
Social Media Analytics for Small, Medium Enterprises

Padma Raghappanavar , Dr.Yerriswamy T.

K. L. E. Institute Of Technology Hubballi, Karnataka, India

ABSTRACT

Social media marketing or digital marketing is now a day a very popular method of marketing the business. Unlike traditional advertising methods the digital marketing is faster, cost effective and helps to reach the more customers in less time. In Indian economy small scale local business occupy important place. The small scale industry contributes to the national GDP, total industry employment, industrial units and export. In this paper, we develop a web application which will derive a social analytics of company or a service or a product to take business decisions

Keywords

Keywords: Marketing, Digital-Marketing, Social-Media.

INTRODUCTION&OTHER HEADINGS (Heading Times New Roman, 12 points, Bold, All caps)

In the era of digital marketing, social media plays a key role in negotiating the human communication and business tactics to build the well know brand in the digital marketing. Social media analytics provides a platform to understand the behavior of the consumer [1]. Social media allow the users to create their own profile, pages to share their feelings, opinions it helps the peoples to communicate with each other [2]. The communication media may be in the form of text, video, photos. All information generated at social media is estimable resource for business peoples to know the human behavior and also for researcher. The generated data is voluminous and unstructured. To know the particular opinions and feelings we must analyze the data with the number of likes for particular post. We can also get the number of page followers, fans, and friends [3]. More and more people are using social media personally and professionally to read the posts and to give the comments on that posts of particular page.

Internet allow the users to generate the unlimited data at different social media sites like Facebook, Twitter ,Google+ ,Youtube etc. Hence, social media analyzer will help to make the qualitative and large scale analysis[3]. Some the available social media analysis tools are hootsuite, keyhole, synthesio [4]. These tools are used for digital marketing. These tools access the data from the social sites based on the keywords. Few business domains require the real time analysis of the data which was accessed from different social networking sites [5].

1.1 RELATED WORK

In this research, the social media analysis is carried out using the Key Performance Indicator (KPI) to know the customer experience, customer interaction, customer activation and customer satisfaction, reach and finance [6]. The social media analysis is necessary to understand the customers, to improve the business and to get good Return of Invest (ROI).

Local small scale businesses are facing the challenges of marketing, crediting, technology up gradation etc. Social media networking sites can be used to extract the information in two different ways; Really Simple Syndication(RSS) feed and access through the API. Accessing the different types of data formats, like, text format, video format or it may be in the form of images. The accessed data is available in the form of JSON or XML which can be stored and processed in the form of CSV files or directly dump into the database to perform analysis and get the inferences [7].

By above discussions, we notice that there are social media analyses tools present for the start-up companies, local business developers and for those who are launching a new brand to the market. To support the small scale business our project statement is

Analysis and derivation of social media analytics of a social media entities like

) The number of friends, fans and followers of the current Facebook and Twitter page and number of shares for that particular post.

1.2 OBJECTIVES OF THIS PAPER

) To develop the social media analysis tool we have access the information from the Facebook and Twitter pages

) Tell the presence of the Twitter and the Facebook pages

) Total count the number of friends for the Facebook page and the number of followers for the Twitter pages.

) Gives the individual page counts for both Twitter and Facebook pages

) Count the number of posts, likes per post in single page of that Facebook page and count the number of Count the number of shares if the post is shared.

1.3 METHODOLOGY

In this paper we discuss about the iterative methodology to develop the social media analyzer, here we are diving the process into the three phases as shown in the fig 1.1 . The three phases are

- Extract data from the social media sites; Facebook and Twitter
- Analyzing the extracted data with TF-IDF vectorization and KNN algorithm
- Presentation : get the users interacted information such as; number of fans , followers ,total number of likes per post , number of comments

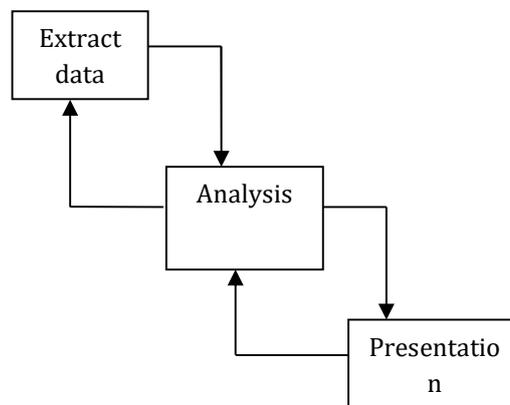


Fig 1 iterative methodology for social media analyzer

2 DATA EXTRACTION

We may extract the data from the social media networking sites in different ways; through RSS [7], API's [7]. In this paper we extract the data from two social media sites i.e., Facebook and Twitter by using the API's; Graph API and Stream API respectively [8].

2.1 Graph Api For Facebook Data Extraction

Graph API is one of the low level HTTP based API, mainly used in; Facebook data extraction, query the data and post the new data to the site. The data of the Facebook is in the form of the graph with nodes and edges, so it is also called as social graph. It is has more features than the rest API. The graph API will provide the three methods to interact with Facebook they are; Request the data by is using the GET method, Post the new feeds to the site by using the POST method and Delete the data

Request data from the social sites needs the authentication of user. Only the authenticated users can access the data. The graph API holds the Facebook data in the form of the nodes, edges and fields.

-) Nodes – is nothing but a user, a photo, a comment around these points the interaction will take place in the Facebook, the communication will build.
-) Edges – the connections between the two or more different nodes.
-) Fields – some of filed like name, id, post_id etc.
-) Every graph API has an object id; these object id's are unique in nature helps to extract the information of particular person or page data. The following instructions show use of the id to request node is

GET graph.Facebook.com/{node-id}[13]

For example: GET graph.Facebook.com/1234567890

-) Post instruction to send the data to the node
-) POST graph.Facebook.com/{node-id}
- Request to the edge is

GET graph.Facebook.com/{node-id}/{edge-name}

Example: GET graph.Facebook.com/me/photo

- Post instruction to send the data to the edge
- POST graph.Facebook.com/{node-id}/{edge-name}

By using the Graph API we have to generate the access token; that gives the data access permission to the user.

2.2 Stream API For Twitter

Twitter gives the different type of API to access the data from its database. The main APIs are; Stream API and REST API.

REST API supports the accessing the data from Twitter with limit rate i.e. up to 150 queries per hour. This limit is the one of the main drawback of REST API because it's not possible to access the large amount of data at a single time, but we can access the data in any angular direction [14].

Stream API is another API, which provides the access to the large amount of data at single time in real time access of Twitter pages. These API's present in the below URL:

<https://dev.Twitter.com/>

Stream API is more flexible than REST API in more application [14]. It allows to-access the huge amount of data in single stretch for variety of data analysis process. Stream API is very helpful when more number of queries is pushing between the server & client to get information. We have to get the access token, access secret, consumer key & Consumer secret.

To access the twitter data we need to require access token, access secret, consumer key and consumer secret.

All these parameters can be fetched only if the admin or developer has the Twitter account with a valid user name & password.

3 DATA ANALYSIS

After accessing the data from the different social media sites we analyze the data by using the algorithms and vectorization methods

To implement the social media analyzer, different types of algorithm and vectorization methods like TF-IDF, algorithms like KNN, Multinomial NB have been used [9].

KNN – K Nearest- Neighbour Algorithm. It's one of the simplest and non-parametric machine learning algorithm, no assumptions are required. This algorithm will give good performance. This is very helpful for the text classification to make the analysis of the data.

The function of the classifier is to combine the large amount of text document in one or more predefined categories [10]. In our work, for the process of the classification KNN algorithm has been used [11]. The work of this algorithm is to classify the object into one or more predefined classes.

3.1 KNN Classifier:

Machine learning the pre-processing and documents preparation will come in the learning phase. The algorithm finds the documents which all contain the searched keyword. If more than one documents contains the word, then use the 'Euclidean distance' to find the nearest word to the searched one.

Euclidean formula:

$$d(p, q) = d(q, p) = \sqrt{(x - a)^2 + (y - b)^2} \dots\dots 1$$

p, q are the co-ordinates.

Ex: Search the key word as Clarks Inn

mainvect = [Clarks Inn]

mainvec [0] = Clarks

mainvec [1] = Inn

In the next process we assign the weightage for each single word in the searched text. This can be done using the TF-idf vectorization method [12].

3.2 TF – IDF vectorization

Term frequency Inverse Document Frequency, this is one of the method used to give the weight-age for a particular word based on their importance in a particular document. Term frequency means, number of times the word appears in the whole document.

The searched keyword contains the term Inverse document Frequency (IDF) ,which is used very often in every document without giving any weight-age to the word, so in such cases we have to ignore such type of words.

Tf(t,d)=ft,dthe number of times the term

't' occurred in document 'd'

t- the term

d- document

$$idf(t, D) = \frac{\log N}{|\{d \in D: t \in d\}|} \dots\dots\dots 2$$

N – Total number of documents

$|\{d \in D: t \in d\}|$ - Number of documents where 't' is appear

$$\therefore tf-idf(t, d, D) = tf(t, d) * idf(t, D)$$

4 DISCUSSION ON RESULTS

For the social media data analysis we took the data from the two different sites; Facebook and Twitter by using the API's. Consider an example of "Clarks Inn" hotel, the analyzer will search for the pages present in both social sites and gives the total number of fans present in the both sites as show in the fig 2. The total numbers of fans in Facebook are 4238 and Twitter is 450. The fig 3 shows the social presence of the Clarks inn hotel in both sites.

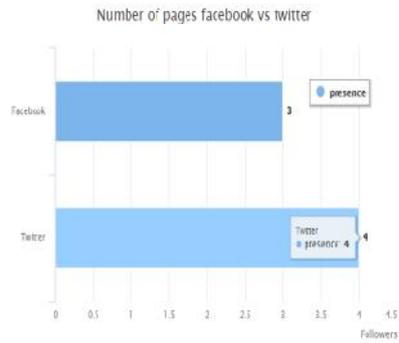


Fig 2 Total number of fans on Facebook and Twitter

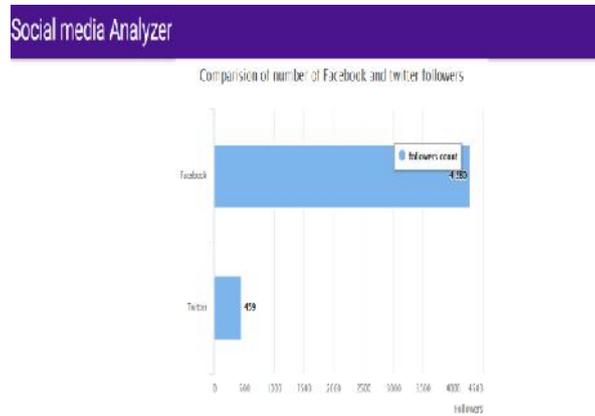


Fig 3 Shows the presence of the Clarks Inn hotel in both sides

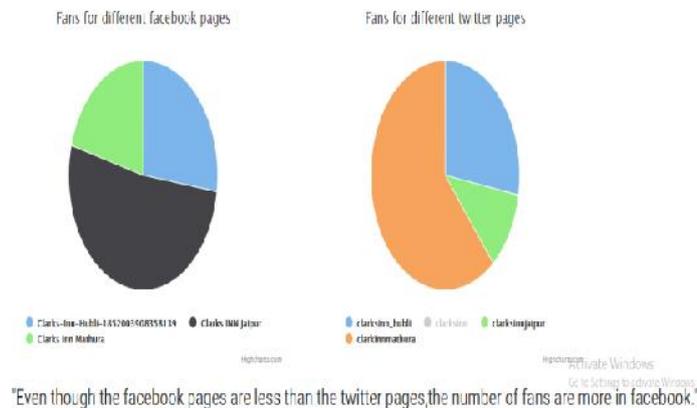


Fig 4 Shows the each page fan count

The fig 4 shows total number of pages present in both sites, for Clarks inn hotel are;

-) clarks in hubli
-) clarks in jaipur
-) clarks in Mathura

5. CONCLUSION

This paper has shown the social media analysis of social media entities like total number of fans, number of likes per post, number of shares from the data; extracted from the two social media sites like Facebook and Twitter with the help of Graph and Stream API respectively. Analyzing the data with TF-IDF vectorization and KNN algorithm and visualized the results with different graphs.

Future scope for this paper; giving the semantic analyses of the extracted data and adding the intelligence to the algorithm.

REFERENCES

1. THE BUSINESS IMPACT OF SOCIAL MEDIA ANALYTICS Kurniawati, University of Melbourne, Doug McDonnell Building, 3010 Victoria, Australia, kurniawati@student.unimelb.edu.au
2. DBus Horiz 53(1):59–68 Lazer D et al (2009) “Kaplan AM, Haenlein M (2010) Users of the world, unite! the challenges and opportunities of social media.” Computational social science. Science 323:721–723
3. “Synthesio - Home - Social Media Monitoring and Engagement.” [Online]. Available: <http://synthesio.com/corporate/en>. [Accessed: 10May-2013].
4. Calderon, Nadya, Richard Arias-Hernandez, and Brent Fisher. "Studying Animation for Real-Time Visual Analytics: A Design Study of Social Media Analytics in Emergency Management." System Sciences (HICSS), 2014 47th Hawaii International Conference on. IEEE, 2014
5. WEINBERG, T., PAHRMANN, C. & W. 2012. Social Media Marketing: Strategien für Twitter, Facebook& Co, Köln, O'Reilly Verlag.
6. Volker Stich1, Roman Emonts-Holley1, Roman Senderek1“Social Media Analytics in Customer Service, 1Institute for Industrial Management, RWTH Aachen University, Campus-Boulevard”
7. Bogdan Batrinca Philip C. Treleaven, “Social media analytics: a survey of techniques, tools and platforms”(2014)
8. Irena Pletikosa Cvijikj, Florian Michahelles Information Management ETH Zurich Zurich, Switzerland “Monitoring Trends on Facebook”
9. Bruno Trstenjaka,Sasa Mikacb, Dzenana Donkoc,”KNN with TF-IDF Based Framework for Text Categorization”,24th DAAAM International Symposium on Intelligent Manufacturing and Automation, 2013
10. Y.Liao, V. Rao Vemuri, Using K-Nearest Neighbor Classifier for Intrusion Detection, Department of Computer Science, University of California, Davis One Shields Avenue, CA 95616
11. Wang , X. Li, An improved KNN algorithm for text classification, 2010
12. <https://en.wikipedia.org/wiki/Tf-idf>
13. <https://developers.facebook.com/docs/graph-api>
14. <https://dev.twitter.com/>