

CHAPTER -1

INTRODUCTION

1.1 Overview

Advances in Internet and web technologies have made significant evolutionary changes in our social, economic and cultural activities. The World Wide Web, which was originally conceived as an information sharing medium is now shifted to more of a communicative or socializing medium. The Web became user-centric, with collaborative technologies, propagating and consuming information. Technology changed our day to day activities and our thought process. Social media emerged as the most powerful communication and connecting media. People all over the world embraced this new revolution, in generating and sharing contents. More and more users started creating new contents, vocalizing and participating in likeminded groups and finding a global voice for them. Cyberspace saw a significant growth, in the scale and richness of online communities and social media.

Social Network, a term coined by the Anthropologist J.A. Barnes in 1954, transformed to online social networks, which became the “Buzzword” through the popular websites like Myspace [17], LinkedIn [18] and Facebook [19]. Online social networks became a fad with the popularity of

such kind of websites. According to Alexa [115], social-networking sites like YouTube and Facebook are among the most visited websites in the cyber world.

Social networking is the practice of expanding social contacts, by making connections through individuals. Online communities allow users to create a web presence, manage their identities and stay connected to friends and colleagues. It relates to the formation of a circuit or a chain of friends who use different tools to interact. In the last few years, the peer dependence has tremendously increased, which has directly impacted the dependence on online networks making it a face to face interaction medium. These sites have gained, immense acceptance over the years, providing global exposure. This makes one's views, thoughts and opinions reach worldwide. It is difficult now to think of a life, without online services. Blogs, chats, scraps etc. have gained so much importance in our day to day life in such a way, that we have developed a habit of it. There are countless sites where people can participate, making us difficult to decide which sites are truly beneficial.

1.2 History, Growth and Application of Social Networks

Social Network is defined [62] as a social structure of people, related (directly or indirectly) to each other through a common relation or interest. Social network analysis (SNA) is the study of social networks to understand their structure and behaviour. It is a network with vertices as people or

group of people. The advent of online social networks is considered as a milestone in the web history. The advances in web sciences have helped in providing comprehensive digital traces of social actions, interactions and transactions that made the analysis a simpler task. The history of social networks dates back to Stone Age where humans gathered around campfires and shared stories or painted history on cave walls. These activities were influential and the tools that they used like words, paints, signals, smoke, stone arts etc. were their social media.



Source:https://c1.staticflickr.com/9/8450/7975205041_7a5e4b65ff_b.jpg

Fig 1.1 Social Networks

Human nature is always to be part of the society and to get connected. The communities evolved based on a common thread of interest. With the fast pace of changes in the computer hardware and communication technologies, this connectivity and networking, became easy and affordable. These virtual networks allowed users to create a web presence, manage their identities

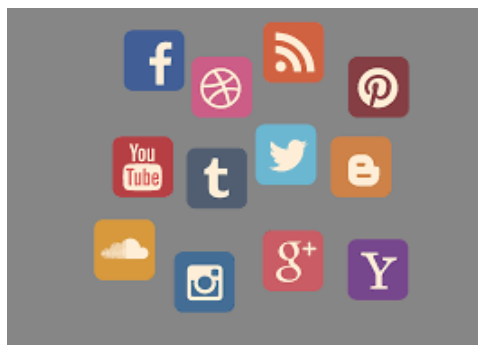
and stay connected with others. The Internet has made it possible to reach around the globe by the mere click of a button. The World Wide Web shifted the information sharing medium, to a communicative medium. Digital social media is obviously a more recent occurrence which began after we found that computers could be used to connect people. CompuServe, AOL and Prodigy wars, the three online services offering group discussions, played a critical role in bringing affordable Internet service and brought online conversations to the mainstream [86].

The online network in its primitive form appeared in 1995 with the site classmates.com which helped the schoolmates to keep in touch and the modern form took off with Friendster[16] in 2002. The popularity of Friendster was so much that its user base grew to 3 million in the first three months. In the year 2003, MySpace [17], a Friendster clone was launched, which had a good customer base for a few years, but had to give way to Facebook which was launched in 2004, as a closed media and became public later in the year 2006 [89].

1.3 Online Social Networking Sites

Social network building starts with the creation of the profile and then adding their social links. The site facilitates users to communicate, interact, and share their interests and ideas attracting millions of users. These user interactions brought in a new socio techno culture of using social networks for socializing, exchanging new information and opinions, building a new

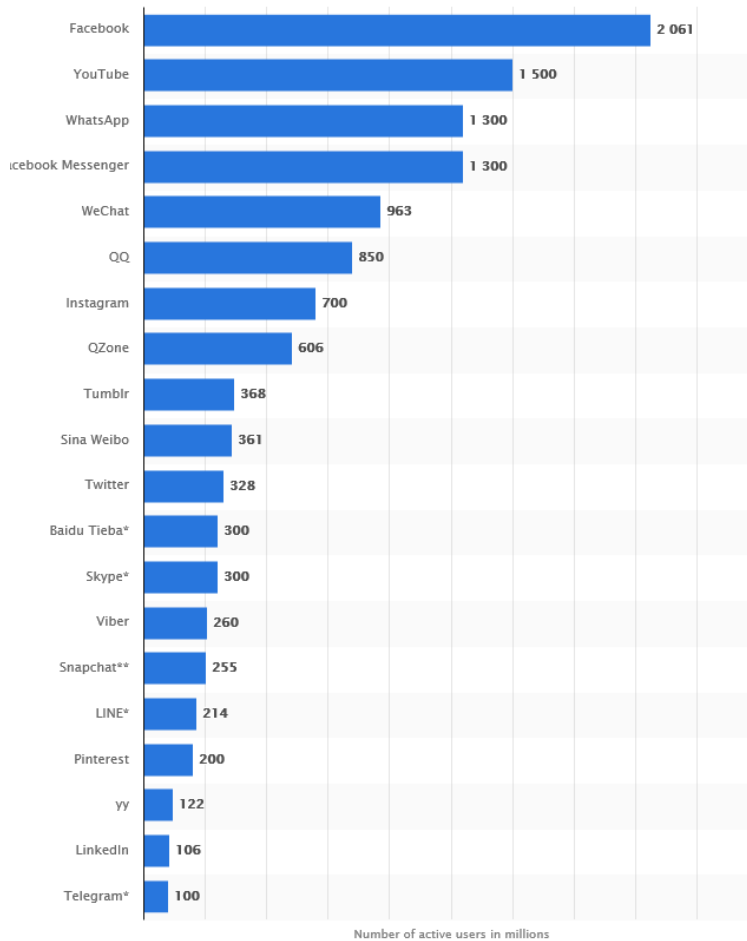
relationship with global audience sharing a common interest, beyond geographic limits and time constraints. Spread of message through social media is being adopted by the organization for brand promotion. The profile is the most important thing in all these social networks. It helps in building relationships with people sharing similar interests, activities, and contacts. Social networks disseminate or send information feeds, to a more targeted community who is profiled. This is the distinguishing feature of social networking sites. This helps business to connect with the potential audiences to promote their brands which are not possible in the traditional forms of marketing or advertising. Social network presence to promote your business, is on the cutting edge and is moving with changing times.



Source:<https://pixabay.com/en/icon-set-social-media-contact-1142000/>

Fig 1.2 Online Social Network Sites

Social network users create and disseminate information at their ease generating a massive amount of data. Figure 1.3 indicates that Facebook is the most popular social networking site with the largest number of active users.



Source: <https://www.statista.com/statistics/272014/global-social-networks-ranked-by-number-of-users/>

Fig 1.3 Popular Social Networks ranked by number of active users

1.4 Information Mining from Social Media

A Social Network can be visualised as a social graph or a collection of

applications that depends on an infrastructure for common messaging and user management. Applications that leverage this infrastructure are called Social Media. The Information and Communication Technologies and Internet technologies facilitate better communication between consumers and commerce. Social media has revolutionized the mode of expression of common people. They now have a platform to openly post their views and opinions towards any social, political events or even respond to marketing campaigns and product strategies. Social media is now being considered as the 'de-facto' platform for interactions and sharing of experiences leading to recommendations. These interactions through channels of expressions like chat, blogs, social network etc. can be considered as the electronic word of mouth, influencing a large number of users. The conversations help to understand personal interests, and more than that the opinions expressed indicate the preferences for a product or service.

Social media savvy business organisations have started using this medium, to build communities and their brand campaigns. More and more business concerns are trying to effectively leverage the burgeoning user base of these sites, as a tool to grow. This engagement helps in building and enhancing the customer relationship, thereby promoting the business and aligning the marketing strategies. This has created an opportunity for the business organizations to reach out to their customers and potential customers.

Traditional one-way conversations are replaced with the two-way flow without any boundary limitations.

Using social media to promote the business has many advantages:

- Easy Reach- Presence in social media helps in finding and connecting, as well as attracting the attention of people, across the globe. This ensures the reach at the shortest possible time, regardless of location.
- Cost Effective – Information diffusion through social media platforms can be done, without much financial overheads in comparison with traditional promotional activities.
- Customer Insights- Through the relevant data available, one can gather understand and share the feelings, including the profile like user's identity, user preferences etc.

Social media opened up a new avenue to accelerate the reach of the brand, communicate with existing and potential customers thereby improving brand identity. This helped the companies to work to gain the confidence of consumers and create new strategies for brand promotion. Social Media Marketing is becoming an indispensable part in promoting consumer engagement.

Talking to friends in our network, navigating the Web and forming opinions by listening to others and to the media have become part of daily life. Observation of their friendship links helps in identifying their common area of interest as well as whom they trust and distrust.

Social media and online contents are much used in brand monitoring platforms. The explosive growth of online contents poses a challenge in listening to the real time conversations, monitoring the sentiments and reactions to a new or existing product online. Scalable natural language processing (NLP) based engine helps in effectively mining large volume of unstructured data available. Text mining softwares with Sentiment Analysing tools are helpful in automatically extracting the information for a business like reviewer comments, customer feedback and competitor strategies.

Social media mining can be applied to large volumes of unstructured information available to extract facts, opinions and relationships to study trends and infer networks of influence. Information from Social Media includes news feeds, RSS, blogs and Twitter, apart from FaceBook posts and similar shared information in the cyberspace.

The information collected may be used in many different ways, such as for identifying current and future trends, creating social profiles, capturing consumer insights or for creating a rich knowledge base from users' clicks across the web. The real time analysis of social media data mining helps in predictive modelling.

AI and cognitive technology combined with social media data mining provide powerful intelligence from the information gathered. The key is its understanding of language, meaning and context. Business organisations, social groups, political parties and others are increasingly listening to the

conversations, often sharing comments. The information and intelligence collected help conduct research on markets, customers, competitors and more. Social Media mining is being applied in many areas.

1.5 Business Intelligence from Social Media

Customer engagement in social media generates massive amount of data. This data gives insights about user interest and their preferences. The ability to collect and process social data, helps researchers to address core questions in social sciences in new ways, which opened up new areas to explore.

The advantage of monitoring by reading up on how the people react and view their brand is attracting more and more marketers. This feedback helps them to gather information, on their target market, what their customers are looking for, and above all, how their competitors are faring. The demographic understanding also helps the business, to better align with the target market. Social media is the best way to get people to talk about your brand. These people are the social influencers. Leveraging the computing power of social media, helps right people talking, about your brand. Customers provide feedback, whether good or bad and share it online. This helps them in the purchase process, in deciding the product, as they trust this non-commercial social channel, more than the company's promotion. This follows the traditional way of brand promotion. If the customer is happy with the product, they engage in positive comments or word of mouth

and vice versa. Companies can leverage these positive feedbacks but the negative word of mouth (WOM) should not be ignored, as it is potentially very damaging and can have far-reaching results. The beauty of the social media is that, unlike the traditional media, the message not only passes to the immediate circle but also transfer to other networks, depending on the network structure. Computational social science helps in investigating social phenomena, through the medium of computing and related advanced information processing technologies. Positioned between the computer and social sciences, this new and emerging interdisciplinary field of Computational Social Science is advanced by :

- Availability of data: With the web, huge volume of social data is now available which enables the study of traces of social interactions on new scales.
- Increasing quantification of social theories: With recent advances in the social sciences, social theories have become increasingly formal and/or mathematical thus amenable to quantification

Taken together, these two developments give rise to a whole range of new and interesting problems, on the intersection between the computer and social sciences. Social networks have already become a bridge to connect our daily physical life and the virtual web space. Social computing and online communities affect the global economy, social interaction, and every aspect of our lives. The emerging field of computational social science, with the help of the social-network giants like Facebook, helps in understanding

the structure of their networks and how the information spreads across them. Social media has efficiently gathered a large volume of timely feedback and opinions, from a diverse global population.

Social computing represents a new computing paradigm and an interdisciplinary research and application field. From both theoretical and technological perspectives, social computing technologies are moving beyond the conventional information processing towards social intelligence. The move from social informatics to social intelligence is achieved by modelling and analysing social behaviour, by capturing human social dynamics by creating artificial social agents and generating and managing actionable social knowledge [57]. Such systems provide an asynchronous platform, for the consumer community to share experiences collectively and influence their purchasing behaviour. They also provide a channel, for giving feedback information, valuable for customer relation management.

1.6 Challenges in Processing Social Media

Challenges have risen recently, with the advent of online social media, which produces large amounts of, both network and natural language data. Thus understanding, predicting, and enhancing human behaviour in networks, poses important research problems for computer and data scientist with practical applications of high impact [104].

- Preferences change over the time. Capturing this in real time is challenging.
- The conventional data mining techniques cannot be applied to social media data due to the nature of difference in data.
- Unlike the conventional database system, this is unstructured, dynamic, noisy and complex in nature.

Hence the importance of social media mining has found applications in various domains like discovering social relationship which helps in behavioural evolution, and e-commerce application like marketing intelligence and many more.

Networks have been studied as graphs in mathematics, physics, biology, sociology, engineering and computer science. Grounded in graph and system theories, this approach has proven to be a powerful tool for studying networks, in physical and social worlds, including the web. A social network is a graph of relationships and interactions within a group of individuals, which plays a fundamental role as a medium for the spread of information, ideas and influence among its members. The individuals or organisations are related to each other, by various interdependencies. Social network expands, by making connections through individuals. Social media has become an integral part of any business promotions today, unlike the earlier times, where it was only considered as a social networking tool. This growing pace of social media and its impact has evolved in such a way, that

it is continuously shifting, leaving marketers constantly challenged, and most businesses overwhelmed with the never-ending changes. Hence social media analysis has become a necessity. The marketers are faced with the issue of finding new ways to capture the attention of consumers, who are immersed in the digital noise or information overload.

Social Network Analytics, provide a great variety of quantitative measures and indicators to help describe the overall relational structure of a social system, as well as the roles of individuals within it. It can be correlated with alternative type's analysis. It is impossible for the establishment of unexpected relationships, but possible to support an interpretation, of a condition of the system.

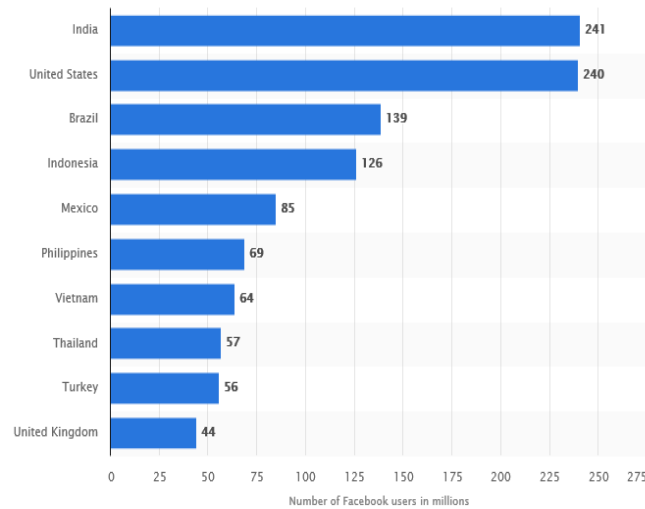
There is a formal mathematical language (graph theory) and a sociological framework, which is the basis for (social) interpretations for different kinds of networks. A Social network can be treated as a graph that consists of a set of nodes, connected through edges of links associating between them, called social graphs.

Various algorithms have been developed for visualising the networks which can also provide good insights, and these are presented in a similar way as social network mapping (SNM) techniques. One of the interesting things about network approaches is that they capture the relationships that the respondents themselves recognize, and because of this, there is also the possibility to obtain information, on informal links in addition to formally sanctioned ones. Other interesting concepts are vertically and horizontally

integrated networks, and the multiple scales involved in social relations. In combination with other methods, Social Network Analysis (SNA) can be very useful, in understanding complex societies.

1.7 Internet Usage Pattern in India

The number of Internet users in India, crossed the population of the US in 2015, becoming the second largest country also in the number of Internet users after China as per the table 1.1. The unprecedented penetration of mobile phones accelerated this growth. Growing internet access accelerated the growth of social media usage in India as indicated in figure 1.4.



Source: <https://www.statista.com/statistics/284436/india-social-network-penetration/>

Fig 1.4 Facebook Users by Country

India, the largest democracy is one of the most economically and socio-culturally diverse and orthodox community distinct from the rest of the world in holding the traditions. Internet penetration in the country and popularity of social networks have brought in a revolution in communication in the Indian community. People associate their identification with social networks like LinkedIn, Facebook etc. for their public profiles.

Media reports state that Indian organisations are increasingly using social media for online campaigns and brand promotion. Many businesses have integrated their marketing, with social media. The challenge here is to measure the effectiveness of their social media engagements.

Internet Users by Country (2014)

#	Country	Internet Users (2014)	Penetration (% of Pop)	Population (2014)	Non-Users (internet less)	Users 1 Year Change (%)	Internet Users 1 Year Change
1	China	675,131,785	49.3 %	1,369,435,670	694,303,885	8.2 %	51,100,254
2	U.S.	279,070,327	87.4 %	319,448,634	40,378,307	4.5 %	12,041,883
3	India	233,152,478	18 %	1,295,291,543	1,062,139,065	20.7 %	39,948,148
4	Brazil	118,700,869	57.6 %	206,077,898	87,377,029	13.9 %	14,446,883
5	Japan	114,850,516	90.6 %	126,794,564	11,944,048	0.8 %	932,305

Internet Users by Country (2015)

#	Country	Internet Users (2015)	Penetration (% of Pop)	Population (2015)	Non-Users (internet less)	Users 1 Year Change (%)	Internet Users 1 Year Change
1	China	705,914,032	51.3 %	1,376,048,943	670,134,911	4.6 %	30,782,246
2	India	354,114,747	27 %	1,311,050,527	956,935,780	51.9 %	120,962,270
3	U.S.	283,712,407	88.2 %	321,773,631	38,061,224	1.7 %	4,642,081
4	Brazil	132,357,306	63.7 %	207,847,528	75,490,222	11.5 %	13,656,437
5	Japan	114,994,210	90.9 %	126,573,481	11,579,271	0.1 %	143,694

Internet Users by Country (2016)

#	Country	Internet Users (2016)	Penetration (% of Pop)	Population (2016)	Non-Users (internet less)	Users 1 Year Change (%)	Internet Users 1 Year Change
1	China	721,434,547	52.2 %	1,382,323,332	660,888,785	2.2 %	15,520,515
2	India	462,124,989	34.8 %	1,326,801,576	864,676,587	30.5 %	108,010,242
3	U.S.	286,942,362	88.5 %	324,118,787	37,176,425	1.1 %	3,229,955
4	Brazil	139,111,185	66.4 %	209,567,920	70,456,735	5.1 %	6,753,879
5	Japan	115,111,595	91.1 %	126,323,715	11,212,120	0.1 %	117,385

Source: <http://www.internetlivestats.com/internet-users-by-country/>

Table 1.1 Internet Users by Country comparison for three years

The above data of table1.1 and figure1.4 clearly indicate that despite the comparatively low Internet penetration rate, social media emerges

powerfully influencing business and consumer behaviour. Social media influence the buying behaviour of customers. The growing popularity of social media has motivated researchers, to utilize this huge data generated for a product recommendation, opinion mining, sentiment analysis etc. Social media listening and mining help in understanding the customer and the market with better accuracy. The traditional recommender system is improvised, with the integration of social network information.

1.8 Motivation -Addressing the Research Gap

India has recorded a fast growth rate in Internet user base. The massive growth of Internet users in India has its reflection in the usage of social media also. It has percolated well into the social strata, leading to a previously unknown phenomenon unique to India providing new opportunities in research in computational social sciences. The growing number of Internet users in India is creating new followership for product brands that are chasing for organic leads. Information of all kinds about human engagements can be collected and analysed. But literature shows that very little or hardly any computational studies have been conducted to address the group dynamics and culture, consumer purchase and consumption patterns of Indian, even though the social media adoption rate is very high. There is an availability of the increased quantity of unstructured data, for training machine learning algorithms. There also exist a lot of opportunities to identify and study such problems unique to the local geographic locations. The current challenges of the few researchers in this

area is building up the culture of inter-disciplinary research. This is the motivation for pursuing this research work. Research done is in the direction of leveraging the available data to measure preferences and explain social behaviour.

Many of the opinion mining tools available in the market are equipped to analyse the feedback of customers, mostly about services and products. They are skewed towards sentiment analysis by interpreting natural language and detecting the emotion as positive or negative.

1.9 Research objective

Online social networks play a major role in the spread of information, at a very large scale. Social network users are connected through predefined online interactions such as Facebook friendship or Twitter following, forming a network. People in such a network can share messages, according to their interests with connected neighbours, propagating information through social links and creating cascades. A social network evolves dynamically and continuously. Such an evolution is coupled with the spread of information on top of the network: the network topology affects the channels of information diffusion; the birth and death of connections in the network are in turn, triggered by the traffic.

The objective of my research is to understand information diffusion process in an online social network, develop a procedure, using semantic text mining techniques to do opinion mining, and utilize the social media data

for business intelligence.

The research work basically focus on three distinct but related contributions of SNA

- Understanding the viral information diffusion patterns through social networks
- Studying and developing opinion mining tools, combining statistical algorithm with ontologies and semantics, coupled with supervised machine learning algorithm.
- Conducting Semantic analysis through lexicon, with known sentiment for sentiment classification.

My research work begins with exploring the flow of information, through a social network, to understand the diffusion pattern, through a unique data set that was studied with respect to Indian consumers and business. This study examines brand promotion relating customer engagement and electronic word of mouth.

The second part involves the engagement of the social media sites, in everyday conversations and promotions. The work investigates the role of contextual information in discovering semantics from the unstructured data, converting to the structured format.

Knowledge graphs are created from the semantic search. The effectiveness is validated, evaluating the real-world data sets, against the available

standard data sets. The different related business domains are identified and studied here. To start with the awareness and adoption of a new product among consumers linked by a social media is considered for promotion of a global chain of supermarket in the local market, the opening of a multiplex, promotion of a movie. The sentiment analysis and predictive analytics are done with the data shared by an eco-resort. An analytical tool is modelled and developed for multilateral initiatives using social media intelligence.

The taxonomy covering three aspects of research are (1) Viral Information Diffusion (2) Crowd Sourcing and Opinion Mining and (3) Social Intelligence.

1.10 Research Questions

User-generated Contents (UGC) are considered important in social media studies specifically, the study attempts to answer the following questions:

1. How does the information diffuse in an online social network?
2. Why and through which paths does the information diffuse?
3. How will the diffusion pattern be, in the future?
4. How can the social media be used to engage with the customers?
5. How to improve brand positioning, using opinion mining techniques to detect hidden patterns in consumers' sentiments?
6. How can sentiment analysis be used to design/redesign the brand promotion strategies?
7. How can the external data be processed to derive Social

Intelligence?

8. How to create a Knowledge Management Framework utilizing Social Media interactions?

1.11 Outline of the Thesis

The thesis is organized into eight chapters.

Chapter 1 introduces computational social sciences, its relevance in the modern world along with its background and history. A brief discussion on the complexity, importance, challenges and research gap is discussed briefly herein. The motivation, research questions and objectives are highlighted in this chapter.

Chapter 2 reviews related research work and the available literature.

Chapter 3 discusses the social media landscape and various scenarios. It introduces Sentiment Analysis and various technologies for the semantic understanding of the sentiment.

Chapter 4 analyses the case studies of three brand promotion activities through social media Facebook and model the mechanics of successful diffusion, in comparison with traditional channels. The influence of social media in opinion formation of the public, is widely used in brand promotion. The casual public posts are classified, sentiments are measured and a feedback loop is created. The theoretical understandings of virality and experiments analyzing social media virality data concluded that the information is diffused in conformity with the standard Bass Model. The

research questions 1 to 4 are addressed in this chapter. Question No. 1 : How does the information diffuse in an online social network? (in section 4.8; pages 82-90) The brand promotion through viral marketing is experimented. Understanding what makes media content go viral is essential to design successful campaigns on social media. Question No. 2 Why and through which paths does the information diffuse? (in section 4.10; pages 90-91) Question No. 3 How will the diffusion pattern be, in the future? (in section 4.10; pages 92) Question No. 4 How can the social media be used to engage with the customers? (in section 4.5; pages 79-80)

Chapter 5 identifies the sentiments behind the casual social media transcripts and converts the general viewers to potential visitors to a confirmed client, applied to an eco-tourist resort property domain. The tool created helps in determining the sentiment orientation of opinions and the feedback, to improve the services offered which in turn increases the traffic. The research questions 5 and 6 of improving the brand positioning using sentiment analysis techniques are addressed in this chapter. Question No. 5 How to improve brand positioning, using opinion mining techniques to detect hidden patterns in consumers' sentiments? (section 5.11; pages 148) Question No. 6 How can sentiment analysis be used to design/redesign the brand promotion strategies? (section 5.11; pages 149)

Chapter 6 discusses a case study with a tool to monitor, study and analyse popular feedback, using forums, social media, surveys and other crowdsourcing techniques. Social media allow real-time feedback for citizens, monitoring developmental initiatives of Governments and

Multilateral agencies. The technology ensures that the consultations and on-going feedback can be captured, analysed, and used in understating the stakeholder reactions, to the project and its implementation. This helps in making necessary course corrections, avoiding costly mistakes and overruns. The feedback is gathered and analysed using, both quantitative as well as qualitative means, to understand the pulse of the crowd. This clearly explains how external data is used for Social Intelligence, which answers research question 7. How can the external data be processed to derive Social Intelligence? (in section 6.14; pages 174)

Chapter 7 discusses how Social media improves the knowledge sharing and the knowledge management practices. This is emerging as a powerful tool for information exchange and a larger community reach. This chapter is addressing the last research question 8. How to create a Knowledge Management Framework utilizing Social Media interactions? (in section 7.10; pages 196)

Chapter 8 includes the summary of the research work carried out, important contributions and suggestions for possible future directions of work in this field.

.....&&.....