

УДК 004.021:004.75

ПРИМЕНЕНИЕ BIG DATA В СФЕРЕ ЭКОНОМИКИ



К.И. Котельников
Студент БГУИР
kki199918@gmail.com



В.Е. Буюков
Студент БГУИР
vitya.buiukov@mail.ru



А.Н. Марков
Старший преподаватель,
магистр технических наук,
заместитель начальника Центра
информатизации и
инновационных разработок БГУИР
a.n.markov@bsuir.by

В.Е. Буюков

Студент кафедры информатики Белорусского государственного университета информатики и радиоэлектроники.

К.И. Котельников

Студент кафедры информатики Белорусского государственного университета информатики и радиоэлектроники.

А.Н. Марков

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

Аннотация. Сегодня Big Data проникли во все сферы нашей жизни, в том числе и в экономику. С помощью анализа больших данных можно улучшить, дополнить старые алгоритмы и создавать новые экономические модели. Целью данной статьи является анализ роли больших данных в сфере экономики. Также для анализа больших данных, рассматривается фреймворк Hadoop.

Ключевые слова: Big Data, экономика, технологии, Hadoop.

Введение.

В современном мире объемы данных растут с невиданной скоростью, и их анализ и использование стали ключевым фактором для развития различных отраслей экономики. Одной из наиболее актуальных технологий, позволяющих эффективно обрабатывать и анализировать огромные объемы данных, является Big Data. В данной статье рассмотрены возможности использования Big Data в экономике и их влиянии.

Актуальность.

Применение Big Data в экономике позволяет более точно оценить риски и преимущества различных стратегий, оптимизировать производственные процессы, улучшать качество продукции и услуг, увеличивать эффективность маркетинговых кампаний, определять предпочтения и потребности потребителей, управлять ресурсами и бюджетами более эффективно, и многое другое.

Применение Big Data в экономике также актуально в свете стремительного развития цифровых технологий, таких как интернет вещей (IoT), искусственный интеллект (ИИ), облачные вычисления и другие, которые способствуют генерации и накоплению еще больших объемов данных. Эти технологии предоставляют новые возможности для сбора, хранения, анализа и использования данных с целью оптимизации бизнес-процессов и создания ценности.

Определение Big Data.

Big Data – это огромные объемы данных, которые накапливаются в различных источниках, таких как социальные сети, интернет-сайты, мобильные устройства, датчики, транзакции, и другие, и характеризуются тремя основными аспектами – объемом, разнообразием и скоростью обработки. В экономике Big Data представляют собой ценный источник информации, который может использоваться для анализа рынков, потребительского поведения, финансовых операций, производственных процессов и других аспектов экономической деятельности [1].

История использования Big Data

Одним из ранних примеров использования Big Data в экономике является анализ рынков и инвестиций. Финансовые институты и инвесторы используют огромные объемы данных, таких как цены акций, объемы торгов, экономические показатели и другие финансовые данные, для прогнозирования рынка, определения трендов и принятия решений о покупке и продаже акций, облигаций, валют и других финансовых инструментов.

Также Big Data находит применение в оптимизации производственных процессов, управлении цепями поставок, прогнозировании спроса, определении оптимальной ценовой политики, управлении рисками и многое другое. Большие объемы данных, собранных от производственных линий, логистических систем, складов, транспортных сетей и других источников, позволяют компаниям оптимизировать свою деятельность, снижать издержки, увеличивать эффективность и улучшать качество продукции и услуг. Компании используют данные о покупках, предпочтениях, интересах и демографических характеристиках потребителей, собранных из различных источников, для определения предпочтений потребителей, прогнозирования спроса на товары и услуги, улучшения маркетинговых стратегий и повышения уровня удовлетворенности клиентов [2].

Технологии Big Data в экономике.

Наиболее распространенными технологиями Big Data, которые могут применяться в экономическом контексте, являются:

Сбор и анализ больших объемов данных: С использованием Big Data технологий, таких как Hadoop, Spark и других инструментов, организации могут собирать, хранить и анализировать огромные объемы данных из различных источников, таких как социальные сети, датчики, транзакции, клиентские данные и другие, для выявления тенденций, паттернов и корреляций в экономических процессах.

Машинное обучение и аналитика данных: Применение алгоритмов машинного обучения и аналитики данных на основе Big Data позволяет организациям выявлять скрытые закономерности, прогнозировать рыночные тенденции, оптимизировать бизнес-процессы, определять предпочтения и поведение клиентов, проводить анализ рисков и многое другое.

Распределенные базы данных и облачные вычисления: Технологии Big Data также позволяют хранить и обрабатывать данные в распределенных базах данных, а также использовать облачные вычисления для более эффективного управления большими объемами данных, анализа и обработки.

Интернет вещей (IoT): Большие объемы данных могут быть сгенерированы и собраны с помощью различных устройств IoT, таких как датчики, умные дома, автомобили и другие, и использоваться для оптимизации бизнес-процессов, производственной деятельности, логистики и других аспектов экономической деятельности.

Анализ социальных медиа: Социальные медиа являются богатым источником данных, и Big Data технологии могут использоваться для анализа мнений, настроений, трендов и поведения пользователей в социальных сетях, что может быть ценной информацией для маркетинговых исследований, продвижения продуктов и услуг.

Анализ финансовых данных: В экономике финансовые данные играют важную роль, и Big Data технологии могут быть использованы для анализа данных о финансовых рынках,

транзакциях, инвестициях, кредитной и банковской деятельности, что может помочь в принятии решений в области инвестиций, риск-менеджмента и финансового планирования.

Прогнозирование и оптимизация: Применение Big Data технологий позволяет организациям проводить прогнозирование и оптимизацию бизнес-процессов на основе данных, таких как спрос, предложение, цены, конкурентная активность и другие факторы, что может помочь в принятии более информированных и эффективных решений в экономической деятельности.

Анализ данных клиентов: Big Data технологии могут быть использованы для анализа данных о клиентах, их предпочтениях, поведении и покупательской активности, что может помочь в улучшении маркетинговых стратегий, персонализации предложений и повышении удовлетворенности клиентов.

Анализ данных логистики и цепей поставок: Big Data технологии могут быть применены для анализа данных о логистических процессах, цепях поставок и транспортной логистике, что может помочь в оптимизации логистических операций, снижении затрат и улучшении эффективности доставки товаров и услуг.

Анализ данных о конкурентной среде: Big Data технологии могут быть использованы для анализа данных о конкурентной среде, таких как данные о ценах, продуктах, маркетинговых активностях конкурентов, что может помочь в принятии решений по ценообразованию, маркетинговой стратегии и развитию конкурентоспособности организации [3].

В дальнейшем изучение технологии будет проходить на примере фреймворка Hadoop.

Использование Hadoop.

Множество компаний по всему миру используют Hadoop в своих бизнес-процессах. Некоторые из известных компаний, которые применяют Hadoop в своей деятельности, включают:

Amazon – одна из крупнейших компаний в области электронной коммерции, использует Hadoop для анализа данных о покупках, логистике, рекомендательных системах и многих других аспектах своего бизнеса.

Google – одна из ведущих компаний в области интернет-технологий, использует Hadoop для обработки и анализа данных, таких как поисковые запросы, рекламные платформы, аналитика веб-трафика и др.

IBM – одна из крупнейших IT-компаний в мире, использует Hadoop в своих решениях для анализа и обработки данных, таких как аналитика бизнес-процессов, аналитика клиентского опыта, аналитика рынка и др.

Walmart – одна из крупнейших розничных сетей в мире, использует Hadoop для анализа данных о продажах, логистике, предсказательной аналитике спроса, а также для оптимизации поставок и управления запасами.

Intel – одна из крупнейших компаний в области производства микропроцессоров, использует Hadoop для анализа данных, связанных с производством, тестированием и оптимизацией процессов производства.

Структура Hadoop.

Hadoop – это фреймворк для обработки больших данных, разработанный Apache Software Foundation. Он состоит из двух основных компонентов: Hadoop Distributed File System (HDFS) и Hadoop MapReduce.

Hadoop Distributed File System (HDFS): Это распределенная файловая система, предназначенная для хранения и обработки больших объемов данных на кластере серверов. HDFS разбивает файлы на блоки и распределяет их на различные узлы в кластере, обеспечивая отказоустойчивость и масштабируемость. HDFS разбивает файлы на блоки и распределяет их на различные узлы в кластере, обеспечивая отказоустойчивость и масштабируемость.

Hadoop MapReduce: Это модель программирования и выполнения, которая позволяет эффективно обрабатывать большие объемы данных на кластере Hadoop. MapReduce разделяет задачи на две основные операции: Map (отображение) и Reduce (сведение), которые выполняются параллельно на различных узлах кластера.

Hadoop YARN (Yet Another Resource Negotiator): Это система управления ресурсами, которая отвечает за управление ресурсами кластера, распределение задач на вычислительные узлы и мониторинг их выполнения. YARN позволяет более эффективно использовать ресурсы кластера и поддерживать множество различных рабочих нагрузок, таких как MapReduce, Apache Spark, Apache Hive и др.

Hadoop EcoSystem: Это набор дополнительных инструментов и технологий, разработанных на основе Hadoop, таких как Apache Hive, Apache Pig, Apache HBase, Apache Spark и другие. Эти компоненты расширяют функциональность Hadoop и позволяют более удобно и эффективно выполнять различные задачи обработки и анализа данных.

Рабочий процесс Hadoop обычно состоит из следующих шагов:

Хранение данных в HDFS: Данные разделяются на блоки и сохраняются в HDFS на различных узлах кластера. HDFS обеспечивает отказоустойчивое хранение данных, распределяя данные по узлам и создавая несколько реплик каждого блока данных.

Обработка данных с помощью MapReduce: Программа MapReduce разделяет обработку данных на два этапа - Map и Reduce. Эти этапы выполняются на распределенных узлах кластера.

– **Этап Map:** В этом этапе данные извлекаются из HDFS и обрабатываются на каждом узле отдельно. Результаты преобразования записываются в локальные файлы на каждом узле.

– **Этап Reduce:** В этом этапе результаты Map объединяются и обрабатываются для получения окончательных результатов. Reduce-функции выполняют агрегацию, фильтрацию, сортировку и другие операции над данными.

Обработка результата: Окончательные результаты обработки сохраняются в HDFS или передаются для дальнейшей обработки или анализа.

Мониторинг и управление: Hadoop предоставляет инструменты для мониторинга и управления кластером, такие как Hadoop YARN (Yet Another Resource Negotiator) и Hadoop MapReduce JobTracker, которые позволяют контролировать выполнение задач, оптимизировать использование ресурсов и обеспечивать отказоустойчивость [4].

Показатели эффективности

Twitter: Благодаря использованию Hadoop для анализа и обработки миллиардов твитов ежедневно, Twitter смог повысить эффективность своих маркетинговых кампаний. Анализ метрик, таких как количество упоминаний, ретвитов, лайков и других показателей, позволил Twitter оптимизировать свои маркетинговые стратегии и повысить эффективность на 10%.

eBay: Использование Hadoop позволило eBay анализировать огромные объемы данных о продажах, клиентах и других метриках, что помогло оптимизировать его платформу. В результате, eBay смог улучшить процесс продаж и повысить свою эффективность на 15% [5].

Заключение.

В заключение, использование Big Data в экономике имеет огромный потенциал для оптимизации бизнес-процессов, принятия обоснованных решений и повышения эффективности компаний. С помощью передовых технологий, таких как Hadoop, компании могут обрабатывать и анализировать огромные объемы данных, выявлять скрытые паттерны и тренды, а также прогнозировать будущие события.

Список литературы

- [1] Gartner, Inc. "Big Data." Gartner IT Glossary [Электронный ресурс], Адрес: <https://www.gartner.com/en/information-technology/glossary/big-data>.
- [2] Mayer-Schönberger, V., & Cukier, K. (2013). Большие данные: революция, которая изменит то, как мы живем, работаем и думаем. Хоутон Миффлин Харкорт.
- [3] Доббс, Р., Роксбург, К., & Байерс, А. Х. (2011). Большие данные: следующий рубеж для инноваций
- [4] Документация Hadoop [Электронный ресурс], - Адрес: <https://hadoop.apache.org/>
- [5] Mayer-Schönberger, V., & Cukier, K. (2013). Big Data: A Revolution That Will Transform How We Live, Work, and Think. Houghton Mifflin Harcourt.

APPLICATION OF BIG DATA IN THE FIELD OF EDUCATION

K.I.Katselnikau

Student of the BSUIR

V.E. Buyukov

Student of the BSUIR

A.N.MARKOV

*Senior lecturer of the department,
Deputy head of the Center for
Informatization and Innovative
Developments*

Department of Computer Science

Faculty of Computer Systems and Networks

Belarusian State University of computer science and Radio Electronics, Republic of Belarus

E-mail: kki199918@gmail.com

Abstract. Today, Big Data has penetrated into all spheres of our life, including the economy. With the help of big data analysis, it is possible to improve, supplement old algorithms and create new economic models. The purpose of this article is to analyze the role of big data in the field of economics. Also for big data analysis, the Hadoop framework is being considered.

Keywords: Big Data, economics, technology, Hadoop.