ISSN: 2582-4686 SJIF 2021-3.261,SJIF 2022-2.889,
2024-6.875 ResearchBib IF: 8.848 / 2024

THE MULTIDISCIPLINARY JOURNAL OF SCIENCE AND TECHNOLOGY

VOLUME-4, ISSUE-3

PROCESSING TECHNOLOGIES BIG DATA

Sharopova Muxayyo Muxtor qizi

Asia International University

General technician Department of Sciences teacher

mukhayyosharopova4@gmail.com

Abstract : This The article is voluminous data flow with work technologies and big voluminous information sources cause passed

Key words : big data, data lake , deep teaching , with a car teaching

Enter

Big data (big data) is a very large amount of non-homogeneous and rapidly falling digital data that cannot be processed by conventional methods. In some cases, the concept of big data also includes the processing of this data. Basically, the object of analysis is called big data. The term big data was born in 2008. Clifford Lynch, the editor of Nature magazine, used the term big data in a special issue devoted to the rapid growth of the world's data volume. However, big data has been around before. According to experts, streams with more than 100 GB of data per day are called big data. Big data analysis helps to identify patterns that are beyond human perception. This makes it possible to further improve all areas of our daily life, government management, medicine, telecommunications, finance, transport, production and other areas, to increase their capabilities, to find alternative solutions to problems. The rapid growth of the volume of information (especially video, audio multimedia data) used in modern information systems presents us with new complex issues of storing and managing this information in the database. Below we will analyze the tools for organizing large multimedia data, as well as working with them. Massive multimedia data has four main characteristics: size, variety, speed, price.

1. Volume. The growing amount of data generated by humans and machines puts new demands on information technology infrastructures for data storage, processing and presentation.

2. Diversity. The information presented in different structures will be different. Whether these are credit card transactions, scientific research results, photographic images, video and audio data, they all require specific handling and storage conditions.

3. Speed. Speed refers not only to the speed at which data enters the database, but also the speed at which data is retrieved from the database.

4. Price. Big data is an expensive resource. The value of information is determined by its relevance, usefulness and content. Examples of big data

The active introduction of Big Data technologies to the market and modern life began after the well-known companies of the world began to use those that have their customers in almost every corner of the world.

It's Facebook and Google, IBM. Social giants such as Master Card, VISA and Bank of America are also financial institutions. For example, IBM applies big data techniques to its ongoing cash operations. With their help, 15% more fraudulent transactions were detected, which made it possible to increase the amount of protected funds by 60%. Also, problems with false alarms of the system were solved - their number was reduced by more than half. VISA used Big Data to track the frauds committed to carry out similar or other operations. This allows them to save more than 2 billion dollars every year. The German Ministry of Labor was able to reduce costs by 10 billion euros by implementing a big data system for unemployment benefits. At the same time, it was found that one fifth of citizens receive these benefits without any reason. Big

ISSN: 2582-4686 SJIF 2021-3.261,SJIF 2022-2.889,
2024-6.875 ResearchBib IF: 8.848 / 2024

THE MULTIDISCIPLINARY JOURNAL OF SCIENCE AND TECHNOLOGY

VOLUME-4, ISSUE-3

Data has not missed the gaming industry. Thus, the developers of World of Tanks studied information about all players and compared the current indicators of their activity. It helped predict the departure of players in the future - based on the predictions made, representatives of the organization communicated more effectively with users. Notable big data organizations include HSBC, Nasdaq, Coca-Cola, Starbucks, and AT&T. The biggest problem with big data is the cost of processing it. This can include expensive equipment and the cost of skilled professionals to handle large amounts of data. Undoubtedly, the equipment must be constantly updated so that it does not lose its minimum performance with the increase of data. The second problem is again related to the large amount of data that needs to be processed. If, for example, the study gives not 2-3, but a large number of results, it is very difficult to remain objective and select from the general flow of information only those that have a real impact on the state of the phenomenon. Big Data Privacy Challenge. As more and more customer services move to using online data, it's easy for cybercriminals to become another target. Even simple storage of personal data without any online transaction can have undesirable consequences for cloud storage customers.

The problem of information loss. Precautions are not limited to a simple single data backup, but require at least 2-3 backups. However, with the increase in volume, the difficulties associated with backup are increasing - and IT professionals are trying to find an optimal solution to this problem. What is Big Data? In fact, the definition of the term emerges: "big data" refers to the management and analysis of very large amounts of data. If you look more broadly, this is data that cannot be processed by classical methods due to its huge volume. The term Big Data itself appeared relatively recently. According to Google Trends, the active growth of this popularity coincides with the end of 2011: in 2010, the first products and solutions directly related to big data processing appeared. By 2011, the largest IT companies, including IBM, Oracle, Microsoft and Hewlett-Packard, are actively using the term Big Data in their business strategies. Gradually, information technology market analysts begin active research on this concept. Currently, this term has gained great popularity and is actively used in various fields.

But it cannot be clearly said that Big Data is a completely new phenomenon, on the contrary, sources of big data have been around for many years. In marketing, they can be called databases of customer purchases, credit histories, lifestyles, etc.

Over the years, analysts have used this information to help companies predict future customer needs, assess risks, form consumer preferences, etc. Currently, the situation has changed in two ways: - analysis of various data sets

and more sophisticated tools and methods for comparison appeared; - Analytical tools have been supplemented by a variety of new data sources associated with the widespread transition to digital technologies, as well as new methods of data collection and measurement. Researchers predict that Big Data technologies will be most actively used in manufacturing, healthcare, commerce, public administration, and various other industries and sectors. Big Data is not a specific collection of data a set of ways to process them.

Data lake (*data lake*) is a large unprocessed data warehouse.

"Lake" stores data from different sources, in different formats. This is cheaper than storing data in a fixed structure in a typical relational database. A data lake allows you to analyze data in its raw state. In addition, "lakes" can be used by several workers at the same time.

Data science (*data science*) is a science that studies problems of analysis, data processing and their presentation in digital form.

ISSN: 2582-4686 SJIF 2021-3.261,SJIF 2022-2.889,
2024-6.875 ResearchBib IF: 8.848 / 2024

THE MULTIDISCIPLINARY JOURNAL OF SCIENCE AND TECHNOLOGY

VOLUME-4, ISSUE-3

This term was born in 1974. In the same year, a Danish computer scientist, Peter Naur, published a book entitled "A Basic Principle of Data Science".

As a result of the proliferation of big data in the early 2010s, this direction has become a very profitable and promising business. And at that time, the demand for specialists working with big data has greatly increased.

The concept of data science includes all methods of data warehouse design and digital data processing. Many experts in his opinion, it is data science that is the modern successor of big data from the point of view of business.

Data mining is the intellectual analysis of data in order to find some pattern . Israeli mathematician Grigory Pyatetsky-Shapiro introduced this term to science in 1989.

Technologies, the process of finding previously unknown and useful unprocessed (raw) data is called data mining. Data mining techniques sit at the intersection of data warehousing, statistics, and artificial intelligence.

Machine learning is the practice and theory of creating self-learning programs, a large part of artificial intelligence .

Programmers teach their algorithms to identify general patterns in specific cases. As a result, the computer makes decisions based on its own personal competence, rather than on the instructions given by a human. Many such training methods can be applied to data mining.

The first definition of machine learning was given by the American computer scientist Arthur Samuel in 1959. He created the world's first self-learning chess game with elements of artificial intelligence.

Deep learning is a type *of* machine learning that creates more complex and self-taught programs. In typical machine learning cases, the computer determines the knowledge with the help of guided competence: the programmer shows certain examples to the algorithm, manually correcting the errors. In deep learning, the system itself designs its own functions, performs multilevel calculations, and makes inferences about the environment.

Deep learning is usually applied to neural networks. This technology is mainly used in image processing, speech recognition, neural machine translation, computing in pharmaceuticals and other modern technologies. Projects are mainly hosted by Google, Facebook and Baidu.

An artificial neural network is an integrated system of simple processors (artificial neurons) that imitates the human nervous system. In return for such a structure, neural networks are not programmed, they are trained. Like real neurons, processors simply receive signals and pass them on to other processors. At the same time, a whole other system performs complex tasks that algorithms cannot perform.

Used books :

1. qizi Sharopova, M. M. (2023). RSA VA EL-GAMAL OCHIQ KALITLI SHIFRLASH ALGORITMI ASOSIDA ELEKTRON RAQMLI IMZOLARI. RSA OCHIQ KALITLI SHIFRLASH ALGORITMI ASOSIDAGI ELEKTRON RAQAMLI IMZO. *Educational Research in Universal Sciences*, *2*(10), 316-319.

2. Sharopova, M. M. qizi . (2023). JAVA TILI YORDAMIDA OB'EKTGA YUNALTIRILGAN DASTURLASH ASOSLARI BILAN TANISHISH. GOLDEN BRAIN, 1(34), 111–119.

3. Sharopova, M. (2023). CHOOSE: COMPOSITION OR INHERITANCE. *Science and innovation in the education system*, *2*(13), 96-102.

4. Sharopova, M. (2023). JAVA PROGRAMMING IN THE LANGUAGE HERITAGE TO DO SYNTAX. *Current approaches and new research in modern sciences*, *2*(12), 82-87.

5. Sharopova, M. (2023). ARRAY AND ARRAYS INSTALLATION. *Development of pedagogical technologies in modern sciences*, *2*(12), 102-107.

6. Sharopova, M. (2023). CLASSES AGAIN APPLY. *Solution of social problems in management and economy*, *2*(13), 106-111.

7. qizi Sharopova, M. M. (2023). INTRODUCING" PROGRAM CONTROL OPERATORS" IN THE JAVA PROGRAMMING LANGUAGE. *Multidisciplinary Journal of Science and Technology*, *3*(5), 222-231.

8. qizi Sharopova, M. M. (2023). Working with folders in the JAVA programming language. *Multidisciplinary Journal of Science and Technology*, *3*(5), 232-236.

9. Sharopova, M. (2024). CREATION OF A DATABASE FOR THE SYSTEM PLATFORM OF NON-GOVERNMENT EDUCATIONAL CENTERS. *Current approaches and new research in modern sciences*, *3*(1), 185-191.

10. Sharopova, M. (2024). DSA ERI STANDARD. ELECTRONIC DIGITAL SIGNATURE OF GOST R 34.10-94. *Theoretical aspects in the formation of pedagogical sciences*, *3*(1), 169-178.

11. Sharopova, M. (2024). COLLECTORS.(OBJECT CONTAINERS). *Development of pedagogical technologies in modern sciences*, *3*(1), 93-101.

12. Sharopova, M. (2024). JAVA PROGRAMMING IN THE LANGUAGE FLOWING INPUT AND RELEASE. *Solution of social problems in management and economy*, *3*(1), 84-93.

13. Latipova, S. (2024). YUQORI SINF GEOMETRIYA MAVZUSINI O'QITISHDA YANGI PEDAGOGIK TEXNOLOGIYALAR VA METODLAR. SINKVEYN METODI, VENN DIAGRAMMASI METODLARI HAQIDA. *Theoretical aspects in the formation of pedagogical sciences*, *3*(3), 165-173.

14. Latipova, S. (2024, February). SAVOL-JAVOB METODI, BURCHAKLAR METODI, DEBAT (BAHS) METODLARI YORDAMIDA GEOMETRIYANI O'RGANISH. In *Международная конференция академических наук* (Vol. 3, No. 2, pp. 25-33).

15. Latipova, S., & Sharipova, M. (2024). KESIK PIRAMIDA MAVZUSIDA FOYDALANILADIGAN YANGI PEDAGOGIK TEXNOLOGIYALAR. 6X6X6 METODI, BBB (BILARDIM, BILMOQCHIMAN, BILIB OLDIM) METODLARI HAQIDA. *Current approaches and new research in modern sciences*, *3*(2), 40-48.

16. Latipova, S. (2024). 10-11 SINFLARDA STEREOMETRIYA OQITISHNING ILMIY VA NAZARIY ASOSLARI. *Академические исследования в современной науке*, *3*(6), 27-35.

17. Latipova, S. (2024). HILFER HOSILASI VA UNI HISOBLASH USULLARI. *Центральноазиатский журнал образования и инноваций*, *3*(2), 122-130.

18. Latipova, S. (2024). HILFER MA'NOSIDA KASR TARTIBLI TENGLAMALAR UCHUN KOSHI MASALASI. *Development and innovations in science*, *3*(2), 58-70.

19. Latipova, S. (2024). KESIK PIRAMIDA TUSHUNCHASI. KESIK PIRAMIDANING YON SIRTINI TOPISH FORMULALARI. *Models and methods in modern science*, *3*(2), 58-71.

ISSN: 2582-4686 SJIF 2021-3.261,SJIF 2022-2.889,
2024-6.875 ResearchBib IF: 8.848 / 2024

20. Shahnoza, L. (2023, March). KASR TARTIBLI TENGLAMALARDA MANBA VA BOSHLANG'ICH FUNKSIYANI ANIQLASH BO'YICHA TESKARI MASALALAR. In *" Conference on Universal Science Research 2023"* (Vol. 1, No. 3, pp. 8-10).

21. qizi Latipova, S. S. (2024). CAPUTO MA'NOSIDAGI KASR TARTIBLI TENGLAMALARDA MANBA FUNKSIYANI ANIQLASH BO 'YICHA TO 'G 'RI MASALALAR. *GOLDEN BRAIN*, *2*(1), 375-382.

22. Latipova, S. S. (2023). SOLVING THE INVERSE PROBLEM OF FINDING THE SOURCE FUNCTION IN FRACTIONAL ORDER EQUATIONS. *Modern Scientific Research International Scientific Journal*, *1*(10), 13-23.

23. Latipova, S. (2024). GEOMETRIYADA EKSTREMAL MASALALAR. B DEVELOPMENT OF PEDAGOGICAL TECHNOLOGIES IN MODERN SCIENCES (T. 3, Выпуск 3, сс. 163–172).

24. Latipova, S. (2024). EKSTREMUMNING ZARURIY SHARTI. B SOLUTION OF SOCIAL PROBLEMS IN MANAGEMENT AND ECONOMY (T. 3, Выпуск 2, сс. 79–90).

25. Latipova, S. (2024). FUNKSIYANING KESMADAGI ENG KATTA VA ENG KICHIK QIYMATI. B CURRENT APPROACHES AND NEW RESEARCH IN MODERN SCIENCES (T. 3, Выпуск 2, сс. 120–129).

26. Latipova, S. (2024). EKSTREMUMLARNING YUQORI TARTIBLI HOSILA YORDAMIDA TEKSHIRILISHI. IKKINCHI TARTIBLI HOSILA YORDAMIDA EKSTREMUMGA TEKSHIRISH. B SCIENCE AND INNOVATION IN THE EDUCATION SYSTEM (T. 3, Выпуск 3, сс. 122–133).

27. Latipova, S. (2024). BIR NECHA O'ZGARUVCHILI FUNKSIYANING EKSTREMUMLARI. B THEORETICAL ASPECTS IN THE FORMATION OF PEDAGOGICAL SCIENCES (T. 3, Выпуск 4, сс. 14–24).

28. Latipova, S. (2024). SHARTLI EKSTREMUM. B МЕЖДУРОДНАЯ КОНФЕРЕНЦИЯ АКАДЕМИЧЕСКИХ НАУК (T. 3, Выпуск 2, сс. 61–70).

29. Latipova, S. (2024). KASR TARTIBLI HOSILALARGA BO'LGAN ILK QARASHLAR. B CENTRAL ASIAN JOURNAL OF EDUCATION AND INNOVATION (T. 3, Выпуск 2, сс. 46–51).

30. Latipova, S. (2024). TURLI EKSTREMAL MASALALAR. BAZI QADIMIY EKSTREMAL MASALALAR. B CENTRAL ASIAN JOURNAL OF EDUCATION AND INNOVATION (T. 3, Выпуск 2, сс. 52–57).

31. Latipova, S. (2024). FUNKSIYA GRAFIGINI YASASHDA EKSTREMUMNING QO'LLANILISHI. B CENTRAL ASIAN JOURNAL OF EDUCATION AND INNOVATION (T. 3, Выпуск 2, сс. 58–65).

32. Latipova, S. (2024). BIRINCHI TARTIBLI HOSILA YORDAMIDA FUNKSIYANING EKSTREMUMGA TEKSHIRISH, FUNKSIYANING EKSTREMUMLARI. B CENTRAL ASIAN JOURNAL OF EDUCATION AND INNOVATION (T. 3, Выпуск 2, сс. 66–72).

33. Bobokulova, M. (2024). IN MEDICINE FROM ECHOPHRAPHY USE. *Development and innovations in science*, *3*(1), 94-103.

34. Bobokulova, M. (2024). INTERPRETATION OF QUANTUM THEORY AND ITS ROLE IN NATURE. *Models and methods in modern science*, *3*(1), 94-109.

35. Bobokulova, M. (2024, January). RADIO WAVE SURGERY. In *Международная конференция академических наук* (Vol. 3, No. 1, pp. 56-66).

36. Bobokulova, M. (2024). UNCERTAINTY IN THE HEISENBERG UNCERTAINTY PRINCIPLE. *Академические исследования в современной науке*, *3*(2), 80-96.

37. Bobokulova, M. (2024). BLOOD ROTATION OF THE SYSTEM PHYSICIST BASICS. *Инновационные исследования в науке*, *3*(1), 64-74.

38. Bobokulova, M. (2024). THE ROLE OF NANOTECHNOLOGY IN MODERN PHYSICS. *Development and innovations in science*, *3*(1), 145-153.

39. ***Boboqulova, M. X. (2023). STOMATOLOGIK MATERIALLARNING FIZIK-MEXANIK XOSSALARI. Educational Research in Universal Sciences, 2(9), 223-228.***

40. Xamroyevna, B. M. (2023). ORGANIZM TO ‘QIMALARINING ZICHLIGINI ANIQLASH. *GOLDEN BRAIN*, *1*(34), 50-58.

41. Bobokulova, M. K. (2023). IMPORTANCE OF FIBER OPTIC DEVICES IN MEDICINE. Multidisciplinary Journal of Science and Technology, 3(5), 212-216.

42. Khamroyevna, M. B. (2023). PHYSICO-CHEMICAL PROPERTIES OF BIOLOGICAL MEMBRANES, BIOPHYSICAL MECHANISMS OF MOVEMENT OF SUBSTANCES IN THE MEMBRANE. Multidisciplinary Journal of Science and Technology, 3(5), 217-221.

43. Bobokulova, M. K. (2024). TOLALI OPTIKA ASBOBLARINING TIBBIYOTDAGI AHAMIYATI. GOLDEN BRAIN, 2(1), 517–524.

44. Boboqulova, M. (2024). FIZIKA O`QITISHNING INTERFAOL METODLARI. B CENTRAL ASIAN JOURNAL OF EDUCATION AND INNOVATION (T. 3, Выпуск 2, сс. 73–82).

45. Boboqulova, M., & Sattorova, J. (2024). OPTIK QURILMALARDAN TIBBIYOTDA FOYDALANISH. B INNOVATIVE RESEARCH IN SCIENCE (T. 3, Выпуск 2, сс. 70–83).

46. Boboqulova, M. (2024). FIZIKAVIY QONUNIYATLARNI TIRIK ORGANIZMDAGI JARAYONLARGA TADBIQ ETISH . B MODELS AND METHODS IN MODERN SCIENCE (T. 3, Выпуск 2, сс. 174–187).

47. Boboqulova, M. (2024). IONLOVCHI NURLARNING DOZIMETRIYASI VA XOSSALARI. B DEVELOPMENT AND INNOVATIONS IN SCIENCE (T. 3, Выпуск 2, сс. 110–125).

48. Boboqulova, M. (2024). KVANT NAZARIYASINING TABIATDAGI TALQINI. B ACADEMIC RESEARCH IN MODERN SCIENCE (T. 3, Выпуск 7, сс. 68–81).

49. Komilov, O. S., & Sayfulloev, S. S. (2024). HORIZONTAL AND VERTICAL LOOPS GEOTHERMAL HEATING SYSTEM. *Educational Research in Universal Sciences*, *3*(2), 384-391.

50. Sayfulloev, S. S. (2023). HEAT-TECHNICAL CHARACTERISTICS OF HEAT PUMP DEVICE FOR HEAT SUPPLY SYSTEMS. *GOLDEN BRAIN*, *1*(34), 91-101.

51. Komilov, O. S., Sayfulloev, S. S., & Rustamova, F. R. (2022). CALCULATION OF THE THERMODYNAMIC CYCLE OF A VAPOR COMPRESSION HEAT PUMP INSTALLATION WITH A SUBCOOLER FOR HEATING AND HOT WATER SUPPLY. *Центральноазиатский журнал образования и инноваций*, *1*(1), 43-47.

52. Sayfulloev, S. S. (2023). HEAT-TECHNICAL CHARACTERISTICS OF HEAT PUMP DEVICE FOR HEAT SUPPLY SYSTEMS. *GOLDEN BRAIN*, *1*(34), 91-101.

53. Komilov, O. S., Sayfulloev, S. S., & Urinov, S. (2021). Analysis Of Energy of Heat Pump Heating System with The Environment. *Texas Journal of Multidisciplinary Studies*, *3*, 230-233.