Engineering Universe for scientific Research and management
ISSN (Online): 2319-3069
Vol. XIV Issue XII
December 2022

A Survey on machine Learning Methods for Movies Review

Ashish Rathor[1], Sumit Sharma[2]
Dept. of CSE, Vaishnavi Institute of Technology & Science, Bhopal, India[1, 2]
er.ashishrathor@gmail.com[1], *sumit_sharma782022@yahoo.co.in*[2]

Abstract

The Web clients for emerging social networks (SNs) like Facebook, Twitter, and WeChat are globally interconnected. WeChat is one of these SNs as an example. By doing research on social networks and analyzing the data gathered from these networks, we were able to get information on the comments that people made regarding a certain product. The analysis of these remarks shows how effective this method is for creating marketing and advertising campaigns. Viral marketing, finding popular bloggers, social advertising, social healthcare, finding experts, individualized praise, citation networks, and other typical applications are some examples. Social media platforms contain interactive tools and recommendations for the creation, distribution, and replacement of user-generated content. The use of social media, particularly social networking sites, has significantly increased over the past several years. As a result, our techniques for systematizing and communicating have had to change to keep up with this development. It does it at a very low cost and at the same time gathers the opinions and evaluations of many different groups. We can learn about the properties of social organizations, assess action prototypes both subjectively and quantitatively, and occasionally predict impending human-correlated events by mining the characteristics and subjects of social media. In this study article, we first examine the areas that can be predicted using the current social media platforms, after which we sketch a broad picture of the predictors and systems that are now available, and finally we analyze issues and potential research areas. Review the areas that could be computed using current social media platforms; sketch a rough understanding of attainable predictors and systems of prediction.

Keywords: *Movie Reviews, Prediction, Sentiment, Machine Learning, Deep learning.*

1. Introduction

A huge organization of conventional time-series models of calculating potential auctions of a product or service have relied on the earlier period chronological and seasonal auctions and often provided untrustworthy calculation conclusions. Significantly, since these analytical models depend on only the earlier period information of auctions, they lean to overlook the dynamic collision of modern events that might have a significant persuade on auctions. Additionally, even though clients' existing judgment and opinion about companies and products (e.g., qualified evaluates of products, individual biased private estimations) would influence buying actions and potential auctions by words-of-mouth causes, the conventional models don't have any source of inputs to judge these sorts of socio-emotional issues.

Social networking (e.g., news gatherings, item audit locales, web journals, twitter and facebook) may fill in as an intermediary of individuals' assessments, for example, their prior period encounters and current assumptions about items or administrations. Consequently, joining new key elements from the investigation of related social networking matter could enable organizations to add another layer to their current prescient models and lift the expectation exactness.

The mix of social networking expectations with existing systematic estimating models is probably going to be superior to both of the two in confinement because of the way that each model spotlights on only one part of the earth which is basic to deciding potential results. By and large, social factors alone are available a deficient photo of the present situation and are in this manner not sufficiently intense to anticipate correct request. Be that as it may, without social networking input, current request figures can't consider troublesome get-togethers and huger societal patterns that will without a doubt impact request too. Along these edges, this mix is probably going to demonstrate advantageous in any space where customer conduct is a vital segment what is to be anticipated. What's more, since existing prescient models of interest are extremely mind boggling and hard to alter, it's desirable over make changes in view of social networking without using specifics of the first request demonstrate detailing.

Social networking is a platform that enables normal people to make and distribute matters. Two overall well known social networking sites, Twitter and Facebook, show its hazardous development and significant impact. Both Twitter and Facebook are in the best 15 most-visited sites on the planet as per Alexa positioning [1]. Facebook has in excess of 1400 million dynamic customers [2] by 31 December 2017, and by December 2017, on Twitter, there were around 140 million data pieces made and exchanged day by day [3]. There is other specific social networking that is centered on stimulation, games, fund and legislative issues.

Since there are numerous customers imparting their insights and encounters by means of social networking, there is accumulation of individual intelligence and

distinctive perspectives. Such collection has confinements as perspectives are liable to change with time. It might be said the social networking predict issue is paralleled by expectation of budgetary time arrangement in light of prior period history, which has its uses in exchanging. By and large, if extricated and examined appropriately, the information via social networking media could prompt helpful predicts of certain human related occasions. Such expectation has extraordinary advantages in numerous domains, for example, fund, item advertising and governmental issues, which have pulled in expanding number of specialists to this subject. Investigation of social networking additionally gives bits of knowledge on social elements and general wellbeing. A review gives us point of view and is useful for completing further research.

Rest of the paper is organized as following: in section 2 we briefly describes about Social Networks, in section 3 overview of social media search and its application is described, different social networking services are presented in section 4, in section 5 offline and real time social media analytics is explained, key techniques to analyze textual data is given in section 6, research challenges and scopes are studies in section 7 and 8, finally we summarized our paper with facts and findings in section 9.

2. Social Network

A social media is a social structure including people or associations, which more often than not are spoken to as hubs, together with social relations, which compare to the connections among hubs. The social connection could express both, for example, family relationship and cohorts, and understood, for instance fellowship and normal intrigue. For example, fig. 1 is a case of undirected informal community in an organization, from open source programming GUESS [4]. In fig. 1, every hub speaks to a worker. The edge between two hubs implies these two representatives have a few interchanges in work and the heaviness of each edge is the correspondence recurrence.

A little social media might be demonstrated by general charts, for example, that of a little world system [5]. For a colossal very much associated arrange, most hubs could achieve each other hub through few connections. The possibility of six level of detachment proposes that, by and large, every two people are connected by six bounces [6]. The circumstance in Social Networking Service (SNS) isn't vastly different. The normal separation on Facebook in 2008 was 5.28 jumps, while in November 2011 it's 4.74. In the MSN delegate organizes, which contains 180 million customers, the middle and the 90th percent level of partition are 6 and 7.8 separately. On Twitter, the middle, normal, and 90th percent remove between any two customers are 4, 4.12 and 4.8, separately [7]. To some things up, the level of partition shifts on various SNS stages as well as on various time yet it's very little. An informal community is a scale free system [8] for which the degree dissemination asymptotically takes after a power law. On Twitter, up to 105 of the quantity of followings/adherents fit the power-law dissemination with the type of 2.276. The quantity of being re-tweeted and specified by customers on Twitter additionally takes after a power law [9].

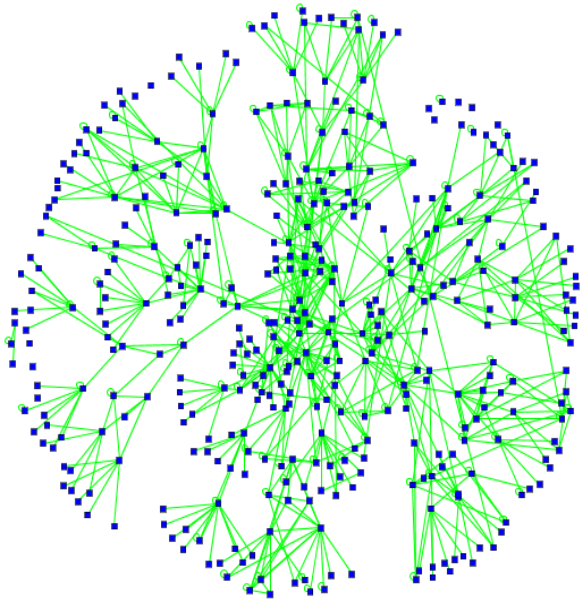

Fig. 1. Social Network Example from GUESS

3. Social media

Social networking contains stages to make and trade customer created matter [10]. Some of the time social networking is called shopper delivered media (CGM). Social networking is not the same as customary media, for example, daily paper, books, and TV, in that nearly anybody could distribute and get to data cheaply utilizing social networking. Conversely, conventional media (which is additionally eluded as old media or inheritance media) requires huge assets to distribute matters. In any case, social networking and conventional media are not totally particular. For instance, significant news channels have official records on Twitter and Facebook.

There are numerous types of social networking that incorporate web journals, person to person communication locales, virtual social universes, communitarian ventures, matter groups and virtual diversion universes [11]. A few types of social networking do not have an informal organization. In this manner in blogspot.com, which is a celebrated blog stage, there are no social connections among bloggers.

Social networking has a few or these seven capacity squares: character, discussions, sharing, nearness, connections, notoriety, and gatherings [12]. Distinctive

types of social networking have diverse purposes of core interest. For instance, cooperative ventures, for example, Wikipedia generally think about sharing and notoriety, And in virtual amusement universes, character, nearness, notoriety, and gatherings are of the best concern.
As of late, social networking assumed critical part in unfurling newsworthy occasions. For instance, in the outcome of the Tohoku Earthquake in Japan individuals utilized social networking to contact companions, trade emergency data, and discover vital assets.

3.1 Social media research and applications

Social networking information is obviously the hugest, wealthiest and most unique confirmation base of human conduct, bringing new chances to comprehend people, gatherings and society. Creative researchers and industry experts are progressively discovering novel methods for consequently gathering, consolidating and breaking down this abundance of information. Normally, doing equity to these spearheading social networking applications in a couple of passages is testing. Three illustrative regions are: business, bioscience and sociology.
The early business adopters of social networking examination were regularly organizations in retail and back. Retail organizations utilize social networking to outfit their image mindfulness, item/customer benefit change, publicizing/promoting systems, arrange structure investigation, news spread and even misrepresentation location. In back, social networking is utilized for estimating market slant and news information is utilized for exchanging. As an outline, [31] estimated conclusion of arbitrary example of Twitter information, finding that Dow Jones Industrial Average (DJIA) costs are related with the Twitter opinion 2– 3 days sooner with 87.6 percent precision. They utilized Twitter information to prepare a Support Vector Regression (SVR) model to anticipate costs of individual NASDAQ stocks, finding 'noteworthy preferred standpoint' at determining costs min in the potential.
In the biosciences, social networking is being utilized to gather information on enormous accomplices for behavioral change activities and effect checking, for example, handling smoking and corpulence or observing maladies. An illustration is Penn State University scholars [32] who have created imaginative frameworks and methods to track the spread of irresistible illnesses, with the assistance of news Web locales, web journals and social networking.
Computational sociology applications include: observing open reactions to declarations, addresses and occasions particularly political remarks and activities; experiences into group conduct; social networking surveying of (difficult to contact) gatherings; early location of developing occasions, as with Twitter. For instance, utilize computational etymology to naturally anticipate the effect of news on people in general impression of political competitors. utilize film audit remarks to contemplate the impact of different methodologies in removing content highlights on the precision of four machine learning techniques—Naive Bayes, Decision Trees, Maximum Entropy and K-Means bunching. Finally found that Facebook's Gross National Happiness (GNH) displays pinnacles and troughs in-accordance with significant open occasions in the USA.

3.2 Social media overview

For this research paper, we amass online networking apparatuses into Social media information social networking information writes (e.g. informal organization media, wikis, websites, RSS channels and news, and so on.) and groups (e.g., XML and JSON). This incorporates informational indexes and progressively imperative ongoing information sustains, for example, monetary information, customer exchange information, telecoms and spatial information.

- Social media programmatic access: information administrations and devices for sourcing and scratching (literary) information from interpersonal interaction media, wikis, RSS channels, news, and so on. These could be helpfully subdivided into:
- Data sources, services and tools: where information is gotten to by instruments which secure the crude information or give basic examination. Cases include: Google Trends, Social Mention, Social Pointer and Social-Seek, which give a surge of data that totals different online networking bolsters.
- Data feeds via APIs: where informational indexes and nourishes are open through programmable HTTP-based APIs and return labeled information utilizing XML or JSON, and so on. Cases incorporate Wikipedia, Twitter and Facebook.
- Text cleaning and storage tools: devices for cleaning and putting away printed information. Google Refine and Data Wrangler are cases for information cleaning.
- Text analysis tools: individual or libraries of apparatuses for dissecting online networking information once it has been scratched and cleaned. These are for the most part characteristic dialect handling, examination and characterization devices, which are clarified beneath.
- Transformation tools: straightforward apparatuses that could change literary information into tables, maps, graphs (edge, pie, disperse, bar, and so on.), course of events or even movement (activity over timetable, for example, Google Fusion Tables, Zoho Reports, Tableau Public or IBM's Many Eyes.
- Analysis tools: further developed investigation instruments for dissecting social information, recognizing associations and building systems, for example, Gephi (open source) or the Excel module NodeXL.

- Social media platforms: conditions that give far reaching online networking information and libraries of apparatuses for investigation. Illustrations include: Thomson Reuters Machine Readable News, Radian 6 and Lexalytics.
- Social network media platforms: stage that give information mining and investigation on Twitter, Facebook and an extensive variety of other social media sources.
- News platforms: stages, for example, Thomson Reuters giving business news files/sustains and related investigation

Social media technique to scrutinize the two noteworthy obstacles to utilizing online networking for scholarly research are right off the bat access to thorough informational collections and besides apparatuses that permit 'profound' information examination

4. Social networking service

Long range informal communication benefit is an arrangement of social destinations and applications, which at any rate comprise of three sections: customers, social connections, and intuitive correspondences. Truth be told, SNS is a subset of social networking, which incorporate the social media.

On SNS, correspondence is intuitive. For example, for unadulterated web journals, a non-SNS social networking, for example, blogspot.com, the customers' significant inspirations could be recording one's everyday life, giving critique and suppositions, communicating feeling, exhibiting thoughts through content, and keeping group. The initial four inspirations are all data sharing. For small scale blogging, a regular SNS, the customer aim could be generally characterized into three classes: data sharing, data chasing, and companionship upkeep [13].

All SNS suppliers have two center concentrations: social relations and customer created matters. Regarding social relations, they may mirror the social media of people, in actuality, manufacture new social associations in light of interests and exercises, or both. For customer delivered matters, they give a simple method to make, offer, and rank and trade data.

5. Social Media Analytics

Social networking examination could be depicted as the way toward gathering information from the social networking sites and breaking down that information to settle on business choices. Social networking investigation is for the most part used to mine customer opinion so as to help promoting and customer benefit exercises. Information examination could be ongoing or disconnected investigation, including elements, for example, impact, reach, and importance of reasonable estimations. Time contemplations are imperative to understanding the setting of information being examined. The significance of social networking examination could be viewed as the scientists at AT&T built up a logical programming to listen in customers organize issue grumblings on Twitter. The groups will be sent to settle the issue by removing time, area, and sort, from the tweet [14]. Association's devotion to serving the mass with this level of need makes it all the more intriguing and makes rivalry among the association. Associations have been concentrating on research and advancement in examination in view of the assets they as of now have.

5.1 Offline Social Media Analytics

Disconnected information investigation alludes to the uninvolved examination of information, the most part of which is to be utilized for advanced showcasing channels. The disconnected information is the particular information caught, which is created by the customer or from disconnected sources, for example, CRM information records. The caught disconnected information of a specific customer from social networking has been extremely valuable and the results of the information examination shed light to the revealed factors. The significance of disconnected examination could be viewed as the greatest presidential race happens in the USA, where hopefuls generally crusade through social networking. Analysts display a solid determining framework for US presidential races and US house race, named Competitive Vector Auto Regression (CVAR) [15]. CVAR analyzes the fame of different contending applicants by consolidating visual data with printed data from the Flicker social networking. This sort of framework could give crusade bits of knowledge to the applicant with the goal that hopeful could chip away at their shortcomings and could additionally enhance their self.

Aside from the races, investigation has been utilized to anticipate securities exchange costs. The stock exchange decides the financial estimation of the nation, numerous individuals every day share good and bad times of securities exchange costs via social networking media. Analysts recommended that securities exchange value developments could be anticipated through social networking examination by proposing an Energy Cascade Model (ECM) [16]. ECM could viably anticipate center term directional securities exchange value developments, accomplishing a normal exactness of 67.7% for upward stock value developments. A comparative approach [17] was utilized to investigate the two noteworthy occasions of securities exchange value change and exchange volume. Trust data separated from the Twitter bunch was contrasted and Dow Jones Industrial Average (DJIA). By keeping trust data into account, the outcomes demonstrate that value change and exchange volume are more related than simply checking the quantity of tweets and exchange volume is more grounded connected than value change.

However another investigation has been helped out through people in general miniaturized scale blogging long range informal communication website Twitter. The execution and mental prosperity of sprinters [18] had been followed, by checking Twitter tweets of sprinters gathering. The 925,825 messages of sprinters who utilized Nike + wellness GPS beacon were investigated. Analysts found that fitness devices were most popular in North America less than 2% runners consistently ran for at least 150 min a week, which is recommended by Centers for Disease Control and Prevention Physical activity lowered on Friday as the clients may need to be relaxed.

The sprinters have been recorded for 3 months in length; this by one means or another demonstrates that the old records could be utilized for the examination purposes. Be that as it may, it might be a major test to change over the records into some helpful frame before investigation.

From the viewpoint of dialect and history, digitization of million chronicled books and an investigation of the prior period 200 years had been finished by the tech goliath Google, demonstrating a difference in dialect utilization, elements of notoriety, oversight, and time pressure of aggregate memory. The puzzles of social networking before the Internet age could be tackled through more endeavors in this field. The enormous disconnected information investigation is some of the time simpler to perform on the grounds that information and commotion introduce in the information are predictable. The colossal volume and high speed of information could be a genuine test, and numerous specialists have done sublime work in the continuous social networking investigation.

5.2 Real-Time Social Media Analytics

Constant examination means the ability to utilize every single accessible datum and assets when required. The examination of information is done progressively and reports are created with no postponement. For the most part continuous examination is utilized for geographic area and following purposes. These days, individuals right away offer via social networking media about circumstances like catastrophic events, henceforth the constant investigation of social networking may give life-sparing data. Ongoing social networking examination of streams and charts called as Milano Design Week (MSW13) [19], it prescribes settings to guests of geo-and transiently limited city-scale occasions in Milan included 681 scenes for facilitating 1,127 occasions went to by 500,000 guests in a single week. By consolidating deductive and inductive stream thinking procedures, this framework dissects Twitter's tweets registered with its delivered excellent connection predicts. As specified before, individuals invest more energy in social networking and offer whatever is going on in the encompassing, regardless of whether it's a seismic tremor, auto crash, tidal wave, or avalanche.

A multilevel issue investigation of Twitter tweets gathered about landfall recommends that noteworthy data was less demanding to discover while seeking along hash labels. The utilization of Twitter in pre-crisis phases of a climate occasion could be useful for the crisis administration office's [20]. Another case of continuous social networking examination is the observing of flare-ups through the intermediary of customers look. Google's Flu Trends and Dengue Trends give evaluations of influenza and dengue in light of pursuit designs. Additionally, Google Trends could precisely anticipate the case disconnected achievement in view of the rating of social notices of individual films and the check of the inquiry made on YouTube. From the above investigations, couples of things are concerned, for example, the responses of the general population to the circumstance. The conduct of people groups differs as indicated by the circumstance, which could choose that on whom they will trust indiscriminately or may go out on a limb to put stock in others.

A thorough and quantitative meta-examination was directed to explore the observational proof of the most compelling elements, trust, and hazard which influence the individual conduct towards social networking stages. The discoveries proposed that both hazard and trust had noteworthy impacts, yet trust had a more grounded impact. The impacts of hazard and trust have been plainly noticeable on the social networking. Trust is firmly identified with the satisfaction of the human conduct and generally more joyful people are more reliable.

At a development rate of ~8%, Internet customers are presently over 40% of whole total populace. Social networking ceaselessly assuming an awesome part to achieve that checks and has touched the numerous parts of human life. With this, the social networking is in charge of a radical new pattern that is important for associations, discovering and developing interesting patterns in human conduct.

6. Key Methods

We start with definitions of some of the key techniques related to analyzing unstructured textual data:

- Natural language processing: (NLP) is a field of software engineering, computerized reasoning and phonetics worried about the collaborations amongst PCs and human (characteristic) dialects. In particular, it's the procedure of a PC extricating significant data from normal dialect input as well as creating characteristic dialect yield.
- News analytics: the estimation of the different subjective and quantitative characteristics of literary (unstructured information) news stories. Some of these characteristics are: assessment, pertinence and curiosity.

- Opinion mining: opinion mining (sentiment mining, opinion/sentiment extraction is the region of research that endeavors to make programmed frameworks to decide human sentiment from content written in normal dialect.
- Scraping: gathering social information from online networking and other Web destinations as unstructured content and furthermore known as website scraping, web collecting and web information extraction.
- Sentiment analysis: assumption examination alludes to the utilization of regular dialect preparing, computational etymology and content investigation to distinguish and remove subjective data in source materials
- Text analytics: includes information recovery (IR), lexical examination to think about word recurrence appropriations, design acknowledgment, labeling/comment, data extraction, information mining systems including connection and affiliation investigation, perception and prescient investigation.

7. Research challenges

Social media scraping and analytics provides a rich source of academic research challenges for social scientists, computer scientists and funding bodies. Challenges include:

- Scraping—albeit online networking information is available through APIs, because of the business estimation of the information, the greater part of the significant sources, for example, Facebook and Google are making it progressively troublesome for scholastics to acquire extensive access to their 'crude' information; not many social information sources give reasonable information offerings to the scholarly community and analysts. News administrations, for example, Thomson Reuters and Bloomberg regularly charge a premium for access to their information. Interestingly, Twitter has as of late reported the Twitter Data Grants program, where analysts could apply to gain admittance to Twitter's open tweets and chronicled information to get bits of knowledge from its monstrous arrangement of information (Twitter has in excess of 500 million tweets every day).
- Data cleansing: cleaning unstructured printed information (e.g., normalizing content), particularly high-recurrence spilled continuous information, still exhibits various issues and research challenges.
- Holistic data sources: analysts are progressively uniting and joining novel information sources: online networking information, constant market and customer information and geospatial information for investigation.
- Data protection: once you have made a 'major information' asset, the information should be secured, proprietorship and IP issues settled (i.e., putting away scratched information is against the vast majority of the distributers' terms of administration), and customers furnished with various levels of access; generally, customers may endeavor to 'suck' all the important information from the database.
- Data analytics: modern examination of social networking information for supposition mining (e.g., assumption investigation) still raises a bunch of difficulties because of outside dialects, remote words, slang, spelling blunders and the normal developing of dialect.
- Analytics dashboards: numerous online networking stages expect customers to compose APIs to get to bolsters or program examination models in a programming dialect, for example, Java. While sensible for PC researchers, these aptitudes are normally past most (sociology) specialists. Non-programming interfaces are required for giving what may be alluded to as 'profound' access to 'crude' information, for instance, designing APIs, blending online networking nourishes, consolidating all encompassing sources and creating expository models
- Data visualization: visual portrayal of information whereby data that has been disconnected in some schematic frame with the objective of imparting data unmistakably and successfully through graphical means. Given the extent of the information included, perception is winding up progressively critical.

8. Research Scope

As a developing examination point, predict with social networking faces numerous difficulties. Here we call attention to some earnest and imperative potential works.

At present, analysts pick indicators utilizing the experimentation technique. We know neither why these indicators are superior to others, nor how these indicators could foresee the outcome. Not knowing the foundation rationale between these measurements and the last expectation result, we simply utilize an accumulation of measurements to be prepared on test information, discover which ones have the most elevated coefficients, and utilize them to make the predict display. Thusly, deficient with regards to a strong supporting hypothesis, we can't make sure that one model, which functions admirably in one case, could be connected to different circumstances with a similar exactness. That is the reason a few models demonstrate great execution in one race predict, however totally explode in another. To ensure our model has great

execution in all cases, we have to know the rationale and hypothesis behind the model.
Most analysts utilize straightforward techniques, for example, direct relapse investigation. These techniques are known to function admirably under a few conditions. Social networking is created on a mind boggling framework and in this manner probably the indicators and predicts results have non-direct relationship. Moreover, mix of strategies may prompt leap forward. In such mix, a surface learning specialist, for example, immediately prepared neural systems, rapidly adjusts to new modes and developing patterns via social networking media. What's more, a profound realizing specialist centers on long haul designs. More or less, we should attempt some non-direct techniques and discover the appropriate strategies or potentially mixes for every predict domains.

9. Conclusions

The phrase "social networking" refers to a collection of computer and mobile Internet-based applications that enable users to create, access, and exchange user-generated content anywhere. The study of social networking is particularly significant in the subject of computational sociology, which examines issues using quantitative methodologies (such as computational insights, machine learning, and complex nature). This is so that information mining and entertainment presenting can be done because social networking sites offer what is known as large amounts of data. Numerous information administrations, tools, and evaluation stages have been produced as a result. Despite this, the numerous marketing demands that are currently present have the potential to substantially change the uncomplicated accessibility of social networking information for academic investigation. With the aid of social networking, we examined expectations in this piece of writing. Additionally, we offered a flowchart outlining forecast aspects and techniques, as well as a list of testing problems and areas that needed more research. Despite the fact that social networking prediction is still a relatively new research area and that its results have a relatively low level of precision, it has provided us with another way to gather, extract, and use the collective intelligence of groups in a focused manner that is both easy and efficient.

References

[1] Alexa Internet Inc, "Alexa Top 500 Global Sites". http://www.alexa.com/topsites. [Accessed Jan 24, 2018].
[2] Facebook, "Statistics". http://www.facebook.com/press/info.php?statistics. [Accessed Jan 14, 2018]
[3] Twitter, "#numbers". http://blog.twitter.com/2011/03/numbers.html. [Accessed Jan 24, 2018].
[4] Eytan Adar, "GUESS, the Graph Exploration System". http://graphexploration.cond.org/. [Accessed Jan 24x, 2018].
[5] Kleinberg, Jon. "The small-world phenomenon: An algorithmic perspective." In Proceedings of the thirty-second annual ACM symposium on Theory of computing, pp. 163-170. ACM, 2000.
[6] Watts, Duncan J. Six degrees: The science of a connected age. WW Norton & Company, 2004.
[7] Kwak, Haewoon, Changhyun Lee, Hosung Park, and Sue Moon. "What is Twitter, a social network or a news media?." In Proceedings of the 19th international conference on World wide web, pp. 591-600. ACM, 2010.
[8] Watts, Duncan J. "The "new" science of networks." Annu. Rev. Sociol. 30 (2004): 243-270.
[9] Cha, Meeyoung, Hamed Haddadi, Fabricio Benevenuto, and P. Krishna Gummadi. "Measuring client influence in twitter: The million follower fallacy." Icwsm 10, no. 10-17 (2010): 30.
[10] Romero, Daniel M., Wojciech Galuba, Sitaram Asur, and Bernardo A. Huberman. "Influence and passivity in social media." In Joint European Conference on Machine Learning and Knowledge Discovery in Databases, pp. 18-33. Springer, Berlin, Heidelberg, 2011.
[11] Kaplan, Andreas M., and Michael Haenlein. "Clients of the world, unite! The challenges and opportunities of Social Media." Business horizons 53, no. 1 (2010): 59-68.
[12] Kietzmann, Jan H., Kristopher Hermkens, Ian P. McCarthy, and Bruno S. Silvestre. "Social media? Get serious! Understanding the functional building blocks of social media." Business horizons 54, no. 3 (2011): 241-251.
[13] Java, Akshay, Xiaodan Song, Tim Finin, and Belle Tseng. "Why we twitter: understanding microblogging usage and communities." In Proceedings of the 9th WebKDD and 1st SNA-KDD 2007 workshop on Web mining and social network analysis, pp. 56-65. ACM, 2007.
[14] Ikeda, Kazushi, Gen Hattori, Chihiro Ono, Hideki Asoh, and Teruo Higashino. "Early detection method of service quality reduction based on linguistic and time series analysis of twitter." In Advanced Information Networking and Applications Workshops (WAINA), 2013 27th International Conference on, pp. 825-830. IEEE, 2013.
[15] You, Quanzeng, Liangliang Cao, Yang Cong, Xianchao Zhang, and Jiebo Luo. "A multifaceted approach to social multimedia-based prediction of elections." IEEE Transactions on Multimedia 17, no. 12 (2015): 2271-2280.
[16] Zhang, Wenping, Chunping Li, Yunming Ye, Wenjie Li, and Eric WT Ngai. "Dynamic business network analysis for correlated stock price movement prediction." IEEE Intelligent Systems 30, no. 2 (2015): 26-33.
[17] Ruan, Yefeng, Lina Alfantoukh, and Arjan Durresi. "Exploring stock market using twitter trust network." In Advanced Information Networking and Applications (AINA), 2015 IEEE 29th International Conference on, pp. 428-433. IEEE, 2015.

[18] He, Qian, Emmanuel Agu, Diane Strong, Bengisu Tulu, and Peder Pedersen. "Characterizing the performance and behaviors of runners using twitter." In Healthcare Informatics (ICHI), 2013 IEEE International Conference on, pp. 406-414. IEEE, 2013.
[19] Balduini, Marco, Alessandro Bozzon, Emanuele Della Valle, Yi Huang, and Geert-Jan Houben. "Recommending venues using continuous predictive social media analytics." IEEE Internet Computing 18, no. 5 (2014): 28-35.
[20] Lachlan, Kenneth A., Patric R. Spence, Xialing Lin, Kristy Najarian, and Maria Del Greco. "Social media and crisis management: CERC, search strategies, and Twitter matter." Computers in Human Behavior 54 (2016): 647-652.
[21] Gruhl, Daniel, Ramanathan Guha, Ravi Kumar, Jasmine Novak, and Andrew Tomkins. "The predictive power of social chatter." In Proceedings of the eleventh ACM SIGKDD international conference on Knowledge discovery in data mining, pp. 78-87. ACM, 2005.
[22] Jansen, Bernard J., Mimi Zhang, Kate Sobel, and Abdur Chowdury. "Twitter power: Tweets as electronic word of mouth." Journal of the Association for Information Science and Technology 60, no. 11 (2009): 2169-2188.
[23] Bhatt, Rushi, Vineet Chaoji, and Rajesh Parekh. "Predicting product adoption in huge-scale social networks." In Proceedings of the 19th ACM international conference on Information and knowledge management, pp. 1039-1048. ACM, 2010.
[24] Zhang, Wenbin, and Steven Skiena. "Improving movie gross prediction through news analysis." In Web Intelligence and Intelligent Agent Technologies, 2009. WI-IAT'09. IEEE/WIC/ACM International Joint Conferences on, vol. 1, pp. 301-304. IEEE, 2009.
[25] Asur, Sitaram, and Bernardo A. Huberman. "Predicting the potential with social media." In Proceedings of the 2010 IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology-Volume 01, pp. 492-499. IEEE Computer Society, 2010.
[26] Lica, Liviu, and Mihaela Tuta. "Predicting product performance with social media." Informatica Economica 15, no. 2 (2011): 46.
[27] O'Connor, Brendan, Ramnath Balasubramanyan, Bryan R. Routledge, and Noah A. Smith. "From tweets to polls: Linking text sentiment to public opinion time series." Icwsm 11, no. 122-129 (2010): 1-2.
[28] Metaxas, Panagiotis T., Eni Mustafaraj, and Dani Gayo-Avello. "How (not) to predict elections." In Privacy, Security, Risk and Trust (PASSAT) and 2011 IEEE Third Inernational Conference on Social Computing (SocialCom), 2011 IEEE Third International Conference on, pp. 165-171. IEEE, 2011.
[29] Zhang, Xue, Hauke Fuehres, and Peter A. Gloor. "Predicting stock market indicators through twitter “I hope it’s not as bad as I fear”." Procedia-Social and Behavioral Sciences 26 (2011): 55-62.
[30] Zhang, Wenbin, and Steven Skiena. "Improving movie gross prediction through news analysis." In Web Intelligence and Intelligent Agent Technologies, 2009. WI-IAT'09. IEEE/WIC/ACM International Joint Conferences on, vol. 1, pp. 301-304. IEEE, 2009.
[31] Bollen, Johan, Huina Mao, and Xiaojun Zeng. "Twitter mood predicts the stock market." Journal of computational science2, no. 1 (2011): 1-8.
[32] Salathe, Marcel, Linus Bengtsson, Todd J. Bodnar, Devon D. Brewer, John S. Brownstein, Caroline Buckee, Ellsworth M. Campbell et al. "Digital epidemiology." PLoS computational biology 8, no. 7 (2012): e1002616.