

CHAPTER 1

INTRODUCTION

The 21st century is the century of Knowledge. For every answer people first use Google. The theme line is connecting people using Internet. When we talk about Internet along with knowledge we have Social sites are specially used for connecting world. People especially youth spend its lot of time on sites like Facebook, Twitter etc... Every current topic due to social media is supported with general public's reviews and comment which makes a clear picture of that particular topic to world. The credit goes to popularity of social media.

According to Com. Score India is now at third position taking place of Japan becoming the world's third largest Internet user after China and United States [14]. As per Internet and Mobile Association of India (IAMAI) in June 2017 India is having more than 243 million internet users; which is 23 % increase in count compared to year 2016. The use of social media is significantly complex. The main challenges with social media are security, privacy, accuracy but still it has given new opportunities for sharing, collaboration, and engagement of users. It provides the platform for a politics. Social media platforms, like Twitter, Facebook and YouTube motivate people to get involved in all the political activities by sharing their view about party and candidates. Elections are having a great impact of social media. It specially used by both general public for getting updates on political events and parties/Candidate to update their strategies and policies based on this social media's valuable feedback.

This paper is organized as follows. Section II reviews existing system used for election based analysis. In Section III, present proposed system. In Section IV, it provides the implementation details. Section V evaluates the dataset and results.

Social media is a web-based and mobile-based internet application that will allow the creation, access and exchange of user-generated content that is ubiquitously accessible [15,16]. Besides social networking media like twitter and face- book, the term _social media' to encompass really simple syndication (RSS) feeds, blogs, wikis and news, all typically yielding unstructured text and accessible through the web. Social media is especially important for research into

computational social science that investigates questions using quantitative techniques for example, computational statistics, machine learning and complexity and so-called big data for data mining and simulation modeling [17]. Social media has led to numerous data services, tools and analytics platforms. The tools available to researchers are either give superficial access to the raw data or non-superficial access. Researchers require to program analytics in a language such as Java. So the proposed work is much better than the available ones with respect to cost, efficient handling Big Data and scalability.

The analytics persons and businesses feel the need to gain new insights from social media; they require the analytics tools and expertise to transform this big data information which will have big volume and variety into the respective strategies so as to draw certain conclusions. Social media analytics is useful tool for getting details of customer sentiments that are distributed across online sources [18].

The Apache Hadoop software library is a framework [19] that allows for the distributed processing of large data sets across clusters of computers using simple programming models

[20]. It provide highly scalable and flexible architecture for parallel processing. Rather than rely on hardware to deliver high- availability, the library itself is designed to detect and handle failures at the application layer, so delivering a highly- available service on top of a cluster of computers, each of which may be prone to failures. Social analytics collects and analyzes consumer opinions and convert them into insights and help businesses in identifying areas of customer satisfaction or any customer grievance for the product. It also provides a quick feedback to marketing campaigns, so as to analyze campaign that will be received well by the consumers. Social analytics acts as a new channel between consumers and industries [21]. It also helps them to provide the review of their influence of product in the market. So the proposed model fulfills the needs of companies by analyzing the data efficiently and delivering results.

Micro blogging today has become a very popular communication tool among Internet users. Twitter, one of the largest social media site receives millions of tweets every day on variety of important issues. Authors of those messages write about their life, share opinions on variety of topics and discuss current issues. These posts analysis can be used for decision making in different areas like government, Elections, Business, and Product review etc. Also sentiment

analysis is one of the important area of analysis of twitter posts that can be very helpful in decision making.

Performing Sentiment Analysis on Twitter is trickier than doing it for large reviews. This is because the tweets are very short (only about 140 characters) and usually contain slangs, emoticons, hash tags and other twitter specific jargon. For the development purpose twitter provides streaming API which allows the developer an access to 1% of tweets tweeted at that time bases on the particular keyword. The object about which we want to perform sentiment analysis is submitted to the twitter API's which does further mining and provides the tweets related to only that object. Twitter data is generally unstructured i.e. use of abbreviations is very high. Also it allows the use of emoticons which are direct indicators of the author's view on the subject. Tweet messages also consist of a timestamp and the user name. This timestamp is useful for guessing the future trend application of our project.