(IJRST) 2024, Vol. No. 14, Issue No. 3, Jul-Sep **e-ISSN: 2249-0604, p-ISSN: 2454-180X**

Big Data Analysis-Based Application Scenarios and Data Science Practice Fundamentals

Vadlamudi Vishnu Vardhan

Narayana Junior College, Andhra Pradesh

DOI:10.37648/ijrst.v14i03.004

[1]*Received: 04 May 2024; Accepted: 10 August 2024; Published: 23 August 2024*

ABSTRACT

One significant aspect of China's urbanisation is the simultaneous rise of industrialisation, informatisation, and urbanisation. The development of big data technology and data science majors in China is currently a trending issue. In light of the "big data" era, the state has adopted data science and big data technology as majors in recent years. The goal is to develop highly skilled big data workers using analytical application technologies, big data thinking, and big data thinking. There are various personnel training programs for data science and big data speciality due to the growing significance of these fields. Combining hardware and software, data science and big data technology are primarily based on computing technology and are distinguished by these two fields. This paper presents the problems and useful approaches that big data brings to data science's data analysis tools, as well as an appraisal of the field's current state and application scenarios.

INTRODUCTION

China's industrialisation, informatisation, and urbanisation have all advanced quickly since the reform and opening up. All three are developing at the same time and support one another. Data is rapidly becoming a basic strategic resource for the country due to the economy and information technology's rapid expansion. Big data has permeated every sector of the economy and has a significant influence on national governing ability, social lifestyle, economic system functioning, worldwide production, and consumption [1]. Data has become a crucial component of manufacturing and has permeated many different sectors and company functional areas. Massive data is being mined and applied by people, which portends the coming of a new wave of consumer surplus and productivity growth [2]. Computer networks and communication networks for the quick and secure transfer of information have started to take shape with the advancement of information collection technologies. A growing variety of real-time data types are being collected, demonstrating the "high volume," "fast speed," "diversity," and "low value density" attributes of big data [3]. Data science and big data technology is a broad major that combines hardware and software, is based on computing technology, and includes both fields of study.

It is a fundamental concept in large data analysis. Building a national big data system is a realistic and practicable path to social governance, economic development, and national security against the backdrop of the major global forces evolving, the big data industry growing, and the uneven development of traditional industries in the field of economics [4].

Though it is still in its infancy, data science is an emerging cross-disciplinary field that encompasses numerous disciplines, including economics, network science, and information science. Based on the most recent developments in artificial intelligence, it appears that big data and artificial intelligence algorithms are being combined to create a new trend in artificial intelligence development. Learning from big data can help classical algorithms overcome their knowledge bottleneck [5]. Computer science, technology, and statistics are the core fields of data science and big data technologies.

The objective is to systematically train students to become experts in statistical concepts and computer science, enabling them to use big data for prediction and analysis of particular application scenarios as a decision control approach. of elite engineering skills [6]. In China, majors in big data technology and data science are becoming the

[1] How to cite the article: Vardhan V.V.; August 2024; Big Data Analysis-Based Application Scenarios and Data Science Practice Fundamentals; *International Journal of Research in Science and Technology*, Vol 14, Issue 3, 28-32, DOI: http://doi.org/10.37648/ijrst.v14i03.004

most sought-after new degrees. All universities, however, deal with the same issue of how to design the major in accordance with regional conditions [7]. Data is developing and expanding quickly, and it will shape how the organisation develops in the future. Universities are entrusted with the pressing responsibility of developing big data skills relevant to industry and areas. In order for students majoring in data science and big data technology to succeed, educational institutions and instructors must have a thorough understanding of the big data major's theoretical underpinnings, methodology, and cognitive and educational laws [8]. People's awareness of the value of data to the business will grow with time. This study presents the problems and useful approaches that big data brings to data science's data analysis tools by analysing the field's development state and application scenarios.

SYSTEM OF KNOWLEDGE FOR DATA SCIENCE AND SPECIALISED IN BIG DATA

Big data, a crucial component of data science, presents both enormous obstacles and fantastic potential for the advancement of science and education. It also puts forwards-thinking scientific endeavours. Due to the widespread usage of the Internet, mobile phones have become broadband and digital, and data is expanding exponentially. The creation of urban public policies and plans, their coordination for implementation, oversight of the process, correction of policy irrationality, and assessment of the policies' outcomes comprise the urban management process.

As we move into the big data era, both the volume and variety of data are growing at a rapid pace. For instance, the data from digital phones will be in voice format, whereas the data from banks and supermarkets would be in text format [9]. The operation and upkeep of big data is essential to the data platform's support. The primary sources of big data are the information produced by electronic devices and the Internet during daily operations, as well as the data gathered about a variety of seen objects in conjunction with the development of the Internet of Things, digital cities, and smart cities.

All levels of urban government have acknowledged the significant benefits that digital cities, intelligent cities, digital city management, and other information technology-dependent urban development and management projects have demonstrated in enhancing urban management. The function of information technology in urban management is becoming more and more evident. Big data storage will require a significant amount of resources if standard data storage technology is used. Consequently, it is essential to integrate big data's features and implement fresh approaches to big data storage. Make sure big data information is stored quickly and reliably. The structure of online financial monitoring is redesigned for big data in Figure 1.

Figure 1 Big data reshapes the structure of Internet financial supervision

Big data applications are currently moulded by Internet firms, which possess vast amounts of data assets, in contrast to other industries. These online businesses are keen to enhance user experience and convert big data insights into a valuable product. The data form comprises original, unstructured, and semi-structured data from emails, social media forums, online pages, search indexes, internet log files, and other sources in addition to the conventional relational data. Determine which product features are most appropriate to meet user needs through big data analysis. This will help with product design and development. Additionally, once the business goes online, it will track and analyse user issues related to online ordering and usage on a regular basis, providing data support for optimising business strategies. Data analysis tools are needed to naturally integrate the techniques for processing structured data with the techniques for handling newly acquired unstructured data because of the volume and variety of data being handled. A lot of big data applications are real-time in nature. In this instance, algorithm accuracy is no longer the key indicator of big data

application. Real-time and precision must be balanced in the algorithm. In the big data era, data collection technology will continue to be developed on the basis of established, workable traditional technology in order to create new intelligent identification technologies and create a system of data collecting technology that balances technological and human intelligence.

CONSIDERING THE APPROACHES OF DATA SCIENCE DEVELOPMENT AND DISCIPLINE CONSTRUCTION

A. Selection of Practical Teaching Mode

The emergence of data science is a result of the growth of large data. Data science is a discipline that is still in its infancy and is always evolving. The three main components of big data application are also the primary research subjects of data science: data extraction, data storage and processing capabilities, and data collection. At its foundation, data science is the automated analysis of large amounts of data to derive insights and patterns. From the perspectives of transmission range and transmission speed, the rising breadth, quantity, and speed of data collecting creates new demands for data transmission technologies.

Computer large data has reached a new technological end and area following constant evolution and upgrading, and a unique computer new technology has even become available. There are an increasing number of tools and techniques that online learning uses to support student learning. As illustrated in Figure 2, a popular Internet learning model can be compared to a "catalytic reaction".

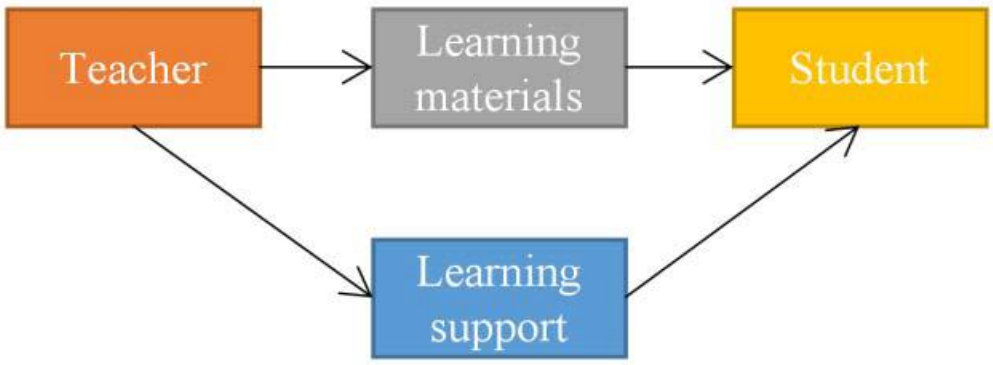

Figure 2 Catalytic reaction mode of Internet learning

Data science's practical applications encompass data gathering, data cleansing, data analysis, data visualisation, and the entire iterative data application process, which in the end aids businesses in making wise development decisions. Data scientists are those who work in the field of data science. Broad vision and complex talents characterise data scientists. They possess a strong foundation in data science, including computer science, statistics, and mathematics, in addition to a broad spectrum of commercial expertise. Data science offers an effective new method for exploring discovery in practically every field of knowledge discovery. This opens up new avenues of insight for businesses that have a lot of data but don't know how to use it to their advantage.

The entire data life cycle—including data collection, analysis, administration, and visualization—as well as the extensive applications of data across all domains are covered by data science. Traditional data types, such as structural data, typically have structure before data. Unstructured data is unable to construct a clear data structure, whereas semi-structured data often consists of data first and structure second.

B. Construction of Practical Teaching System

The primary source of massively organised, semi-structured, and unstructured data for big data collection is the Internet of Things, social networks, mobile Internet, and other channels that tens of thousands of users can access and utilise concurrently. The general education curriculum, fundamental subject curriculum, professional core curriculum, and centralised practice curriculum are all included in the curriculum framework for the major in data science and big data technology. Hundreds of huge servers are insufficient to address an enterprise's data storage demands in the big data scenario due to the rapid growth of data volume and the heightened demand for storage capacity [11].

The fundamental and operational data of every city component serves as the foundation for the comprehensive analysis of the city system and is the source of big data. The primary knowledge structure of data science and engineering, a new multidisciplinary field in the big data era, originates from computer science, applied mathematics, information systems, and information management. Because a huge number of users' needs must be met

simultaneously by the big data system, and it must offer effective services for the processing and analysis of massive data sets. In order to design a big data information system, we must address the premise's real demands. For system delay information for various network settings, see Table I.

Network parameter	Minimum value	Maximum value	Median	Average value
0.6	5.94	6.89	6.42	6.59
0.8	3.38	7.26	5.51	4.82
0.1	1.87	6.82	4.76	5.18

To increase the topology's compactness, each service circuit set must transit as few communication nodes as possible. The relationship between the original and anticipated data is depicted in Figure 3.

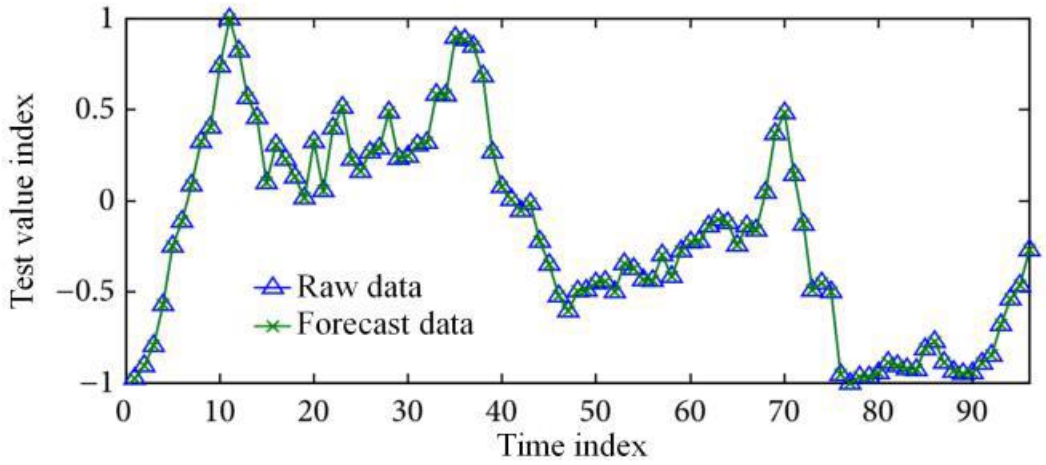

Figure 3 Raw data and forecast data

A distributed real-time data management system made of multi-dimensional tables and orientated to column storage is required to organise and manage data because traditional single-table data description cannot match the demands of the big data era. The key feature of data science and engineering is its great comprehensiveness, which necessitates that students acquire knowledge via in-depth research and build a fundamental system architecture using expert knowledge and skill [12]. The most important step in the big data processing process is data analysis. One way to determine the value of large data is to calculate and analyse its content. A practical teaching system must be built using a variety of professional skills and real-world experience while adhering to the relevant teaching guidelines. Data mining in big data scenarios requires that structured, semi-structured and unstructured data can be analyzed together. When the amount of data increases, only distributed service nodes are required to be added.

CONCLUSION

In the big data era, people will have greater opportunity to play the important role that science and technology play in fostering social progress if data can be better organised and utilised. With time, more and more people come to understand the value of data to businesses, even though many may not be aware of the hidden risks posed by the rapid growth of data. The rise of the Internet of Things and artificial intelligence is greatly aided by big data technology, which connects, transforms, and merges everything into digital codes. Colleges and universities should create a curriculum that is both scientific and forward-looking in order to meet the demands of practical application. They should also take into account the needs of the industry and use the interdisciplinary basis as support. Finally, they should adjust to the new era of young students who need to understand the laws of objective things. A practical teaching system must be built with professional expertise and hands-on experience in various areas, adhering to the relevant teaching guidelines. Big data scenarios involving data mining necessitate the joint study of structured, semi-structured, and unstructured data. All nations must cooperate in the big data future. We must examine the wide development opportunities presented by big data while keeping data security and privacy protection in mind.

REFERENCES

[1] Zhou Aoying, Qian Weining, Wang Changbo. Data Science and Engineering: Emerging Interdisciplinary Disciplines in the Era of Big Data. Big Data, no. 02, no. 90-99, 2015.

[2] Yang Jing, Wang Xiaoyue, Bai Rujiang, et al. Status and development trend of data science analysis tools in the context of big data. Information Theory and Practice, vol. 038, no.003, no.134-137, 144, 2015 .

[3] Zhu Yangyong, Xiong Yun. Training of data scientists in the era of big data. Big Data, vol. 000, no. 003, no. 106-112, 2016.

[4] Guo Yike. Big data; data science; data-driven scientific research. Bulletin of the Chinese Academy of Sciences, no. 6, no.599-607,2016.

[5] Liu Lei. From Data Science to the Fourth Paradigm: The Scientific Origin of Big Data Research. Advertising Concept: Theoretical Edition, vol. 000, no. 002, no. 44-52, 2016.

[6] Li Hui, Zhang Biao. Thinking on the cultivation of data science and big data technology professionals in agricultural universities. Higher Engineering Education Research, no. 5, no. 16-22, 2019.

[7] Yang Qianqian, Wang Long. An Analysis of the Talent Cultivation Model of Data Science and Big Data Technology Specialty in Independent College Undergraduate Colleges. Bohai Rim Economic Observation, vol. 000, no. 005, no. 116-117, 2019.

[8] Yang Yin, Huang Yunqing, Liu Shaoyue. Research on the training model of data science and big data technology professionals in local colleges and universities. Education Modernization, vol. 6 (04, no. 23-25, 2019.

[9] Wang Zhong, Shi Peibei. Research on Data Science and Big Data Technology Specialty Construction. Journal of Anqing Normal University (Natural Science Edition), vol. 25, no. 01, no. 123-126, 2019.

[10] Cui Chen, Yao Li. Research on the Course System Construction of Data Science and Big Data Technology in Public Security Colleges. Computer Age, vol. 312, no. 06, no. 92-94 + 97, 2018.

[11] Liu Minrui. Data Science and Engineering: Emerging Interdisciplinary Disciplines in the Era of Big Data. Smart City, vol. 004, no. 019, no. 123-124, 2018.

[12] Geng Xiurong. Guiding ideology and basic measures of data science and big data technology specialty construction. Journal of Mudanjiang Education College, vol. 193, no. 10, no. 22-23, 2018.