IMPLEMENTATION OF BIG DATA ANALYSIS ON POPULATION GROWTH IN GOVERNMENT SECTOR

[1]S. Hardika, [2]S. Akshaya, [3]G. Anitha

[1,2] I MSc DA, Department of computer Science (PG), PSGR Krishnammal College for Women, Coimbatore, India,

[3]Assistant Professor, Department of computer Science (PG), PSGR Krishnammal College for Women, Coimbatore, India.

ABSTRACT

In this paper we are going to predict the population growth of India for the year 2021. This journal is a kind of survey that makes predictions of the population in 2021 using the available big data. Also, the paper has incorporated various data preparation, survey and also has some critical applications for government sector. In general, the census survey in India is taken in a period of once in 10 years, so this journal aims to predict the population approximately on the basis of big data, well in advance of the government survey.

I INTRODUCTION

Data are the facts and references collected together for analysis. The most common myth associated with Big Data is that it is just about the size or volume of the data. But actually, it is not just about the big amounts of data being collected. Big Data refers to the large amounts of data which is obtained from various data sources and are available in different formats. Even before there was huge data which were being stored in databases, but because of the changed nature of the informations, the traditional relational database systems are incapable of handling this Data. Big Data is a variety collection of datas with different format, it is an important asset which can be used to obtain numerous benefits. Various big-data analytics tools and techniques have been developed for handing this huge amount of data. The three different big data formats are structured, semi-structured and unstructured. The big data is used in various fields such as education, transportation industry, banking sector, government sector, media and entertainment industry, weather patterns.

III GROSS DOMESTIC PRODUCT

Gross Domestic Product (GDP) is that the total finance value of all the finished products and the services made within a border of nation in a particular fundamental quantity. As a broad measure of overall domestic production, it functions as a extensive catalogue of the country's productive health. There are four components of GDP ,its growth rate is driven by the, The main driver of gross domestic product growth is personal consumption which includes the condemnatory sector of retail sales . The second element is business investment, as well as construction and inventory levels. The third component of growth is the Government spending. Its largest classes area unit Social Security advantages, defence spending and Medicare benefits. The government typically will increase defrayment to jump-start the economy throughout a recession. Fourth is by net [8].

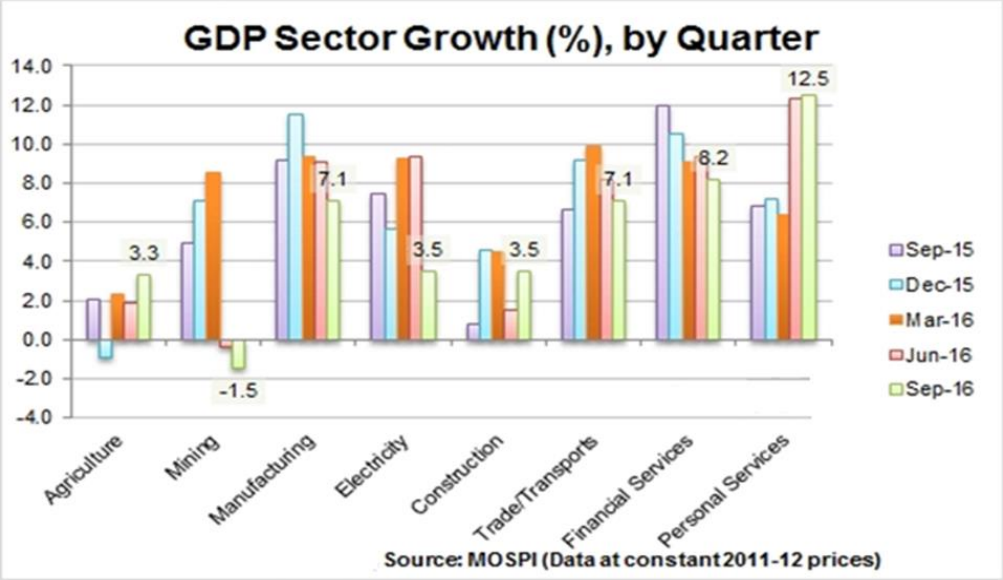

Fig 1.3: Gross domestic product

(It gives the GDP sector growth in percentage for various fields)

III DECADAL GROWTH DATA

Table 1: Data's of population growth

Category	Country/ States/ Union Territories Name	Growth Rate - Total 1991-01	Growth Rate - Total 2001-11	Growth Rate - Rural 1991-01	Growth Rate - Rural 2001-11	Decadal Growth Rate - Urban 1991-01	Decadal Growth Rate - Urban - 2001-11
Country	India	21.5	17.7	19.2	12.3	32.6	31.8
State	Andhra Pradesh	14.6	11	13.9	1.7	16.3	35.6
State	Gujarat	22.7	19.3	17.3	9.3	32.9	36
State	Karnataka	17.5	15.6	12.3	7.4	29.1	31.5
State	Kerala	9.4	4.9	10.1	-25.9	7.6	92.8
State	Tamil Nadu	11.7	15.6	-5.1	6.6	44.1	27
Union Territory	Puducherry	20.6	28.1	12	21.3	25.5	31.5

This table gives data about urban and rural decadal growth rate for the years 1991-2001,2001-2011 for different states and union territories in India.[3]

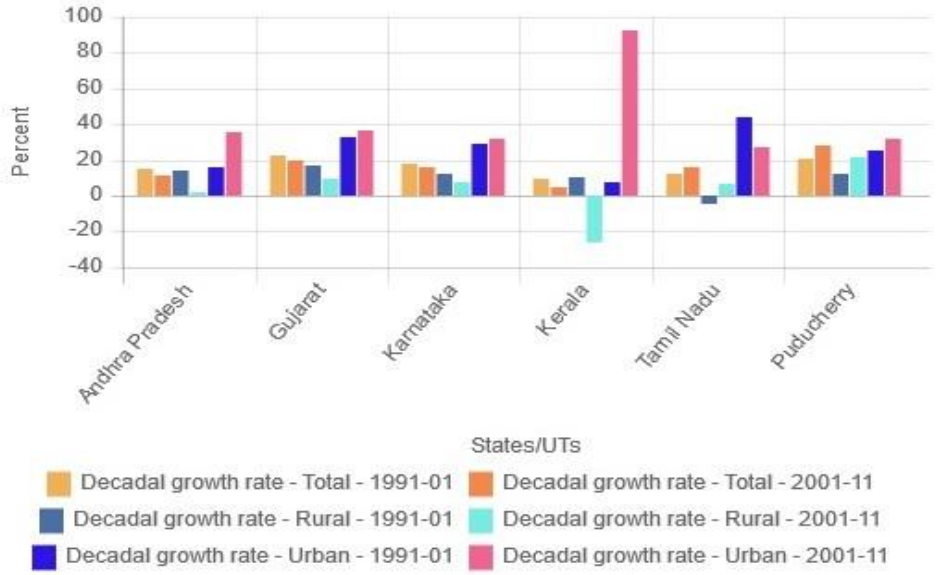

Fig 1.1- Decadal growth rate

(This is the corresponding graphical representation of the above mentioned data)

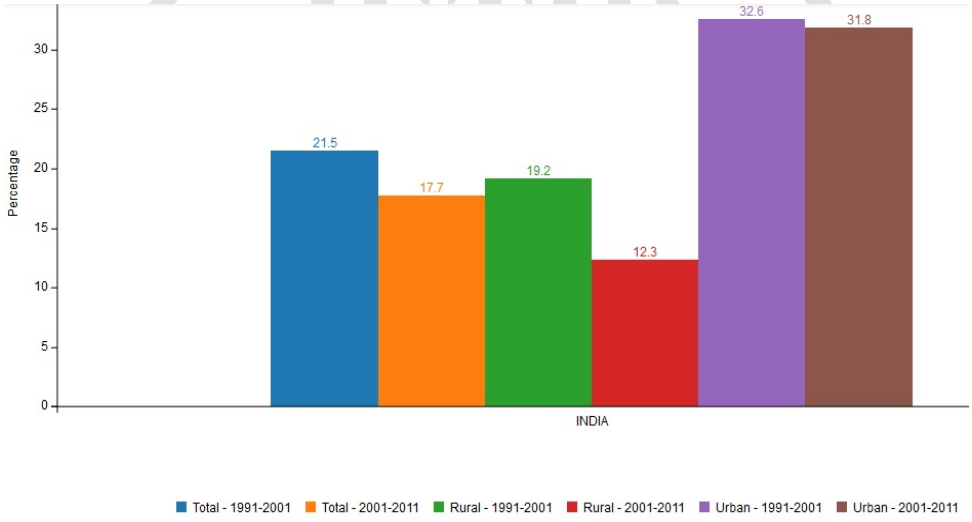

Fig 1.2-Percentage of growth

(It gives the percentage of rural growth rate ,urban growth rate and total growth rate for two decades)

IV BIG DATA STORAGE

Another aspect of BDE is the Big Data storage, which is provided either by the distributed file systems or emerging NoSQL databases. These technologies are briefly discussed in the following subsections.

1)Distributed File Systems: In this subsection, we are discussing two competing distributed file systems, namely HDFS and Tachyon. **Hadoop Distributed File System (HDFS)—**HDFS is suitably designed for managing the larger datasets. It is designed specifically to work with a cluster of commodity servers. Since the chances of hardware failure are higher in such settings, it provides greater fault tolerance for hardware failures. Data distribution and replications are the key traits of HDFS to achieve the fault tolerance and high availability. There are however situations when the usage of HDFS degrades performance, particularly in applications requiring low-latency data access .**Tachyon** is the BDAS flagship distributed file system that extends HDFS and provides access to the distributed data at memory speed across the cluster. Some of the features where Tachyon has outsmarted HDFS include: In-memory data caching for enhanced performance and Backwards compatibility to work seamlessly.

2) NoSQL Databases: Relational databases are found unsuitable for these applications due to their specialized storage and processing needs. Consequently, new systems came into being called "Not only SQL" (NoSQL) systems to fill this technological gap. NoSQL systems improved traditional data management in

numerous ways. More importantly, NoSQL systems eschew the rigid schema. Oriented storage in favour of schema-less storage to achieve. Flexibility Today these systems are prevalent in myriad Data-intensive applications in many industries. [7]

V BIG DATA IN GOVERNMENT SECTOR

1.HEALTH CARE INDUSTRY

The field of big data analytics is growing and has the potential to provide useful information for the healthcare system. Big data can improve healthcare reduce its cost, while supporting advanced patient care, improving patient outcomes, and avoiding unnecessary costs. Big data analytics is currently used to predict the outcomes of decisions made by physicians, the outcome of operations based on patients age, current condition, and health status. Essentially, we can say that the role of big data in the health sector is to manage data sets related to healthcare, which are complex and difficult to manage using current hardware, software, and management tools. The storage of healthcare related data occurs in a various form. A sudden increase in data related to healthcare informatics has also been observed in the field of bioinformatics, where many terabytes of data are generated by genomic sequencing. The frameworks available for the analysis of healthcare data are as follows:

Predictive Analytics in Healthcare: For the past two years, predictive analysis has been recognized as one of the major business intelligence approaches, but its real-world applications extend far beyond the business context. Big data analytics includes various methods, including text analytics and multimedia analytics. However, one of the most important categories is predictive analytics which includes statistical methods like data mining and machine learning that examine current and historical facts to predict the future. This data can help doctors to make important patient care decisions. Predictive analysis requires an understanding and use of machine learning, which is widely applied in this approach.

Electronic Health Records: EHR represents the most widespread health application of big data in healthcare. Each patient has his/her own medical records, with details that include their medical history, allergies diagnosis, symptoms, and lab test results. Patient records are shared in both public and private sectors with healthcare providers through a secure information system. These files can be modified, and the doctors can make changes any time and add new medical test results, without the need of any paper work or duplication of data.[1]

2.AGRICULTURE

Big data in agriculture is used for the following purposes:

Yield prediction: Yield prediction uses mathematical models to analyse data related to yield, weather, chemicals, leaf and biomass index among others, while machine learning is used to crunch the stats and make powerful decisions. Predicting yields in this way allows a farmer to extract insight on what to plant, where to plant and when to plant. The collection of data related to humidity, temperature and chemicals will present a picture of health around smart agricultural businesses.

Ideal Crops and Consumer Expectation: Using big data farmers can precisely see how much they produced in previous years, what was the customer impact, how this affected supply and demand and the tips for ways to improve their operations. They could cut excess waste by producing fewer crops for a lower demand season to save both money and space to grow alternatives. List of areas where Big Data technologies can impact an agribusiness is, Weather data, Improved forecasting of yields and production, Better optimized seeds and livestock and new methodologies that improve yields and production ,Faster delivery of goods produced to distribution centre and consumers, Real-time decisions and alerts based on data from fields and equipment, Integrated production and business performance of data for improved decision making .Rationalized performance data across multiple geographies.[2][4]

3. INTELLIGENT TRANSPORTATION SYSTEM

Big data is becoming a research focus in intelligent transportation systems around the world. Intelligent transportation systems will produce huge amount of data. The produced big data will have huge impacts on the design and application of the intelligent transportation systems, which makes intelligent

transport system safer, more efficient, and profitable. Studying big data analytics in Intelligent transport system is a flourishing field .

Big data collection in intelligent transport systems is done through numerous ways are as follows ,Big Data from Smart Cards ,Big Data from GPS ,Big Data from Videos ,Big Data from Sensors ,Big Data from CAV and VANET ,Big Data from Passive Collection

Application of big data in intelligent transport system are Road Traffic Accidents Analysis ,Road Traffic Flow Prediction, Public Transportation Services Planning ,Rail Transportation Management and Control ,Personal Travel Route Planning[5].

4. BANK DATA MANAGEMENT

Banks takes the most benefit from big data as they now can extract information quick and easy from their data. Internationally banking is beginning to harness the power of data in order to derive utility across various spheres of their functions, ranges from sentiment analysis, product cross selling, regulatory compliances management, risk management, financial crime management and much more .Indian banks area unit catching up with their international counterparts, however a lot of scope remains.

4.1 Impact on Banking Institutions: Big data is the tool which allows an organization to create, manipulate, and Manage very large datasets in a given timeframe and the space required to store the data is characterized by variety, volume and velocity. Their enterprise risk management framework helps to improve transparency, auditability, and executive Oversight of risk.

4.2 Customer Centric: Next offer is Real time allocation based on the offerings. Sentiment-enabled strategy management, lead/referral services. Quality of lead analytics Micro-segmentation.

4.2 Risk Management: Following are the ways in which data analysis is being used to find out and evaluate financial crime management (FCM) by early detection of the correlation between the series of transactions**.** a) MIS/ Regulatory reporting – Disclosure reporting B) Real time keyboard conversation tracking- Anti-money laundering**.**

4.3 Transactions: Transactions and mercantilism, when followed over a period of time, tend to reveal a lot of information about the, Nature of trade, log analytics, trading sentiments and other aspects**.** Banks and other financial institutions leverage.

4.4 Future scope: Big data analytics helps them to Deliver higher services to their customers, each internal and external, also serving to them improve on their active and passive security systems .Banks strengthen data security and prevent any type of attack. Bank reaped after their implementation of Big Data Analytics and predict the improvements in financial statements of the bank. This work may be extended to hide the varied data processing techniques that may be employed by banks to enhance the analysis quality**.**[6]

CONCLUSION

Thus from this paper and from the available previous big data record of surveys we have approximately predicted the population growth. India's decadal growth rate was 21.5% and 17.7% during 1991-01 and 2001-11 respectively. During 1991-01, there was a decadal growth rate of 19.2% in rural part of India. During 2001-11, there was a decadal growth rate of 12.3% in rural part of India. During 1991-01, there was a decadal growth rate of 32.6% in urban part of India. During 2001-11, there was a decadal growth rate of 31.8% in urban part of India. There were various big data applications discussed on the available government sectors. On Concluding with the key aspects like the security, services, storage infrastructure, networking, data centre infrastructure, discovery tools, data applications and analytics this Big data plays a vital role in internet world today.

REFERENCE

[1]. Kumar, Sunil, and Maninder Singh. "Big data analytics for healthcare industry: impact, applications, and tools." *Big Data Mining and Analytics* 2.1 (2018): 48-57.

[2]. https://eleks.com/blog/practical-uses-of-big-data-in-agriculture/

[3]. https://data.gov.in/

[4]. Rubrics, Pal. "Big Data and its impact on agriculture." *Eco cycles* 2.1 (2016): 33-34.

[5]. Zhu, Li, et al. "Big data analytics in intelligent transportation systems: A survey." *IEEE Transactions on Intelligent Transportation Systems* 20.1 (2018): 383-398.

[6]. Srivastava, Utkarsh, and Santosh Gopalakrishnan. "Impact of big data analytics on banking sector: Learning for Indian banks." *Procedia Computer Science* 50 (2015): 643-652.

[7]. Sagiroglu, Seref, and Duygu Sinanc. "Big data: A review." *2013 International Conference on Collaboration Technologies and Systems (CTS)*. IEEE, 2013.

[8]. Leimbach, Marian, et al. "Future growth patterns of world regions–A GDP scenario approach." *Global Environmental Change* 42 (2017): 215-225.