IS 601
Term Research Paper

Impact of Big Data on Social Media Marketing

By
Arjun Kumar
arjunk2@umbc.edu

Instructor
Dr. Carlton Crabtree
University of Maryland, Baltimore County

Table of Content **Page No.**

1. INTRODUCTION

In the digital world of today, almost each and every small device now or soon to be connected with IOT (Internet of things) and will be able to collect data. These collection of data helps the businesses to understand the customer behaviour and their buying patterns. Big Data provides the ability to target the right set of audiences within a group. Using Machine learning techniques data are segregated in more useful and clearer format and these data are used by marketers to target customers on social media platforms. Big Data allows companies to directly interact the with right set of audience and market the products on social media.

Social media refers to the interactive Web based applications and websites which facilitate the users in the creation and sharing of generated content through the formation of virtual communities and networks. In the past decade, social media has emerged as major marketing platform. As per 2019, the global penetration of social media is at about 45% with around 3.5 billion users while the number of social media users in India stood at 326.1 million as per 2018. With such a large user base being present on Social Media, it becomes a favourable platform for businesses to market their products and increase their consumer base. Digital life has been the life saver of each individual in the present computerized world. Where there are individuals, there are potential clients, to draw in with that potential client base online digital advertising is used. Social media plays a imperative part in marketing and making connections with clients. (Bhor, Koul, Malviya, & Mundra, 2018). In todays social media driven environment, it is fundamental for Small and Medium business to get it Facebook, Twitter, and the techniques behind using social media for developing their commerce. With restricted obstructions to section, little businesses are starting to use social media as a implies of showcasing. Numerous little businesses battle to utilize social media and have no technique going into it. As a result, without an essential understanding of the focal points of social media and how to utilize it to lock in clients, endless openings are missed. To understand how smaller scale and Medium business can utilize social media marketing strategies to interact with customer and market product. (Adegbuyi & Akinyele,2015)

Market is defined as the platform for exchange of goods and service between buyer and seller. Marketing is the process of informing the consumers about the various goods and services being offered in the market and influencing their buying decisions. The use of social media by marketers to pass information about commodities in order to attract consumers led to the presence of social media marketing. It has grown as an effective yet inexpensive marketing tool. Such a tool comes in handy for new and small businesses with little money to invest in marketing. It acts as a medium to connect with a business's probable consumer base and helps gauge the reaction of consumers to the various products in market.

Many countries like India, China, United States, Singapore, UAE with a large and continuously growing number of small businesses. These businesses could largely benefit from the use of social media for marketing and many are already utilizing the benefits of the reach of social media. The major social media apps that are used for the purpose of marketing are Facebook, Instagram, WhatsApp, Twitter and Snapchat. These

apps have grown in popularity and visibility over the years and tend to connect large populations.

Impact of Big Data on Social Media Marketing- With the use of Big data on social media allows the data to be collected and segregated in much useful and meaningful format. It allows the companies to market right product to right set of customers on social media, which increases the profits for company by increasing the visibility of the product and boosting sales while decreasing the advertisement cost. In detail, if a customer is interested to buy an iPhone product and the company advertise the customer android phone products. The customer will never buy the product as he is not interested in android and will add a marketing cost to the company, with no result from such advertisement. So, with big data allows the companies to understand the user behaviour using machine learning algorithms and data are segregated in different groups as required by the companies for better decision making and advertisment of right product to right set of audience. (Bansal, Kumar&Choudhury, 2018).

2. Literature Review

Traditional medium of advertisement like banners, radio, newspaper, boards, yellow pages, prints have lost the power create brand value and brand name in the market. These methods of marketing are no longer creating value for the companies to advertise the products or service. In this respect, company and business have to pay a lot for better decision making, to acquire customers and to create brand awareness. With the advancement of technologies, companies have moved far beyond traditional mediums to reach the customers with the help social media platforms. (Srinivasan,Bajaj&Bhanot, 2016).

Macro phenomenon, such as, social mobility, political organisation, social cohesion and diffusion can be studied through the use of network analysis and can be linked to small scale interaction between people. The flow of information tends to be faster and covers greater social distance when people are connected through weak ties. There is greater social mobility and higher risk seeking behaviour observed for people having large number of weak ties. Modern day social media networking connects people having such interpersonal weak ties and thereby could be used for the spread of information. The capacity of different types of media to convey messages and cues differs from each other based on the number of mediums used, such as video, audios or both, the language variety and the personalisation and feedback speed. This is widely known as the media richness theory. (Adegbuyi & Akinyele,2015)

Small scale businesses have more marketing opportunities nowadays and technology and networks end up giving support to these businesses. Networking is very important for the establishment, development and growth of such businesses and they rely heavily on personal contacts as well as positive word of mouth from customers. Social media helps form such networks and online word of mouth spreads very quickly on social media.

The use of social media by businesses helps companies obtain feedback regarding their products by monitoring customer’s online communications. It also helps promote word of mouth and at the same time companies have limited control over the content and distribution of information and there is large impact of user generated content (Srinivasan,Bajaj&Bhanot, 2016). The informative and persuasive powers of user

generated content and marketer generated content on Brand Community and consumer behaviour was analysed and the finding was that User generated content had a larger impact on purchase behaviour of customers.

The effectiveness of social media marketing was tested on four grounds, that is, the impact on organizational brand, the impact on consumer purchase behaviour, its effectiveness as a competitive tool and impact on organisation efficiency. Three out of four indicators gave positive results and supplemented that social media marketing acts as an effective tool for generating higher brand attitudes and purchase intentions. (Adegbuyi & Akinyele,2015)

2.1 Social Media as a Marketing Tool

Social media is the most trending word of the market. Marketing and advancement is exclusively depending upon social media which was a sideline media amid its viable beginning a decade back. Print media and tv media misplaced its pertinence to social media, since of numerous variables counting fetched adequacy and wider reach. Nearly all areas have received Social media for its effective and productive working. Social media viability is mostly utilized in all statistic sections as the unused age millennial cannot be disconnected from social media. Social media is being received by different areas. Promoting product through social media is superior than the traditional promotion approach, but too makes a difference firms to get it the essential driver for adjustment of social media for their marketing advancement methodologies. Modern day social media networking connects people having such interpersonal weak ties and thereby could be used for the spread of information. It has become so strong that every business uses social media platform to attract the customers and grow the business. (Suresh&Mohan,2016).

2.2 Big Data and Social Media

Many a times, businesses make use of social media marketing without having proper information and marketing strategy, which leads to inefficient utilization of the marketing strategy (Adegbuyi & Akinyele,2015). Social media creates large data about the user behaviour and user interests, and it becomes very difficult for the companies to use these complex data for advertising.

Big Data is used to understand the customer needs and customer behaviour. It segregates the data into digital form and transforms the information into useful data which is used by the companies to target the customers. Datafication using Big Data on social media helps them to analyse and understand the human behaviour. It helps to create and utilize new advertising practices based on social media for companies.

Among different social media sites, Facebook is the most commonly used social media app for the purpose of social media marketing followed by Instagram. (Srinivasan,Bajaj&Bhanot, 2016).

3. Technical Details

The web and Social Media are used by almost every companies to interact with the sellers and buyers to increase the company visibility about the product and services on web. It helps the companies to understand customers opinion about the products and services offered by the company. GolbaliFusion(GiF) brings small size business, Medium size

business and Large scaled business together with its customers by collecting data from social media and converting the data into valuable form which is used for planning marketing strategies and to increase the growth in business by using technology innovation . GiF is based on Information and communication technology (ICT) to manage the Big Data in Social Media. ICT based platform uses keyword to search the content on social media. (Sand,Tsitoura, Dimitrakopoulos & Chatzigiannakis , 2014) GiF represents a mixture of technologies which described as follows:

I. **Big Data Aggregation and Processing**: Social Media is a platform where people interact with others and exchange information in virtual network. Social Media is used by companies to reach large audience, target customers, to understand the trend, generate traffic on products, customer acquisition, promote special offers, to increase brand recognition and brand value, increase the visibility on social network. Enormous amount of data is generated every day on social media like twitter, Facebook, LinkedIn. GiF collects the data of interaction from different social media platforms and presents the information in an organised way which are used by companies to reach the audience on social media. (Sand,Tsitoura, Dimitrakopoulos & Chatzigiannakis , 2014)

II. **New Generation Sentimental Analysis:** A social media profile is the first step for collection of the data based on real time basis from various social media platform. Opinion Mining is the also know as Sentimental Analysis which focus not much on content of the data but analyses the opinion from the information. “Pn Polarity” analyses if the opinion analysed on sentimental analysis are positive or negative. Machine translation is used to extract the data in various language. Data extracted are transformed into English before extracting Sentiment. GiF converts the data extracted into valuable form for the companies. (Sand,Tsitoura, Dimitrakopoulos & Chatzigiannakis . 2014)

3.1 Big Data as A Service (BDaaS)

Big Data is very large volume of data and as data is increasing day by day it is getting very difficult to store and analyse the data. Big is not only data but helps to analyse the useful information for the company. Big data as a service is a tool which analyses the data and provides useful insights to gain competitive strategies. Big Data consist of three layers Big Data Infrastructure as a service, Platform as service, Big Data Analytics Software as service. Big data Analytics software as a service is a combination of Big Data Infrastructure as service and Big Data Platform as a service. It provides web-based environment and allow sharing of resources. Web based environment is used to post their own information. (Park, Nguyen, Won.2015)

3.2 Tools Used for Analytics

Tools such as Apache Hadoop, Apache Spark, Elastic Search are mostly used for data analytics and presentation of data.

I. **Apache Hadoop:** Hadoop framework is used to store and process large amount of data sets. It stores the data into different segments which becomes easy for computation. For high performance and easy access of data, Hadoop uses sorting feature and sort the information into structured, unstructured and semi structured

data. It uses Map Reduce Framework as a software tool to process large volume of data in distributed systems. (Hanamanthrao, Thejaswini S. 2017)

II. **Apache Spark:** Apache Spark is an open source venture planned to perform in Large memory-based cluster data processing. Moreover, it gives integrated system for a diverse sort of data processing necessities which incorporate chart information, content data, and other sources such as group and real-time information. In memory cluster computing highlight of Start increments the processing speed of Start applications. It supports Machine-learning algorithms, Group and Spilling applications, Interactive SQL questions etc. It also supports programming languages such as Java, Python, R and Scala. (Hanamanthrao, Thejaswini S. 2017)

III. **Elastic Search:**Elastic Search is one of the real-time conveyed frameworks used for data analytics and web lookup. Able to effortlessly investigate information at a scale and at a speed at no other time conceivable. Basically, used for analytics and search such as inline, organized and structured look, and by combining all three. It is one of the real-time conveyed system utilized for web lookup and data analysis. Able to effortlessly investigate information at a scale and at a speed at no other time conceivable. (Hanamanthrao, Thejaswini S. 2017)

4. Risks, Limitations and Strategies/solutions:

Analysing the social media data sets to reach out the customers is not easy, it includes a huge risks and limitations. Social Media sites are now used to sell the products and services, which also includes sharing personal information, transferring funds over a network through social media sites like Facebook and LinkedIn. A social media is used by the companies to advertise their products and services, and interaction occur between consumer to consumer, consumer to business, business to consumer and business to business. Business to analyse the site insights to decide the promotion of a product methodology and offer items to client based on the earlier site visit insights is not easy, as data are not structured and cannot be easily analysed. Determining based on data sets of customers on social media can risk and sometimes can cost you very high. There can be lot of concerns like security, data breach, high cost, understanding market trend for a business. (Sharma, Sawai&Surve.2017)

4.1 Big Data Privacy in Social Media

Big data security related concerns with social media can appear to be secure for customers and business. As a security of big data on social media looks very secure for social media users, but in real its not. There are several layers of data that is being collected. Data Layers can be considered as Visible Layer and Non-Visible Layer. It is separated in two parts one is obvious to all and other is undetectable for all. Unmistakable parts make others to depend on it through account settings, security settings, etc because it the same data which they contribute by themselves. But the undetectable portion is where the user's information is put away and collected for distinctive investigation utilizing different information mining calculations focusing on promotion, showcasing or other employments. (Yin & Zhao, 2015)

Facebook is one of the foremost greatest sources of social media and enormous information supporter. It is already analysed that Facebook is exceedingly criticized over security issues while, most of its locale's subtle elements the sort of data utilized and collected however, their usage isn't detailed within the arrangements. All the data which is collected by it such as user behaviour, related articles, messages, contact list and other data are been collected totally by Facebook and twitter. Facebook has already been criticized as user data were leaked by a group of hackers. Apart from security there are lot of issues such duplicate data, false data, especially when it comes to Big Data. (Bansal, Kumar&Choudhury.2018)

4.2 Analysed Problems and Proposed Solution

Problem: Is the information available on social media are completely secure and these data cannot be accessed by any other 3rd party members?

Solution: Large amount of data is created every day, and no one can say if the data is completely secure but yes, the superior security or an improved level can be given by analysing and executing routinely. As the number of programmers and fake clients are expanding which not as it were collect the information but moreover abuse the same by making fake accounts so, in arrange to spare clients information there's must that screenshot alternative for social destinations ought to be precluded and at the same time client get to caution that xyz tries to require screenshot of your information. Cautioning makes client mindful that this individual is collecting its information through this client can complaint or piece that individual and the level of security is upgraded. This handle ought to be construct in program whereas advancement. In expansion of it one choice ought to be included in which client makes a list of their expensive one who have the consent to spare their information but that as well in information of client (Sharma, Sawai&Surve.2017)

Problem: As the large amount of data is generated everyday on social media. Hoe to determine if the data is useful and how it can be used. Also, how to control duplicate data to be created?

Solution: Most of the information transferred through social media are important information, garbage, copy information, shared information, or information of no utilize. But it doesn't cruel that its information is futile. It is utilized for different examination, enhancements, notice, etc. Particularly in human behaviour and interaction computations and its different investigation social media and its collected information are majorly taking portion. Due to such examination it often appears what client needs that leads to assist client to discover different things of their choices. Not as it were this but these days the overwhelming recommendations are been taken through social media locales which is much simpler than other implies since here most of the individuals cherish to answer or comment and post related distinctive issues and grant their recommendations too. (Bansal, Kumar, 2018)

But in other cruel looking towards such millions of tall persistent creating data-set capacity space gotten by social destinations in enormous information. In arrange to this issue the boundless space in social locales to each single client ought to be constrained and for more utilization it pay-per-use benefit is pertinent i.e. Each client is given with a few great sums of information and the time it is close to wrap up it caution the client. As before long as client conclusion up with the accessible information client must be pay for the information capacity benefit in case client need to transfer any other post separated from shared post. This will keeps up the level and on the off chance that somebody surpasses than it is paying at a few level so, by one means or another it'll keep up as ordinarily client as of now make its reinforcement at different places such as cloud, one drive, etc., moreover put each and each picture or video as web utilization is about free these days by which a part of copy information is made. So, in a result client will post the chosen things. (Bansal & Kumar, 2018)

4.3 Data Mining Complexity

Earlier used data mining methods are troublesome to apply. Single processor isn't enough to do mine. So that huge numbers of processors are required. This increments the fetched too. Information from diverse applications are display in Huge information. Social Media such as Facebook, Twitter, etc. features an expansive number of clients. It can be very complex task to extract data from data mining. Big data is combination of information from several sources. Aggregation of data will be very difficult with large data. he data extracted may be uncertain, it might contain errors and duplicate data, false data etc. which should be solved before the extraction of the data. As solving these issues post mining the data can increase the overall cost. Different algorithms result in different data, and hence it might be very hectic to combine and get insights from the resulted data. Data storage and its management is getting also the biggest challenge as data are increasing very fast pace. (Sebastian, Babu & Kizhakkethottam.2015)

5. Implications for Society and Industry

Analysing the expansive information set on climate & harmful emissions helps in understanding the dangers relating to environment and after that to ensuring the same. Data that leads to data is the modern weapon of defence. Huge information is collected from satellites, capturing messages and unmanned aircraft. Health care is another imperative section with gigantic data challenges, primarily unstructured or connected. It has changed and various partners, counting the government, patient’s medical items businesses and pharma. (Goyal,2017)

Social media posting is where huge information is playing a crucial role. Estimations investigation, political see point, psychological aspect brings a common agreement of a community. Sports moreover utilizing enormous information analytics through video analytics, instruments and sensor innovation. Execution of players are monitored and followed to deliver superior results. Nutrition and rest orders of players are followed for social wellbeing. The computational control of enormous information may also be connected to nearly each set of information, opening new avenues to researchers. To

induce way better pictures of our wellbeing and social sciences census and government collected information can more effectively be gotten to and analysed by researchers. (Goyal,2017)

The information produced through worldwide area following are growing rapidly, reflecting the prospering appropriation of smart phones and other applications helps in preparing personal location information, and usually the space that incorporates law enforcement, child security, travel arranging and tracking terrorists. Big information computing is barely utilized in making strides the productivity of open segment whereas the governments have access to expansive pools of computerized information. Data inferred from data ought to be shrewdly utilized to bring changes in this division. (Goyal,2017)

6. Suggested Course of Action

Traditional methods to extract the information of users to understand the user behaviour have lot of technical issues based on security, reliability. The information also provides the incomplete information on user behaviour and data about user interest. Thus accuracy, data duplication, security, data breach are still the problems and major issues with extracted information. To minimise the listed issues, several future studies and framework should be taken in account. (Qu,Li,Wang,Yin & Zheng, 2018)

6.1 Deep Graph-Based Neural Network (DGBNN)

Despite of technical advancement efficiency and accuracy are main concern in data extraction. Deep Graph-Based Neural Network is applied to solve these complex problems. DGBNN system consists of two primary parts. The primary portion is highlight extraction which is divided into two steps. The primary step is that utilizing Latent Dirichlet Allotment to induce the subject dispersion of each user and at that point to decrease the information measurement. The second step is to construct a low-dimensional representation for each vertex to extricate highlights of social joins. Within the second part, the information is set into the input layer and classified by Back-Propagation Neural Network. The system accomplishes more productive and accurate recommendation by anticipating in case there exists a social link between two clients. Many Studies, with the fast increment within the sum of information, more efficient deep learning strategies will be planned to solve complex and abstract issues. (Qu,Li,Wang,Yin &Zheng, 2018)

6.2 Efficient Classification of Social Media Data

Prediction and classification of data is trending area of research. As data is growing at very pace on social media sites, but extraction and collection of useful data is still a major issue. Better classification of data will help the business analyst and business to take better decisions for business. Classification of data can help the business to understand the customer interests and behaviour. Thus, integration of algorithm is required for better and useful classification of Big data. Below is algorithm which can be used to remove unwanted data and classification for useful data. (Desai & Patil. 2015).

I. Decision Tree Algorithm

Decision tree is exceptionally valuable classification method. Trees are exceptionally adaptable, simple to understand, and simple to investigate. They will work with

classification issues and relapse issues. So, if you are attempting to foresee a categorical esteem like (ruddy, green, up, down) or if you're attempting to foresee a nonstop esteem. Main point about the Decision Tree is that it just requires a set of data in tabular form and it will be building a label classifier automatically from the data given without any extra effort to be done. Decision tree is classification of data in tree structure which is easy to predict. Decision tree consists of three elements.

Decision Node - Holds the decision from given set of data.
Leaf Node – Holds the value of a given attribute
Path – Holds the value attribute of decision node to conclude the final decision.

Decision trees classify occurrences or illustrations by beginning at the root of the tree and moving through it until a leaf node. The tree can develop tremendous number of decisions depending on the data given. These trees simplify the big data and makes it easy to read and understand. (Desai & Patil. 2015).

6.3 Future Trend of Social Networking Data

The upcoming ways in social organizing will emphasize on the different ways to share big data. In the current used technologies, it is watched that the full number of important contents generated is expansive, but it will get to be indeed bigger as more clients connect these online socials organize destinations. Although useful data is an alluring include, numerous of the things appeared there are essentially not significant enough for any client to need to spend a part of time perusing. Also, relevant upgrades from closer companions are effectively neglected since there's such an expansive volume of data to sort through**.** Also, how data are shared on network has improved a lot in order to keep the user data private. (Qu,Li,Wang,Yin &Zheng, 2018)

6.4 Improving Security

Cybercrime costs $118 billion every year and this figure is anticipated to develop altogether growing. (Arora, Malik 2015). With simply get to information accessible online, modern cybercrimes are occurring at a disturbing rate due to which conventional security solutions are now not adequate to guard against these heightening dangers. Instances of hacking, character robbery and stealing credit card information from retailers and banks are within the news quite regularly but later modern and organized breaches at sony including an unreleased movie have shaken the world. While a part still must be done to avoid cyberterrorism, Big information analytics in security presently offers promising solutions towards proficient discovery of suspicious exercises over the network. It is anticipated that huge information analytics will impact various viewpoints of data security such as organize observing, client confirmation and control, authorization, identity management, extortion location, information misfortune anticipation and control.

Number of cyber detection techniques are being proposed to implement with Big Data extraction. As mentioned in earlier section, techniques like deep graph based neural networks, decision tree can be efficiently used to detect the unusual behaviour on big data. Several attacks on data like denial of service (DOS), malware can also be detected

by classification model. Many companied like IBM are working on to protect the data from being hacked by users using malware and other unusual activities. (Arora & Malik, 2015)

7. Conclusion

The enormous growth in generation of data allows the business to take better decisions for business. Social media being powerful tool have helped the business to reach and interact with millions of users. Social media data collection helps the businesses to understand the customer behaviour and their buying patterns. Big Data provides the ability to target the right set of audiences within a group. Using Machine learning techniques data are segregated in more useful and clearer format and these data are used by marketers to target customers on social media platforms. With the use of Big data on social media allows the data to be collected and segregated in much useful and meaningful format. It allows the companies to market right product to right set of customers on social media, which increases the profits for company by increasing the visibility of the product and boosting sales while decreasing the advertisement cost. Understanding data can be complex and difficult, so many data classification techniques are integrated with Big Data such as Decision Tree Algorithm. Decision Tree Algorithm efficient segregates the data into different sectors and this segregation makes the data more useful for the business decision. Various analytics tool such as Spark, Hadoop allows the data and information to be presented in graphical way which enhances the knowledge of the business to market the product based on user interest. Visualization is presentation of data which can help the industries and business to take powerful decisions and be competitive in the market. Despite of many positive impacts there are many other concerns related to big data and social media which should not be overshadowed. Data Security is major concerns that is deploying the world in the today's era. It causes huge losses to the industry and the society. The above-mentioned algorithms are not widely used, using them can help the organization to improve their efficiency and helps for reliability of the society.

8. Annotated Bibliography

Adegbuyi, O. A., Akinyele, F., & Akinyele, S. (2015). Effect of Social Media Marketing on Small Scale Business Performance in Ota-Metropolis, Nigeria. *International Journal of Social Sciences and Management*, *2*(3), 275–283. Retrieved on October 21,2019 from https://www.nepjol.info/inde x.php /IJSSM/ article/view/12721

This article describes the effect of social media marketing on small scale business. Social media plays vital role for business in marketing and interactions with the customers. For small and medium business, it's a huge platform to advertise the product and attract the customers. Social media sites like Facebook, Twitter and Instagram makes it possible for the small business to understand the customer behaviour and interest. Through these platforms, business also understands about the trending products with can help them grow in digital commerce.

The use of social media helps the business to grow their sales online. On a theoretical construct, social media participation helps the business to establish a brand name in the market and gain brand trust by providing authentic product and service. Brand awareness will directly help to gain more customers and Brand trust will help in customer retention. These all will result in increasing the market share and more profit for a business.

Amrita, S. M., & Mohan, R. (2016). Application of social media as a marketing promotion tool — A review. *2016 IEEE International Conference on Computational Intelligence and Computing Research (ICCIC)*. Retrieved on November 10,2019 from https://ieeexplore-ieee-org.proxy-bcResearchportumd.edustamp/stamp.jsp?tp=&ar number=7919587

This paper discusses about the social media being powerful tool for the business to interact and promote the products online. Social media is the most trending word of the market. Marketing and advancement are exclusively depending upon social media which was a sideline media amid its viable beginning a decade back. Print media and tv media misplaced its pertinence to social media, since of numerous variables counting fetched adequacy and wider reach. Nearly all areas have received Social media for its effective and productive working. Social media viability is mostly utilized in all statistic sections as the unused age millennial cannot be disconnected from social media.

Promoting product through social media is superior than the traditional promotion approach. Modern day social media networking connects people having such interpersonal weak ties and thereby could be used for the spread of information. It has become so strong that every business uses social media platform to attract the customers and grow the business.

Arora, D., & Malik, P. (2015). Analytics: Key to Go from Generating Big Data to Deriving Business Value. *2015 IEEE First International Conference on Big Data Computing Service and Applications*. Retrieved on November 14,2019. from https://ieeexplore-ieee-org.proxy-bc.researchport. umd.edu/ document /7184914.

In this paper, different techniques and concerns related to security has been addressed. These techniques are widely used for the classification of data which provides useful insights to business for better understanding and decision making. Although, the security on big data applications seems to be improving, there are still many instances of data leakage by hackers. This impact the business most as trust between the parties are totally lost and regaining again the trust is very big challenge.

Number of cyber detection techniques are being proposed to implement with Big Data extraction. Techniques like deep graph based neural networks, decision tree can be efficiently used to detect the unusual behavior on big data. Several attacks on data like denial of service (DOS), malware can also be detected by classification model. Many companied like IBM are working on to protect the data from being hacked by users using malware and other unusual activities.

Bajaj. R, Bhanot. R, Srinivasan. R, (2016), Impact of Social Media Marketing Strategies used by Micro Small and Medium Enterprises (MSMEs) on Customer acquisition and retention, *IOSR Journal of Business and Management.* Retrieved on October 25,2019 from https://pdfs.Semanticscholar.org/ac3a/0716ea01a9829cdabc32 b148 5992a057a13d

In this paper, the author discusses about the impact of social media marketing strategies on small and medium sized business. Small scale businesses have more marketing opportunities nowadays and technology and networks end up giving support to these businesses. Networking is very important for the establishment, development and growth of such businesses and they rely heavily on personal contacts as well as positive word of mouth from customers. Social media helps form such networks and online word of mouth spreads very quickly on social media.

The use of social media by businesses helps companies obtain feedback regarding their products by monitoring customer's online communications. It also helps promote word of mouth and at the same time companies have limited control over the content and distribution of information and there is large impact of user generated content. Marketing right product to right customer helps the business to gain more profit while decreasing the cost of marketing.

Bansal, S., Kumar, P., Rawat, S., & Choudhury, T. (2018). Analysis and Impact of Social Media and its Privacy on Big Data. *2018 International Conference on Advances in Computing and Communication Engineering (ICACCE).* Retrieved on Oct 25,2019 from https://ieeexplore-ieee-org.proxy-bc.researchport.umd.edu/document/8458066.

In this article, the authors mainly discuss about the security concerns on private data. Social Media as a platform to connect and interact with millions allows users to upload the data and without much information users also upload personal data online which becomes security threat. This personal are also being monitored and captured by various platforms which might affect user privacy. The paper focus on big data privacy related issues on social media.

To support an example of security threat, Facebook which one of the foremost greatest sources of social media and enormous information supporter. It is already analyzed that Facebook is exceedingly criticized over security issues. In last few years there are news that hackers have stolen the data from Facebook, which is one of the major security concerns.

Bhor, H. N., Koul, T., Malviya, R., & Mundra, K. (2018). Digital media marketing using trend analysis on social media. *2018 2nd International Conference on Inventive Systems and Control (ICISC)*. Retrieved on Oct 17,2019 from https://ieeexplore-ieee-org.proxy-bc.researchport.umd.edu/document/8399038

This article explains how social media is lifeline in today's digital world where there are lot of users. Using social media allows companies to find potential customer base and for interaction social media marketing is used to showcase the products. Facebook, Instagram, Twitter are most used social media platforms for digital marketing. Social Media marketing has become so strong that every business uses social media platform to attract the customers and grow the business.

The main focus of this paper is on trend analysis which allows the business to have personalised interaction with the customer. This helps the business to create much more in-depth marketing campaign with increasing sales. Marketing without strategy is just waste of money without any output. For tend analysis, profile is classified into different hasgtags and this hashtag allows the companies to target specific set of the targeted audience or customers. Digital marketing has made it very easy for the business to showcase the product catlog to large audience without much effort.

Desai, S., & Patil, S. (2015). Efficient regression algorithms for classification of social media data. *2015 International Conference on Pervasive Computing (ICPC)*. Retrieved on November 10,2019 from https://ieeexplore-ieee-org.proxy-c.researchport.umd.edu/stamp/stamp.jsp?tp=&=7087040.

In this paper, the author discusses about the various algorithm that can be used for further classification of data sets accurately and provide results based on need

accurately. It is found that one of the algorithm Decision Tree algorithms on random data sets classify the data in more impactful way and provide insights about the data. Prediction and classification of social media data are still in research for better accuracy of resulted data. The analysis will help the business to get feedback on offered products and services which can further help the business to improve the product and service.

As data is growing at very pace on social media sites, but extraction and collection of useful data is still a major issue. Better classification of data will help the business analyst and business to take better decisions for business. Classification of data can help the business to understand the customer interests and behaviour. Thus, integration of algorithm is required for better and useful classification of Big data.

Goyal, Abhishek. (2017). Impact study of big data on capitalism and socialism. *2017 3rd International Conference on Computational Intelligence & Communication Technology (CICT)*. Retrieved on November 3,2019 from https://ieeexplore-ieee-org.proxy-bc.researchport.umd.edu/ document /7977301

In this paper, Big Data is considered as next big thing. The paper discusses about the implementation of the big data on social life. Big data analytics is present in every field to help the business to take better decisions. The author in this paper mainly focuses on field like health and education data sets which mainly help to industry in health and education sector to take better decisions. There are lot of other fields yet to adapt the big data analytics. These analytics help the companies to create several predictive models to better understand its customer, interest and behaviour. Big data on web is also used to optimize the business based on predictive data the business can refill their stock.

The author describes the impact on society from the use of big data analytics on social media. Big data is so powerful that it can be used in any field of business. The data extracted should be used properly to bring changes in the society.

Hanamanthrao, R., & Thejaswini, S. (2017). Real-time clickstream data analytics and visualization. *2017 2nd IEEE International Conference on Recent Trends in Electronics, Information & Communication Technology (RTEICT)*. Retrieved on Oct 16,2019 from https://ieeexplore-ieee-org.proxy-bc.researchport.umd.edu/document/8256978

This paper discusses about the clickstream data generation. Every time a user on social media clicks a new data is recorded and this recorded data is used to understand the customer insights and behaviour. It refers to analyzing every clicked data and its optimization for useful use by the companies. Data analytics and Visualization is very important which presents the data in much easier form to understand and gives deep insights about the data. Many tools and technologies such as Apache Spark, Hadoop and Elastic Search are used for Data Analytics.

In this paper, Apache Hadoop, Apache Start, ElasticSearch, are the parts which are joined together to to analyse the data. This real time clickstream data analytics allows the companies to predict the business. Also, this will help the companies to take real time decision. ElasticSearch is more powerful tool than Hadoop and Spark for real time analysis of the click streamed data.

Park, K., Nguyen, M. C., & Won, H. (2015). Web-based collaborative big data analytics on big data as a service platform. *2015 17th International Conference on Advanced Communication Technology (ICACT)*. Retrieved on October 16,2019 from http ://ieeexplore.ieee.org /stamp/stamp.jsp?tp arnumb er=722 4859&isnumber=7224736

This article published by many authors discuss about the new challenge to store and process large chunk of data. Due to social networks and cloud computing an enormous amount of data are being generated every day. The paper focus on solving the problem related to storing and analyzing the Big Data. A Big data analytics platform is mixture of algorithms and analytical tools which process the data much faster. It supports to manage the large chunk of data and algorithms are developed to process these data easily which are used by the business to understand the user behavior and interest.

Big data as a service is a tool which analyses the data and provides useful insights to gain competitive strategies. Big Data consist of three layers Big Data Infrastructure as a service, Platform as service, Big Data Analytics Software as service. Big data Analytics software as a service is a combination of Big Data Infrastructure as service and Big Data Platform as a service. It provides web-based environment and allow sharing of resources. Web based environment is used to post their own information.

Qu, Z., Li, B., Wang, X., Yin, S., & Zheng, S. (2018). An Efficient Recommendation Framework on Social Media Platforms Based on Deep Learning. *2018 IEEE International Conference on Big Data and Smart Computing (BigComp)*. Retrieved on November 10,2019from https: //ieeexplore-ieee- org. proxy-bc.researchport. umd.edu/stamp/ stamp.jsp?tp=&arnumber=8367184

The paper describes about the traditional methods to extract the information of users to understand the user behaviour have lot of technical issues based on security, reliability. The information also provides the incomplete information on user behaviour and data about user interest. Thus accuracy, data duplication, security, data breach are still the problems and major issues with extracted information. To minimise the listed issues, several future studies and framework should be taken in account

Efficiency and accuracy are main concern in data extraction. Deep Graph-Based Neural Network is applied to solve these complex problems. DGBNN system consists of two primary parts. The primary portion is highlight extraction which is divided into two steps. The primary step is that utilizing Latent Dirichlet Allotment to induce the

subject dispersion of each user and at that point to decrease the information measurement. The second step is to construct a low-dimensional representation for each vertex to extricate highlights of social joins.

Sand, G., Tsitouras, L., Dimitrakopoulos, G., & Chatzigiannakis, V. (2014). A big data aggregation, analysis and exploitation integrated platform for increasing social management intelligence. *2014 IEEE International Conference on Big Data (Big Data)*. Retrieved on October 10,2019 from https://ieeexplore-ieee-org.proxy-bc.researchport. umd.edu/stamp/stamp.jsp?tp=&arnumber=7 00 4411

This paper published by many authors mainly discuss the technical aspect on how Big data and social media has changed the way people used to interact earlier. Web and Social media are playing very crucial role to connect the people on virtual network where users share their thoughts and ideas. This exchange of information leads to enormous collection of data which allows companies to understand the customer behaviour and to take better decision. This collected data allows the companies to target right group of users to market their products. The web and Social Media are used by almost every companies to interact with the sellers and buyers to increase the company visibility about the product and services on web. It helps the companies to understand customers opinion about the products and services offered by the company. This paper includes the technical several technologies and marketing practices.

To understand the big data and social media interaction technology GlobaliFusion is to explain the technical details in depth. GolbaliFusion(GiF) brings small size business, Medium size business and Large scaled business together with its customers by collecting data from social media and converting the data into valuable form which is used for planning marketing strategies and to increase the growth in business by using technology innovation

Sebastian, L. R., Babu, S., & Kizhakkethottam, J. J. (2015). Challenges with big data mining: A review. *2015 International Conference on Soft-Computing and Networks Security (ICSNS)*. Retrieved on November 10,2019 from https://ieeexplore-ieee-org.proxy-bc.researchport.umd.edu/stamp/st amp.jsp?tp=&arnu mb er=7292371

This paper focuses on challenges with Big Data extraction. As data are growing day by day, the main challenges are data storing, transferring data and capturing data. It can be very complex task to extract data from data mining. Big data is combination of information from several sources. Aggregation of data will be very difficult with large data.

The data extracted may be uncertain, it might contain errors and duplicate data, false data etc. which should be solved before the extraction of the data. As solving these issues post mining the data can increase the overall cost. Different algorithms result in different data, and hence it might be very hectic to combine and get insights from the resulted data. Data storage and its management is getting also the biggest

challenge as data are increasing very fast pace. Accuracy of data after data mining can cab help the business in target based digital marketing. This will also minimize the overall cost for the company.

Sharma, N., Sawai, D., & Surve, G. (2017). Big data analytics: Impacting business in big way. *2017 International Conference on Data Management, Analytics and Innovation (ICDMAI)*. Retrieved on Oct 24,2019. From https://ieeexplore-ieee-org.proxy-bc.researchport.umd.edu/ tamp /stamp .jsps? tp= & arnumber= 8073494

This paper discusses about the use cases on buying pattern, customer behavior, user interest in social media. IOT devices are used to capture clickstream data from web and social media based on location, purchase history. With the use of internet and intelligent machines like sensors for information to capture the data and process this data through analytical tool gives insights about the customer. As data are generated in large scale, Big Data framework is used to store and process the data for success of business.

In this paper, Concerns of big data are discussed. Business to analyse the site insights to decide the promotion of a product methodology and offer items to client based on the earlier site visit insights is not easy, as data are not structured and cannot be easily analyzed. Determining based on data sets of customers on social media can risk and sometimes can cost very high. There can be lot of concerns like security, data breach, high cost, understanding market trend for a business.

Yin, J., & Zhao, D. (2015). Data confidentiality challenges in big data applications. *2015 IEEE International Conference on Big Data (Big Data)*. Retrieved on Oct 25,2019. From https://ieeexplore-ieee-org.proxy-bc.researchport.umd.edu/stamp /stamp.j sp? tp =&arnumber=736 4111.

In this paper, Security challenges in data analytics are discussed. Big data security related concerns with social media can appear to be secure for customers and business. As a security of big data on social media looks very secure for social media users, but in real its not. There are lot of data are collected which user are not aware. Every user click on social media is recorded for the behavioural study. Security is one the major concern in digital world.

There are several layers of data that is being collected. Data Layers can be considered as Visible Layer and Non-Visible Layer. It is separated in two parts one is obvious to all and other is undetectable for all. Visible layer includes the basic profile which anyone can see but the invisible layer is the data collected of which user is not aware. This can include data related to transaction, types of content viewed, and every click.