

ПРИМЕНЕНИЕ BIG DATA АНАЛИТИКИ В ОПТИМИЗАЦИИ ПРОИЗВОДСТВЕННЫХ ПРОЦЕССОВ

Басипов С.Н.

Научный руководитель: Самохина Н.С.

Россия, г. Тольятти,
Поволжский государственный университет сервиса

Аннотация. Статья рассматривает большие данные, способы их обработки и возможность их применения в аналитике производственных процессов. Затрагивается и создание виртуальной модели предприятия для дальнейшей его оптимизации.

Ключевые слова: Big Data, MapReduce, реляционные базы данных, NoSQL базы данных, цифровые двойники, сбор данных, бизнес-процессы, техническое обслуживание, моделирование ситуаций, стриминговая обработка данных, свойства больших данных.

С каждым днём объёмы данных в цифровой среде растут. Миллионы компаний по всему миру собирают различную информацию об окружающем нас мире. Это влечёт за собой необходимость придумывать новые способы хранения и обработки больших данных.

О больших данных.

Big Data – это большие объёмы неоднородных данных, требующих нестандартных способов их обработки. Зачастую эти данные являются неоднородными, что также усложняет их анализ традиционными методами.

У Big Data есть несколько основных признаков, по которым можно определить, относится ли текущий массив данных к большим или нет. Среди них выделяют следующие:

- velocity (скорость), для обработки массива в реальном времени требуется высокая скорость анализа данных;
- volume (объём), данные имеют значительные объёмы, измеряемые десятками гигабайтов;
- variety (разнообразие), информация в одном наборе может быть представлена как текстовая, видео, аудио и пр.

Хранение больших данных производят в SQL и NoSQL базах данных. При этом реляционные БД работают значительно быстрее за счет того, что они обладают четкой структуризацией, целостностью данных и нормализованностью данных, которой не имеют NoSQL БД. Тем не менее нереляционные базы данных (например, графовые, столбцовые, документные), также имеют свои преимущества, а именно:

- возможность хранить разнообразные виды данных;
- легкость горизонтального масштабирования;
- повышенная отказоустойчивость в сравнении с SQL-моделями.

Способы обработки Big Data.

Для обработки больших данных используется несколько основных подходов, каждый из которых имеет свои преимущества и недостатки.

Первый из подходов – это стриминговая обработка данных, исходя из названия, которое является транскрипцией от слова «streaming» – «поточковый», понятно, что в этом случае обработка происходит в режиме реального времени.

Преимущества:

- данные обрабатываются по мере их поступления;
- можно работать со «скоропортящимися» данными.

Недостатки:

- высокая сложность настройки.

Вторым подходом является MapReduce, происходящий от слов «Map» – «карта» и «Reduce» – «сокращать», он, как можно понять из названия, сначала создает карту с парами «ключ-значение» и после этого сортирует, и производит другие операции, заложенные разработчиком, после которых происходит свертка данных и подготовка результата.

Преимущества:

- удобство работы и настройки;
- хорошая отказоустойчивость.

Недостатки:

- затруднительное горизонтальное масштабирование.

Цифровые двойники предприятия и Big Data аналитика.

Цифровой двойник – это виртуальная модель физического объекта, описывающая его основные параметры, по которым можно воспроизвести его копию в цифровой среде. Основываясь на цифровом двойнике объекта, можно провести исследование с помощью автоматизированных методов обработки данных, таких, как аналитика больших данных.

Методика исследования включает в себя следующие шаги:

- проведение технического интервью с заказчиком;
- выявление проблемных и узких мест на предприятии;
- создание технического задания;
- согласование ТЗ;
- разработка системы сбора и хранения данных, а также создание цифрового двойника;
- внедрение разработанной системы на предприятие;
- анализ собранных больших данных;
- применение полученных результатов.

Использование методов Big Data аналитики в совокупности с цифровыми двойниками позволяют моделировать различные ситуации на производстве, не оказывая влияния на реальные бизнес-процессы и, соответственно, значительно снижая вероятность уменьшения эффективности и получения убытков предприятия.

Методика исследования.

В ходе исследования было проведено подключение к системам мониторинга оборудования, что частично облегчило дальнейший сбор таких данных, как:

- нагрузка на оборудование;
- время простоя;
- количество полученной продукции за определенный промежуток времени;
- периодичность обслуживания;
- сведения о поломках;
- рабочая температура узлов агрегатов.

На основе собранных данных был проведен автоматизированный анализ с помощью инструментов Big Data, а также машинного обучения. Были выявлены различные зависимости, например, зависимость поломок от нагрузки, периодичности обслуживания и прочих параметров.

Результаты исследования.

Полученные данные помогли переработать график технического обслуживания и улучшить загруженность оборудования. В итоге время бездействия снизилось на 89%, а количество поломок сократилось на 14% в сравнении с показателями до оптимизации бизнес-процессов.

Заключение.

Разработанный программный комплекс может быть использован в дальнейшем в качестве системы оптимизации некоторых производственных процессов на предприятии. Если продолжить развитие данного подхода, то можно будет оптимизировать множество разнообразных процессов, например, документооборот между отделами, закупку материалов, согласование документов, прием сотрудников и т.д. Улучшение этих процессов повлечет за собой повышение эффективности, а следовательно, и снижение издержек на поддержание работоспособности компании.

Библиографический список

1. Gartner Says Solving 'Big Data' Challenge Involves More Than Just Managing Volumes of Data [Электронный ресурс]. URL: <https://webcitation.org/67j1UMds0?url=http://www.gartner.com/it/page.jsp?id=1731916> (дата обращения: 13.02.2024).

2. Manyika J., Chui M., Brown B. et al. Big data: The next frontier for innovation, competition, and productivity [Электронный ресурс]. URL: https://webcitation.org/6ComRQdf1?url=http://www.mckinsey.com/~media/McKinsey/dotcom/Insights%20and%20pubs/MGI/Research/Technology%20and%20Innovation/Big%20Data/MGI_big_data_full_report.ashx (дата обращения: 10.02.2024).

3. The Missing V's in Big Data: Viability and Value [Электронный ресурс]. URL: <https://www.wired.com/insights/2013/05/the-missing-vs-in-big-data-viability-and-value/> (дата обращения: 12.02.2024).

APPLICATION OF BIG DATA ANALYTICS IN THE OPTIMIZATION OF PRODUCTION PROCESSES

Basipov S.N.

Scientific adviser: Samokhina N.S.

Volga Region State University of Service, Togliatti, Russia

Abstract. *The article examines big data, methods of processing it, and the possibility of using it in the analytics of production processes. The creation of a virtual model of the enterprise for its further optimization is also touched upon.*

Keywords: *Big Data, MapReduce, relational databases, NoSQL databases, digital twins, data collection, business processes, maintenance, situation modeling, streaming data processing, properties of big data.*