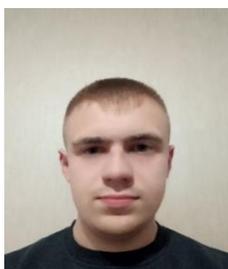


УДК [611.018.51+615.47]:612.086.2

АНАЛИТИКИ BIG DATA. ТИПЫ АНАЛИТИКИ



И.В. Шилов
Студент БГУИР



С.Н. Нестеренков
Кандидат
технических наук,
доцент, декан
факультета
компьютерных
систем и сетей



А.Н. Марков
Старший
преподаватель,
магистр технических
наук, заместитель
начальника Центра
информатизации и
инновационных
разработок БГУИР



В.Т. Кучеренко
Студент БГУИР

Белорусский государственный университет информатики и радиоэлектроники
E-mail: prettuonex@gmail.com, s.nesterenkov@bsuir.by, a.n.markov@bsuir.by
vova.kucherenko.00@mail.ru

И.В. Шилов

Студент 4 курса ФКСиС специальности «Программное обеспечение информационных технологий» Белорусского государственного университета информатики и радиоэлектроники.

С.Н. Нестеренков

Кандидат технических наук, доцент, декан факультета компьютерных систем и сетей Белорусского государственного университета информатики и радиоэлектроники, доцент кафедры программного обеспечения информационных технологий. Автор публикаций на тему машинного обучения, алгоритмов принятия решений, искусственных нейронных сетей и автоматизации.

А.Н. Марков

Магистр технических наук, старший преподаватель кафедры ПИКС, заместитель начальника Центра информатизации и инновационных разработок Белорусского государственного университета информатики и радиоэлектроники.

В.Т. Кучеренко

Студент 4 курса ФКСиС специальности «Программное обеспечение информационных технологий» Белорусского государственного университета информатики и радиоэлектроники

Аннотация. В настоящее время предприятия и компании начинают осознавать важность наличия больших объемов данных для принятия правильных решений и поддержки своих стратегических планов. С развитием Интернета, новых технологий и социальных сетей производство цифровых данных постоянно растет. Понятие «Big Data» или «Большие данные» относится к разнородной массе цифровых данных, создаваемых компаниями и частными лицами, характеристики которых (большой объем, скорость обработки) требуют специальных и все более сложных компьютерных средств хранения и анализа. Цель данной работы заключается в определении типов аналитики, а также важности анализа больших данных.

Ключевые слова: Big Data, аналитика больших данных, Интернет, безопасность

Введение.

Аналитика Big Data – процесс изучения больших данных для выявления информации, такой как скрытие закономерности, корреляция, тенденции рынка и предпочтения

клиентов, которая может помочь организациям принимать обоснованные бизнес-решения. В широком масштабе технологии и методы анализа данных дают организациям возможность анализировать наборы показателей и собирать новую информацию, в то время как бизнес-аналитики отвечают за основные вопросы о бизнес-операциях и производительности. Аналитика больших данных часто относится к форме продвинутой аналитики, которая включает в себя сложные приложения с такими элементами, как прогнозные модели, статистические алгоритмы и анализ «что, если», основанный на аналитических системах [1-2].

Типы аналитики Big Data.

Существует 4 основных типа аналитики Big Data: диагностическая, описательная, предписывающая и прогностическая. Все они используют различные инструменты для таких процессов, как интеллектуальная обработка данных, интеграция, визуализация, чтобы улучшить процесс анализа информации и обеспечить получение компанией выгоды от собираемых данных.

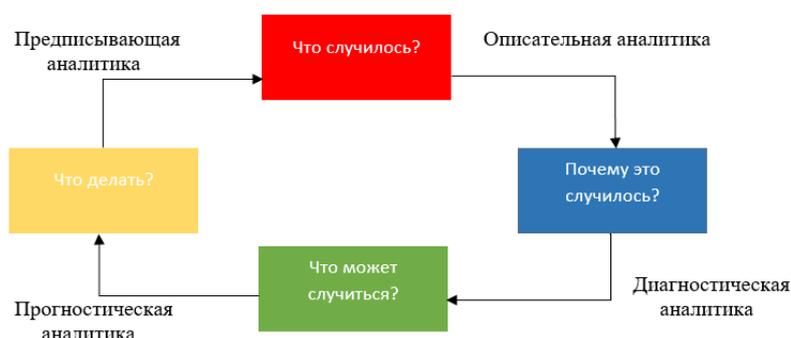


Рисунок 1 – Основные вопросы, рассматриваемые каждым типом аналитики

Описательная аналитика – самая распространенная форма аналитики, которую используют компании, чтобы быть в курсе текущих тенденций и результатов своей деятельности. Это один из первых шагов анализа необработанных данных путем выполнения математических операций и создания отчетов о различных измерениях. После определения тенденций с помощью описательной аналитики можно использовать другие типы анализа, чтобы узнать больше о причинах этих самых тенденций [3].

Описательная аналитика предоставляет множество преимуществ. Однако одним из его главных достоинств является то, что данный тип помогает компаниям и организациям разобраться в больших объемах, поступающих и необработанных данных, которые они собирают, сосредоточив внимание на более важных областях. В настоящее время компании используют данный метод аналитики, чтобы лучше понять свою текущую бизнес-ситуацию и положение по сравнению с прошлым. Это важнейший шаг в анализе данных, и без него было бы невозможно предвидеть какие-либо будущие тенденции или принимать решения, основанные на текущих данных.

Следующим типом аналитики является диагностическая – один из наиболее продвинутых видов анализа Big Data, который можно использовать для исследования данных и показателей. С помощью такого рода аналитики полученные знания можно применить, чтобы ответить на вопрос: «Почему это случилось?». Таким образом, анализируя данные, можно понять причины определенного поведения и событий, связанных с компанией и ее клиентами, сотрудниками, продуктами и многим другим [4].

Преимущества диагностической аналитики включают лучшее понимание поступающих данных и различные способы поиска ответов на вопросы компании. Этот тип аналитики позволяет компаниям понимать своих клиентов, используя инструменты для

поиска, фильтрации и сравнения данных, полученных отдельными лицами.

Наиболее важным типом аналитики является прогностический. Как следует из названия, данный вид анализа показателей предназначен для прогнозирования будущих результатов на основе поступающей информации. Для получения наилучших результатов прогностическая аналитика использует множество сложных инструментов и моделей предсказания, таких как машинное обучение и статистическое моделирование. На сегодняшний день, прогностический подход анализа – один из наиболее широко используемых видов изучения данных. Согласно прогнозу, объем рынка прогностической аналитики достигнет 10,95 миллиардов долларов к 2024 году, увеличившись на 21% в течение шести лет [5].

Главное преимущество прогностической аналитики – надежный и точный прогноз будущего. Благодаря предсказаниям, сделанным с помощью такого рода анализа, компании могут находить способы экономии и заработка средств для управления графиками поставок и быть в курсе потребностей в имеющихся запасах. Использование прогностической аналитики в отделе маркетинга может помочь организациям привлечь новых и удержать старых клиентов. Также благодаря такой форме анализа компании могут предвидеть потребности клиента, анализируя данные и текущие рыночные тенденции [6].

Последним и самым сложным с точки зрения реализации типом аналитики является предписывающий. Данная форма анализа использует результаты описательного и прогностического подхода, а также находит решения для оптимизации бизнес-процессов с помощью различных методов и симуляций. Некоторые высокотехнологичные автомобильные концерны используют данный подход при проектировании своих самоуправляемых автомобилей, которые в свою очередь анализируют данные в режиме реального времени и принимают решения на основе предписывающей аналитики [7].

Определенная часть преимуществ предписывающей аналитики включает улучшение процессов, кампаний, стратегий, производства и обслуживания клиентов. Используя статистику и моделирование, данный тип аналитики помогает производителям лучше понимать рынок и прогнозировать его состояние в будущем. Это также способствует организациям в правильной расстановке приоритетов и понимании, какой курс действий может привести к большей финансовой выгоде [8].

Заключение.

Аналитика Big Data – это развивающаяся технология и процесс изучения информации с целью выявления скрытых закономерностей, тенденций рынка, предпочтений клиентов и других полезных данных для принятия правильных бизнес-стратегических решений. С каждым днем аналитика принимается все большим числом различных компаний и в то же время постепенно становится самостоятельным направлением, инструменты которой используются организациями в тех случаях, когда необходимо знать, что произойдет в будущем и преобразовывать данные прогностические показатели и метрики в универсальный язык.

Список литературы

- [1] Wang, Yandong; Goldstone, Robin; Yu, Weikuan; Wang, Teng (October 2014). "Characterization and Optimization of Memory-Resident MapReduce on HPC Systems". 2014 IEEE 28th International Parallel and Distributed Processing Symposium. IEEE. pp. 799–808.
- [2] Pllana, Sabri; Janciak, Ivan; Brezany, Peter; Wöhrer, Alexander (2016). "A Survey of the State of the Art in Data Mining and Integration Query Languages". 2011 14th International Conference on Network-Based Information Systems. 2011 International Conference on Network-Based Information Systems (NBIS 2011). IEEE Computer Society. P. 341–348.
- [3] Starodubtsev, I.E. Fractal dimension as a characteristic of biological cell AFM images // Computer Data Analysis and Modeling: Theoretical and Applied Stochastics: Proceedings of the Eleventh International Conference (September 6-10, 2016, Minsk). –Minsk, 2016. – P. 304-307.
- [4] Кукареко, А.В. Способы машинного обучения для выявления ошибок выполнения упражнений на smart-тренажере / А.В. Кукареко, С.Н. Нестеренков // BIG DATA and Advanced Analytics = BIG DATA и анализ

высокого уровня : сб. материалов VI Междунар. науч.-практ. конф. (Республика Беларусь, Минск, 20-21 мая 2020 года): в 3 ч. Ч. 2 / редкол. : В. А. Богуш [и др.]. - Минск : Бестпринт, 2020. - С. 214-224.

[5] Беляк, А. А. Анализ производительности технологии Hadoop / А. А. Беляк, С. Н. Нестеренков // BIG DATA and Advanced Analytics = BI DATA и анализ высокого уровня: сб. научных статей VII Междунар. науч.-практ. конф. (Республика Беларусь, Минск, 19-20 мая 2021 года): / редкол. : В. А. Богуш [и др.]. – Минск : Бестпринт, 2021. – С. 343–346.

[6] Щербаков, М. А. Проектирование масштабируемого графического пользовательского интерфейса на основе Redux / М. А. Щербаков, С. Н. Нестеренков // BIG DATA and Advanced Analytics = BI DATA и анализ высокого уровня: сб. научных статей VII Междунар. науч.-практ. конф. (Республика Беларусь, Минск, 19-20 мая 2021 года): / редкол. : В. А. Богуш [и др.]. – Минск : Бестпринт, 2021. – С. 347–349.

[7] Калоша, А.Л. Система анализа качества текстовых коллекций / А.Л. Калоша, М.А. Медунецкий, М.П. Хоронеко, А.А. Александров, А.И. Гридасов, С.Н. Нестеренков // BIG DATA and Advanced Analytics = BIG DATA и анализ высокого уровня : сб. материалов VI Междунар. науч.-практ. конф. (Республика Беларусь, Минск, 20-21 мая 2020 года): в 3 ч. Ч. 2 / редкол. : В. А. Богуш [и др.]. - Минск : Бестпринт, 2020. - С. 369-375.

[8] Dedic, N.; Stanier, C. (2017). "Towards Differentiating Business Intelligence, Big Data, Data Analytics and Knowledge Discovery". Innovations in Enterprise Information Systems Management and Engineering. Lecture Notes in Business Information Processing. Vol. 285. Berlin ; Heidelberg: Springer International Publishing. pp. 114–122.

BIG DATA ANALYTICS AND CONCEPTS. TYPES OF ANALYTICS

I.V.SHILOV
Student of BSUIR

S.N. NESTERENKOV
*PhD, Associate
Professor Dean of the
Faculty of Computer
Systems and Networks*

A.N. MARKOV
*Master of
Engineering and
Technology, Senior
Lecturer, Deputy
Head of the CIIR*

V.T. KUCHERENKO
Student of BSUIR

Republic of Belarus

Belarusian State University of Informatics and Radioelectronics

E-mail: prettuonex@gmail.com, s.nesterenkov@bsuir.by, a.n.markov@bsuir.by, vova.kucherenko.00@mail.ru

Abstract. Currently, enterprises and companies are beginning to realize the importance of having large amounts of data to make the right decisions and support their strategic plans. With the development of the Internet, new technologies and social networks, the production of digital data is constantly growing. The concept of "Big Data" or "Big Data" refers to a heterogeneous mass of digital data created by companies and individuals whose characteristics (large volume, processing speed) require special and increasingly complex computer storage and analysis tools. The purpose of this paper is to define Big Data, its concepts, problems, and the importance of big data analysis.

Key words: Big Data, Big Data Analytics, Internet, Security.