

ПРИМЕНЕНИЕ БОЛЬШИХ ДАННЫХ И МАШИННОГО ОБУЧЕНИЯ В РИТЕЙЛЕ (НА ПРИМЕРЕ ОНЛАЙН-МАГАЗИНА OZON)

В. Д. Ермак,

студент экономического факультета

БГУ, г. Минск,

vladislav.yermak1@gmail.com

А. М. Самаль,

студентка экономического факультета

БГУ, г. Минск,

alnsml2101@gmail.com

Аннотация. В данной статье кратко рассмотрены результаты исследования, которое проводилось с целью лучшего понимания возможностей применения «больших данных» в ритейле. Исследование включало изучение понятия «больших данных» и технологий, применяемых для их анализа и прогнозирования, а также знакомство с практикой успешного применения этих технологий интернет-магазином OZON.

Ключевые слова. «Большие данные»; машинное обучение; прогнозная аналитика, предсказательные модели; интернет-магазин OZON.

APPLICATION OF BIG DATA AND MACHINE LEARNING IN RETAIL (ON THE EXAMPLE OF OZON ONLINE STORE)

Abstract. This article reviews the results of research that was conducted with the aim of better understanding the possibilities of using "big data" in retail. The scientific article includes the study of the concept of "big data" and the technologies used for the analysis and forecasting, as well as an acquaintance with the practice of successful application of these technologies by the OZON online store.

Kew words. «Big Data»; machine learning; predictive analytics; predictive models; OZON online store.

С ростом популярности сети интернет, пользователями которой в настоящее время являются 63,2 % мирового населения появлением крупных интернет-сервисов значительно возрос объем генерируемых в сети данных [1]. К ним относятся, например, размещаемый пользователями контент и данные об их действиях в сети. Аналитика и прогнозирование на основе «больших данных» дают бизнесу информацию для принятия более эффективных управленческих решений. Например, результаты прогнозной аналитики «больших данных» позволяют компаниям, работающим в ритейле, лучше понимать своих покупателей, их запросы и на основании этой информации эффективнее строить рекламные кампании.

С целью лучшего понимания возможностей применения «больших данных» в ритейле нами проведено исследование по изучению понятия «больших данных» и технологий, применяемых для их анализа и прогнозирования, а также знакомству с практикой успешного применения этих технологий интернет-магазином OZON.

Под большими данными принято понимать совокупность технологий обработки больших объемов (сотни терабайтов и более) данных [2;3].

Наряду с анализом и прогнозированием, к задачам, которые решаются применительно к «большим данным», также относят накопление, верификацию, хранение, обеспечение информационной безопасности, поиск и визуализацию.

Процесс прогнозирования на основе «больших данных» включает такие этапы, как:

1) очистка данных (data cleaning), которая включает поиск и исправление ошибок в первичном наборе данных, например, ошибки ручного ввода (опечатки), некорректные значения с измерительных приборов, замена пропусков в данных;

2) генерация предикторов (feature engineering), которая состоит в отборе переменных из числа представленных в данных и/или генерации новых переменных для включения их модель;

3) моделирование, в результате чего создаются модели для предсказания значений целевой переменной на основе значений предикторов;

4) оценка качества построенных моделей и выбор лучшей модели;

5) применение лучшей модели для прогнозирования.

На основании отчета к технологиям анализа больших данных относят следующие.

- Краудсорсинг — категоризация и обогащение данных силами большого количества людей.
- Смешение и интеграция данных — интегрирование данных разного рода из разнообразных источников для возможности приведения данных к одному виду, уточнения и дополнения данных.
- Машинное обучение, включающее обучение с учителем и без учителя, — построение прогнозных моделей на основе собранных данных.
- Имитационное моделирование — создание моделей для помощи проведения эксперимента виртуально, без влияния на объекты реального мира.
- Статистический анализ — проверка статистических гипотез.
- Data Mining — извлечение новой и практически полезной информации из данных.
- Визуализация — представление «больших данных» и результатов их анализа в виде удобных и наглядных диаграмм, понятных человеку.
- Другие[5].

Интернет-магазин OZON широко применяет в своей деятельности технологии машинного обучения, сутью которой является поиск закономерностей в массиве данных и представление найденных закономерностей в виде прогнозной или предсказательной модели.

Предсказательные модели — новый метод использования больших данных в онлайн-рекламе. Благодаря предсказательным моделям маркетологи пытаются спрогнозировать поведение потребителя, или пользователя, задолго до того, как этот пользователь захочет купить определенный товар. В связи с этим «большие данные» — это основа, на которой держится маркетинг компании OZON. Эта технология позволяет понять, какую аудиторию необходимо таргетировать и с какой целью. Основным фокусом в отделах онлайн-маркетинга и ценообразования является принятие решений на основе «больших данных».

Компания также использует «большие данные» для создания предсказательных моделей онлайн-рекламы, что позволяет дифференцировать ставки компании в зависимости от ожидаемого эффекта какой-то определенной аудитории.

У OZON существует метод разделения базы клиентов на определенные сегменты. Сущность этого метода заключается в том, что компания собирает объясняющие переменные про покупателей: какой товар покупал, как давно покупал, через какие каналы приходил, в каком городе/стране находится пользователь и другие. Суммарное количество таких переменных составляет около трёхсот. Затем модель машинного обучения, которая обучается на исторических данных, по этим трёмстам показателям предсказывает вероятность будущей покупки по каждому из посетителей сайта OZON, который посещал сайт за последние три месяца. Далее можно всех посетителей отсортировать и распределить на специальные сегменты.

Эти сегменты загружаются в системы рекламы — Google, «Яндекс», другие площадки и далее работа осуществляется по модели cost per click (CPC). Конверсия в сегментах существенным образом отличается во всех рекламных площадках.

В политике ценообразования OZON также используют «большие данные». Основной задачей компании является осведомление о ценах конкурентов на большую часть их ассортимента. Каждый день получается информация о четырёх сотнях тысяч товаров, которые продает компания, по более чем сотне конкурентов. Это большой массив информации, который складывается в базах данных и доступен для того, чтобы им пользоваться [6].

Список использованной литературы

1. World internet usage and population statistics 2020 Year [Electronic resource] // Internet World Stats. — Mode of access: <https://www.internetworldstats.com/stats.htm>. — Date of access: 10.02.2021.

2. Абрамов, Р. Что такое «Big Data»? [Электронный ресурс] / Р. Абрамов // Хабр: Блог компании ProductStar, Big Data. – 2020. – Режим доступа: <https://habr.com/ru/company/productstar/blog/503580/>. – Дата доступа: 10.02.2021.
3. Big data [Electronic resource] // Wikipedia, Big data. – 2010. – Mode of access: https://en.wikipedia.org/wiki/Big_data. – Date of access: 10.02.2021.
4. Manyika J., Chui M. Big data: The next frontier for innovation, competition, and productivity [Electronic resource] / J. Manyika, M. Chui // McKinsey Digital, Our Insights. – 2011. – Mode of access: <https://www.mckinsey.com/business-functions/mckinsey-digital/our-insights/big-data-the-next-frontier-for-innovation#>. – Date of access: 14.02.2021.
5. Машинное обучение [Электронный ресурс] // Сквозная аналитика CallTouch. – Режим доступа: <https://www.calltouch.ru/glossary/mashinnoe-obuchenie/>. – Дата доступа: 14.02.2021.
6. Леонова, А. Большие данные и машинное обучение в ритейле: интервью с директором по онлайн-маркетингу OZON [Электронный ресурс] / А. Леонова // Информационный портал Cossa. – 2017. – Режим доступа: <https://www.cossa.ru/211/162081/>. – Дата доступа: 18.02.2021.

ТЕНДЕНЦИИ И ПЕРСПЕКТИВЫ ТРАНСФОРМАЦИИ БАНКОВСКОГО СЕКТОРА В УСЛОВИЯХ ЦИФРОВИЗАЦИИ

О. Ю. Жуковская,

канд. экон. наук, доцент, доцент кафедры
инноватики и предпринимательской деятельности
Белорусского государственного университета, г. Минск

Ян Цзе,

магистрант кафедры
инноватики и предпринимательской деятельности
Белорусского государственного университета, г. Минск

Аннотация. В данной статье рассматриваются некоторые аспекты трансформации банковского сектора под влиянием развития цифровых технологий. Исследуются потенциальные возможности расширения и преобразования банковских услуг, предложения новых банковских продуктов, а также совершенствования взаимоотношений с клиентами в современных условиях.

Ключевые слова. Цифровые технологии, информационные услуги, банковский сектор, бизнес-процессы, ИТ-инфраструктура.

THE TRENDS AND PROSPECTS OF THE BANKING DEVELOPMENT UNDER THE CONDITIONS OF DIGITALIZATION