JOURNAL OF APPLIED SCIENCE AND SOCIAL SCIENCE
eISSN 2229-3113 pISSN 2229-3205

Volume 15 Issue 12, December 2025
Impact factor: 2019: 4.679 2020: 5.015 2021: 5.436, 2022: 5.242, 2023: 6.995, 2024 7.75

ANALYTICAL METHODS USED IN BIG DATA

Mustafoyev Abror Omonjon ugli
Alfraganus University, Tashkent, Uzbekistan
abrorbilolbek05@gmail.com

Annotation: The article defines the term "big data", identifies the features of big data, describes the methods of their analysis, reveals the differences between traditional data processing methods and big data technologies, determines the importance of using big data by enterprises. Big data has become a critical asset in modern organizations due to the rapid growth of digital technologies and data-generating systems. The ability to analyze large, complex, and diverse datasets enables organizations to extract valuable insights, improve decision-making processes, and gain competitive advantages. This paper explores key techniques for analyzing big data, including data mining, machine learning, statistical analysis, and distributed computing frameworks. The study highlights the importance of scalable and efficient analytical methods in handling the volume, velocity, and variety of big data. The findings emphasize how advanced analytical techniques contribute to predictive analytics, pattern recognition, and real-time data processing across various industries.

Keywords: big data, big data analysis methods, big data processing technologies.

INTRODUCTION

At the turn of the decades of the era of digital transformation of the economy, which marked the stage of maturity of the information society, big data technologies radically modify production and management processes, replace many hours of human intellectual labor and develop artificial intelligence, accelerate the processing of ever-increasing huge amounts of data on human life and economic activity.The activities of enterprises find previously hidden dependencies, simplify the perception of the results of human data processing, and provide recommendations for decision-making. In the conditions of total abundance of a post-industrial society, it is big data technologies that make it possible to individualize production, anticipate and satisfy the desires of consumers, while ensuring reasonable prices, and produce public goods without the participation of government funding. The scientific study of this topic will allow us to further develop methodological materials for teaching not only technical specialists, but also philologists, financiers, marketers, managers.

In recent years, the exponential increase in data generated from social media, sensors, mobile devices, and online transactions has led to the emergence of big data. Traditional data processing tools are often insufficient to manage and analyze such massive and complex datasets. As a result, new techniques and technologies have been developed to efficiently extract meaningful information from big data. Techniques for analyzing big data play a crucial role in transforming raw data into actionable knowledge. These techniques include data mining, machine learning algorithms, statistical modeling, and distributed processing systems such as Hadoop and Spark. By applying these methods, organizations can uncover hidden patterns, predict future trends, and

JOURNAL OF APPLIED SCIENCE AND SOCIAL SCIENCE
eISSN 2229-3113 pISSN 2229-3205
Volume 15 Issue 12, December 2025
Impact factor: 2019: 4.679 2020: 5.015 2021: 5.436, 2022: 5.242, 2023: 6.995, 2024 7.75

enhance operational efficiency. This paper aims to provide an overview of the most widely used techniques for big data analysis and their significance in modern data-driven environments.

Definition and features of big data. The term "big data" is considered in terms of data volumes and methods of processing the for effective use. It designates data sets whose size exceeds the capabilities of typical databases for recording, storing, managing and analyzing information, taking into account the growth of these data volumes .Researchers note the uselessness of the data itself until it turns into information necessary for specific purposes. Therefore, the use of the term "big data" implies and sometimes implies means analyze this data to obtain useful information. Enterprises and organizations create huge amounts of data, but most of it is presented in a poorly structured format (web logs, video recordings, text documents, machine code or geospatial data), and the speed of updating this data increases. They are stored in a variety of different repositories, sometimes outside the organization. Traditional methods of data analysis in this form do not allow them to be used fully and most effectively.

The concept of big data implies working with data of a huge volume and diverse composition, frequently updated and located in different sources, in order to increase the efficiency of their use, obtain useful information for creating new products, improve the efficiency of production processes and the competitiveness of the enterprise.The term "big data" combines both quantitative and qualitative characteristics of the data, as well as the specifics of their processing for specific significant purposes of economic activity. Wikipedia gives the following definition: "In a broad sense, big data is referred to as a socio-economic phenomenon associated with the emergence of technological capabilities to analyze huge amounts of data, in some problematic areas – the entire global volume of data and the transformational consequences resulting from this." Company analysts IBS "the entire global volume of data" was estimated by such values:

JOURNAL OF APPLIED SCIENCE AND SOCIAL SCIENCE
eISSN 2229-3113 pISSN 2229-3205
Volume 15 Issue 12, December 2025
Impact factor: 2019: 4.679 2020: 5.015 2021: 5.436, 2022: 5.242, 2023: 6.995, 2024 7.75

Fig.1. Big Data Analysis

The Forrester consulting company provides a brief formulation: "Big data combines techniques and technologies that extract meaning from data at the extreme limit of practicality". "Big Data (Big Data) is a designation structured and unstructured data of huge volumes and significant diversity, efficiently processed by horizontally scalable software tools that appeared in the late 2000s, alternative to traditional database management systems and Business class solutions Intelligence».By 2020, research has identified the following criteria for big data (Fig. 1): volume (Volume), rate of increase (Velocity), Variety (Variety), reliability (Veracity), viability (Viability), Value (Value), variability (Variability), visualization (Visualization).

Principles of working with big data. Horizontal scalability implies an increase in the number of computing nodes over which this data is distributed, with the ability to process it without compromising node performance Fault tolerance. Methods of working with big data should take into account the possibility of failure of computing nodes and provide for preventive measures. The locality of the data and its processing. In order to save the cost of transferring large amounts of data, it is desirable to process them as close as possible to the place of their accumulation. The principles of working with big data differ from traditional methods of processing centralized databases with their characteristic vertical storage model.

Characteristic	Traditional database	Big Data Database
The amount of information	From gigabytes to terabytes	From petabytes to exabytes
Storage method	Centralized	Decentralized
Data structuring	High-level	Weak or lack thereof
Data storage and processing model	Vertical	Horizontal
The relationship of data	Strong	Weak

Table – Differences between traditional databases and big data databases

structured data. Therefore, new approaches and technologies are being developed to work with big data.The traditional economic analysis used:

- the collected data with a large lag, the constructed factor models could become outdated by the timethe results were evaluated and forecasts were made;

- the factors in the model were limited by the ability to assemble an acceptable database of regular accounting of certain indicators. If it was necessary to include a new factor in the model , it was necessary to restructure the accounting of indicators at the level of accounting and statistical data enterprises and industries, which caused great difficulties;

- simplified models and averaged indicators, as complex mathematical methods without automation required a lot of processing time.

The use of modern methods of processing big data allows:

JOURNAL OF APPLIED SCIENCE AND SOCIAL SCIENCE
eISSN 2229-3113 pISSN 2229-3205
Volume 15 Issue 12, December 2025
Impact factor: 2019: 4.679 2020: 5.015 2021: 5.436, 2022: 5.242, 2023: 6.995, 2024 7.75

• analyze data in real time and constantly monitor the dynamics of more accurate indicators, rather than average values;

• find the influence of new factors on the final result;

• predict the behavior of the model based on the most recent data, including previously overlooked factors;

• visualize the results of modeling and forecasting.

The combination of big data analysis technologies carries out three main groups of processing operations: 1) fast incoming data in very large and constantly increasing volumes 2) both structured and poorly structured data in parallel; 3) a wide variety of scenarios for the interconnections of incoming data. It is believed that these capabilities of big data processing technologies make it possible to reveal hidden patterns that escape the limited human perception. This opens up unprecedented opportunities for optimizing many areas of our lives: public administration, medicine, telecommunications, finance, transport, manufacturing, and also creates a basis for for artificial intelligence technologies, which most accurately predicts the needs and wishes of people, and will allow in the future to individualize the production of products and services at the lowest cost.

Methods of big data analysis. Methods of the Data Mining class (data mining, data mining, deep data analysis) are a set of methods for detecting previously unknown, non–trivial, practically useful knowledge in data that is necessary for decision-making. Such methods, in particular, include teaching associative rules (association rule learning), classification (categorization), cluster analysis, regression analysis, detection and analysis of deviations, etc. Crowdsourcing is the classification and enrichment of data by a wide, indefinite range of forces persons who perform this work without entering into an employment relationship.

Predictive analytics – statistical methods, methods of data mining, game theory, which, based on the analysis of the current state and past experience, identify a pattern of behavior of a subject or group and the probability of deviation from it to make predictions about future events, making decisions taking into account risks. Simulation is a method that allows you to build models describing processes the way they would actually take place. Simulation modeling is possible to be considered as a kind of experimental testing. Spatial analysis is a class of methods that use topological, geometric, and geographic information extracted from data. Statistical analysis – time series analysis, A/B testing. A/B testing, split testing is a marketing research method; when using it, a control group of elements is compared with a set of test groups in which one or more indicators have been changed to find out which of the changes improve the target indicator. Data fusion and integration is a set of techniques that allow you to integrate heterogeneous data from various sources in order to conduct in–depth analysis (for example, digital signal processing, natural language processing, including tonal analysis, etc.) Machine learning, including learning with and without a teacher, is the use of models built based on statistical analysis or machine learning, to obtain complex forecasts based on basic models.

JOURNAL OF APPLIED SCIENCE AND SOCIAL SCIENCE
eISSN 2229-3113 pISSN 2229-3205

Volume 15 Issue 12, December 2025

Impact factor: 2019: 4.679 2020: 5.015 2021: 5.436, 2022: 5.242, 2023: 6.995, 2024 7.75

Artificial neural networks organized by principles of genetic algorithms, – heuristic search algorithms used to solve optimization and modeling tasks by randomly selecting, combining and varying the desired parameters using mechanisms similar to natural selection in nature. Pattern recognition is a method of classifying and identifying objects, phenomena, processes, signals, situations, and similar objects that are characterized by a finite set of certain properties and features. Visualization of analytical data is the presentation of information in the form of drawings, diagrams, using interactive features and animation both to obtain results and to apply them in as initial data for further analysis. A very important stage of big data analysis, which allows you to present the most important results in the most convenient way for perception

CONCLUSION

Big data processing methods allow enterprises to digitally transform production and management processes, more accurately predict the behavior of consumers, staff, investors, plan repair or replacement of equipment, software updates, optimize energy consumption, traffic, money and information flows, improve the efficiency of decisions made due to higher accuracy of forecasts of the impact of the external and internal environment on the activities of the enterprise.

REFERENCE:

1. O'.I.Begimov, T.M.Bo'riboyev / Extracting tagging from exocardiographic images via machine learning algorithmics // Analysis of world scientific views International Scientific Journal Vol 2 Issue 1 IF(Impact Factor)8.2 / 2023

2. O'.I.Begimov, T.M.Bo'riboyev / General Theory About the Traditional Methods and Algorithms of Machine Learning // AMERICAN Journal of Public Diplomacy and International Studies Volume 02, Issue 04, 2024 ISSN (E):2993-2157.

3. T.M.Bo'riboyev / Hetnet tizimi asosida avtonobillaring harakat trafigini boshqarish va tahlil qilish // Nejmettin, 03-06 Ekim 2023 tarihlerinde Erbakan Üniversitesi ve Alfraganus üniversitesi öncülüğünde düzenlenen "ipek Yolunun Ötesinde kongreler dizisi: Bir Yol, Bir Kuşak: Göç, turizm ve ekonomi politik Kongresi (SIRCON 2023)" programına sertifika almak için katıldı. (Sayfa 320-324)