

Научная статья
УДК 339.133.4:338.001.36+332.15

Выбор наиболее приоритетного региона для продвижения товаров или услуг в поисковых системах при помощи скоринга семантики с применением статистических методов

Александр Александрович Любаков

Самарский государственный экономический университет, Самара, Россия, alex@lyubakov.ru

Аннотация. В статье построена скоринговая модель ранжирования поисковых запросов. Для проведения скоринга вводятся коэффициенты популярности, коммерческой и конкурентности запросов. Популярность оценивается на основании базовой и фразовой частотности за последние 24 месяца. Коммерческая, или коммерциализация, запроса оценивается на основании прогнозируемого бюджета в «Яндекс.Директ» и стоимости за клик. Конкурентность оценивается на основании количества документов по запросу, количества главных страниц, участвующих в ранжировании по запросу, количества тегов Title, содержащих ключевое слово запроса, количества слов в запросе. При помощи скоринга SEO-оптимизаторы и предприниматели смогут определить те направления бизнеса и/или регионы, которые могут обеспечить максимальный экономический результат.

Ключевые слова: семантическое ядро, SEO-продвижение, интернет-маркетинг, поисковый маркетинг, статистические методы в маркетинге

Основные положения:

- ♦ основная задача настоящего исследования состоит в определении тех запросов, которые имеют наибольший экономический потенциал с точки зрения поискового маркетинга: обладают высокой популярностью и коммерческой одновременно со сравнительно низкой конкурентностью;
- ♦ представлена аналитическая модель скоринга семантического ядра с применением статистических методов для нормализации разнородных массивов данных.

Для цитирования: Любаков А.А. Выбор наиболее приоритетного региона для продвижения товаров или услуг в поисковых системах при помощи скоринга семантики с применением статистических методов // Вестник Самарского государственного экономического университета. 2024. № 10 (240). С. 54–58.

Original article

Selecting the most prioritized region for promoting products or services in search engines through semantic scoring using statistical methods

Alexander A. Lyubakov

Samara State University of Economics, Samara, Russia, alex@lyubakov.ru

Abstract. The article proposes a scoring model for ranking search queries. To conduct the scoring, coefficients for query popularity, commercial intent, and competitiveness are introduced. Popularity is assessed based on basic and phrase frequency over the past 24 months. Commercial intent or commercialization of the query is evaluated based on the projected budget in "Yandex.Direct" and the cost per click. Competitive-

© Любаков А.А., 2024

ness is assessed based on the number of documents for the query, the number of main pages involved in ranking for the query, the number of Title tags containing the query's keyword, and the number of words in the query. Through this scoring method, SEO optimizers and entrepreneurs can identify business areas and/or regions that can ensure maximum economic results.

Keywords: semantic core, SEO, Internet marketing, search engine marketing, statistical methods in marketing

Highlights:

◆ the primary objective of this study is to identify queries that have the highest economic potential from the perspective of search engine marketing: those that exhibit high popularity and commercial intent while maintaining relatively low competitiveness;

◆ as a result, the article presents an analytical scoring model of the semantic core, employing statistical methods to normalize heterogeneous data sets.

For citation: Lyubakov A.A. Selecting the most prioritized region for promoting products or services in search engines through semantic scoring using statistical methods // Vestnik of Samara State University of Economics. 2024. No. 10 (240). Pp. 54–58. (In Russ.).

Введение

Search Engine Optimization (SEO) – один из основных инструментов, которыми пользуются современные маркетологи, зачастую он играет ключевую роль в успехе бизнеса онлайн. Цель SEO заключается в том, чтобы сделать сайт более привлекательным для поисковых систем (например, Google и «Яндекс»), что в конечном счете приводит к улучшению ранжирования портала в поисковой выдаче по ключевым запросам.

В том случае, если бизнес ведет операционную деятельность в нескольких регионах или рассматривает для себя новые субъекты в будущем, встает проблема приоритизации работ. Нюанс заключается в том, что поисковое продвижение – инертный и долгий процесс, до первых результатов проходит приблизительно от 6 до 12 месяцев, поэтому ошибка в выборе приоритетного региона может быть критической.

Учитывая, что семантическое ядро (полный перечень ключевых запросов) в зависимости от типа и размера бизнеса может исчисляться от нескольких сотен до нескольких тысяч ключевых фраз на каждый регион, ручная приоритизация такого массива данных не только затруднительна, но и малоэффективна. Целесообразным будет использование скоринга семантического ядра, основанного на статистических методах.

Для ранжирования ключевых запросов, кластеров ключевых запросов и, в дальней-

шем, регионов следует опираться на объективные данные, такие как популярность поисковых запросов, структура выдачи и уровень конкуренции.

Методы

Популярность. Д.С. Яковлев в своей работе предлагает разделить запросы по частотности на высокочастотные, среднечастотные и низкочастотные [1]. Под популярными предлагаю понимать те регионы/кластеры/запросы, в которых общая частотность выше. Фокусировка на высокочастотных запросах позволяет привлекать большое количество посетителей на сайт, увеличивая трафик и потенциальные возможности для конверсии. Однако следует учитывать, что высокая популярность запросов часто сопряжена с высокой конкуренцией

Структура выдачи. Для оценки структуры выдачи предлагаю использовать термин «коммерческость» («коммерциализация»). Коммерческие запросы указывают на намерение пользователя совершить покупку или воспользоваться услугой. Такие запросы могут включать слова, явно имеющие коммерческую направленность, например, "купить", "цена", "заказать" и т.д. Такие запросы с более высокой вероятностью приведут к конверсии, с одной стороны. С другой стороны, такие запросы характеризуются полностью коммерческой выдачей, т.е. по результатам поиска не будут представлены информационные сайты. Часть

запросов характеризуется смешанной выдачей, например запрос «тимбилдинг». Часть ресурсов, представленных в выдаче в топ-10, будут коммерческими, часть – информационными. Таким образом, и вакантных мест не 10, а меньше.

Конкуренция. Низкоконкурентные запросы можно оценить по количеству документов, проиндексированных поисковыми системами и отнесенных ими к релевантным по текущему запросу. Фокусирование усилий на таких запросах позволяет быстрее достигать результатов в поисковом маркетинге. Так, регион с относительно небольшим количеством локальных документов может оказаться привлекательным для запуска SEO именно по нему.

Каждый из введенных параметров скоринга следует оценить.

Оценку популярности стоит делать при помощи анализа частотности запросов в Яндексе. Для оценки возьмем базовую частоту (она учитывает как точные совпадения запросов, так и близкие по смыслу фразы) и фразовую (она целиком фиксирует искомое словосочетание) [2]. Яндекс предоставляет статистику за 2 года. Накопленный массив данных со значениями базовой и фразовой частот с выделением медианы за каждый месяц стоит нормализовать. Это можно сделать при помощи логарифмирования, такой шаг позволит уменьшить разброс данных и сделает их более сопоставимыми.

Формула для расчета коэффициента популярности Pop будет выглядеть так:

$$\text{Pop} = \text{LN}24 \times k_{\text{LN}24} + \text{LNPhrase} \times k_{\text{LNPhrase}} + \text{LNFull} \times k_{\text{LNFull}}.$$

Здесь LN24 обозначает логарифмированное значение базовой частоты за 24 месяца.

LNPhrase обозначает логарифмированное значение фразовой частоты.

LNFull представляет собой логарифмированное значение «полноты», иллюстрирующее вхождение фразовой частоты в базовую.

Коэффициенты значимости k (в диапазоне от 0,8 до 3) позволяют изменить влияние каждого фактора на итоговый коэффициент популярности запроса. Данные коэффициенты устанавливаются на основании экспертной оценки таким образом, чтобы сохранилась

корреляция каждого слагаемого с итоговым значением коэффициента популярности.

Представленная формула позволяет объединить разноразмерные данные о частотности запросов в единый показатель популярности, который можно использовать для ранжирования запросов по их популярности для приоритизации работ по SEO-продвижению.

Оценка коммерческой позволит понять, является ли выдача коммерческой, смешанной или информационной. Таким образом, можно оценить, насколько контент сайта соответствует ожиданиям поисковой системы, реально продвижение по тому или иному запросу в том или ином регионе. Выдачу считают информационной, если доля коммерческих результатов составляет от 0% до 35%, смешанной – при доле от 35% до 60% и коммерческой, когда эта доля превышает 60% [3].

Для скоринга семантики с точки зрения коммерческой стоит оценивать такие параметры, как бюджет на рекламные кампании и стоимость за клик (CPC), которые можно найти в сервисе «Яндекс.Директ».

Высокая стоимость за клик и большой бюджет могут указывать на то, что запрос обладает высоким коммерческим потенциалом и конкуренцией среди рекламодателей.

Формула для расчета коэффициента коммерциализации Ecom выглядит следующим образом:

$$\text{Ecom} = \text{LNBudget} \times k_{\text{LNBudget}} + \text{LNCPC} \times k_{\text{LNCPC}}.$$

Здесь LNBudget обозначает логарифмированное значение прогнозируемого бюджета в Яндекс.Директе.

LNCPC отражает логарифмированное значение стоимости за клик по данным Яндекс.Директа.

Коэффициенты значимости k (в диапазоне от 0,8 до 3) позволяют изменить влияние каждого фактора на итоговый коэффициент коммерциализации запроса. Данные коэффициенты устанавливаются на основании экспертной оценки таким образом, чтобы сохранилась корреляция каждого слагаемого с итоговым значением коэффициента коммерциализации.

Оценка конкурентности запросов возможна при анализе оптимизированности документов в выдаче. Оптимизированность вклю-

чает в себя такие показатели, как количество документов по запросу, количество главных страниц, оптимизация заголовков и количество слов в запросе.

Количество документов по запросу. Высокое количество документов означает, что запрос является достаточно явным, выдача по нему насыщена, а конкуренция по запросу высока.

Количество главных страниц. Главные страницы обладают высоким (более высоким, чем внутренние) авторитетом и доверием со стороны поисковых систем [4]. Соответственно, чем больше сайтов используют именно главную страницу для продвижения по тому или иному запросу, тем выше конкуренция по нему.

Оптимизация заголовков. Количество тегов Title (одного из наиболее влиятельных тегов с точки зрения текстовой оптимизации в SEO), содержащих ключевое слово запроса является показателем высокого внимания со стороны SEO-специалистов.

Количество слов в запросе. Короткие запросы (одно-двухсловники), как правило, лежат на поверхности и характеризуются высокой конкуренцией. Например, запрос «отдых в Карелии из Самары на теплоходе в июле» будет менее конкурентным, чем «туры в Карелию».

Формула для расчета коэффициента конкурентности $Comp$ выглядит следующим образом:

$$Comp = LNDocs \times k_{LNDocs} + Main \times k_{Main} + Title \times k_{Title} + LNWordsCount \times k_{LNWordsCount},$$

где $LNDocs$ – прологарифмированное количество документов по запросу;

$Main$ – балльная оценка количества главных страниц, участвующих в ранжировании по запросу;

$Title$ – количество тегов Title, содержащих ключевое слово запроса;

$LNWordsCount$ – логарифмированное количество слов в запросе;

K – коэффициенты значимости (в диапазоне от 0,8 до 3), которые позволяют изменить влияние каждого фактора на итоговый коэффициент конкурентности запроса. Данные коэффициенты устанавливаются на основании экспертной оценки таким образом, чтобы сохранилась корреляция

каждого слагаемого с итоговым значением коэффициента конкурентности.

Результаты

Заключительный этап скоринга – объединение ранее рассчитанных показателей в одну формулу для оценки общего приоритета запросов. Итоговая формула оценки ранга (Rank) запросов выглядит следующим образом:

$$Rank = Pop \times k_{Pop} + ECom \times k_{ECom} - Comp \times k_{Comp},$$

где Pop – коэффициент популярности запроса;

$ECom$ – коэффициент коммерциализации запроса;

$Comp$ – коэффициент конкурентности запроса;

k_{Pop} , k_{ECom} , k_{Comp} – коэффициенты значимости для каждого из показателей. Данные коэффициенты устанавливаются на основании экспертной оценки таким образом, чтобы сохранилась корреляция каждого слагаемого с итоговым значением коэффициента конкурентности.

Отдельно стоит отметить тот факт, что показатель конкурентности указывается в формуле с отрицательным знаком. Это логично, так как более высокая конкурентность запроса снижает его приоритет, особенно если целью является выход в новый регион или выбор наиболее значимого из существующих с точки зрения маркетинговых усилий.

Обсуждение

В результате исследования была получена аналитическая модель скоринга семантического ядра при помощи оценки ранга Rank. При выборе приоритетного направления бизнеса и/или региона с точки зрения поискового маркетинга семантика может достигать более нескольких тысяч запросов. Ключевое преимущество этой модели заключается в предоставленной возможности быстро обрабатывать массивные объемы данных. Модель позволяет определить те регионы и те кластеры запросов в них, которые обладают наименьшей сложностью продвижения и максимальным потенциальным трафиком, которые они могут дать. Ограничение текущей модели заключается в том, что она позволяет расставить приоритеты в собранном семантическом ядре, но

не позволяет построить прогноз по количеству приведенного трафика на горизонте 12–24 месяцев. Соответственно, будущие исследования данной темы должны концентрироваться вокруг прогнозирования таких параметров, как возможная позиция по запросу / кластеру запросов / региону; кликабельность данной позиции; трафик, который может быть приведен по данному запросу в разный момент времени; итоговая отдача от поискового продвижения по запросу / кластеру запросов / региону.

Заключение

Предложенная аналитическая модель скоринга семантики включает в себя важнейшие аспекты поискового продвижения – популярность, коммерческую и конкурентность запросов, чтобы обеспечить SEO-специалисту мощный инструмент для приоритизации усилий. При помощи скоринга SEO-оптимизаторы и предприниматели смогут определить те направления бизнеса и/или регионы, которые могут обеспечить максимальный экономический результат.

Список источников

1. Яковлев Д.С. SEO-оптимизированный контент веб-сайта и использование сервиса Яндекс.Вордстат // Вестник магистратуры. 2015. № 10 (49). С. 19–22.
2. Yandex.Wordstat – Key Collector. URL: <https://www.key-collector.ru/docs/data-collection/yandex-wordstat/> (дата обращения: 19.03.2024).
3. Проверка коммерциализации запросов в Яндекс и Google, определить коммерческую запроса, локализацию и геодепанденцию. URL: <https://arsenkin.ru/tools/commerce/> (дата обращения: 19.03.2024).
4. Неелова Н. SEMbook. Энциклопедия поискового продвижения Ingate. 2-е изд. ИП Андросов, 2017. 541 с.

References

1. Yakovlev D.S. SEO-optimized website content and the use of the Yandex service.Wordstat // Bulletin of the Magistracy. 2015. No. 10 (49). Pp. 19–22.
2. Yandex.Wordstat – Key Collector. URL: <https://www.key-collector.ru/docs/data-collection/yandex-wordstat/> (date of access: 19.03.2024).
3. Checking the commercialization of requests in Yandex and Google, determine the commerciality of the request, localization and geo-dependence. URL: <https://arsenkin.ru/tools/commerce/> (date of access: 19.03.2024).
4. Neelova N. SEMbook. The encyclopedia of Ingate search engine promotion. 2nd ed. IP Androsov, 2017. 541 p.

Информация об авторе

А.А. Любаков – аспирант Самарского государственного экономического университета.

Information about the author

A.A. Lyubakov – graduate student at Samara State University of Economics.

Статья поступила в редакцию 14.06.2024; одобрена после рецензирования 21.06.2024; принята к публикации 28.06.2024.

The article was submitted 14.06.2024; approved after reviewing 21.06.2024; accepted for publication 28.06.2024.