Social Media Analytics

Riya M. Hate and Vaishnavi T. Naik

Department of Computer Science

Shri Bhagubhai Mafatlal Polytechnic, Mumbai, Maharashtra, India

riyahate9@gmail.com and vaishnavinaik2003@gmail.com

Abstract: *Social media is a broad spectrum of process like data collection, data storage into databases and preparation of data which is utilized for research, decision making, marketing campaign with the aid of various tools and algorithms to measure the performance of a product. Analyzing social media to collect data and planning strategy around them has become one of the major business activity. Topics covered in this papers would be applications and impact of social media, challenges faced while retrieving data and preparing data for business intelligence, it also mentions some tools and algorithm.*

Keywords: Business Intelligence, Decision Making, Social Media, Data Analysing, Tools, Data Preparation.

I. INTRODUCTION

Social media is a creative platform where people create, share their interests, ideas, communicate and form their own internet community. It is both mobile, web based interactive site and in this era almost everyone has access to it. Social media is also used to track user's behaviour to advertise the products based on the tracked information. Social media analytics is referred to the process of gathering data from social media platforms like Instagram, Facebook, Twitters and using these data to make various business decisions. Social media analytics is a bigger than just likes, follows, retweets and impressions. It provides services which can support various marketing campaigns. Social media analytic is a designed software which monitors data on web search tools, here data about certain keywords are retrieved, stored and then worked upon to derive reliable decisions. There are various tools available in the market to monitor user's data and give appropriate results and based on which you can make decisions related to your business. Social media produces ton of content every single day which is utilized by organisation to keep a watch on trends, seek in to collect data in order to achieve insights. This data is often collected using some tool with their respective social media network. There are three main domains where social media is important and generates profit to organisation:

1. In business
2. In crisis management
3. Journalism

II. APPLICATION AND IMPACT

The web and various mobile technologies have been the foundation for the growth of social media that serve as a medium for content generation, information dissemination and interactive communication. social media includes user generated content or consumer-generated media that majorly define the Web 2.0. Considering the application aspect, social media websites are the most popular and ones Wikipedia (collective knowledge generation), Instagram and Facebook (social networking), Digg and Delicious (social browsing, news ranking, and bookmarking), YouTube (social networking and multimedia content sharing) and Twitter (social networking and microblogging) are some of them. Profit organizations are engaging into social media to obtain information, execute business plans related to product design and innovation, manage consumer and stakeholder relations, operate their brand awareness analytic and fraud detection. In finance, social media is used for measuring market sentiment and news data is used for trading. The Commercial industry sees social media as a boon for business intelligence. For example, Bollen et al. (2011) measured sentiment of random sample of Twitter data, finding that Dow Jones Industrial Average (DJIA) prices are correlated with the Twitter sentiment

2–3 days earlier with 87.6 percent accuracy. Social media plays a vital role in politics. Politicians, political parties, and governments seek help from social media to gauge public opinion on policies and political positions, also to build community support for candidates running for public offices. For example, Lerman et al. (2008) use computational linguistics to automatically predict the impact of news on the public perception of political candidates In the biosciences, social media is being used to collect data about large groups of people for behavioral change initiatives and impact monitoring, such as tackling smoking and obesity or monitoring diseases. An example is Penn State University biologists (Salathe et al. 2012) who ´ have developed innovative systems and techniques to track the spread of infectious diseases, with the help of news Web sites, blogs and social media. Social media can also be used to receive early clues about disease outbreaks and to provide feedback on public-health policies and response measures. social media provides abundant opportunities to study terrorist group behaviour, their recruiting and public relation schemes. Even think tanks and social science and business researchers are conceptually use social media as sensor network and a medium for natural experimentation, to test hypotheses about social production and interactions and their economic, political, and societal impact. Generally, for netizens social media is a unique information source to deal with information, find answers to questions, and discover opportunities for social and economic exchange. It has become a platform for people to network and contribute to all kinds of dynamic dialogues by sharing their expertise and opinions.

III. CHALLENGES

We first start with explaining some terms

1. **Natural Language Processing (NLF):** It is computer science or artificial intelligence department which is involved with communication between computers and human languages.
2. **Scrapping:** It is a process of collecting data from various social media channels in form of unstructured text also known as web scrapping or site scrapping or web data extraction.
3. **Text Analytic:** It is the process of drawing meaning out of written communication, it is mainly used for deeper insights, such as identifying a pattern or a trend from unstructured text.

Challenges There are 4V's to identify the type of challenges, the concept of 4 V's was proposed by McAfee and Brynjolfsson(first three V's in 2012)

1. **Volume:** The amount of data produced on a daily basis is mesmerizing and requires huge amount of storage to save this data. This acts as one of the biggest obstacle for Manual Inspection. Challenges occurred in discovery steps are mostly because of volume.
2. **Velocity:** This term refers to the frequency of incoming data that needs to be processed. With millions of users using social media at the same can produce a large amount of data.
3. **Variety:** Data has many different forms, it is often unstructures or its structure depending on the source of data. Structured data is augmented by unstructured data, anything can be captured but it does not have a meta model. Whereas with Unstructured data, there are no such rules it could be anything like a tweet, photo, a voice recording but which express ideas and thoughts based on human understanding. Here the challenge is to take technology to filter data out and make sense out of it.
4. **Veracity:** This refers to the reliability of the data, challenge is to check the quality of the data whether it could be useful or not. Unstructured data and uncertain nature of this data creates a new challenge that how to appraise the quality of data and also filter out the unnecessary data.

IV. SOLUTIONS

Problem: Data growth issues (Volume) One of the greatest problem is storing such huge amount of data properly in such way that it is easily managed. As these data grow with time, it becomes extremely difficult to handle data. Most of the data is unstructured, meaning it comes from documents, videos, text files etc. And you cannot find such data in databases.
**Solution:** In order to handle data, companies are opting for modern approach such as compressing, avoiding redundancy and tiering. Compressing is used to reduce the number of bits from data. Avoiding redundancy is to remove all unwanted and duplicate data. Tiering is to store data in different storage tiers. It ensures that data is residing in proper storage space.

Data tiers can be any cloud deployment model (public, private and hybrid) or flash storage depending on it's size and importance.

Problem: Securing Data In current world, data is one of the biggest asset and securing it one of the daunting challenge of company. Companies are sometimes so busy understanding, manipulating data that they push data security to later stages.

Solutions: Companies should hire more cyber security professionals to take steps like: data encryption, data segregation, identity and access control, implementing endpoint security.

V. SOCIAL MEDIA ANALYTICS PROCESS
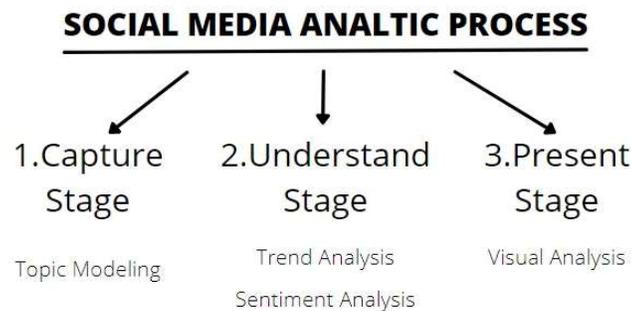
The process of social media is divided into 3 stages "capture"," understand" and "present".

1. **Capture stage:** here the relevant data is collected using social media listening. Social Media can be performed by organizations or third party vendors. Not all data that is collected is useful therefore we move to the next step. For a business engaged in social media analytics, this stage to classify discussions related to its interests. This is done by collecting a huge amount of data from all social media sources such as APIs, feeds etc. This capture includes all popular platforms such as Facebook, Twitter, Instagram, YouTube, Pinterest, Google, LinkedIn, Tumblr etc. along with smaller sources such as blogs, microblogs, Wikipedia, news sites, etc. To prepare this collected data for the next stage, it performs various steps such as data modeling, linking of data from different sources, feature recognition and other operations to support analytic.

2. **Understand stage:** This stage filters the data for modeling, cast away low-quality data. Once the organization has collected conversations related to its products and campaign, then it must analyze the data to make future business decisions. Since the Capture stage gathers data, a sizeable part of data needs to be removed using text classifiers or advanced text classifiers to do the cleaning function. Deriving meaning from cleaned data may include various statistical methods and techniques derived from data mining, machine translation, and network analytic. This stage provides information about how users feel about the company's products. This stage is the core of the entire social media analytic process.

3. **Present stage:** This field displays data filtered in Stage 2 in a meaningful way. Here in the last stage the data collected in Understand Stage is summarized, calculated, and presented to the user in a meaningful format. One of the most commonly used techniques is Data Visualization. There are certain tools that help with data visualization are Tableau, Looker, Zoho Analytics, Sisense, etc.

VI. SOCIAL MEDIA ANALYTIC TECHNIQUES

Social Media Analytic is a growing field with lots of modeling and analytic techniques. These techniques are useful in various social media analytic processes.

1. **Opinion Mining or Sentiment Analytic:** It is an approach to Natural Language Processing that identifies tone of the message. This is one of the most successful way of sort out the review, service or idea about the product. Sentiment analytic help Companies to collect information from unstructured text that comes from online social media channels such as mails, blogs , feed, chats, etc. Algorithms replace manual data processing by

implementing automatic or hybrid methods. Rule based system performs sentiment analytic based on predefined rules. Automatic system perform sentiment analytic with the help of Machine Learning techniques. A Hybrid sentiment analytic is combination of both Rule based and automatic. Vendors who offer sentiment analytic are: Hootsuite, NetBase, Zoho, Sysomos, Sprout Social.

2. **Topic Modeling:** It is an unsupervised ML technique that's capable of scanning texts, documents and detecting phrases, words, patterns within them and grouping them accordingly. It is called unsupervised because it does not require pre-fed data to train the model. That is why it is quick and easy way to start analyzing your data. Classifying topics require training, therefore, it is called supervised. Topic classification needs to know the topic of of certain text or document before grouping them. Using these topics, data is manually tagged. Although topic classification requires more effort is=t gives accurate results.

3. **Trend Analytic:** Brands use social media Trend analytic to predict future trends but tracking the trend movement, future outcomes and behaviours based on previously collect data . This includes framing marketing strategies and positioning your brand with current and emerging trends. Application of trend analytic include growth of customers or sales, predict the success of campaign.

4. **Visual Analytic:** It is a tool used to analyze data using visual presentation of data in form of graphs, charts, maps and flowcharts which help to identify the pattern in them, hereby developing an actionable insights. These insights help organization with data driven decisions. Few benefits of visual analytics are: 1. Finding and Tracking progress: Interactive reports help users track, organize and share performance across an organization. 2. Fast Decisions: With the help of interactive reports, users quickly understand the data and it makes it easier to work with data sets when they are in visual formats. 3. Encourage Data Literacy: Making data easier for everyone to understand and get people involved across the organization.

VII. TOOLS

An ideal analytics tool contains an intuitive, easy-to-use interface which enables transparency in a campaign; it should also streamline the social media marketing processes and workflows.

1. Sprout Social Sprout Social it's a cross-channel social media analytics, that helps businesses to deep dive in to their performances on single or multiple networks without delay. Be it to tally up Twitter link, Track Facebook Page impressions, measure Instagram follower rise, evaluate LinkedIn engagement or to check content performance on Pinterest all of it will be done with the assistance of sprout from the identical location. It further also organizes data into an presentation ready reports that may be understood by anyone easily.it facilitates creating custom reports tailored to the organization's KPIs with the premium analytics facility. Sprout comes with various reporting options that can be templated or customizable, so you've got multiple ways to not only gather important social metrics, but thoroughly understand their interpretations and also present them to others.

2. **HubSpot:** HubSpot's analytics tools is that the easy method to bind social media to business and revenue growth. To publish directly through HubSpot, or publish using other tools, organizations is report the success of their social strategy across every channel. It also enables to match the performance of multiple social channels to live campaign performance. HubSpot's social analytics tools offer vast graphs and analytic-visuals that aid to interrupt down the numbers by particular platform features, like audience, session lengths, and impressions.

3. **Google Analytics:** Google analytics isn't solely a social media analytics tool, it's one in all the most effective ways to trace social media campaigns and even helps to live the social ROI. Companies can access and make reports specifically for social media tracking. To be precise, what proportion traffic involves the website from every social network independently are often found out. UTM parameters are often accustomed track specific social media campaigns.

VIII. THE BUSINESS VALUE OF SOCIAL MEDIA ANALYTICS

Social media has changed our perspective about products and services but not the business activities underlying them. There are a spread of business applications of social media analytics. Here we discuss the life cycle analytic framework.

A life cycle analytic perspective looks upon the lifetime of a product or service from its design to disposal, and also the support activities that run in parallel with these activities. Though there are various versions of the merchandise life cycle with different levels of description.

One that suffices for our purposes contains these four stages:

1. Design-development
2. Production
3. Utilization
4. Disposal

With these stages we discuss how social media analytics help firms to grasp their business environment, suppliers, and competitors precisely. Product design-development: Product design-development The stage covers the conceptual, preliminary, and detailed design of a product. Various risks threaten success during this stage like, risks involving technology change may create gaps in technology among different products. Design risks arise due to poor selection of product features, improper differentiation with other products. Trend analytic and other social media analytic tools can help to bring changes in tastes, behaviours, and other sentiments which will affect product design and development. These tools can allow features to be added or adjusted, and build sufficient time interval for creating "next generation" products.

1. **Product Production:** Product Production Social media analytics can diminish risks of supply chain responsiveness. companies can anticipate significant changes in demand by being attuned to changing tastes and behaviours, then in keeping with the analytic they'll ramp up or down production. Visual analytics help in pointing out correlations, outliers, geographic patterns, or other trends that support smoother functioning.

2. **Product Utilization:** Product Utilization social media analytics is most useful during the merchandise utilization stage. During this stage, there are three main social media objectives: brand awareness, brand engagement, and word of mouth. Brand awareness gives knowledge of the merchandise to customer and increases their familiarity. Brand engagement increases connection of the user with a brand. Word of mouth encourages users' attempts to positively persuade other users' to shop for the merchandise. influencer profiling uses social media to develop a deep understanding of different categories of users to make better customer segmentation. Segmentation assists a business in effectively recognizing product utilization in step with the kind of users.

3. **Product Disposal:** The customers must have complete knowledge of disposal or replacement of the merchandise. it's the company's duty to provide this information and ensure harmless disposal of their products. For variety of consumers, ecological disposal of a product may influence their overall impression of a company and its products. Social media analytics can track conversations of the general public covering disposal. Through this information the can adapt to new techniques that ensure their goodwill. Competitive Intelligence up to now we discussed business values of social media analytics with regard to the companies their products/services and their customers. Social media analytics also provides businesses with value by helping it understand its environment, suppliers, competitors, and overall business trends to remain competitive. This value is termed as competitive intelligence. It enables the firms to be told about their suppliers, other brands within the market etc.

IX. CONCLUSION

Social media analytics is the new area of interest to the information systems community. Through this article we tried to describe multiple aspects of social media analytics. We have covered the basic idea of social media analytics, applications and impact, the challenges faced in this process the solutions to them, detailed description of the process and the business value of social media analytics. Lastly, this research also includes tools like sprout social that helps businesses to deep dive into their performances by organizing data into reports.

REFERENCES

[1]. Ria Andryani, Edi Surya Negara, and Dendi Triadi. "Social Media Analytics: Data Utilization of Social Media for Research". In: Journal of Information Systems and Informatics 1.2 (2019), pp. 193–205.

[2]. Bogdan Batrinca and Philip C Treleaven. "Social media analytics: a survey of techniques, tools and platforms". In: Ai & Society 30.1 (2015), pp. 89–116.

[3]. Weiguo Fan and Michael D Gordon. "The power of social media analytics". In: Communications of the ACM 57.6 (2014), pp. 74–81.

[4]. Gabriele Pierantoni et al. "Flexible Deployment of Social Media Analysis Tools." In: IWSG. 2018.

[5]. Stefan Stieglitz et al. "Social media analytics–Challenges in topic discovery, data collection, and data preparation". In: International journal of information management 39 (2018), pp. 156–168.

[6]. Daniel Zeng et al. "Social Media Analytics and Intelligence". In: IEEE Intelligent Systems 25.6 (2010), pp. 13–16. DOI: 10.1109/MIS.2010.151