

Social Media Analytics using Machine Learning

Manasvi Varshney

Department of Computer Science and Engineering,
School of Engineering and Technology, Sharda
University, India.

Sparsh Singh

Department of Computer Science and Engineering,
School of Engineering and Technology, Sharda
University, India.

Vudit Kumar

Department of Computer Science and Engineering,
School of Engineering and Technology,
Sharda University, India.

Abstract:- Social Media is one of the platforms that produces large amount of data. According to a survey there are around 250 million post per hour from billion users. Data retrieve from social media is to analyze behavior and perception of the community based on particular event or phenomenon. Social media is a platform where users upload their data in different forms like images, text, videos, etc., These large volumes of data are analyses to extract meaningful information through tools and technologies which comes under social media analytics. One of the biggest data producing sources are YouTube, Facebook, Instagram, and Facebook. Social media Analytics tools and technologies are used for sentimental analysis of these platforms to extract insights from it. The goal of this paper is to perform social media analytics through machine learning.

1. INTRODUCTION

To define social media is nothing but a web-based services that allows individual to share post, update their profiles, enhance their connection, and also allow people to connect with outer community and collaborate with each other [1]. Social media is widely used as a communication tool for politics, business, fashion, tourism and so on [2]. In recent years, people are highly influenced by social media and it was expected that social media users will be 2.77 billion in 2019 [3] due to which it alone produces large amount of data. According to recent survey Facebook produced Four pentabytes of data every day in 2020 and the number is increasing every year. SMA is used to developed framework to evaluate, analyze, visualize, and collect data from social media to ease the conversation and extract useful information [4].

SMA can be used to extract information from social media that can be used in various fields like politics [5], agriculture [6], business and so on. The process of SMA is done by collecting, storing, processing and analyzing [7]. Facebook is one of the most popular social media networks whose data can be retrieved for analysis. It allows their users to create an account, make friends, chat with people, share videos and many more which is one of the major sources of data. Facebook produces data that is qualitative in nature that shows some emotions, feeling, perception of users, messages, and so on [8]. SMA can also be used to analyze social issues through analyzing comments and post on social media sites [9]. During SMA processing, the major challenge occur is data from individual site which uses its own platform, it has its own volume, complexity and is unstructured [10]. This problem is solved by providing framework and tools for collecting, storing, processing, and analyzing data of social media [11]. Organization needs high quality data to retrieve data from social media to perform analysis and to gain insights about it [12]. The main aim of the paper is to provide insights about social media analytics using machine learning.

This paper is organized as follows. Section 2 defines about the social media analytics. Moving on to section 3 we conversed about the social media analytics process. Section 4 describes challenges in social media analytics. Section 5 devotes itself to the use of machine learning in social media analytics. Section 6 calls the capabilities of social media analytics. Section 7 labels about the application of social media analytics succeeded Future research direction and conclusion in section 8.

2. SOCIAL MEDIA ANALYTICS (SMA)

Social media Analytics (SMA) is nothing but collecting valuable information from different social media sites and directing it for different conclusion. Social media contains tools, technologies, and techniques that is used to ease the communication in tech world. Social media sites include Blogspot, Twitter, Facebook, Wikimapia, AIM, YouTube, and many others. Till now social media is undergoing through many researches which focuses on three main topics: group/network analysis (defines how a group of people together interacts with each other and identify dominant users), content analysis (analysis moods/sentiments and hot popular topics in social media), prediction of real-world characteristics or events [13]. Main role of SMA is analyzing and collecting audience data which is the backbone of today's social media networks. It gives us the deep insights about user's interest which helps in regulating the market.

SMA is an easy and efficient way to analyze what audience expects and wants from others contents which is easily shown in trending searches. Though, SMA is an efficient method but it is far more complex and critical method. In recent years, social media experience the huge amount of growth in its sector. Despite of which analytics solutions is heavily utilized. Many startups

crashed in short period of time as they are not able to take advantage of analytics in their growth period [14]. Social media is proving its worth in recent years. It has become an essential part to:

- Make better business strategies.
- Track marketing team efficiencies.
- Measure and prove of marketing impact.
- Comparing performance with others.

2.1 Overview

SMA is usually used for competitive analytics, influencer analytics, audience analytics, performance analytics, community management and customer service analytics, sentimental analytics and so on. SMA can be professionally used through media monitoring and collecting social media metrices to achieve desired goals. The recent capabilities of SMA are listed below:

2.1.1 Group/Network Analysis

Group/Network analysis helps to characterize users in a group who interacts with each other to find significant users.

Group Identification: Through SMA analysts are researching about different groups present in social media like popular groups, new groups etc. by exploring their characteristics and how groups can be differentiated by each other and understanding the core concept of the same. This can be done by studying Twitter, Instagram followers or Facebook friends through mutual activities such as posting on online handles, commenting, and sharing post with their mutual groups. Certain research communities' studies social media groups by following their strengths and weaknesses and the impact of circulation of particular columns on social media [15]. Other research communities observed that way of communication can differ for different groups like their communication styles such as using hashtags, emoticons, punctuations, timing, etc. [16]. Recent researches are based on identifying popular users from the groups by analyzing users' reliability and bandwidth [17] and message distribution through followers/audience forwarding for instance retweets on twitter [18].

Characteristics of Users: There are many research communities which studies content of social media by observing their features such as aggressiveness of social media groups [19]. There are many other methods to understand the user's personality like their media profiles, their self-description, their interest, photos they share and many more [20].

2.1.2 Content Analysis

Content analysis helps to analyze audience opinions and trending common topics in social media.

Analysis of Emotions: Emotion analysis is done through analyzing conversation certain post and history of chat rooms [21]. Researchers are looking for new ways that are more accurate to differentiate users' emotions from social media [22-24]. There are many researches aiming towards users' emotions such as humor, sarcasm, irony etc. [16].

Analysis of social multimedia: Analysts are expanding traditional approach for analyzing different social multimedia such as twitter and Facebook. Online multimedia content like photos, videos, audios that are posted by individual users helps to promote discussion and reuse of archiving contents [25]. SMA also focuses on analysis of content like photos of popular landmarks, video clips of tourist place.

2.1.3 Prediction of real-world characteristics or events

Prediction uses social media standards to predict real world events and future. Social media prediction uses sentiments analysis, combination of scale of social media coverage and influence analysis [26]. This prediction method in social media is being recently used in financial market activity [27], prediction of movie revenues [21], popularity of song in billboard weekly chart [28], outcomes of political elections and many more [29].

3. SMA Process

Social Media Analytics (SMA) provides developing and evaluating tools of information and framework to gather, analyze, conclude, and visualize the data [4][30]. The SMA process is carried out in three phases namely the Capture phase, the Understand phase and the Present phase [30]. The explanation of each phase is given through analysis of trending videos on YouTube.

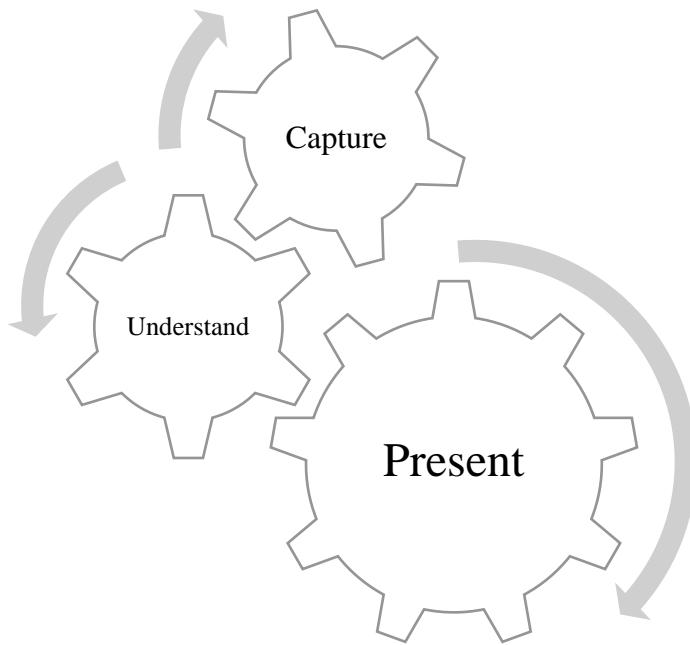


Figure 1: Social Media Analytics Process

3.1 Capture

To start the process of analyzing SMA, first comes the capture stage where relevant data across various social media platforms is gathered. This stage covers all relevant platforms such as Twitter, YouTube, Tinder, WhatsApp, LinkedIn, etc. Including smaller and more specialized sources such as forums, blogs, Wikis, podcasts and social bookmarking sites. But gathering so much data can lead to inclusion of irrelevant and redundant data, that is why the capture stage must keep the balance between inclusivity of the data that is gathering information from all sources and exclusivity that is focusing on the more relevant part of the information [4]. Capturing data for YouTube trending videos analysis as shown in Figure 2.

Data Exploration																																																											
In [4]:	df.describe()																																																										
Out[4]:	<table><thead><tr><th></th><th>category_id</th><th>views</th><th>likes</th><th>dislikes</th><th>comment_count</th></tr></thead><tbody><tr><td>count</td><td>40949.00</td><td>40949.00</td><td>40949.00</td><td>40949.00</td><td>40949.00</td></tr><tr><td>mean</td><td>19.97</td><td>2360784.64</td><td>74266.70</td><td>3711.40</td><td>8446.80</td></tr><tr><td>std</td><td>7.57</td><td>7394113.76</td><td>228885.34</td><td>29029.71</td><td>37430.49</td></tr><tr><td>min</td><td>1.00</td><td>549.00</td><td>0.00</td><td>0.00</td><td>0.00</td></tr><tr><td>25%</td><td>17.00</td><td>242329.00</td><td>5424.00</td><td>202.00</td><td>614.00</td></tr><tr><td>50%</td><td>24.00</td><td>681861.00</td><td>18091.00</td><td>631.00</td><td>1856.00</td></tr><tr><td>75%</td><td>25.00</td><td>1823157.00</td><td>55417.00</td><td>1938.00</td><td>5755.00</td></tr><tr><td>max</td><td>43.00</td><td>225211923.00</td><td>5613827.00</td><td>1674420.00</td><td>1361580.00</td></tr></tbody></table>						category_id	views	likes	dislikes	comment_count	count	40949.00	40949.00	40949.00	40949.00	40949.00	mean	19.97	2360784.64	74266.70	3711.40	8446.80	std	7.57	7394113.76	228885.34	29029.71	37430.49	min	1.00	549.00	0.00	0.00	0.00	25%	17.00	242329.00	5424.00	202.00	614.00	50%	24.00	681861.00	18091.00	631.00	1856.00	75%	25.00	1823157.00	55417.00	1938.00	5755.00	max	43.00	225211923.00	5613827.00	1674420.00	1361580.00
	category_id	views	likes	dislikes	comment_count																																																						
count	40949.00	40949.00	40949.00	40949.00	40949.00																																																						
mean	19.97	2360784.64	74266.70	3711.40	8446.80																																																						
std	7.57	7394113.76	228885.34	29029.71	37430.49																																																						
min	1.00	549.00	0.00	0.00	0.00																																																						
25%	17.00	242329.00	5424.00	202.00	614.00																																																						
50%	24.00	681861.00	18091.00	631.00	1856.00																																																						
75%	25.00	1823157.00	55417.00	1938.00	5755.00																																																						
max	43.00	225211923.00	5613827.00	1674420.00	1361580.00																																																						

Figure 2: Data Capturing of YouTube Trending videos

3.2 Understand

After pre-processing is done and the data is cleaned, extracting meaningful information using various statistical methods and techniques from NLP (natural language processing), data mining and network analysis [31]. Many kinds of parameters and metrics are generated in this stage covering the background, interests and network of relationships of a business, the company and its products. These metrics are also referred as sentiments. We understand that the likes and views are strongly positively correlated using the Correlation map in Figure 3.

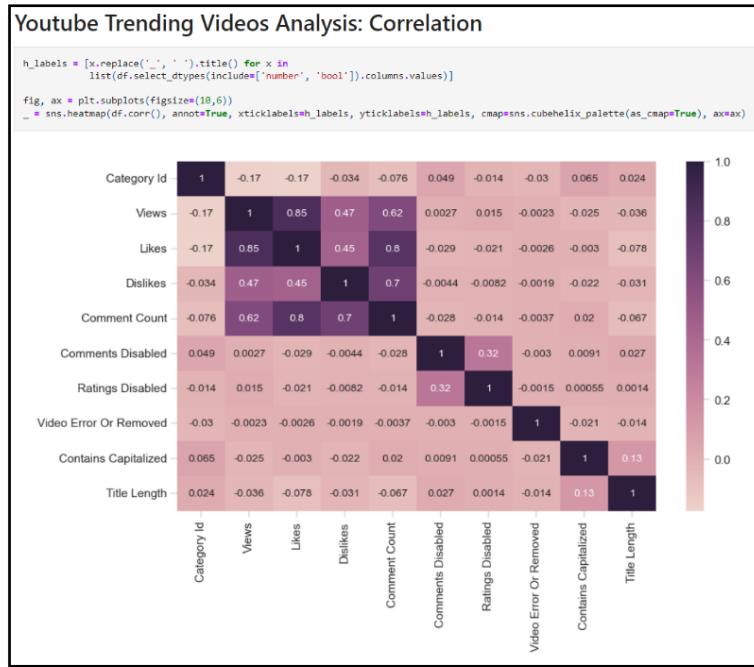


Figure 3: Understanding the Correlation between the likes and views of YouTube Trending Videos.

3.3 Present

This is the third stage in the SMA process. The goal of this stage is to represent the analytics gathered and show to users in a format that is easy to read and understand. There are a lot of visualization techniques that can be used in this stage. Data visualization of YouTube trending videos analysis representing how many capitalized words are there in the titles. We can see that 44% of trending video titles contain at least one word in all caps as shown in Figure 4.



Figure 4: Percentage showing YouTube videos contain capitalized word.

4. CHALLENGES IN SMA

Social media platforms have changed the way customers and brands connect and interact. Truth be told, it is viewed as one of the most troublesome advancements of the 21st hundred years. The capacity to represent customers' feelings and views in the form of basic, quantifiable information has made the medium social media turns into an exceptionally powerful business improvement

instrument [41]. Be that as it may, web-based entertainment checking programs are not without flaws. There are huge impediments in precisely deciphering on the web perceptions into valuable business experiences. We have listed some limitations below:

- *Unbiased Data*: The unaffected environment of social media platforms encourages individuals to use familiar and distinctive material in their own language. And often, this is outside the ability to comprehend of mechanized opinion investigation programs. This makes it difficult for them to understand the context and emotions in which the brand is being discussed.
- *The numbers don't grow*: Likes on a brand's page don't precisely reflect genuine commitment or changes. In any case, since they are unmistakable in friendly profiles, most examination crusades end up with an end goal to dissect and increment pages likes. This frequently brings about an absence of unmistakable ROI, in any event, for effective missions.
- *The incomplete picture*: The variance between individuals in the social media ecosystem are as striking as they are around the globe. While some have a functioning voice and that's just the beginning proactive interactions, most are simply perusing. Commitment can likewise fluctuate in view of socioeconomics and crusades. While you might get additional intrigued thoughts from more youthful crowds, the more seasoned segment might decide to stay quiet.
- *Relevancy and quality of data*: The standard of analyzed online data has always been a concern for companies. Virtual entertainment stages are overwhelmed with phony and copy profiles. Adding to the issue is that entrance limitations on most setups make it hard to check their legitimacy. In addition, social channels give next to zero data about the client venture. Actually, the way from SMA to transformation is frequently messy and out of pattern. This makes it difficult for companies to search the right uses for the collected analysis of data and to identify the feature that drive clients to associate with a brand. Despite all this, social media is still the most successful way for brands to cooperate straightforwardly with their end clients, figure out them, and understand what they want. Harnessed effectively, it tends to be an incredible asset for brands to gather information that can be utilized to produce experiences for long haul business development.

5. USES OF MACHINE LEARNING IN SMA

Artificial Intelligence (AI) and Machine Learning techniques are being used by corporations for their marketing techniques. Use of machine learning in SMA are listed below:

5.1 Sentiment Analysis

Sentiment Analysis, often known as opinion mining or emotion AI, is the practice of monitoring data with a desired outcome of judging the opinion of the text data. It connects social media data with predetermined sentiment classifications such as positive, negative, or neutral using Natural Language Processing (NLP) and machine learning. The analysis includes methods such as; *Word Counts*: a product is assumed to be more liked if it is mentioned more; *Polarity Lexicons*: a means of analyzing sentiments and opinions in text in an automated way [32]; *Semantic Methods*: studying relationships between a generic term and two extreme values such as “poor” and “excellent”, and then determine the sentiments about it [33]. Following is the resultant sentiment analysis of a WhatsApp group chat. As shown in figure 5a, 5b, 5c every line of the chat is rated among three categories: Positive, Negative and Neutral. Since the Neutral category is more dominant, hence the resultant sentiment is neutral as well.

```
df = pd.DataFrame(data, columns=["Date", "Time", "Author", "Message"])
df['Date'] = pd.to_datetime(df['Date'])

data = df.dropna()
from nltk.sentiment.vader import SentimentIntensityAnalyzer
sentiments = SentimentIntensityAnalyzer()
data["Positive"] = [sentiments.polarity_scores(i)["pos"] for i in data["Message"]]
data["Negative"] = [sentiments.polarity_scores(i)["neg"] for i in data["Message"]]
data["Neutral"] = [sentiments.polarity_scores(i)["neu"] for i in data["Message"]]
print(data.head(30))
```

Figure 5a: Preparing the data collected from WhatsApp chat.

	Positive	Negative	Neutral
2	0.245	0.000	0.755
3	0.000	0.000	1.000
4	0.000	0.000	1.000
5	0.000	0.000	1.000
6	0.000	0.000	1.000
7	0.000	0.000	1.000
8	0.000	0.000	1.000
9	0.000	0.000	1.000
10	0.000	0.000	1.000
11	0.000	0.324	0.676
12	0.000	0.000	1.000
13	0.000	0.000	1.000
14	0.000	0.000	1.000
15	0.000	0.000	1.000
16	0.000	0.000	1.000
17	0.000	0.000	1.000
18	0.000	0.000	1.000
19	0.000	0.000	1.000
20	0.000	0.000	1.000
21	0.000	0.000	1.000
22	0.000	0.000	1.000
23	0.000	0.000	1.000
24	0.000	0.000	1.000
25	0.000	0.000	1.000
26	0.000	0.000	1.000
27	0.000	0.000	1.000
28	0.000	0.000	1.000
29	0.000	0.000	1.000
30	0.000	0.000	1.000
31	0.000	0.000	1.000

Figure 5b: Shows list form of Positive, Neutral, and Negative messages.

```
x = sum(data["Positive"])
y = sum(data["Negative"])
z = sum(data["Neutral"])

def sentiment_score(a, b, c):
    if (a>b) and (a>c):
        print("Positive 😊 ")
    elif (b>a) and (b>c):
        print("Negative 😥 ")
    else:
        print("Neutral 😐 ")
sentiment_score(x, y, z)

Neutral 😐
```

Figure 5c: Result the chat between the participants are Neutral.

5.2 Topic Modeling

Topic Modeling follows unsupervised classification method which classifies media or documents such as blog posts or news articles and divide them into groups so they can be studied separately. Topic modeling utilizes various machine learning techniques to identify the dominant topics from large bodies of captured text. For example, a variety of algorithms analyze word co-occurrence frequencies in a single communication [34] or across subjects and communities of users [35] to detect "latent" topics. It is also possible to examine information regarding the placement of words within communications. [36]. One of the most used algorithms for topic modelling is Latent Dirichlet Allocation. Without delving into the arithmetic, we can deduce that the model is guided by the following rules:

Every document is a mixture of different topics: We believe that each paper will contain words from a variety of themes in different amounts. For instance, in a two-topic model "Document 1 is 90% topic A and 10% topic B, while Document 2 is 30% topic A and 70% topic B".

Every topic is a collection of words: Now for example, consider a two-topic document such as an American news format, with one for "politics" and the other for "entertainment." "President," "Congress," and "government" are probably the most commonly used words in the politics issue. While the entertainment topic could include terms like "film," "television," and "actor," Importantly, words can be shared between topics; for example, the word "budget" may appear in both.

LDA is a mathematical method for determining both of these simultaneously: the mixture of words associated with each topic as well as the mixture of themes that describe each text.

5.3 Social Media Marketing

Companies employ the results of machine learning to obtain a deeper grasp of consumer views and to improve their marketing strategy. Digital marketers can benefit from machine learning tools since they allow them to better reveal and comprehend data. One can predict online customer behavior by keeping track of client preferences and delivering necessary information.

LinkedIn employs algorithms that can forecast which individuals are the greatest fit for a certain position. It identifies prospects who are more likely to reply or who are looking for new opportunities using machine learning techniques. [37]. Twitter has updated its service to allow users to trim an image using facial identification or build a thumbnail from a complete image using artificial intelligence. [38-39].

6. CAPABILITIES OF SMA

According to a recent Technology Advice research, around 52% of the grown-up populace has at least one social media account. Social media is attracting a lot of attention from people of all ages, and if you haven't used it to build your business, you're already 2-3 years behind. However, not everything is lost. You can simply harness the immense potential of social media by utilizing a simple social media marketing program that has all of the necessary characteristic while remaining within your budget [42]. Here is a list of the things you should look for in a social networking platform.

- *Analytical Centralization:* Once you've begun with your marketing activities, it's critical to maintain track of the information you've collected. You must track all of your campaigns as they arrive at additional clients and get more offers. Measurements like remarks, likes, shares, retweets, and other similar items must be tracked and gathered in order to make judgments based on changing patterns. Analytics panels are embedded into significant stages like LinkedIn, YouTube, Facebook, and Twitter. The panel monitors all of your platform postings and gives you with useful data that you can use to optimize future marketing efforts.
- *Planning:* Your firm may have many social media accounts, and posting on each site independently consumes a significant amount of your valuable time. All things being equal, the web-based entertainment showcasing apparatus gives you a characteristic that appears on all of your various profiles at the same time. As a result, the tool handles all of your repetitive and monotonous duties for you. Tools like Buffer, in addition to auto-scheduling, allow you to add labels to your articles to carry setting to the discussions. You may try out several timetables to observe which one turns out best for your crowd.
- *Participation:* social media is a popular tool for interaction. Simply publishing is not enough to develop a trustworthy and devoted client base; you must also respond to comments and address the user's questions. The majority of social networking platforms available now allow you to simplify things for easy control. You can view your Facebook account's likes, shares, comments, and everything else, as well as your Twitter account's retweets and mentions. And if you have any fresh product ideas, you may reach out to your audience and solicit comments.
- *Consistent Collaboration:* Every social media product should include a unified dashboard for safe communication. Every team member in a company should have a simplified method for reporting back on their allocated duties. A useful tool, like the live action tracker, permits you to observe when a teammate responds to your message progressively. This function comes in when you're working with colleagues from various departments. A centralized location for reporting to and communicating with other colleagues makes a straightforward working air.
- *Intelligent Reporting:* Your social media marketing software should have an automatic reporting function to save you time. There are several solutions available that give you with tailor-made layouts that you may edit and utilize as needed. You only need to set it up once and afterward forget about it until the following time you really want another format. If your social media marketing software includes built-in templates, it makes it simple to work on several reports that highlight different features.
- *Services for Locating Locations:* Geo-targeting social media postings is critical for company. Your virtual entertainment advertising programming ought to have the option to help you with the accompanying:
 - When should you publish to garner the most engagement?
 - How to Find New Customers in a Specific Location
 - What local variables should you consider while interacting with your customers?
 - The majority of marketing platforms available today can categorize profiles, postings, and messages based on geography. You may also browse other profiles in your area and interact with them directly. To begin building new relations, it is best to begin with adjacent profiles. This will provide them with a feeling of trust, and you might have another committed client.
- *Security:* When interacting with many persons across several social media accounts, it is best to ensure that they all have the proper rights. Every group member must be given a different ID so that if someone uploads an unauthorized message, they may be held accountable. Security measures must be up to date in order to safeguard you from any doubtful behavior on their social media sites.

7. APPLICATION OF SMA

The marketing strategy entails outlining the precise actions that must be carried out, determining the intended audience each of its operations, providing metrics for assessing performance, remaining flexible enough to allow for revisions if necessary, and automating the process [42]. Some application of SMA is enumerated below:

- *Automate the procedure:* When a marketing approach still up in the air to be compelling, the whole interaction is mechanized to week-by-week post or on a regular basis about product and business changes.
- *Defining measures to measure success:* The indicators used to assess the performance of a marketing campaign include the number of snaps, traffic to the site, number of devotees, likes, shares, remarks, and so on. Compare the company's existing analytics versus analytics one month after implementing a new marketing plan. Apart from these indicators, social media reports (Page Insights for Facebook; Twtrland for Twitter; Klout for all social media) demonstrate how effective the firm's social media engagement has been, allowing the company to determine whether or not it is connecting with genuine prospects.
- *Specifying the exact operations that must be carried out:* Choosing the sort of material that will be used to advertise the company's goods on social media. Content can take numerous forms, including simple text, links, photos, videos, quotations, and re-shares. In addition to having a significant content type, providing a diverse sort of material on a regular basis may serve to alleviate any prospective boredom within potential customers. Aside from the sort of material, the firm should also work on building a decent social media profile that is consistent over several platforms and clearly informs followers about what they can anticipate from the brand.
- *Being adaptable to accommodate for changes:* Customers' input may be gathered in a very short period of time through social media. Any errors in the business strategy may be recognized promptly depending on client input, and appropriate steps will be made to correct them. Being adaptable allows the whole cycle to turn into nimble, allowing it to respond to ever-changing client demands swiftly.
- *Determining the intended audience for each of its operations:* Because various sorts of people are available on different social media sites, all satisfied kinds won't prevail on all of them. To create a successful marketing plan, it is critical to understand what sort of content can provide substantial results on which social media platform. Aside from that, Pew research data and Google analytics demographics data may assist in determining the sort of people living in a specific location, which will support determining the target audience for each activity (or posts).

8. FUTURE RESEARCH DIRECTION & CONCLUSION

The study aims on providing social media analysis on different social media platforms to gather and find meaningful data from it. It helps analyst to collect data and perform strategic decision. Different analysis uses different tools and technologies. According to a survey social media analysis is preferred over other analysis. Choosing tools and technologies greatly depends upon the type of analysis one wants to perform on retrieve data from social media. In this paper, the social media platforms that we used for analysis are YouTube, Instagram, WhatsApp, and Facebook. YouTube analysis was done on trending videos. Sentimental analysis of WhatsApp and Facebook was done, and reach in case of Instagram was analyzed through SMA. This paper primarily focuses on showing the importance of social media analysis in different fields and to understand the data of social media, its advantages and other aspects. Social media analysis is a growing field and will continue to boost up in near future. It will arise as an advantageous field to boost up a business by understanding the preference of the audience. It will put a great impact on digital marketing and politics and many other fields. Thus, will emerge as one of the most prosperous fields in utilizing social media platforms.

9. REFERENCES

- [1] Das, M., & Das, G. (2015). Structured analytics in social media. *Proceedings of the VLDB Endowment*, 8(12), 2046–2047. <https://doi.org/10.14778/2824032.2824135>
- [2] Stieglitz, S., Dang-Xuan, L., Bruns, A., & Neuberger, C. (2014). Social Media Analytics. *Business & Information Systems Engineering*, 6(2), 89–96. <https://doi.org/10.1007/s12599-014-0315-7>
- [3] Drus, Z., & Khalid, H. (2019). Sentiment Analysis in social media and its application: Systematic Literature Review. *Procedia Computer Science*, 161, 707–714. <https://doi.org/10.1016/j.procs.2019.11.174>
- [4] Zeng, D., Chen, H., Lusch, R., & Li, S.-H. (2010). Social Media Analytics and Intelligence. *IEEE Intelligent Systems*, 25(6), 13–16. <https://doi.org/10.1109/mis.2010.151>
- [5] Stieglitz, S., & Dang-Xuan, L. (2012). Social Media and Political Communication: A Social Media Analytics framework. *Social Network Analysis and Mining*, 3(4), 1277–1291. <https://doi.org/10.1007/s13278-012-0079-3>
- [6] Saravanan, M., & Perepu, S. K. (2019). Realizing social-media-based analytics for Smart Agriculture. *The Review of Socionetwork Strategies*, 13(1), 33–53. <https://doi.org/10.1007/s12626-019-00035-3>
- [7] Kumar, S. (2013, November 12). Twitter data analytics. Google Books. Retrieved May 1, 2022, from https://books.google.com/books/about/Twitter_Data_Analytics.html?id=yO8LswEACAAJ
- [8] Togyer, J. (2013, September 27). Research notebook: The discipline of machine learning. Carnegie Mellon School of Computer Science. Retrieved May 1, 2022, from <https://www.cs.cmu.edu/link/research-notebook-discipline-machine-learning>
- [9] Dias, D. S., Welikala, M. D., & Dias, N. G. J. (2018). Identifying racist social media comments in Sinhala language using text analytics models with machine learning. 2018 18th International Conference on Advances in ICT for Emerging Regions (ICTer). <https://doi.org/10.1109/icter.2018.8615492>
- [10] Stieglitz, S., & Dang-Xuan, L. (2012). Social media and political communication: a social media analytics framework. *Social Network Analysis and Mining*, 3, 1277-1291.
- [11] Zeng, D., Chen, H., Lusch, R., & Li, S. H. (2010). Social media analytics and intelligence. *IEEE Intelligent Systems*, 25(6), 13-16. [5678581]. <https://doi.org/10.1109/MIS.2010.151>
- [12] He, W., Tian, X., Tao, R., Zhang, W., Yan, G., & Akula, V. (2017). Application of social media analytics: A case of analyzing online hotel reviews. *Online Information Review*, 41(7), 921–935. <https://doi.org/10.1108/oir-07-2016-0201>
- [13] Gastelum, Z. N., & Whattam, K. M. (2013). State-of-the-art of Social Media Analytics Research. <https://doi.org/10.2172/1077994>
- [14] Social Media Analytics: The complete guide. Socialbakers.com. (n.d.). Retrieved January 29, 2022, from <https://www.socialbakers.com/blog/social-media-analytics-the-complete-guide>

[15] Grabowicz, P. A., Ramasco, J. J., Moro, E., Pujol, J. M., & Egiluz, V. M. (2012). Social features of online networks: The strength of intermediary ties in online social media. *PLoS ONE*, 7(1). <https://doi.org/10.1371/journal.pone.0029358>

[16] Paris, C., Thomas, P., & Wan, S. (2012). Differences in Language and Style Between Two Social Media Communities. *ICWSM*.

[17] wu, M. (n.d.). [web log]. Retrieved from <https://community.khoros.com/t5/Khoros-Communities-Blog/The-6-Factors-of-Social-Media-Influence-Influence-Analytics-1/ba-p/5708>.

[18] Romero, D. M., Galuba, W., Asur, S., & Huberman, B. A. (2010). Influence and passivity in social media. *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.1653135>

[19] Gawron, Jean Mark, Dipak Gupta, Kellen Stephens, Ming-Hsiang Tsou, Brian Spitzberg, and Li An, Using Group Membership Markers for Group Identification, Presented at the Sixth International AAAI Conference on Weblogs and social media, 04-08 June 2012, Dublin, Ireland

[20] Golbeck, J., Robles, C., & Turner, K. (1970, January 1). [PDF] predicting personality with social media: Semantic scholar. undefined. Retrieved May 1, 2022, from <https://www.semanticscholar.org/paper/Predicting-personality-with-social-media-Golbeck-Robles/3f1442472a674968d6a6f40bb9bcd9e019606e49>

[21] Kilian Thiel, Fabian Dill, Tobias Kotter, and Michael R. Berthold". (n.d.). Retrieved May 1, 2022, from https://www.uni-konstanz.de/bioml/bioml2/publications/Papers2007/Towards_visual_exploration.pdf

[22] Coping with the document frequency bias in sentiment classification. (n.d.). Retrieved May 1, 2022, from <https://hal.archives-ouvertes.fr/hal-00878275/document>

[23] Happy, nervous or surprised? Classification of Human Affective States ... (n.d.). Retrieved May 1, 2022, from https://www.researchgate.net/profile/Michael-Gamon/publication/255564370_Happy_Nervous_or_Surprised_Classification_of_Human_Affective_States_in_Social_Media/links/56705ecc08ae5252e6f1e218/Happy-Nervous-or-Surprised-Classification-of-Human-Affective-States-in-Social-Media.pdf

[24] External links. Grief-Stricken in a Crowd: The Language of Bereavement and Distress in social media - researchr publication. (n.d.). Retrieved May 1, 2022, from <https://researchr.org/publication/BrubakerKTH12>

[25] Naaman, M. (2010, May 21). Social Multimedia: Highlighting opportunities for search and mining of multimedia data in social media applications - multimedia tools and applications. SpringerLink. Retrieved May 1, 2022, from <https://link.springer.com/article/10.1007/s11042-010-0538-7>

[26] Asur, S., & huberman, B. A. (2010). predicting the future with social media. - references - scientific research publishing. (n.d.). Retrieved May 1, 2022, from <https://www.scirp.org/reference/ReferencesPapers.aspx?ReferenceID=1300552>

[27] Bollen, J., Mao, H. and Zeng, X. (2011) twitter mood predicts the stock market. *Journal of Computational Science*, 2, 1-8. - references - scientific research publishing. (n.d.). Retrieved May 1, 2022, from <https://www.scirp.org/reference/ReferencesPapers.aspx?ReferenceID=2444907>

[28] Bae, Y., & Lee, H. (2015, June 18). Sentiment analysis of Twitter audiences: Measuring the positive or negative influence of popular tweeterers. Korea University. Retrieved May 1, 2022, from <https://koreauniv.pure.elsevier.com/en/publications/sentiment-analysis-of-twitter-audiences-measuring-the-positive-or>

[29] What's In your tweets? I know who you supported in the UK 2010 general ... (n.d.). Retrieved May 1, 2022, from <https://hal.inria.fr/file/index/docid/702390/filename/main.pdf>

[30] Stieglitz, S., Dang-Xuan, L., Bruns, A., & Neuberger, C. (2014, February 15). Social Media Analytics - Business & Information Systems Engineering. SpringerLink. Retrieved May 1, 2022, from <https://link.springer.com/article/10.1007/s12599-014-0315-7>

[31] Fan, W., Wallace, L., Rich, S., & Zhang, Z. (2006, September 1). Tapping the power of text mining. Arizona State University. Retrieved May 1, 2022, from <https://asu.pure.elsevier.com/en/publications/tapping-the-power-of-text-mining>

[32] Vicente, I. S., & Saralegi, X. (n.d.). Polarity lexicon building: To what extent is the manual effort worth? *ACL Anthology*. Retrieved May 1, 2022, from <https://aclanthology.org/L16-1149>

[33] Turney, P.D. (2002) thumbs up or thumbs down semantic orientation applied to unsupervised classification of reviews. *proceedings of the 40th annual meeting of the association for computational linguistics*, 417-424. - references - scientific research publishing. (n.d.). Retrieved May 1, 2022, from <https://www.scirp.org/reference/ReferencesPapers.aspx?ReferenceID=1483011>

[34] Probabilistic latent semantic analyses (PLSA) in ... - researchgate. (n.d.). Retrieved May 1, 2022, from https://www.researchgate.net/publication/45491735_Probabilistic_Latent_Semantic_Analyses_PLSA_in_Bibliometric_Analysis_for_Technology_Forecasting

[35] Stieglitz, S., Dang-Xuan, L., Bruns, A., & Neuberger, C. (2014). Social Media Analytics. *Business & Information Systems Engineering*, 6(2), 89-96. <https://doi.org/10.1007/s12599-014-0315-7>.

[36] Topic modeling: Beyond bag-of-words - hanna wallach. (n.d.). Retrieved May 1, 2022, from <http://dirichlet.net/pdf/wallach06topic.pdf>

[37] Schwartz, H. A., Eichstaedt, J. C., Kern, M. L., Dziurzynski, L., Ramones, S. M., Agrawal, M., Shah, A., Kosinski, M., Stillwell, D., Seligman, M. E. P., & Ungar, L. H. (2013, September 25). Personality, gender, and age in the language of social media: The open-vocabulary approach. *PloS one*. Retrieved May 1, 2022, from <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3783449/>

[38] Pang, B. and Lee, L. (2008) opinion mining and sentiment analysis. *Foundations and trends® in Information Retrieval*, 2, 1-135. - references - scientific research publishing. (n.d.). Retrieved May 1, 2022, from <https://www.scirp.org/reference/ReferencesPapers.aspx?ReferenceID=2442500>

[39] Liu, B. (1970, January 1). [PDF] sentiment analysis and subjectivity: Semantic scholar. undefined. Retrieved May 1, 2022, from <https://www.semanticscholar.org/paper/Sentiment-Analysis-and-Subjectivity-Liu/29366f39dc4237a380f20e9ae9fb7fa645200c9>

[40] 4 challenges of using social media data for Analytics. Official GMI Blog. (n.d.). Retrieved May 1, 2022, from <https://www.globalmediainsight.com/blog/4-challenges-using-social-media-data-analytics/>

[41] Solution, S. I. T. (2021, February 9). 7 must have features in social media app. Medium. Retrieved May 1, 2022, from <https://sparxitsolution.medium.com/7-must-have-features-in-social-media-app-e5fc982ff21b>

[42] Journal, I. R. J. E. T. (2015, May 25). IRJET-applications of social media analytics in designing an effective marketing strategy. *Academia.edu*. Retrieved May 1, 2022, from https://www.academia.edu/12594777/IRJET_APPLICATIONS_OF_SOCIAL_MEDIA_ANALYTICS_IN_DESIGNING_AN_EFFECTIVE_MARKETING_STRATEGY