Proceedings of International Conference on Advances in Computer Technology and Management (ICACTM)
In Association with Novateur Publications IJRPET-ISSN No: 2454-7875
ISBN No. 978-81-921768-9- 5
February, 23rd and 24th, 2018

BIG DATA ANALYSIS SYSTEMS FOR BUSINESS POLICIES

MAYUR P DESHMUKH[1]
Department Of Computer Science
Balbhim Art's, Science &
Commerce College, Beed, (MS)
India.
maydes1@gmail.com

PRAPTI D DESMUKH[2]
Principal, MGM's Dr. G Y Pathrikar
College, of CS & IT, Cidco,
Aurangabad, (MS) India.
prapti.research@gmail.com

SUHAS B KALE[3]
Department Of Computer Science
Balbhim Art's, Science &
Commerce College, Beed, (MS)
India.
ctekbeed@gmail.com

ABSTRACT
Big Data arrangement concentrated on whether enormous information was ideal for your business, and gave a few cases of how huge information is being utilized as a part of the undertaking [1]. This study focuses on the tools that we can utilize for the analysis related to big data issues that occurs into business. As India's big data continues to growing and is facing challenges to enhance the data. In this article we will introduce how to really embrace huge information into your business, laying out seven systems that can enable you to find significant solutions from your information.

KEYWORDS*:* Big Data, Data Protection, Big data Analysis Systems.

I. INTRODUCTION TO BIG DATA

This 'Big Data' is the application of specialized techniques and technologies to process very large sets of data. These data sets are often so large and complex that it becomes difficult to process using on-hand database management tools. Examples include web logs, call detail records, medical records, military surveillance, photography archives, video archives and large-scale e-commerce [1]. By 'very large' we're talking about datasets that require at least one terabyte – if not hundreds of petabytes – of storage. (Note that 1 petabyte = 1024 terabytes!). Facebook is estimated to store at least 100 petabytes of pictures and videos alone. Big data is first and foremost, data; an elusive, yet valuable, corporate asset quite unlike any other that needs to be properly managed.

Big Data is defined from the McKinsey Global report from 2011: Big Data is data whose scale, distribution, diversity, and/or timeliness require the use of new technical architectures and analytics to enable insights that unlock new sources of business value [2]. McKinsey's definition of Big Data implies that organizations will need new data architectures and analytic sandboxes, new tools, new analytical methods, and an integration of multiple skills into the new role of the data scientist. Figure 1 highlights several Sources of the Big Data deluge.

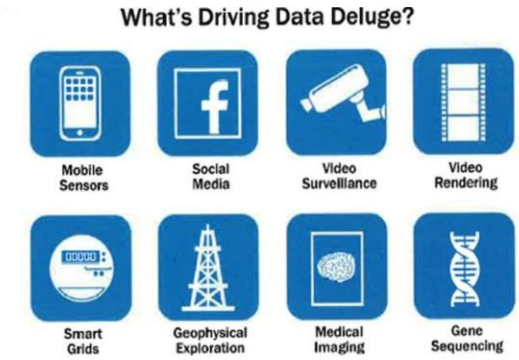

Fig 1. Sources of Big Data deluge

The rate of data creation is accelerating, driven by many of the items in Figure 1. Social media and genetic sequencing are among the fastest-growing sources of Big Data and examples of untraditional sources of data being used for analysis. Unique characteristics of data:

- It can be copied perfectly at very low cost
- It can easily be combined with other data to uncover valuable insights
- It can be used by more than one person at the same time

By managing big data effectively, businesses are better able to capitalize on its value. Data is everywhere. In fact, the amount of digital data that exists is growing at a rapid rate—in fact, more than 2.7 zettabytes of data exist in today's digital universe, and that is projected to grow to 180 zettabytes in 2025[3][4][5]. All this data—from your photos to the Fortune 500's financials—has only recently begun to be analyzed to tease out insights that can help organizations improve their business. That's why more organizations are seeking professionals who can make sense of all the data.

II. METHODS USED FOR BIG DATA ANALYSIS

Their are many methods being used to analyze datasets. In this article, we provide a list of some techniques applicable across a range of industries. This list is by no means exhaustive. Indeed, researchers continue to develop new methods and improve on existing ones, particularly in response to the need to analyze new combinations of data. However, all of the techniques we list here can be applied to big data and, in general, larger and more diverse datasets can be used to generate more numerous and insightful results than smaller, less diverse

ones. According to IDC Canada, a Toronto-based IT research firm, Big Data is one of the top three things that will matter in 2013[6]. With that in mind, there are 7 widely used Big Data analysis techniques that we'll be seeing more of over the next 12 months [7]:

1) Association rule learning
2) Classification tree analysis
3) Genetic algorithms
4) Machine learning
5) Regression analysis
6) Sentiment analysis
7) Social network analysis

1) ASSOCIATION RULE LEARNING

Are people who purchase tea more or less likely to purchase carbonated drinks? Association rule learning is a method for discovering interesting correlations between variables in large databases. A set of methods for discovering interesting relationships, i.e., "association rules," among variables in large databases. These techniques consist of a variety of algorithms to generate and test possible rules. It was first used by major supermarket chains to discover interesting relations between products, using data from supermarket point-of-sale (POS) systems.

Association rule learning is being used to help:

- Place products in better proximity to each other in order to increase sales
- To extract information about visitors to websites from web server logs
- Aanalyze biological data to uncover new relationships
- To monitor system logs to detect intruders and malicious activity
- To identify if people who buy milk and butter are more likely to buy diapers.

2) CLASSIFICATION TREE ANALYSIS

Which categories does this document belong to? Statistical classification is a method of identifying categories that a new observation belongs to. It requires a training set of correctly identified observations – historical data in other words.

Statistical classification is being used to:

- Automatically assign documents to categories
- Categorize organisms into groupings
- Develop profiles of students who take online courses

3) GENETIC ALGORITHMS

The Which TV programs should we broadcast, and in what time slot, to maximize our ratings? Genetic algorithms are inspired by the way evolution works – that is, through mechanisms such as inheritance, mutation and natural selection. These mechanisms are used to "evolve" useful solutions to problems that require optimization.

Genetic algorithms are being used to [8]:

- Schedule doctors for hospital emergency rooms
- Return combinations of the optimal materials and engineering practices required to develop fuel-efficient cars
- Generate "artificially creative" content such as puns and jokes

4) MACHINE LEARNING

Which movies from our catalogue would this customer most likely want to watch next, based on their viewing history? Machine learning includes software that can learn from data. It gives computers the ability to learn without being explicitly programmed, and is focused on making predictions based on known properties learned from sets of "training data."

Machine learning is being used to help [9]:

- Distinguish between spam and non-spam email messages
- Learn user preferences and make recommendations based on this information
- Determine the best content for engaging prospective customers
- Determine the probability of winning a case, and setting legal billing rates

5) REGRESSION ANALYSIS

How does your age affect the kind of car you buy? At a basic level, regression analysis involves manipulating some independent variable (i.e. background music) to see how it influences a dependent variable (i.e. time spent in store). It describes how the value of a dependent variable changes when the independent variable is varied. It works best with continuous quantitative data like weight, speed or age.

Regression analysis is being used to determine how [10]:

- Levels of customer satisfaction affect customer loyalty
- The number of supports calls received may be influenced by the weather forecast given the previous day
- Neighborhood and size affect the listing price of houses
- To find the love of your life via online dating sites

6) SENTIMENT ANALYSIS

How well our new is return policy being received? Sentiment analysis helps researchers determine the sentiments of speakers or writers with respect to a topic.

Sentiment analysis is being used to help:

- Improve service at a hotel chain by analyzing guest comments
- Customize incentives and services to address what customers are really asking for
- Determine what consumers really think based on opinions from social media

7) SOCIAL NETWORK ANALYSIS

Social network analysis [11] is a technique that was first used in the telecommunications industry, and then quickly adopted by sociologists to study interpersonal relationships. It is now being applied to analyze the relationships between people in many fields and commercial activities. Nodes represent individuals within a network, while ties represent the relationships between the individuals.

Social network analysis is being used to:

- See how people from different populations form ties with outsiders
- Find the importance or influence of a particular individual within a group
- Find the minimum number of direct ties required to connect two individuals
- Understand the social structure of a customer base

III. Conclusion:

A smart phone these days has additional computing power than all of independent agency once it place a person on the moon in 1969. There square measure 5 billion telephone users and 4000 tweets generated per second. There square measure over 900 million Face book users either sharing knowledge or using alternative people's knowledge. This unstructured knowledge is not any longer controlled by organizations; it's in people's hands. The problem is the way to mine, store and transfer, refine, synthesize, and build secure systems that use this knowledge. Big knowledge presents a striking nevertheless alarming image of the longer term for world. The striking side is that it provides United States of America a wider perspective and newer discoveries that successively can do modification to our life, work, and mental attitude.

There is conjointly association in nursing absence of a legal framework on problems related to knowledge manipulation at the side of the growing fields of information mental image, prophetical and link analyses, and assortment methods. The iniquitousness of huge knowledge needs the proper combination of individuals, tools and technology to leverage its unlimited potential.

In these circumstances Asian country conjointly desires info Security Policy for enjoyment of massive knowledge. Whether your business wants to discover interesting correlations, categorize people into groups, optimally schedule resources, or set billing rates, a basic understanding of the seven techniques mentioned above can help Big Data work for you.

IV. REFERENCES:

1) Debbie Stephenson, "Does Big Data Make Sense For Your Business?", Dec 14, 2012 Web Reference- https://www.firmex.com/thedealroom/does-big-data-make-sense-for-your-business/J.
2) Willey,"Data Science &Big Data Analytics: Discovring, Analyzing,Visualizing and Presenting Data", Published by John Wiley & Sons, Inc. 10475 Crosspoint Boulevard Indianapolis, IN 46256. www. wiley. comMcKinsey Global report from 2011
3) Mark Mulcahy – Waterford Technologies, "Big Data Statistics & Facts for2017", Posted on 22 Feb, 2017. https://www.waterfordtechnologies.com/big-data-interesting-facts/
4) Avantika Monnappa, "Data Science vs. Big Data vs. Data Analytics", Last updated December 13, 2017, https://www.simplilearn.com/data-science-vs-big-data-vs-data-analytics-article
5) Mayur P Deshmukh, Prapti Deshmukh, Ratnadeep Deshmukh, " Big Data analytics sensing to your business " , State level seminar on "Big data & Business Intelligence", organized by Dept. of Computer Science and sponsored by BCUD Savitribai Phule Pune University, Pune, held at New arts Commerce and Science College, Shevgaon, dist. Ahemadnagar 414502 (MS) India, Proc. pg 47-51, on Jan, 2018.
6) Ryan B. Patrick,"The three things IDC Canada predicts will matter in 2013", December 12, 2012.
7) Web Reference https://blog.allstream.com/the-three-things-idc-canada-predicts-will-matter-in-2013/
8) Debbie Stephenson, "Big Data Techniques That Create Business Value", Jan 18, 2013.Web Reference - https://www.firmex.com/thedealroom/7-big-data-techniques-that-create-business-value/
9) Brainz, "15 Real-World Applications of Genetic Algorithms" August 2017. Web reference http://www.brainz.org/15-real-world-applications-genetic-algorithms/
10) http://www.forbes.com/sites/netapp/2012/10/31/big-data-law-cloud-analytics/
11) http://www.creatalytics.com/data-driven-marketing/eharmonys-use-of-regression-analysis
12) CRM Innovation Editors**, "**Social network analysis surges into vertical markets**",**2012-04-05. https://www.enterpriseinnovation.net/article/social-network-analysis-surges-vertical-markets