

Повышение полноты и точности поиска товаров на торговых интернет-площадках

Ф. В. Краснов^{1*}

¹ООО «ВАЙЛДБЕРРИЗ СК» на базе Инновационного центра «Сколково», Москва, Россия

*krasnov.fedor2@wb.ru

Аннотация. Экономический эффект торговых интернет-площадок напрямую зависит от того, как быстро будут находить покупатели нужные товары. Низкая скорость поиска приводит к отказам от покупок, а отсутствие нужного товара на первой странице с результатами поиска снижает количество добавлений товаров в корзину. Время отклика системы поиска и наличие нужного товара на первых позициях поисковой выдачи характеризуются показателями полноты и точности системы поиска. Проблема низкого показателя полноты поиска, то есть всех подходящих под поисковый запрос товаров, обостряется в связи с ростом нагрузки и количества предлагаемых магазином товаров. Множество исследований направлено на поиск возможностей улучшения показателей полноты и точности путем более эффективного использования аппаратной части системы поиска и разработки моделей машинного обучения с новыми «слоями» искусственных нейронных сетей, обучающимися данными и функциями потерь. В настоящей работе предложен подход, который для увеличения полноты поиска использует знания о товарных категориях. Универсальные методы приближенного поиска используют искусственные кластеры товаров для сокращения времени обработки поискового запроса. При этом полнота найденных товаров-кандидатов ухудшается на 10–20%. Такой подход оправдан, когда данные не имеют естественной структуры. В случае торговой интернет-площадки данные хорошо известны, ими являются товары и их модальности – название, описание, товарная категория, изображения товара, что дает возможность использовать эти знания о данных для улучшения приближенного поиска. Экспериментальные оценки результатов применения данного подхода выявили улучшения показателей полноты и точности извлеченной информации о товарах более чем на 50% без снижения скорости обработки поискового запроса.

Ключевые слова: информационный поиск, ANN, IR, приближенный поиск ближайших соседей, Sentence-BERT, Dual Encoder

Для цитирования: Краснов Ф. В. Повышение полноты и точности поиска товаров на торговых интернет-площадках // Прикладная информатика. 2024. Т. 19. № 2. С. 118–136. DOI: 10.37791/2687-0649-2024-19-2-118-136

Research into improving the recall and precision of product searches on e-commerce platforms

F. Krasnov¹

¹Research Center of WILDBERRIES SK LLC based on the Skolkovo Innovation Center, Moscow, Russia
^{*}krasnov.fedor2@wb.ru

Abstract. The speed, precision and recall of information search in e-commerce are critical indicators for business success. A large number of academic studies are aimed at increasing these indicators through more efficient utilization of hardware and the development of machine learning models with new “layers”, training data and loss functions. In this study, the author focused on the practical task of speeding up the search by using knowledge about the nature of the load and data. Widely used Approximate Nearest Neighbors methods use artificial clusters to reduce the processing time of a search query. At the same time, the recall of the list of candidates found worsens. This approach is justified by its universality in relation to data. But in the case of an electronic online trading platform, the data are products and their modalities – name, description, product class, images, which makes it possible to use this knowledge about data to create more effective search structures and algorithms. In conditions of high dynamics of changes in product data, it is also necessary to take into account the speed, accuracy and completeness for offline and online processes. Therefore, the author considered the task of forming the completeness and accuracy of search results within the framework of an end-to-end process, and not only as an retrieval phase. As a result, the author received an improvement in the recall and precision of the retrieved product information by more than 50% without reducing the speed of search query processing.

Keywords: Information Retrieval, ANN, IR, Approximate Nearest Neighbors, Sentence-BERT, Dual Encoder

For citation: Krasnov F. Research into improving the recall and precision of product searches on e-commerce platforms. *Prikladnaya informatika*—Journal of Applied Informatics, 2024, vol.19, no.2, pp.118-136 (in Russian). DOI: 10.37791/2687-0649-2024-19-2-118-136

Введение

Внешняя нагрузка на торговых интернет-площадках постоянно растет: на сегодняшний день она измеряется в сотнях тысяч запросов в секунду при одновременном росте количества товаров в каталоге до миллиардов. Однако при увеличении интенсивности запросов на поиск нужных товаров на торговой интернет-площадке качество поиска снижается.

С другой стороны, с приходом архитектур, которые основаны на искусственных

нейронных сетях с глубоким обучением, существенно изменился состав операций, выполняемых в рамках информационного поиска. Одновременно с этим трансформировалась и аппаратная часть систем поиска. В настоящей работе на основе исследований возможных подходов к улучшению показателей качества поиска поставлена задача совершенствования алгоритмов поиска путем добавления возможности приближенного поиска товарных категорий.