International Journal of Recent Technology and Engineering (IJRTE)
ISSN: 2277-3878, Volume-8 Issue-4, November 2019

Mining Social Networking Sites: A Specific area of Facebook

Sonam, Surjeet Kumar

Abstract: ***Now a days facebook is the specific social network site for communicating more people, to develop learning process & research area. Although mining and analysis are needed area of social network so, we are showing useful part in this paper that are related to communication with people, collection of data and uses of social network's tool and techniques .During our research work we found several tools are available for collecting data and analyzing fan pages. For the purpose of our research work, we used social network site such as Facebook, in this platform we created one page related to blog who did the panacea for us. In this research paper we have applied online social media network for extracting data. Uniform node detection has used for finding common properties among audience as well as Regular equivalence nodes are using for calculating uniformity. Effective user detection, Graph structure, sampling framework are an efficient way that are explained in our paper. Breadth first search & Uniform are best approaches that has used .Graph API has major role to collect data and for the purpose of analyzing data, we used Facebook crawling process. Whenever for growing audience and know about audience information , we have explained about Facebook Insight, Like Alyzer , Sociograph , Agorapulse, Quintly , Brandwatch, So Trender , Brand 24 , Social Bakers, Rival IQ, Unmetric etc. that are more appropriate and accurate tools .***

Keywords: Sociograph, Agorapulse, Quintly, Brandwatch, Web Data Mining, Face book

I. INTRODUCTION

With advancement of technologies and computerization, we are able to collect and analyze data. From data we can extract and get information but without analyze and interpret data, this collected information is not valuable.An uncertain need for new methods and automation to transform large stored data into useful knowledge. For this, the demand for tools is increasing which is proving effective for data analysis. Data mining applications are so profitable is because data leads to knowledge, and knowledge leads to power. This is why a vast range of companies, organizations, and even political campaigns purchase targeted marketing in order to gain an advantage in selling a product or supporting a political candidate.

Manuscript published on November 30, 2019.
* Correspondence Author
Sonam*, Research Scholar , MCA Dept. VBSPU, Jaunpur, India Email:sonamjha1980@gmail.com
Dr Surjeet Kumar, Assistant Professor. MCA Dept. VBSPU, Jaunpur, India Email: surjeet_k_yadav@yahoo.co.in

Because of this, almost 90% of Fortune 500 companies are using big data initiatives to help them predict consumer behavior. Social networking sites SSN plays an important and crucial role in determining the consumer behavior. Facebook is a great location for data mining because of the variety and speed of which new data is added. Another reason why Facebook is a perfect place to mine data is the type of data Facebook's users provide.

II. BACKGROUND AND RELATED WORK

We studied several literatures and got techniques that are useful for collecting and analyze specific information. We are discussing some techniques such as:

A. **OSMNs Dataset:** Mining related data from the Web mining platform by using tools such as OSMN's Web extraction technology. Since the OSMN dataset resides in the backend server, it is not available in the public domain and can therefore only be accessed through the web interface. FB uses many access algorithms, such as Random Walk or BFS [25]. Uniform FB map sample has generated by this tool.

B. Uniform Node Detection (UND): Used to obtain degree for consistency of two users or nodes in the SM diagram. The Jacquard coefficient similarity measure used for calculates uniformity among finding users who has common attributes but it has certain drawbacks. The first drawback is that it does not consider global information. Second, it shows the similarity between nodes, because the fact is that although there is no real similarity between them, the possibility of sharing with a large number of acquainted nodes very high. If one node has less communication with other nodes, the consistency between the two nodes will increase.

In this regard, many other methods have also been explored, such as "rule equivalence" (two users are the same or similar, if they also have the same acquaintance), many authors use the Katz coefficient, Simrank[23], etc to provide the method of iterating fixed points, in which the author gives the node homogeneity as an optimization problem, they studied the directed graph and adopted the iterative method. Formal concept analysis is a different approach analysis to node on social network. This method depends on the related nodes and then calculates the similarities among the nodes. It is difficult to calculate because it follows the concept of singular value decomposition (SVD) between common partners and nodes (using linear algebra-based techniques).

C. **Effective User Detection (EUD):** This is the procedure of finding users for attracting other users to share discussions, event or activities.

In order to analyze blogs, few algorithms are designed that are HITS Algorithm, Random walk technology and HP labs also used for analyze user's behavior.

III. EXPLORING THE GRAPHIC STRUCTURE OF FACEBOOK

Researchers are interested in the large-scale for using of the features and attributes of social network. According to the graphical representation of FB, [G = V, E] where [V] = end user [E] =edge that denotes relationship. The structure related to two forms, first is in sighted graph that depict all relationships within the network have the similar value i.e. identified by undirected graphs.

A. Survey of statistical network models: According to researchers, Machine Learning supports social network model like facebook which does not cycle simple unweighted undirected graphs. Unlike the facebook, other online social network sites are more complex during arrangement or configuration, trying to explore and retrieve the data mining process.

B. How to access the graphical structure of FB: There are several ways to show information on the FB map. For example, first method is to get data directly from company related to social network. This is no appropriate worthwhile result. Another option is to access data from social network platform that is necessary to recreate the network structure. Due to the dynamic nature of graphic structure, the resulting sample is only a Polaroid of the graph structure during the data collection process. Several data sampling algorithms are available to accomplish the above tasks, but we only consider two methods, first the "broadness first (BFS) (bias method)", followed by "unified (unified) - partial Set method".

C.How to mine facebook: During collecting data, it can be used to compare and analyze its attributes, styles and quality. The parameters that can be evaluated for the quality for the data collection, sample are:

a) By using with mathematical or statistical models

b) Agree to the quality of the results of other similar research work.

Due to privacy and data protection in social network sites, companies running suitable social networking services that cannot share data about users. Information can be accessed through a graphical user interface, but there are some technical issues, such as using asynchronous scripts, we can grab a list of friends.

The FB developer team provided some other online services in 2010, such as the "Graphic API (Application Programming Interface)". By the end of 2011, we could access the structure of the facebook by using technology such as web data mining.

IV. THE SAMPLING FRAMEWORK OF FB:

According to Web data mining process architecture, following components are related i.e. 1) Execute a proxy to mine 2) Java-based platform that is an independent application web server, application executes these proxy code 3) The Apache HTTP Server Control Interface that is designed to help administrator control to manage information. During execution of process, the data mining agent checks the facebook server to obtain a collection of web pages about friends who connected to the specific requesting user, there by reforming the friendship structure between them. Finally, the collected data are stored on a web server and then post for processing and passed.

A. Facebook Crawling Process:

The facebook Crawler architecture has been designed as a Java-based cross-platform proxy that crawls the Facebook (front-end platform) GUI and key part of the web data mining. For the execution of the crawl agent, the first process of the data mining is to prepare two steps, the first is to select the sampling algorithm, and the second is to set some technical parameters, such as

existing standards, maximum execution time, and so on. During execution, the java-based crawl agent has accessed the friend list related to requested user and sampling algorithm instructions are used to searching social networks or charts.

Data mining provide us the knowledge but we cannot use this type of knowledge until we analyze the data monitoring metrics and analyzing the performance is crucial part of a successful social media strategy. As facebook is the largest network available presently, it is often the first site which has turned to brand. There are many tools are available in which facebook analytics has used to grow its potentiality that is beneficial for researchers.

However multiple tools offer analytics in social media sites, that helps the brands in analytics as 91% of brands have multiple social accounts. Now, there are free as well as paid facebook analyze tools are available, which are defining.

Facebook Insights: If Page's admin have 30 fans, facebook insight as a tool is available. Facebook insight can provide to detail about metrics related to posts and the engagement of audience. That means it offers the fan pages analysis, including reached post among people as well as reports according to like comments and share and overall activities of post. For promoting pages paid and organic options are available so page's admin can use for promote posts, video views, actions taken on the page and the reach among more audience.

Sociograph.io: A free facebook analytical tool. Once one has been authenticated to analyze page for facebook accessibility. It displays posts, about authors, comments and likes on the base of graphical structure. Pages can extract about overall appearance related to posts and on the base of posts, this tool is sufficient for displaying reports of audience performance for improving pages quality.

Agorapulse: There are two unbound tools are available. First part can perform as a standard page to tell about page's measurement is up to average and calculated metrics that has to be considered. Second part allows for competitive area, questionnaire and figures out on the timeline. The particular area is a participatory as well as management part for multiple social network accounts. Agorapulse tracks response speed and response time. The mechanism includes the most authoritative audience and users who care most about the page. It shows analysis of page and timeline. One can analyze paid, natural and viral transmission. With this tool, we can see reports according to work and helpful to see performance of marketing on facebook platform that are customized to meet user needs.

Like Alyzer : Meltwater offers a free tool i.e. Like Alyzer that allows access to any page that are created on facebook platform. It is specific tool for measurement and performance analysis. It sets the page level up to 100 and compares it to other pages. This means we can also snoop on competitors' pages. The report includes a number of sections, including indicators of participation, time and length of posts, and recommendations for improvement. Output measures to help drive more engagement.

Quintly: Quintly covers several social media accounts including Facebook. It covers the users social profiles and that of the competitors, visualizes the content to help for measuring best data and related information for improve quality. This analytical tool covers over 250 metrics and reports.

Brandwatch Analytics: Free facebook analytical tool offers social media analytics covering over 90 million sites. It covers several social media accounts.

SoTrender: Insights about audience, demographics and activities. Post reached among people and engagement with different types of post. Post can be easily managed scheduled and published. This tool allows tracking any brand, for the analysis of the competitor and industry benchmarking.

Brand24:Shows recently report about fan pages and audience performance in statistical graph. It can perform sentiment analysis and impact scoring and permits teams to work together for similar or meaningful purpose. Have an alert to keep track of any issues. It covers multiple social networks, including facebook.

Social Bakers: Tracking facebook as well as Twitter and YouTube pages as well as competitors' pages means that reports can focus on their own pages only, as well as comparing facebook metrics with competitors' pages. Download it as a PDF. Includes the number of fans and the growth of fan lists and locations. Engagement includes usual breaks, engagement types and response rates.

Rival IQ: Rival IQ provides extensive channel analysis including a range of social media and SEO, SEM and web analytics. The tool helps analyze your own data and it creates an industry landscape to keep a close watch on competitors. The area covers a wide range of metrics such as audience engagement and other activities. It includes features that monitoring all competitors updates .We can customize the report. This tool has alerts to help us to keep up to date with popular posts.

Unmetric : It tracks rival pages and compare these to an industry benchmark .It provides a breakdown of types of stories like event, content, advert, employee stories, etc posted to account or competitor's accounts. A strong point of this tool is categorization related to different campaigns and level of post. Features are available to provide searching keyword.

User SonamJha

Profile Pic

Page Social Network Analysis And Mining

Date Of Report 08/19/2019

POST TYPES	
Like	1K
Photos	21
Videos	13
Links	52
Posts	91

Sociograph.io – Report Page

ANALYZE FAN PAGE ON SOCIOGRAPH

Using sociograph.io/dashboard/page for showing information about pages and groups such as https://sociograph.io/page/1328261797324978 select page and display related data. Fig 1 has shown about graph related page's information

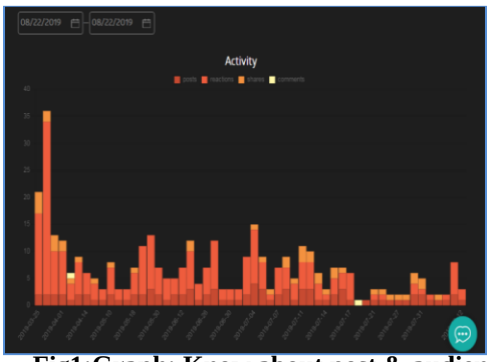

Fig1:Graph: Know about post & audience

V. CONCLUSION

After review systematic literature , we find out appropriate terms like facebook social graph- graph API reference , Fb sdk, graph explorer tools & techniques for collecting data, audience interest, mining opinion, track & prepare data, tag suggestions etc.Some of my concern is, just whether a social site can be given such privy to capture and use of privacy to any extent where information can change opinion, views, candidacy and future in broader extent. All these statements might be open to interpretation but yes it is indeed a strong domain of the new sphere. The topic has been nicely touched with the glimpse of its usage both for the general usage & its revenue model. Use of tools and its ever evolving platforms designed by face book is continuously taking attention as its approach is pragmaticand useful for mining and analyzing data and beneficial for researchers.

REFERENCES

1. Matthew A. Russell. Mining the Social Web, 2nd Edition.chapter2: MiningFacebook: Analyzing FanPages, Examining Friendships and More.2014: August; ISBN 13:978-93-5110-311-0.

2. https://www.simplilearn.com/how-Facebook-is-using-big-data-article
3. Kmieckowiak T. Facebook Analytics – How to Analyze Your Facebook Performance.2018; Jan 25.
4. https://techcrunch.com/2019/03/06/mark-zuckerberg-discovers-privacy/
5. https://bigdata-made simple.com/12-common-problems-in-data-mining/
6. Newton C.Facebook's pivot to privacy has huge implications — if it's real Mark Zuckerberg is fond of grand pronouncements.2019; March 6 .
7. https://en.wikipedia.org/wiki/Facebook
8. https://developers.Facebook.com/tools/explorer/
9. https://developers.Facebook.com/apps
10. https://developers.Facebook.com/docs/graph-api/reference/.
11. https://developers.Facebook.com/policy/
12. Stieglitz S^1, Mirbabaie^2, Ross B^3, Neuberger C^4.Social media analytics-challenges in topic discovery, data collection and data preparation.2018;April:156-168.
13. https://www.pythonforbeginners.com/requests/using-requests-in-python
14. https://www.researchgate.net/post/what_are_the_most_ effective_tools_to_mine_Facebook_data
15. Said A. Salloum1, 2 *, Mostafa Al-Emran3, Azza Abdel Monem4, Khaled Shaalan1. A Survey of Text Mining in Social Media: Facebook and Twitter Perspectives
16. https://www.nytimes.com/2018/04/11/technology/Face book-privacy-hearings.html/What You Don't Know About How Facebook Uses Your Data/By Natasha Singer.
17. Ranjan R.How to use Facebook Graph API and extract data using Python .2016; Oct 3.
18. https://github.com/xbwei/Data-Mining-on-Social-Media.
19. https://www.studytonight.com/network-programming-in-python/mining-Facebook-data
20. Batrinca B, Philip C, Treleaven.Social media analytics.A Survey of techniques, tools and platforms.2015; Feb: 89-116.
21. https://developers.Facebook.com/apps/3243708250725 28/add/
22. https://www.Facebook.com/analytics/32437082507252 8/?__aref_src=devsite&__aref_id=app_dashboard
23. Jeh,G.Widom and J : Simrank: a measure of structural – context similarity In:Proc. Of the 8^{th} SIGKDD international conference on knowledge discovery and data mining , pp.538-543. ACM(2002)
24. Mathioudakis,M.Koudas,N.: Efficient identification of starters & followers in social media. In : Proc. of the International Conference on Extending database Technology , pp.708-719.ACM(2009)
25. Madan M,Chopra M.: Social Media Networks :An eye to envision and extract information.2014

AUTHORS PROFILE

Sonam is Research Scholar in the department of Computer Applications, Veer Bahadur Singh Purvanchal University, Jaunpur. She has received her M.C.A. degree in computer applications as well as MSc(IT) from Sam Higginbottom Institute of Agriculture , Technology & Sciences. Her research area is Big Data On Social Network. She has published some review articles in international journals and conferences. Her interests include data mining, python, R programming languages and educational technology. She has more than five years work experience in the field of computer science such as Java programming, PhP, Java Script and .Net Platform.

Surjeet Kumar is an Assistant Professor in the department of Computer Applications, Veer Bahadur Singh Purvanchal University, Jaunpur. He has received his M.Sc. degree in computer science from Marathwada University, Aurangabad, Maharastra and Ph.D. from Sri Venkateshwara University, Gajraula Uttar Pradesh, India. His major research interests are data mining, algorithm, network security and knowledge discovery. He has published more than twenty seven research articles in international reputed journals and conferences. He has authored a book on Internet and Java Programming. He has efficient work experience in the field of research.