образом для функции `CMNSRL` использовался так же низкий температурный коэффициент равный 0.01. Оптимизатором выступает `AdamW`, скорость обучения (от англ. `learning rate`)  $5.0e-5$ . Коэффициент регуляризации (от англ. `weight_decay`) составляет  $2.0e-2$ . Размер пакета (от англ. `batch_size`) равен 64, а размер кэшированного пакета – 65536.

После обучения модели были получены следующие результаты, представленные на рисунке 2.

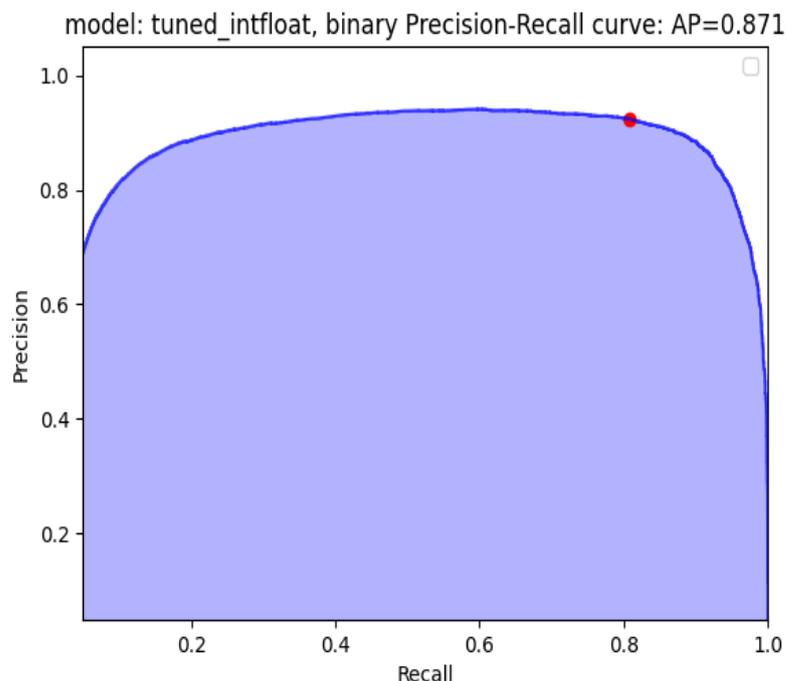


Рис. 2. – Результаты обучения модели `intfloat/multilingual-e5-large`  
Показатель  $AP = 0.871$ . Лучшее пороговое значение – 0.927 при

значении  $F_{1/2}$  меры = 0.892.

Таким образом, при использовании данный нейросетевой классификатор с пороговым значением 0.93, предоставляет высокое качество сопоставления текстовых описаний товаров различных торговых сетей.

### **Выводы**

Проведенные исследования продемонстрировали высокую эффективность использования языковых нейронных сетей, в частности архитектуры BERT, для сопоставления текстовых описаний товаров. Полученные показатели, в том числе  $AP = 0.871$  и  $F_{1/2}$  мера = 0.892, демонстрируют высокую эффективность модели в задаче сопоставления текстовых описаний товаров.

Таким образом, полученная модель может быть полезной для различных приложений в области электронной коммерции, автоматизируя процесс сопоставления товаров и обеспечивая точность и качество в аналитике данных.

## Литература

1. «Лента» тестирует технологии видеораспознавания товаров и количества покупателей в магазине. URL: [lenta.com/o-kompanii/news/-----30](https://lenta.com/o-kompanii/news/-----30). (дата обращения: 25.02.2025).

2. «Магнит» до конца года масштабирует технологию распознавания товаров на полке на всю сеть. URL: [dsmedia.pro/news/magnit-do-konca-goda-masshtabiruet-tehnologiju-raspoznavanija-tovarov-na-polke-na-vsju-set](https://dsmedia.pro/news/magnit-do-konca-goda-masshtabiruet-tehnologiju-raspoznavanija-tovarov-na-polke-na-vsju-set). (дата обращения: 25.02.2025).

3. Christian H., Agus M. P., Suhartono D. Single document automatic text summarization using term frequency-inverse document frequency (TF-IDF) // ComTech: Computer, Mathematics and Engineering Applications, 2016, Т. 7., №4. URL: [journal.binus.ac.id/index.php/comtech/article/view/3746](http://journal.binus.ac.id/index.php/comtech/article/view/3746) (дата обращения: 25.02.2025).

4. Qaiser S., Ali R. Text mining: use of TF-IDF to examine the relevance of words to documents // International journal of computer applications. 2018. Т. 181. №1.

5. Аксенов К.А. Сунь Л. Эволюция и современное состояние систем ответов на вопросы: технологии распознавания намерений и именованных сущностей с использованием модели BERT // Инженерный вестник Дона, 2024, №7. URL: [ivdon.ru/ru/magazine/archive/n7y2024/9371](http://ivdon.ru/ru/magazine/archive/n7y2024/9371) (дата обращения: 25.02.2025).

6. Зенков В.А., Панищев В.С., Желанов А.Л., Волков Д.А. Сверточная нейронная сеть для сопоставления изображений товаров // Инженерный вестник Дона, 2024, №8. URL: [ivdon.ru/ru/magazine/archive/n8y2024/9444](http://ivdon.ru/ru/magazine/archive/n8y2024/9444) (дата обращения: 25.02.2025).

7. Krasnodebska K., Goch W., Uhl J.H., Verstegen J.A., Pesaresi M. Advancing Precision, Recall, F-score, and Jaccard Index: An Approach for

Continuous, Ratio-scale Measurements // SSRN preprint. URL: [ssrn.com/abstract=4865121](https://ssrn.com/abstract=4865121) (дата обращения: 25.02.2025).

8. Conneau A., Khandelwal K., Goyal N., Chaudhary V., Wenzek G., Guzmán F., Grave E., Ott M., Zettlemoyer L., Stoyanov V. Unsupervised Cross-lingual Representation Learning at Scale // arXiv preprint arXiv:1911.02116. 2019. URL: [arxiv.org/abs/1911.02116](https://arxiv.org/abs/1911.02116) (дата обращения: 25.02.2025).

9. Sanh V., Debut L., Chaumond J., Wolf T., et al. DistilBERT: A distilled version of BERT // arXiv preprint arXiv: 1910.01108. 2019. URL: [arxiv.org/abs/1910.01108](https://arxiv.org/abs/1910.01108) (дата обращения: 25.02.2025).

10. Clark K., Luong M., Le Q., Manning C. ELECTRA: Pre-training Text Encoders as Discriminators Rather Than Generators // arXiv preprint arXiv: 2003.10555. 2020. URL: [arxiv.org/abs/2003.10555](https://arxiv.org/abs/2003.10555) (дата обращения: 25.02.2025).

11. Hugging Face. URL: <https://huggingface.co> (дата обращения: 25.02.2025).

12. Gao L., Zhang U., Han J., Callan J. Scaling Deep Contrastive Learning Batch Size under Memory Limited Setup // arXiv preprint arXiv: 2101.06983. 2020. URL: [arxiv.org/pdf/2101.06983](https://arxiv.org/pdf/2101.06983) (дата обращения: 25.02.2025).

## References

1. «Lenta» testiruet tekhnologii videoraspoznavaniya tovarov i kolichestva pokupatelej v magazine [The Lenta company is testing technologies for video recognition of goods and the number of customers in the store]. URL: [lenta.com/o-kompanii/news/-----30](https://lenta.com/o-kompanii/news/-----30). (date accessed: 25.02.2025).

2. «Magnit» do konca goda masshtabiruet tekhnologiyu raspoznavaniya tovarov na polke na vsyu set [The Magnit company will scale the technology of recognizing goods on the store shelf to the entire network by the end of the year].



URL: [dsmedia.pro/news/magnit-do-konca-goda-masshtabiruet-tehnologiju-raspoznaniya-tovarov-na-polke-na-vsju-set](https://dsmedia.pro/news/magnit-do-konca-goda-masshtabiruet-tehnologiju-raspoznaniya-tovarov-na-polke-na-vsju-set). (date accessed: 25.02.2025).

3. Christian H., Agus M. P., Suhartono D. ComTech: Computer, Mathematics and Engineering Applications, 2016, T. 7, №4. URL: [journal.binus.ac.id/index.php/comtech/article/view/3746](http://journal.binus.ac.id/index.php/comtech/article/view/3746) (date accessed: 25.02.2025).

4. Qaiser S., Ali R. International journal of computer applications. 2018. T. 181. №1.

5. Aksenov K.A. Sun' L. Inzhenernyj vestnik Dona, 2024, №7. URL: [ivdon.ru/ru/magazine/archive/n7y2024/9371](http://ivdon.ru/ru/magazine/archive/n7y2024/9371) (date accessed 25.02.2025).

6. Zenkov V.A., Panishchev V.S., Zhelanov A.L., Volkov D.A. Inzhenernyj vestnik Dona, 2024, №8. URL: [ivdon.ru/ru/magazine/archive/n8y2024/9444](http://ivdon.ru/ru/magazine/archive/n8y2024/9444) (date accessed: 25.02.2025).

7. Krasnodebska K., Goch W., Uhl J.H., Verstegen J.A., Pesaresi M. SSRN preprint. URL: [ssrn.com/abstract=4865121](https://ssrn.com/abstract=4865121) (date accessed 25.02.2025).

8. Conneau A., Khandelwal K., Goyal N., Chaudhary V., Wenzek G., Guzmán F., Grave E., Ott M., Zettlemoyer L., Stoyanov V. arXiv preprint arXiv:1911.02116. 2019. URL: [arxiv.org/abs/1911.02116](https://arxiv.org/abs/1911.02116) (date accessed: 25.02.2025).

9. Sanh V., Debut L., Chaumond J., Wolf T., et al. arXiv preprint arXiv:1910.01108. 2019. URL: [arxiv.org/abs/1910.01108](https://arxiv.org/abs/1910.01108) (date accessed: 25.02.2025).

10. Clark K., Luong M., Le Q., Manning C. arXiv preprint arXiv:2003.10555. 2020. URL: [arxiv.org/abs/2003.10555](https://arxiv.org/abs/2003.10555) (date accessed: 25.02.2025).

11. Hugging Face. URL: [huggingface.co](https://huggingface.co) (date accessed: 25.02.2025).

12. Gao L., Zhang U., Han J., Callan J. arXiv preprint arXiv: 2101.06983. 2020. URL: [arxiv.org/pdf/2101.06983](https://arxiv.org/pdf/2101.06983) (date accessed: 25.02.2025).

**Дата поступления: 4.12.2024**

**Дата публикации: 10.03.2025**

---