

# Big Data in Social Media: Introduction, Methodology, Applications.

**Akash K Salián**

Institute of Distance and Open Learning, Mumbai, Maharashtra, India  
salianakash133@gmail.com

**Abstract:** *In this paper, we have discussed about how Big-Data has influenced the working, modeling and behavior of social media on a huge scale. Not only use but also storing and restructuring of data have created a new model and pattern of data on its own. Big-data is used and managed in a lot many ways by the different social media applications and platforms according to the need and methodology accepted. We also saw how we can use this available data can help us to grow our business*

**Keywords:** big-data, social media, structure, management, pattern, processing

## I. INTRODUCTION

Coming to today's world security has been a huge concern of people world while. However, we don't look into it as a major concern when it comes to the social media. Social media platforms are in a big hype for data collection and social networking. All the population that has been associated with social media is a part of Big Data which involves their personal information such as name, number, email which might put them to a risk. Dealing with big data has turned out to be a major issue as well as a sensitive one.

Connecting one's profile or linking it through different mediums increase the share and spread of data. Social media account can be linked with your email account, game account, bank account and even the official account which might not be a good idea to do. With increase in laziness the single sign on function has gained a great importance.

### 1.1 PROBLEM STATEMENT

Social media is often considered as a virtue of wasting time. Each and every platform that links a social group or id is a part of this tag. But it can rather be used in a better way to generate revenue and expand one's business. We can see various types of data available on social sites which can be processed and used for our own business. We will be talking about revenue in the later half but for now let's understand the base of social media and the data generated by the use of these devices and applications.

### 1.2 EASE OF USE

#### A. Daily use of social media

Use of social media is increasing day by day. Updating the social life has become an integral part of people's routine schedule. Uploading images, tagging people, liking pictures to videos has influenced them on a great level. But why are people being so flexible towards it, let's dig deeper and get to know about this. Social media is the best way to create friends online. The curiosity of knowing new things and new people, exploring new cultures, creating communication links are the major advantages of social media. With the ease of networking the social network can now be spread through large areas within seconds of time. With the increase in communication and ease of technology the data that is generated is also huge and is increasing day by day.

#### B. Maintaining the Integrity of Data

Data is something that can be found or shared in any form and can be used in multiple ways which may or may not be safe. Giving away any form of information related to personal life can increase the risk of data theft as well as other life threats

Before you begin the day you need to make sure that you don't spread-out any kind of personal data in any form such as giving personal information over a call, sharing information to any unknown person or sending it over an email. All these information gathered by the sources are stored in one place to combine and form a big junk of information and by using it any one can be targeted depending upon the type of information revealed.

### **C. Abbreviations and Acronyms**

We are going to look at some of the basic words that we have heard but don't exactly know about it. Starting with the first one as MPP - Massively Parallel Processing, NLP - Natural Language Processing, OLAP - Online Analytical Process, OLTP - Online Transactional Processing.

The storing of data is always a big concern. While using our phones we never think of how the data is stored in our cell phones, at most what we do is sort them into folders but this is not the case in terms of bigdata. All these data coming from various sources cannot be stored in various folders. The data is stored in huge data centers consisting of huge number of storage devices e.g. hard drives and also for new network topology and services such as SAN (Storage Area Network), ESX Servers etc. The capacity of data centers is in thousands of terabytes or even petabytes. Each user created has his/her separate identity by which the account and the data of that account can be extracted through various means depending upon the need.

This stored data can be used for many purposes such as storage, data recovery, data backup, networking etc. The process of managing data can be obtained by the DBMS rules of data storage. The relation between data, format of storage, means of storage can all be defined in the instruction.

### **D. Forms of Data**

Qualitative and Quantitative data: Any kind of data that can be measured in terms of quantity is called as quantitative data while those who can be measure in terms of quality is called is qualitative data. For example we can count the number of names of places, users, companies etc.. Also in addition to that we can measure the number of pixels in an image to determine its quality

.Continuous and Discrete data: Discrete data can be numeric like numbers of apples but it can also be categorical like red or blue, or male or female, or good or bad. Continuous data are not restricted to defined separate values, but can occupy any value over a continuous range.

Primary and Secondary Data: The data collected from direct sources is a primary data while the data collected from secondary sources such as from consulting or reading is called as secondary data.

Apart from this there are also many other forms such as images, videos and gifs. But the twist here is a very different form which is related to our actions we perform on these social media platforms.

## **II. METHODOLOGY**

We are often more than happy to fill our data at any place may it be a shopping center, a movie theater, a lottery center or even on social media platforms. We don't realize how our data can be utilized or be exploited. Let's discuss some situations that happen to us in day to day life.

Whenever we open a new bank account, after some period of time we get to start getting calls for loan, credit-card, new account etc. Where do they get our numbers from? How do they know our name? All these questions are answered by sharing of our data. It completely depends on where do we share our data, how do we share our data, and with whom do we share our data and what type of data.

So how do these companies use our data? Let's take an example of any social media page. First of all we have the basic profile page where the user finds their identities by updating the data which might be real or fake. This is the first step of data collection. Once you are logged in these websites use Machine learning algorithms to study the pattern of our use of these platforms. The way we use these platforms can create a pattern. Every click that we make on a screen or for every sec the screen open has a track to it. The logs created from this activity can be used to create the pattern for that particular user.

**A: Inputs and Outcomes**

- The main purpose of making a social media platform is business and money. This being a base priority of all the start-ups and organizations a social app or website is made
- We saw the base priority in the above point and the next point is network. Social media is the best, fastest and most continent way to grow the network of a company as well as people all over the world.
- So how does they earn from this? We see a lot of ads while browsing through these platforms. So where does this ads come from, and how specific are they?
- The data that we talked about in the previous sections is used here to target the audience over the whole medium.
- The data is used to select target audience and hence the ads are displayed to the user and some of them end up buying the product displayed.

Let's get a deeper idea on how does it actually works. Whenever a user is signed-up with the social account the details and device information is used by the social media app. This data which includes the browsing, searching and activities similar to this. So how does this tracking actually help?

Take an example of a teenager. Consider he's very active on all social media platforms and now he's on a shopping platform looking for some good headphones. This will be tracked by the social media and stored in the database. Next time when he browses through the social media he will get an ad regarding headphones showing him some good recommendations or discounts on the same product that he searched for. This may remind him of his recent activity and he might get convinced by the ad and buy that product.

This is the basic activity that happens with all the users and hence the cycle of data gets updated with every search and the process continues. The ads keep on changing and updating depending upon the user's activity and search activity. The revenue collected from the ads run the overall expenses of the organization. The more the popularity of the app the more people will use it and the more it can cost for ads. This is the basic working of all social apps and the process can be modified depending upon the company policy.

Here the main role is creation of patterns. As we saw earlier each and every user may have a different pattern based on his/her preferences and so does the marketing. The daily routine of a user can be traced through the browsing data and pattern created by these machine learning algorithms. The main concern can now change to marketing once the pattern is created. Regular updates make sure the user gets ads regarding his/her need of product or commodities.

**III. APPLICATIONS**

So we saw how our data is gathered, sorted and used. Now let's understand how we can use this data for our benefit. Social media provides a huge platform for personal usage such as social networking but along with this there is also another space for social business. Yes, along with the personal space we can also use it for business. The services provided for business are a bit advance in use of tools and performance. Using normal social account and business account has a huge difference in its own. The business account helps to interact with the client or customer in a new and more organized manner. This helps to organize the business and helps by giving it a new boost. Social media campaigns helps to create informative idea about the product.

The business models have changed. Globalization, social media, networking, and technology all these components have marked their presence in changing the business model and make it more and more effective in modern as well as the social world.

Let's take an example of a cake shop owner. Suppose he is really doing well with it. But now he's facing loss in the business. He comes to know about another cake shop which is located just in the next lane through a social media platform. He thinks this of a good idea to start his own social marketing. So now let's understand how this can help his business grow. The first thing that comes into existence is creating a profile for the shop. Once this is done the next step comes here is making people aware about the shop and finding target audience for the shop. Marketing works the same as advertising. Just the difference is the marketing can be more organized and managed. The main component in marketing is targeting. Selecting target customer and selling the product to them is the main goal of any business. Social media makes it easy for us to know the people around us and gives us a media to interact with them. Knowing

the audience and turning them into a customer is the most important consent of the business. Social Media can help this cake shop to be known by the people in the local as well as remote locations and grow its business.

Let's take one more example of the same cake shop. Now if the owner comes to know that there's an event happening near the cake shop this might be a good chance to increase the customers. The information related to the event can be known through social media and he can apply some coupons or deals on cakes on or around the event date. More the number of people coming in the event more can be the chances of people becoming his customers.

Social media and big data together come along to form a major component of the digital world. Each and every media today generates a huge amount of data that can be sorted as well as combined to form the overall information which includes your name to your browsing history. Companies collect these data to optimize the use of social technologies and improve the digital as well as social experience.

#### **IV. CONCLUSION**

Big data and social media is an immense part of our day to day life influencing the speed of digital era. It is important for us to make sure that this data generated by us should be secured and be used in right manner. Along with this it is also important for us to see that how we can use this data for our goodwill and profit. With this immense data available we can make full use of it to boost our day to day understanding about technology, data and its use.

#### **ACKNOWLEDGEMENT**

I sincerely thank my guides for the guidance and encouragement in carrying out this and rectified the mistakes in the same for the betterment of this work.

#### **REFERENCES**

- [1]. M. K.Kakhani, S. Kakhani and S. R.Biradar, Research issues in big data analytics, International Journal of Application or Innovation in Engineering & Management, 2(8) (2015), pp.228-232.
- [2]. Gandomi and M. Haider, Beyond the hype: Big data concepts, methods, and analytics, International Journal of Information Management, 35(2) (2015), pp.137-144.
- [3]. X. Jin, B. W.Wah, X. Cheng and Y. Wang, Significance and challenges of big data research, Big Data Research, 2(2) (2015), pp.59-64.
- [4]. R. Kitchin, Big Data, new epistemologies and paradigm shifts, Big Data Society, 1(1) (2014), pp.1-
- [5]. 12. [5] C. L. Philip, Q. Chen and C. Y. Zhang, Data- intensive applications, challenges, techniques and technologies: A survey on big data, Information Sciences, 275 (2014), pp.314-347.
- [6]. K. Kambatla, G. Kollias, V. Kumar and A. Gram, Trends in big data analytics, Journal of Parallel and Distributed Computing, 74(7) (2014), pp.2561-2573.
- [7]. MH. Kuo, T. Sahama, A. W. Kushniruk, E. M.Borycki and D. K. Grunwell, Health big data analytics: current perspectives, challenges and potential solutions, International Journal of Big Data Intelligence, 1 (2014), pp.114-126
- [8]. T. K. Das and P. M. Kumar, Big data analytics: A framework for unstructured data analysis, International Journal of Engineering and Technology, 5(1) (2013), pp.153-156.
- [9]. T. K. Das, D. P. Acharjya and M. R. Patra, Opinion mining about a product by analyzing public tweets in twitter, International Conference on Computer Communication and Informatics, 2014.
- [10]. A. Jacobs, The pathologies of big data, Communications of the ACM, 52(8) (2009), pp.36-44.
- [11]. Chen, H., Chiang, R., Storey, V.C.: Business intelligence and analytics: from big data to big impact. Miss. Q. 36(4), 1165–1188 (2012)