



# Big Data and its implication on Social media and Electoral Process

Rachitha M V<sup>1</sup>, Manish Parasher<sup>2</sup>, Mohan Krishna Ponduri<sup>3</sup>

<sup>1</sup>Assistant Professor, <sup>2,3</sup> Students,  
Dept. of Computer Science & Engineering  
Vemana Institute of Technology, Bangalore

**Abstract** — Big data has expanded its reach to such an extent that the traces in facebook or even Smartphone's has been used to analyze human behaviour and even their decisions for voting in elections. It has enabled us to analyse the psychometric traits of humans conveniently, reaching out to the remotest location and coming up with more accurate results as compared to other methods of electoral forecasting. Results for BREXIT and US presidential election amply testify the same.

**Keywords** — Big data, Volume, Velocity, Variety, Variability, Veracity, OCEAN model, Artificial Intelligence, Data analysis, Data mining, psychometric traits, voter data model.

## I. INTRODUCTION

The term Big data has been in use since the 1990s. John Mashey is given the credit for coining the term or at least making it popular. Big data is extremely large data sets that may be analysed computationally to reveal patterns, trends, and associations, especially relating to human behaviour and interactions. Much of the present IT investment is going towards managing and maintaining big data.

Big data characteristics were originally defined by the three Vs but nowadays additional Vs have been added to its characteristics. The characteristics are:

**Volume:** Organizations collect data from several sources, including business deals, social media and information from machine to machine data. In the past, storing it would've been a problem but new technologies (such as Hadoop) have relieved the burden.

**Velocity:** Data streams in at an unparalleled speed and must be dealt with in a prompt manner. RFID tags, sensors and smart metering are driving the need to deal with torrents of data in near-real time

**Variety:** Data comes in all types of formats from structured, numeric data in old databases to unstructured text files, email, video, audio, and monetary proceedings.

**Variability:** Irregularity of the data set can hamper processes to handle and govern it.

**Veracity:** The quality of captured data can vary greatly, influencing precise analysis.

Big data is the new artificial intelligence. Artificial Intelligence is a programming style, where programs operate on data according to the rules in order to accomplish goals. To a chess playing program, the data are the positions of man on board. Rules are the moves permitted in the game of chess. The goal is to win the game. Now let us take a very good example which will help us to understand it better.

### Deep blue vs Gary Kasparov

The first match was played in Philadelphia in 1996 and won by Kasparov. The second match was played in New York city in 1997 and won by Deep Blue. Deep blue's win was seen as very significant, a sign that AI was catching up to human intelligence and could defeat one of humanity's great intellectual champions. Big data has a crucial role in modern day enterprises. When combined with AI, it can become a robust tool for proactive and robust management. AI has seeped with our daily life through the internet. For instance, when we select our shopping cart on an Ecommerce website, it automatically shows suggested products depending on our likes and what we have purchased previously.

## II. BIG DATA AND SOCIAL MEDIA

Big data is a Driving Factor Behind Every Marketing Decision. Just like any other industry, social media companies find big data useful for examine markets and foretell consumer behavior. In 2012, Jay Parikh, engineering VP at Facebook, revealed that Facebook supervises over 500 terabytes 'likes' and 2.5 billion content uploads. All this data is managed in mere minutes giving Facebook insight into user reactions and the ability to modify its offering, correlating content to statistics of users age, gender, relationship status, geographic location, income levels, educational achievement, inclination to purchase certain products allows a company to know more about the people it's dealing with. This is very beneficial as advertisers can react in near real time and adjust campaigns to make more revenue. Now for example, let us consider Facebook and see how it is using big data. Facebook is the most popular social networking site with over 1.86 billion monthly active users worldwide, Facebook stores a gigantic amount of user data. Every day users feed Facebook with mounds of data. At first this data may seem to be unrelated. But with data like this, Facebook knows your preferences, your likes and dislikes and so much more. the more users use Facebook, the more information they get from them. Apart from examining this user data, facebook has other ways of ascertaining user behaviour.

- 1) **Tracking Cookies:** Facebook tracks its users over the Web with the use of tracking cookies. If a user is logged into Facebook and simultaneously browses the Web, Facebook can track the sites they are visiting.
- 2) **Tag Suggestion:** Tag suggestions can be made on user photos due to image processing and facial recognition.
- 3) **Analyzing the 'Likes':** A recent study conducted showed that Facebook is able to predict data accurately on a range of personal attributes that are highly sensitive just by analyzing the 'Likes' that have been clicked by a user on Facebook. one research study shows how the patterns of Facebook 'Likes' can very precisely predict your intelligence, satisfaction with life, emotional solidity, religion, alcohol use and relationship status, age, gender, race and political views among various others."

One example that shows how Facebook use Big Data is The Flashback. Celebrating its 10th anniversary, Facebook offered its users the option of watching and sharing a movie which traces the course of their social network activity from the date of registration till the present. Called the 'Flashback', this video is a combination of photos and posts that have the most comments and likes. Facebook today cannot exist without AI. One of facebook biggest assets is that when it has a billion people interacting with its product everyday it collects a lot of data and it stores endless examples of user behaviour. This allows the ads team to ship several models every week thereby increasing the durability. Every time we use Facebook, Instagram and messenger our experiences are being powered by AI. Currently the biggest competitor for Google is Bing. The AI that is used in Bing has been conceptualized by a team at microsoft led by Joaquin Candela. Candela made sure to do the work where multiple teams were involved. It's a three step process:

- FOCUS ON PERFORMANCE
- FOCUS ON UTILITY
- BUILD A COMMUNITY

CANDELA BREAKS DOWN THE APPLICATIONS OF AI IN FOUR AREAS:

- VISION
- LANGUAGE
- SPEECH
- CAMERA EFFECT

## III. BIG DATA AND ELECTORAL PROCESS

### A. CAMBRIDGE ANALYTICA

It combines data analysis and data mining for electoral process. It was involved in US election campaign almost two years ago. The basic idea was to work based on people's interest. Like other campaigners so far relied on demographics Cambridge Analytica was using psychometrics. Success of Cambridge Analytica was due to the following reasons

- *Behavioral science using OCEAN model*
- *Big Data analysis*
- *Ad targeting (personalized advertising aligned as accurately as possible to personality of an individual consumer)*

## B. PSYCHOMETRIC ANALYSIS USING BIG DATA

Psychometrics is data driven sub branch of psychology. It was a revolutionary approach which played an extraordinary role in Trump's win. It enabled digital revolution. Big data leaves digital traces. Every activity of an individual can be traced and analysed using big data for analytical and result oriented purpose. Psychometrics includes 5 personality traits:

1. OPENNESS (HOW OPEN WE ARE TO NEW EXPERIENCES?)
2. CONSCIENTIOUSNESS (HOW MUCH OF A PERFECTIONIST WE ARE?)
3. EXTROVERSION (HOW SOCIABLE ARE WE?)
4. AGREEABLENESS (HOW CONSIDERATE AND COOPERATIVE WE ARE?)
5. NEUROTICISM (ARE WE EASILY UPSET?)

Psychological traits of individuals are properly assessed and relative accurate assessment of the kind of assessment of the kind of person can be made by using big data. Used data set enabled the researchers to make accurate prediction about various personality traits of a human being like intelligence ,religion, alcohol ,cigarettes used etc. Even the smart phone which is constantly used gives a lot of inputs for creation of psychological profiles as well as to search for specific profiles .A company named SCL(Strategic Communication Laboratories) used these tools for influencing elections.SCL created a new company named Cambridge Analytica to participate in US election. They created "Voter data modeling" wherein personality of every adult was profiled. After proper analysis a conclusion was drawn with regard to age, address, interests and political inclinations. The same was used by the agency to target the person with appropriate political message. Every message ad was data driven. The process was used by the agency to target the American population. The proper analysis and the reach to individual led to get support of the people who had not decided which party to vote. Individual messages were transmitted through apps and the reaction flowed back for the campaign. Final analysis resulted in victory of Trump and falsifying the electoral forecast.

## IV. CONCLUSION

Psychological traits of individuals are properly assessed Analysis of psychological traits of individuals using big and relative accurate assessment of the kind of data has resulted better electoral forecasting. Initially it was assessment of the kind of person can be made by evident during BREXIT poll and now US presidential using big data. Used data set enabled the researchers to elections results have showed the importance of "voter data make accurate prediction about various personality model". The new methodology of electoral forecasting is traits of a human being like an eye-opener and is worth emulating in our country. Only intelligence, religion, alcohol, cigarettes used etc. Even we require robust systems to facilitate the same with the Smartphone which is enabling intention of the political class of our country.

## REFERENCES

- [1] (2015) The simplilearn website. [online]. Available: <https://www.simplilearn.com/>
- [2] (2017) Wikipedia website .[online]. Available: [https://en.wikipedia.org/wiki/Big\\_data/](https://en.wikipedia.org/wiki/Big_data/)
- [3] (2014) Medium website.[online]. Available: <https://medium.com/>
- [4] (2017) Motherboard website. [online] . Available: [https://motherboard.vice.com/en\\_us/article/how-our-likes-helped-trump-win/](https://motherboard.vice.com/en_us/article/how-our-likes-helped-trump-win/)
- [5] Robert X. Crungley (2017) Insidebigdata website [online] . Available: <https://insidebigdata.com/>
- [6] (2014) BetaNews website [online] . Available: <https://betanews.com/>
- [7] Steven Levy, Editor of Backchannel, (2017) Backchannel website [online] . Available: <https://backchannel.com/>
- [8] (2017) Underhood reputations of companies website [online] . Available: <https://blog.underhood.co/>