УДК 65.011.56

АНАЛИЗ БОЛЬШИХ ДАННЫХ НА ПРОМЫШЛЕННОМ ПРЕДПРИЯТИИ

Дементьев С.Ю., [1]Роза М.П.
ФГБОУ ВО «Сибирский государственный университет науки и технологий имени академика М. Ф. Решетнева», Красноярск, Россия (660037, Красноярский край, город Красноярск, проспект имени газеты «Красноярский рабочий», д. 3), e-mail: [1]super.wark@mail.ru

Анализ больших данных становится все более важным в современном промышленном секторе, поскольку предприятия стремятся использовать возможности больших объемов данных, генерируемых из различных источников. В этой статье рассматривается применение анализа больших данных на промышленном предприятии и выделяются его потенциальные преимущества. В исследовании рассматриваются проблемы, с которыми пришлось столкнуться, используемые методологии и результаты, достигнутые благодаря анализу больших данных. Используя передовые методы аналитики, промышленные предприятия могут принимать решения на основе данных, оптимизировать операции и повышать общую эффективность. Полученные результаты демонстрируют преобразующий потенциал анализа больших данных и его способность стимулировать инновации и конкурентоспособность в промышленном секторе.

Ключевые слова: Анализ больших данных, промышленное предприятие, решения на основе данных, оптимизация, эффективность.

BIG DATA ANALYSIS IN AN INDUSTRIAL ENTERPRISE

Dementev S.Y., [1]Roza M.P.
Reshetnev Siberian State University of Science and Technology, Krasnoyarsk, Russia (660037, Krasnoyarsk Krai, Krasnoyarsk city, prospect named after the newspaper "Krasnoyarsk worker", 3), e-mail: [1]super.wark@mail.ru

Big data analysis has become increasingly vital in the modern industrial landscape, as enterprises seek to harness the power of large volumes of data generated from various sources. This article explores the application of big data analysis in an industrial enterprise and highlights its potential benefits. The study investigates the challenges faced, methodologies employed, and results achieved through big data analysis. By leveraging advanced analytics techniques, industrial enterprises can make data-driven decisions, optimize operations, and improve overall efficiency. The findings demonstrate the transformative potential of big data analysis and its ability to drive innovation and competitiveness in the industrial sector.

Keywords: Big data analysis, industrial enterprise, data-driven decisions, optimization, efficiency.

In today's digital age, industrial enterprises generate vast amounts of data from various sources, such as sensors, machines, supply chains, and customer interactions. This explosion of data offers unprecedented opportunities for enterprises to gain valuable insights and improve their operations. Big data analysis, which involves extracting actionable information from large and complex datasets, has emerged as a powerful tool for industrial enterprises. By leveraging big data analytics, enterprises can optimize processes, enhance productivity, and unlock new avenues for

growth. This article delves into the application of big data analysis in an industrial enterprise, examining its benefits, methodologies, and outcomes.

One of the important aspects of big data analysis in industrial enterprises is to increase the efficiency of the decision-making process. Big data analysis enables industrial enterprises to make data-driven decisions by providing comprehensive insights into operational processes, market trends, and customer preferences. This data-driven decision-making approach helps in identifying opportunities, mitigating risks, and improving overall business performance [1].

The next aspect of big data analysis in industrial enterprises is Optimization of Operations. By analyzing large datasets, industrial enterprises can identify bottlenecks, inefficiencies, and areas for improvement in their operations. Through advanced analytics techniques, such as predictive modeling and machine learning algorithms, enterprises can optimize production processes, reduce downtime, and enhance resource utilization.

The end aspect of big data analysis in industrial enterprises is Improved Supply Chain Management. Big data analysis facilitates real-time monitoring of the supply chain, enabling enterprises to identify potential disruptions, forecast demand, and optimize inventory levels. By integrating data from suppliers, manufacturers, and distributors, enterprises can achieve greater supply chain visibility, streamline logistics, and improve customer satisfaction [2].

The Methodologies for Big Data Analysis in Industrial Enterprises in industrial enterprises consists of 4 stages. These steps are Data Collection and Integration, Data Preprocessing and Cleansing, Exploratory Data Analysis and Advanced Analytics Techniques. Let's look at these steps in more detail [2]:

1. Data Collection and Integration. Industrial enterprises need to collect data from various sources, including sensors, production systems, and external databases. This data is then integrated into a centralized repository for further analysis. Advanced data integration techniques, such as data virtualization or data lakes, allow enterprises to merge heterogeneous datasets and eliminate data silos.

2. Data Preprocessing and Cleansing. Before analysis, raw data often requires preprocessing and cleansing to remove errors, outliers, and irrelevant information. Techniques like data normalization, outlier detection, and missing data imputation ensure the quality and integrity of the dataset.

3. Exploratory Data Analysis. Exploratory data analysis involves performing initial investigations to uncover patterns, relationships, and anomalies in the data. Visualizations, statistical summaries, and data profiling techniques aid in understanding the characteristics of the dataset and guide further analysis.

4. Advanced Analytics Techniques. Industrial enterprises employ various advanced analytics techniques, such as machine learning, predictive modeling, and optimization algorithms, to extract insights from big data. These techniques enable enterprises to predict equipment failures, optimize maintenance schedules, and forecast demand more accurately.

Several technologies are associated with big data analysis in an industrial enterprise. Here are some of the key technologies commonly used [3]:

1. Distributed File Systems. Distributed file systems, such as Apache Hadoop HDFS (Hadoop Distributed File System) and Apache HBase, are fundamental for storing and managing large volumes of data in a distributed computing environment. These file systems allow enterprises to scale storage capacity and handle massive datasets across clusters of commodity hardware.

2. Data Processing Frameworks. Data processing frameworks provide the computational power needed to analyze and extract insights from big data. Apache Hadoop, Apache Spark, and Apache Flink are widely used frameworks that offer distributed processing capabilities. These frameworks allow enterprises to parallelize computations and execute complex data analysis tasks on large datasets efficiently.

3. Data Integration and ETL Tools. Data integration and Extract, Transform, Load (ETL) tools are essential for ingesting, cleansing, and transforming data from various sources before analysis. Tools like Apache Nifi, Talend, and Informatica enable enterprises to extract data from different systems, integrate it into a unified format, and prepare it for analysis.

4. Distributed Databases. Distributed databases, such as Apache Cassandra, Apache HBase, and Amazon DynamoDB, are designed to handle large-scale data storage and retrieval in a distributed environment. These databases offer high scalability, fault tolerance, and low-latency access, enabling efficient storage and retrieval of big data for analysis.

5. Machine Learning and Predictive Analytics. Machine learning algorithms and predictive analytics techniques play a crucial role in extracting insights and making predictions from big data. Technologies such as Python's scikit-learn, TensorFlow, and Apache Mahout provide libraries and frameworks for implementing machine learning models and predictive analytics algorithms on large datasets.

6. Data Visualization Tools. Data visualization tools help in presenting complex data analysis results in a visually appealing and understandable manner. Tools like Tableau, Power BI, and D3.js allow enterprises to create interactive charts, graphs, and dashboards, enabling stakeholders to explore and interpret big data analysis outcomes effectively.

7. Real-time Stream Processing. Real-time stream processing technologies, such as Apache Kafka and Apache Flink, enable enterprises to process and analyze streaming data in real-time. These technologies support the continuous ingestion, processing, and analysis of data as it flows, allowing for timely insights and quick decision-making.

8. Cloud Computing. Cloud computing platforms, including Amazon Web Services (AWS), Microsoft Azure, and Google Cloud Platform (GCP), provide scalable and on-demand computing resources for big data analysis. These platforms offer various managed services, such as Amazon EMR, Azure HDInsight, and Google Cloud Dataproc, which simplify the deployment and management of big data analytics infrastructure.

These technologies, along with others, form the foundation for performing big data analysis in an industrial enterprise. The selection and integration of these technologies depend on the specific requirements, scale, and complexity of the enterprise's big data analytics needs.

There are several real cases of big data analysis being successfully implemented in industrial enterprises. Here are a few examples [4]:

1. Predictive Maintenance in Manufacturing. Industrial enterprises have started leveraging big data analytics to implement predictive maintenance strategies. By collecting and analyzing data from sensors embedded in machinery and equipment, enterprises can identify patterns and anomalies that indicate potential failures. This allows them to schedule maintenance proactively, reducing downtime, optimizing maintenance costs, and extending the lifespan of assets. For example, General Electric (GE) uses big data analytics to monitor the performance of jet engines in real-time, predicting maintenance needs and optimizing aircraft operations.

2.   Quality Control and Defect Detection. Big data analysis helps industrial enterprises improve quality control processes and detect defects in products. By analyzing data from various stages of the manufacturing process, including sensors, machine logs, and inspection records, enterprises can identify deviations from normal operating conditions and detect potential quality issues. This enables them to take corrective actions promptly, reducing waste and improving overall product quality. For instance, a semiconductor manufacturer might analyze sensor data to detect anomalies during the production of microchips, ensuring high-quality output.

3.   Supply Chain Optimization. Big data analysis plays a crucial role in optimizing supply chain operations. By integrating data from suppliers, manufacturers, logistics providers, and customers, enterprises can gain real-time visibility into their supply chain. This enables them to identify bottlenecks, optimize inventory levels, streamline logistics, and improve demand forecasting. Amazon, for instance, uses big data analytics to analyze customer order patterns, optimize warehouse operations, and facilitate faster deliveries.

4.   Energy Management and Optimization. Industrial enterprises consume significant amounts of energy in their operations. Big data analysis can help optimize energy usage, reduce costs, and minimize environmental impact. By analyzing energy consumption data from sensors and meters, enterprises can identify energy inefficiencies, implement energy-saving measures, and optimize resource utilization. For example, Google employs big data analytics to optimize the cooling systems in its data centers, resulting in substantial energy savings.

5.   Customer Analytics and Personalization. Big data analysis enables industrial enterprises to gain insights into customer behavior, preferences, and buying patterns. By analyzing customer data from multiple touchpoints, such as sales transactions, social media interactions, and customer feedback, enterprises can personalize their offerings, tailor marketing campaigns, and improve customer satisfaction. For instance, automotive companies analyze customer data to understand preferences for specific features, enabling them to design and market vehicles that meet customer expectations.

These examples demonstrate the diverse applications of big data analysis in industrial enterprises, showcasing its potential to drive operational efficiency, enhance decision-making, and improve overall business performance. As technology continues to advance, we can expect further innovation and expansion of big data analytics in the industrial sector.

While big data analysis offers numerous benefits, it is essential to consider potential negative effects and challenges associated with its use. One example of a negative effect of big data analysis is the concern related to privacy and data security.

Big data analysis involves the collection, storage, and analysis of vast amounts of personal and sensitive information. As more data is gathered and analyzed, there is an increased risk of privacy breaches and data security incidents. If not handled properly, the misuse or unauthorized access to sensitive data can result in privacy violations, identity theft, or other forms of cybercrime. Moreover, the aggregation of diverse datasets can lead to unintended identification or re-identification of individuals, potentially compromising their privacy rights [5].

For instance, consider a scenario where an industrial enterprise collects and analyzes data from various sources, including employee records, customer transactions, and public data. If the enterprise fails to implement robust security measures, it becomes vulnerable to data breaches, exposing personal information to unauthorized individuals or malicious actors. Such incidents can erode customer trust, damage the reputation of the enterprise, and lead to legal consequences [6].

To mitigate this negative effect, organizations must prioritize data privacy and security. They should implement robust security protocols, employ encryption techniques, and comply with relevant data protection regulations such as the General Data Protection Regulation (GDPR) or the California Consumer Privacy Act (CCPA). Additionally, enterprises must be transparent about their data collection and usage practices, obtain informed consent from individuals, and establish clear policies for data handling and retention [7].

It is important to recognize that while big data analysis offers tremendous potential, organizations must strike a balance between extracting valuable insights and protecting individual privacy and data security. Responsible data governance practices and ethical considerations are essential to address this negative effect and maintain public trust in the use of big data analysis.

Results of Big Data Analysis in Industrial Enterprises include Operational Efficiency. Through big data analysis, industrial enterprises can optimize their production processes, reduce energy consumption, and minimize waste. This optimization leads to improved operational efficiency, increased productivity, and cost savings. The next result is Predictive Maintenance. By leveraging big data analytics, enterprises can implement predictive maintenance strategies. By analyzing sensor data, historical maintenance records, and external factors, enterprises can predict equipment failures, schedule maintenance proactively, and avoid costly unplanned downtime. One of the advantages is Customer Insights and Personalization. Big data analysis helps industrial enterprises gain a deeper understanding of customer behavior, preferences, and buying patterns. By analyzing customer data, such as purchase history, feedback, and social media interactions, enterprises can personalize their offerings, target specific customer segments, and enhance customer satisfaction. Last but not least, the advantage is Risk Mitigation. Big data analysis enables industrial enterprises to identify and mitigate potential risks. By analyzing historical data and external factors, enterprises can develop risk models, detect anomalies, and implement proactive measures to prevent accidents, quality issues, or supply chain disruptions.

Big data analysis holds immense potential for industrial enterprises, offering valuable insights, optimized operations, and improved decision-making capabilities. By leveraging advanced analytics techniques, industrial enterprises can unlock the power of big data and gain a competitive edge in the market. However, the implementation of big data analysis requires careful consideration of data privacy, security, and infrastructure challenges. As technology advances and more sophisticated analytical tools become available, industrial enterprises must adapt and embrace the transformative potential of big data analysis.

Список литературы

1. Chen, M. Big data: A survey / M. Chen, S. Mao, Y. Liu // Mobile Networks and Applications. – 2014. – No. 19(2). – pp. 171-209
2. Davenport, T. H. Big data at work: Dispelling the myths, uncovering the opportunities. / T. H. Davenport. – Harvard Business Review Press, 2014. – 121 p.
3. McAfee, A., Big data: The management revolution / A. McAfee, E. Brynjolfsson // Harvard Business Review. – 2012. – No. 90(10). – pp. 60-68.
4. Provost, F. Data science for business: What you need to know about data mining and data-analytic thinking / F. Provost, T. Fawcett. - Sebastopol: O'Reilly Media, 2013. - 409 p.
5. Tan, P. N. Introduction to data mining / P. N. Tan, M. Steinbach, V. Kumar. – Pearson, 2018. – 173 p.

6. Dement'ev, S. Ju. Metody avtomaticheskoj gruppirovki dlja povyshenija jekonomicheskoj jeffektivnosti v industrii 4.0 / S. Ju. Dement'ev, M. P. Roza // Jekonomika i predprinimatel'stvo. – 2022. – № 10(147). – pp. 1390-1393. – DOI 10.34925/EIP.2022.147.10.278. – EDN YDFFDH.

7. Wu, X. Data mining with big data / X. Zhu, G. Q. Wu, W. Ding // IEEE Transactions on Knowledge and Data Engineering. – 2014. – No. 26(1). – pp. 97-107

References
1. Chen, M. Big data: A survey / M. Chen, S. Mao, Y. Liu // Mobile Networks and Applications. – 2014. – No. 19(2). – pp. 171-209

2. Davenport, T. H. Big data at work: Dispelling the myths, uncovering the opportunities. / T. H. Davenport. – Harvard Business Review Press, 2014. – 121 p.

3. McAfee, A., Big data: The management revolution / A. McAfee, E. Brynjolfsson // Harvard Business Review. – 2012. – No. 90(10). – pp. 60-68.

4. Provost, F. Data science for business: What you need to know about data mining and data-analytic thinking / F. Provost, T. Fawcett. - Sebastopol: O'Reilly Media, 2013. - 409 p.

5. Tan, P. N. Introduction to data mining / P. N. Tan, M. Steinbach, V. Kumar. – Pearson, 2018. – 173 p.

6. Dement'ev, S. Ju. Metody avtomaticheskoj gruppirovki dlja povyshenija jekonomicheskoj jeffektivnosti v industrii 4.0 / S. Ju. Dement'ev, M. P. Roza // Jekonomika i predprinimatel'stvo. – 2022. – № 10(147). – S. 1390-1393. – DOI 10.34925/EIP.2022.147.10.278. – EDN YDFFDH.

7. Wu, X. Data mining with big data / X. Zhu, G. Q. Wu, W. Ding // IEEE Transactions on Knowledge and Data Engineering. – 2014. – No. 26(1). – pp. 97-107.