



STUDY OF ARTIFICIAL INTELLIGENCE TECHNIQUES TO PERMIT THE AUTOMATED BIG DATA ANALYSIS IN EGOVERNANCE PROJECTS TO PROMOTE DIGITALIZATION

Computer Science

Dr. Bankim L. Radadiya

Associate Professor, Navsari Agricultural University, Navsari, Gujarat, India

Dr. Parag C. Shukla*

Assistant Professor & HEAD, Department of MCA, Atmiya Institute of Technology & Science, Rajkot, India *Corresponding Author

ABSTRACT

Our today's journey is towards digital India. Our government is also promoting the same. We know the days when bank used to store customer information such as names, photograph and signature in individual postcard like data sheet. That was the era where thick registers were used in different government offices like post office, property tax collection centers etc. to store customer's details or maintain the daily attendance records of employees. If employee had to update any of the registered customer's details, the task could take up the whole day. Hours were wasted searching for the customer's details and then creating a new record to replace the old one. The customers too, had to wait for hours for such minor tasks to be completed. Apart from the tediousness of searching for data from piles of ledgers, such paper files could be lost at any time due to disaster like floods or fire, apart from the degradation of the very paper on which the data was recorded. Above is the only one of the example rather than every government sector unit, public sector unit, semi government sector unit all was used manual recorded system. The data has also become too difficult to be stored, handled, analyzed and managed with normal data tools. So, different techniques of big data are now analyzed by computer systems to reveal specific trends, patterns and associations, especially those relating to human behaviors and interaction. So, making our machines smart enough to 'think and decide' on what action need to be performed and when by using Big Data techniques like predictive analytics, user behaviors analytics etc.

KEYWORDS

e-Governance Project, digitalization, Elicitation technique, Artificial Intelligence, Big data, Machine Learning

INTRODUCTION

Today, our journey towards a digital India, all government offices are switching towards to digitization instead of manual record keeping. As we proceed along this path, we see a tremendous increase in the size of data. There are around 230 billion tweets posted on Twitter every day, every hour 112 million comments and likes posted to Facebook, and around 10 hours video uploaded to YouTube each second. All this leads to about 2.5 Exabyte's of data being generated daily by different online application, transactional data source, IoT devices etc. The term that encapsulated such immense volumes of information is Big Data. Existing hardware and software systems are unable to handle such volumes of dissimilar types of data being generated at such enormous speed. The data has also become too difficult to be stored, handled, analyzed and managed with normal data tools. So big data is now analyzed by computer systems to reveal specific trends, patterns and associations, especially those relating to human behaviors and interaction.

Our today's journey is towards digital India. Our government is also promoting the same. We know the days when bank used to store customer information such as names, photograph and signature in individual postcard like data sheet. That was the era where thick registers were used in different government offices like post office, property tax collection centers etc. to store customer's details or maintain the daily attendance records of employees. If employee had to update any of the registered customer's details, the task could take up the whole day. Hours were wasted searching for the customer's details and then creating a new record to replace the old one. The customers too, had to wait for hours for such minor tasks to be completed. Apart from the tediousness of searching for data from piles of ledgers, such paper files could be lost at any time due to disaster like floods or fire, apart from the degradation of the very paper on which the data was recorded. Above is the only one of the example rather than every government sector unit, public sector unit, semi government sector unit all was used manual recorded system. The data has also become too difficult to be stored, handled, analyzed and managed with normal data tools. So, different techniques of big data are now analyzed by computer systems to reveal specific trends, patterns and associations, especially those relating to human behaviors and interaction. So, making our machines smart enough to 'think and decide' on what action need to be performed and when by using Big Data techniques like predictive analytics, user behaviors analytics etc.

WHY ARTIFICIAL INTELLIGENCE?

Major organizations have promoted from the automation of big data analysis. Subject to technology applied, it can take a few days to

process, evaluate, and understand any volume of big data. In all these regards, automation has advantages like decreasing the operational costs, improving operational effectiveness, better self-service modules, and better the scalability of big data technologies. For example, it can be gathering as a numerical identifier successful across the data tables in industry of e-commerce. Also, it looks for categorical data to generate the set of features having interrelated values. Automated analytics is destined to go further faster. Automated recommendations are the latest development, and these will continue to gather speed and sophistication in the months and years to come. Decision-makers will no longer have to comb through data and make subjective decisions based on their own interpretations. Instead, they'll get a list of recommendations based on evident trends in the data. That will further reduce the time it takes to travel from problem to solution and will mean marketers will be able to create and implement campaigns now of peak relevance. It's the next logical step in the evolution of automated analytics tools and will mean brands can reach consumers within ideal timeframes.

Automatic data analysis is become demand or primary need for any organization. One of the solution for implementation of automatic data analysis is used for artificial intelligence. So next question is why is AI is important? Artificial intelligence works by merging huge volumes of data with repeated processing and intelligent algorithms, permitting the software to learn automatically from data patterns or features available in the huge volumes of data.

Following are the reason why AI is important.

1. AI automates repetitive learning and discovery through data.
2. AI adds intelligence
3. AI adapts through progressive learning algorithms
4. AI analyzes more and deeper data
5. AI achieves incredible accuracy
6. AI gets the most out of data

WHY BIG DATA?

"Market size of big-data services increased over 58%, to around 18.6 billion dollars, in 2013. Companies and organizations have been competing intensively to develop technologies for analyzing big data" [1][2]. Big data is analyzed, the number of assumptions to be considered rises exponentially, so all possible cases are difficult to analyze. Hence, it is necessary to develop a computer system that can automatically analyze all the cases at one time.

In recent times, several kinds of "analysis support systems" have been

developed. For example, technologies for supporting decision making that used to be done by people (by analyzing big data collected in the company) have been developed by IBM, GE, Siemens, and other organizations. [3][4].

In the United Kingdom, a project for developing a system that understand & recognizes the meaning of requests (posted in a natural language by end users) to extract data from a database has been initiated [5]. "This system is higher in terms of handling data using natural language". The Center for Mathematics and Computer Science (CWI) is developing a system, called "Data Cell," which can analyze data online [6]



Figure-1 Benefits of Big Data

DESIGN, METHODS & PROCEDURE

Today businesses are grappling with a big question: How can we manage our data? While collecting it is easy, analyzing and understanding big data is more complicated. There's internal data, which could be structured data coming from charts and databases, or unstructured data, like metadata or text from emails. There's social media data, data collected from wearable's or other IoT, data from web traffic—the list goes on and on. Finally, there's "big data"—the sweet spot that gathers all the various strands together for a complete picture of what's going on.

AI is a wide area of study that consist of many theories, methods and technologies, as well as the succeeding key subfields.

MACHINE LEARNING:

It automates building an analytical model. It uses methods from physics, operations research, statistics and neural networks to search hidden intuitions in data without clearly being programmed for where to look or what to conclude.

NEURAL NETWORK:

It is a type of machine learning that is made up of interrelated elements (like neurons) that processes data by responding to outside inputs, imparting information between every element. The process needs several passes at the data to find relations and derive meaning from indeterminate data.

DEEPLARNING:

It uses massive neural networks with several layers of handling units, taking benefit of advances in computing power and enhanced training techniques to understand difficult patterns in huge quantities of data. Common applications include speech recognition and image processing.

COGNITIVE COMPUTING:

It attempts for a natural, human-like interaction with machines. Using cognitive computing and artificial intelligence, the ultimate objective is for a machine to simulate human procedures over the ability to understand speech and images and then speak logically in response.

COMPUTER VISION:

It's depend on deep learning and pattern recognition to identify what's in image or video. While machines can process, examine and recognize images, they can capture images or videos in real time and understand their environments.

NATURAL LANGUAGE PROCESSING:

It is the capability of computers to analyze, recognize and make human language, including speech. The next step of NLP is natural language

interaction, which allows persons to communicate with computers using usual, daily language to execute tasks.

In addition, many technologies allow and support artificial intelligence:

GRAPHICAL PROCESSING UNITS:

This is key to artificial intelligence because they make available the huge compute power that is required for iterative processing. Neural networks need compute power plus big data.

INTERNET OF THINGS:

IoT generates huge quantities of data from associated devices, most of data unstructured as well as unanalyzed. Automating models with artificial intelligence will permit us to procedure additional of it.

ADVANCED ALGORITHMS:

It is being developed and joint in different ways to analyze additional data quicker and at many levels. This intelligent processing is fundamental key to classifying and redacting exceptional events.

APIS, OR APPLICATION PROCESSING INTERFACES:

APIs are transferrable packages of code that make it possible to add artificial intelligence functionality to current products and software packages. They can create titles and headlines, or call out interesting patterns and insights in data, Q&A abilities that describe data, add image recognition abilities to home security systems.

Several tools that enable and support AI in automatic big data analysis.

SISENSE:

It is a cutting-edge tool that is great for data scientists who are interested in incorporating predictive models. It contains everything an organization will require to prepare and analyze complex data sets, Its In-Chip technology optimizes the use of computational resources, making it highly efficient and cost effective. The tool will also "allow for advanced analytics and statistics with their library of R functions. Sisense is one of the tools that is used some artificial intelligence techniques for data analysis.

TIBCO SPOTFIRE

This data visualization and analytics tool can handle data transformation efficiently and simply, it incorporates statistical functions to allow for deeper analysis and exploration of patterns and trends within the data." TIBCO Spotfire is a clever, safe, administrated, creativity class analytics platform with built in data dispute that delivers artificial intelligence driven, graphical, geo, and streaming analytics.

GROW:

This new tool is powerfully simple and ideal for small to medium organizations looking to get started with data visualization and analytics. It offers a simple way to pull scattered information together.

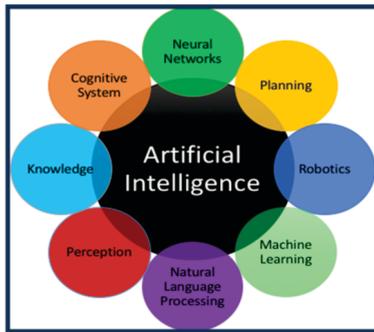
SAS VISUALANALYTICS:

It is interactive reportage. Graphical data detection. Scalability and governance. Self-serve analytics. All from are in distinct, powerful in memory environment. Visually discover critical drivers for taking better decisions. Find out why something take place. Observe all alternatives and find opportunities hidden deep in information. Automatically highlight outliers, collection, important relationships and more, illuminating dynamic visions that stimulate action. Build stunning cooperative reports and dashboards. Speedily encapsulate key routine metrics and share them through the web application and mobile devices. Managers and frontline staff can speedily interact with and it to find out their personal answers, and better realize industry performance. It's give predictive analytics that enables to measure possible results and make cleverer, data driven decisions, no programming knowledge is required. Intelligence algorithms reduce the prerequisite for manual research.

CELONIS

Celonis Process Mining is a smart big data technology that visualized and analyzes each

process. It makes known weaknesses and makes processes faster, clearer, and more cost effective.



CONCLUSION

We came to conclusion that artificial intelligence techniques can be used for automated big data analysis in e-Governance Projects to promote digitalization. Human like interactions with system, software and give support to make decision for specific tasks will provide by artificial intelligence, but it's not that machine take place of human means or replacement for humans. It is not about the replacement of humans but it a digitalization and use of the intelligent tools for automated task. We conclude that in era of technology, we can take maximum benefit of intelligence technique in digitalization and automation.

ACKNOWLEDGMENT

We wish to thank our mentor for inspiring to do research work.

REFERENCES

- [1] C. C. F. L. F. Almeida, "The main challenges and issues of big data management," *International Journal of Research Studies in Computing*, vol. 2, no. 1, 2012.
- [2] L. Liu, "Computing infrastructure for big data processing," *Frontiers of Computer Science*, vol. 7, no. 2, pp. 165-170, 2013.
- [3] S. Guerreiro, "Decision-making in partially observable environments in Business Informatics (CBI)," in *IEEE 16th Conference*, 2014.
- [4] P. K. R. K. a. C. W. N. Roedder, "Decision making based on incident data analysis in Business Informatics (CBI)," in *IEEE 16th Conference*, 2014.
- [5] D. B. I. H. Y. E. I. E. J.-R. E. K. M. K. a. D. Z. H. Killapi, "Distributed query processing on the cloud: the optique point of view," in *OWLED, Citeseer*, 2013. [6] S. I. S. M. a. M. K. E. Liarou, "Monetdb/datacell: online analytics in a streaming column-store," vol. 5, no. 12, pp. 1910-1913, 2012.
- [7] E. Gabrilovich, "Feature generation for textual information retrieval using world knowledge," *ACM SIGIR Forum*, vol. 41, no. 2, pp. 123-123, 2007.
- [8] P. E. Utgoff, "Automatic feature generation for problem solving systems," 1992.
- [9] D. O'Leary, "Building and Evolving Data Warehousing and Business Intelligence: The Case of Sysco," *Intelligent Systems in Accounting, Finance and Management*, vol. 18, no. 4, pp. 195-213, 2011.
- [10] D. O'Leary, "Artificial Intelligence and Big Data," *IEEE Intelligent Systems*, vol. 28, no. 2, pp. 9699, 2013.
- [11] A. K. D. G. a. D. A. F. J. Fan, "Automatic knowledge extraction from document," *IBM Journal of Research and Development*, vol. 56, no. 3, pp. 1-5, 2012.
- [12] K. A. N. M. a. H. K. K. Yano, "Measurement of human behavior: creating a society for discovering opportunities," vol. 58, no. 4, p. 139, 2009.
- [13] <https://www.wherescape.com/resources/white-paper-the-case-for-big-data-automation/>, [Online].
- [14] "http://news.mit.edu/2015/automating-big-data-analysis-1016," [Online].
- [15] "http://www.dataversity.net/big-data-analytics-role-automation/," [Online].