International Journal of Engineering and Advanced Technology (IJEAT)
ISSN: 2249-8958 (Online), Volume-8 Issue-3, February 2019

Analyzing Geosocial Media for Decision Making

G. Vishnu Murthy, K. Priya Darshini, G.Balakrishna

Abstract: ***Web-based social networking has radically meddled in another methodology by enabling individuals to distribute their information alongside their areas to convey advantages to the network and the nation overall. The use of geo-social networks is growing significantly, and these networks are generating huge amounts of data. As a result, current geosocial media can be used in real-time as databases for countries, governments and other organizations. By breaking down the social conduct of a network in a specific zone, individuals can suggest a shop, lodgings, shabby markets, managing an account frameworks, promotions, and so forth dependent on their inclinations and impediments. Breaking down such a lot of information and settling on continuous choices is a testing errand. In this way, an effective framework for gathering information is proposed to settle on ongoing choices while distinguishing different occasions. In this paper we analyse the foursquare data which is a geosocial networking site and then detect the most populated venue or a place by analysing the tips given by users. The analysis is done on the New York city check-is and tips. The decision making is done based on the results to predict the most popular place or venue in a given location based upon the tips given by users.***

Keywords: Geosocial Network; Sentiment Analysis; VGI

I. INTRODUCTION

To react progressively, associations must channel, connect and process vast volumes of quick moving information in numerous arrangements. Occasion based choice administration frameworks utilize leaders, not simply basic leadership administrations, to react to rapid information continuously. They combine pattern recognition with business decisions and recognize that business decisions need to be shared between event-oriented and process-oriented solutions. Social media are driving their role dramatically day by day as they change from long range informal communication (i.e. social networks) to geo social systems.

This has prompted expanded utilization of geo social systems, offering clients the capacity to voice feelings, report occasions and trade perspectives, outrage or love while drawing in with others unbelievable in the pre-web period was. The data shared over all media is geosocial on the grounds that:

1) The posts have rich substance that speaks to geographic data with specific areas that are either entered expressly (at registration) or included verifiably (by geographic directions, for example, scope or longitude), and
2) Views shared via web-based networking media uncover social information and reinforce connections and correspondence.

With the advancements in the mobile networks and the process of inventing new mobile devices made the people attracted to the location based services very rapidly. The geosocial networking sites have the location based service as distinguish feature. Foursquare which is one of the popular location based networking site enables it's registered users to explore the places nearby. Innovative advances have empowered the utilization of GPS frameworks in cell phones, making area information overwhelming. Where individuals post, remark or transfer to web based life is recorded. By amassing such sorts of area information from all system clients, interpersonal organizations hence make geo social information stores. Another technique for creating geo-social information is publicly supporting, while self-created applications are accommodated an assortment of purposes or causes. It recovers geo-social information from volunteers or paid clients who give information or data to this reason. This sort of online information assembled through publicly supporting is named "Volunteer Geographic Information" (VGI) [2].Today, numerous stages and programming have been produced to use publicly supporting for geo social information collecting purposes to expand business, advance causes, or adventure other business purposes. A case of such a product stage called Ushahidi, which enables reports to be created by coordinating a particular catchphrase in a geo-interpersonal organization that compares to particular areas. The report would then be able to be utilized for instruction and help with crises or catastrophes [3]. There are additionally different applications with constrained usefulness, for example, Hootsuite[4] and 140kit[5].

Social networking data information could be advantageous to numerous areas if all are well analyzed. By investigating the social conduct of a network in a specific zone through filtering and customization, individuals can suggest a store, lodgings, shabby markets, managing an account frameworks, promotions, and so forth dependent on their inclinations and impediments. Thus, in view of traveler and vehicle developments, open specialists can make better city maps and prescribe appropriate transport courses to individuals based on current conditions.

Manuscript published on 28 February 2019.
***** Correspondence Author (s)
Dr. G. Vishnu Murthy*, Professor and HOD, Dep. Of CSE Anurag Group Of Institutions Hyderabad, Telangana, India
K. Priya Darshini, Post Graduate Student, Dep. Of CSE Anurag Group Of Institutions Hyderabad, Telangana, India
G.Balakrishna, Assistant Professor and HOD, Dep. Of CSE Anurag Group Of Institutions Hyderabad, Telangana, India

What's more, the investigation of interpersonal organization information, for example, Twitter is utilized in numerous medicinal applications to screen and battle lethal illnesses and diseases. There are a couple of various papers that have been done with Geosocial Networks that usage simply limited geographic information. Be that as it may, every current framework don't consider the handling parts of geosocial ongoing movement at fast. There are couple of other rapid enormous information preparing frameworks that expect to fabricate a social insurance framework, process M2M information [26] utilizing an information combination show, set up Smart City, and so forth. On the other hand, these structures are not sensible for taking care of relational association data geosocial nature.

Geosocial media

Geosocial systems are a kind of interpersonal organization that utilizes geographic administrations and highlights, for example, geocoding and geo tagging to encourage extra social elements. Through client submitted area information or geolocation strategies, interpersonal organizations can interface clients to neighbourhood individuals or arrange occasions with them as per their interests. Finding online informal communication administrations can be IP-based or utilize hotspot trilateration. In portable interpersonal organizations, area data or the area of cell phones with area based administrations can improve informal organizations. Interpersonal organizations enable clients to connect with respect to their present areas. Web mapping administrations with geocode information for spots (roads, structures, and stops) can be utilized with geo-labelled data (Meetups, Concerts, Nightclubs, or Restaurant Reviews) to coordinate clients with a place, occasion, or neighbourhood gathering to contacts make or interface gathering of clients to settle on the gathering movement.

Foursquare has pulled in as the most recent fever in long range informal communication. Some state this wellknown application may even be the following Twitter or Facebook. Regardless of whether you are visiting a neighborhood bar or looking at the most recent eatery, the foursquare application causes you associate with companions or find new activities in the place where you grew up. The application utilizes the implicit GPS of the cell phone to show eateries, bars, parks, and different attractions in your city. When you visit any of those areas, you "check in" on the Four Square application, which communicates your area to your companions. We can likewise observe where our companions have checked in, which causes you get together with them or find new activities. Description of foursquare check-in and tips is as follow:

Venue Check-in: Foursquare allows the registered users to make their presence or a visit recorded at a venue. It is done by making an active selection through the website or a mobile application. The check-in recorded made by the users can be public or private. Users have a choice to display their check-in to their friend's foursquare sites and can even post their check-in on Twitter and Facebook accounts. Each check-in made will be rewarded points or coupons.

Venue Tipping: Foursquare users can add "Tips" which are generally the reviews, to venues which are available to other users to read. These tips serve as suggestions to the users to do, see, or eat at the location. After you've checked in, you can write reviews for the location, which will be available to other foursquare users. These tips include everything about a venue can be a positive or negative review. Analysis of these tips will help to know the most popular venues. This can be done by analysing the tips given by users and then group the mostly tipped venues. Analysis of these tips is very much useful o both the business (i.e. the venue owner) and the users. A venue owner will be able to know the user interest and their opinions about the venue. If a venue has more no of negative tips the venue owner can take the necessary steps to make his venue popular. This analysis also helps a user to know about a particular venue whether it is worthy to visit or not.

II. LITERATURE SURVEY

At present, analysts are increasingly intrigued by geo social systems since they think they are new information resources as proposed in the paper [6, 7]. Crooks et al. [8] utilized Twitter information in their work to delineate in the United States. Also, Chow in [9] and Papadimitriou in [10] proposed designs that utilized geo-label suggestions while shaping new informal community. Geo Life 2.0 [11] performed comparability identification, circle investigation.

Table 1: comparative study of literature survey

Done by	Alexis Papadimitriou	Brian Eriksson	Laura Ferrari
Knowledge base	Data collection and recommending users	Location analysis	Pattern analysis
Approach	Higher Order Singular Value Decomposition (HOSVD) technique.	Learning based IP geolocation approach	Topic model based approach
Objectives	To get recommendations on friends ,locations and activities	To pinpoint the geographic location of the IP accurately	To automatically extract urban patterns and recurrent behaviour from location based network
Advantages	New users, locations, or activities can be easily inserted due to the use of tensor algorithm	Need to study machine learning problem is reduced Geolocation estimates were very accurate	Accurately analyses urban patterns
Algorithm/protocol	Tensor algorithm	Naives Bayes estimation algorithm	Latent Dirirchlet Allocation

Future	Conducting user study on recommendations	Improving geolocation estimates accuracy	Applying the approach on other kind of data and combining it with other techniques for better results
People	Required	Not required	Minimal requirement

Table 2: comparative study of literature survey

Done by	Bobadilla .J	Yaronkanza	Yanhua Li∗ , Moritz Steiner$
Knowledge base	Recommender systems	Geospatial data analysis	Popularity detection
Approach	Content based filtering and model based	intuitive approach	Categorising venues and then predicting popularity
Objectives	To explain the evolution of recommender systems and provide an original classification.	To analyse geo-tagged posts to discover places that were jointly visited by many people	To detect the popular categrie of venues and the predicting popular venue individually
Advantages	Describes the Recommender system trends to implicitly collect data	Improves and facilitate the work of urban planners and of policy makers,	Can know the most popular places in the city and it's descendings
Algorithm/protocol	Bayesians ,genetic algorithms	Clustering ,algorithms, agglomerative clustering	-
Future	Improve the quality of recommender systems predictions and recommendations.	To compare with other parameters	To improve the prediction results to be more accurate
People	Minimal requirement	Required	Minimal

There are couple of different frameworks that perform examination of the tweets, their unique circumstance, e.g. B. [12] - a few employments. Some work has likewise been done to identify hotspots and hyper neighbourhood occasions utilizing tweeter information [3, 14] in a city. Social networking organization information could be useful to numerous territories if very much examined. By separating the social direct of a system [15] in an explicit region through filtering and profile planning, one can endorse a store, lodgings, unobtrusive markets, dealing with a record structures, advancements, etc subject to their tendencies and imperatives [16]. So also, in view of traveller and vehicle developments, open experts can complete better city maps [17-19] and prescribe street transport to appropriate street clients dependent on current conditions [20, 21]. Furthermore, the investigation of interpersonal organization information, for example, Twitter is utilized in numerous social insurance applications [22-24] to screen and battle savage illnesses and diseases. There are a couple of various papers that have been done with Geo social Networks that usage simply obliged geographic information. Be that as it may, every single existing framework don't consider the handling parts of continuous fast social transport. There are couple of other fast information handling frameworks for vast volumes of information that expect to manufacture a wellbeing framework [25], process M2M information utilizing an information combination display [26], a built up city [27], and so forth.

III. EXISTING SYSTEM

Geo social Network data is not only beneficial to governments , it can also affect human life from ordinary citizens to business people. Most of the existing system ideas used the check-in parameters to detect the popular place or a venue. There are limited studies based upon the micro review (tips) about the venue. Former studies dealt only with the check-in counts. They did not consider the tips given by user about a venue. Even if the tips were considered the prediction was like the more the tips the more the popularity despite analysing the user view about the venue.

IV. PROPOSED SYSTEM

The proposed work is done taking the following statements into consideration that is "a venue or a place being every now and again checked-in indicates that a venue is popular in a sense that numerous people visit it and are likely to share their visit" and "a venue being very frequently reviewed or tipped indicates that the venue is popular in a sense that really a person is interested in the venue and would like to share their views".

Here in the proposed work both the tips and venue check-in related data are analysed to find the most popular venues. Includes the in depth analysis of the tip. The micro reviews or the tips are well analysed to detect if the review was given positively or negatively. Also the ironic tips are detected. Once the sentiment analysis is done venues are given ranking based on the positive count .This analysis is useful for both the users and the business owner (i.e. the venue owner) for further decision making. The proposed work is broadly categorised into three categories:

A. DATA COLLECTION

Data Collection is the foremost thing to be done. Foursquare search application programming interface (API) will return a list of venues in a region. The bound or the region is specified by the latitudes and longitudes of the region bounding box. Both time and space constraints are laid by the Foursquare API i.e., it returns only 500 queries per hour from an authenticated account, and a space constraint, i.e., only 50 venues returned for every requested query.

Published By:
Blue Eyes Intelligence Engineering and Sciences Publication (BEIESP)

To ensure that we get all the venues in a given region, we keep the bounding region small enough so that the number of venues in it does not exceed the return limit. The dataset on which analysis performed is about the New York City restaurant data. This data set includes the venue details, the check-ins made, the user's details and the tips given by users. The data set consist of about 10,130 tips generated by different users on the venues in the city.

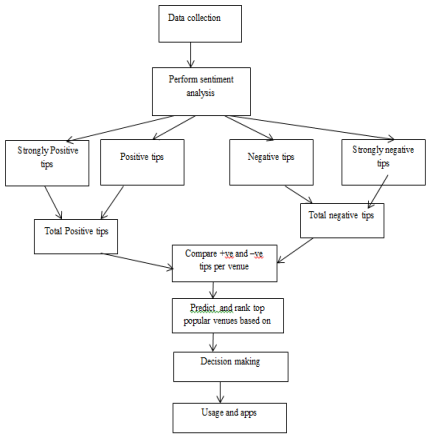

Fig : Work flow of the proposed implementation

B. VENUE POPULARITY PREDICTION

This phase includes performing sentiment analysis on the tips then categorising the mostly tipped venues. The results includes displaying the mostly tipped five venues with their sentiment reports. The one with more positive polarity tips is the first most popular venue. And also the check-ins are analysed to get the top popular venues. Unfortunately, most of the existing work, such as [13, 18] have not demonstrated an unbiased and comprehensive picture of tips in Foursquare. Therefore, in this paper, we choose tips in Foursquare for a thorough study.

Sentiment analysis on tips:

For analysing sentiments of the tips a Sentiment analysis API provided by meaning cloud is used. Meaning cloud which is an excel add-in is a free add-in which has to be downloaded and installed. For analysis purpose the data must be imported into the excel file and then select the meaning cloud add-in. This API can be even included in the java script to get sentiment analysis done. Select the sentiment analysis option and then give the range of the tips to be analysed. The response would be the polarity and some attributes. The polarity of the sentiment can be one of the five categories. They are as follows: Positive (P), Negative (N), Neutral (NEU), Strongly Positive (P+) and Strongly Negative (N+). This add-in is very much useful to the people who are not so aware of programming or a naive user to programming, since this add-in does not require any coding.

Popularity Prediction:

This phase includes the process of detecting the most popular venue based on the tips. We cannot say that the venue tipped most is the popular venue because the tips may be positive or negative. For example let us consider two venues A and B. Venue A is mostly tipped and Venue B is medium tipped with more positive tips. Here we can't infer that A is most popular because even though , B is a bit low tipped venue than the other it has more positive tips. So, the venue B is more popular venue in terms of its service. For this first the mostly tipped venues are to be identified then sentiment analysis is performed on the corresponding mostly tipped venues. The popularity of venue is detected by reviewing the sentiments reports , the venue with most positive tips is the most popular venue in the city. The venues are then ranked based on the results.

Table 3: Top five tipped venues

Venue id	Total tips	P	N	NEU
82569	234	162	14	10
72	144	16	85	20
196154	101	74	11	3
33560	72	37	7	9
25807	56	32	14	4

Table 4: Top five popular venues ranking

Venue id	Total tips	P	N	NEU	Ranking
82569	234	162	14	10	1
196154	101	74	11	3	2
33560	72	37	7	9	3
25807	56	32	14	4	4
72	144	16	85	20	5

C. DECISION MAKING AND USAGE

Decision making and usage here can be made by the different users in different perspectives. As a person interested in visiting a location or a venue in and around the city decision making can be done by viewing the results produced that is we can get an overview of the popular venues. The results produced can also be used by the venues owner's to get an overview of how well the venue is popular and then decision making will be done so as to improve his business operations.

V. RESULTS:

This phase includes all the results of the experiment done on the dataset. the results given are polarity of the tips given by user about the venue, the top five mostly tipped venues , the top five popular venues in terms of positive tips ant the top five checked-in venue.

55238	5639	Three words: Pork Belly Taco (sounds dirty? It is! And delicious!)	P
124076	5639	Get the Ricky- Margarita and Dos Equis. Killer combo for a killer night.	N+
54561	5639	Very loud. Perfect if you're in the mood for unlimited drinks (included in the $22 Brunch Special)	P+
2221	5648	The chips were great with blue chz- extra chips was $3 more. Salmon fri special was delish. Bourbon dessert was light and perfect even on full stomach. Dirty martini kettel1 $14 and was like 3 shots.	P
53783	5648	Dear Bobby, enough with the tv crap and it's about time you attain to your restaurant. Service blows ad food is worse... The bread is the best thing about this hell hole. 'nuff said	N+
61001	5648	The New York strip is outstanding! The creamed corn (when available) and mashed potatoes are also super delish.	P+

Fig: sentiment analysis of the tips.

The collected tips are very deeply analysed to find the sentiment of the tip given by the user regarding the venue. Then all the tips are grouped to their venues ID's respectively.

Published By:
Blue Eyes Intelligence Engineering and Sciences Publication (BEIESP)

That means all the tips of venue id x are grouped together for further analysis. Then the positive tips are gathered .The graph shown below depicts the number of total tips per venue and the corresponding categories of tips (i.e. positive tips, negative tips and the neutral tips). The x-axis in the graph represents the venue id's and the y-axis represents the number of tips.

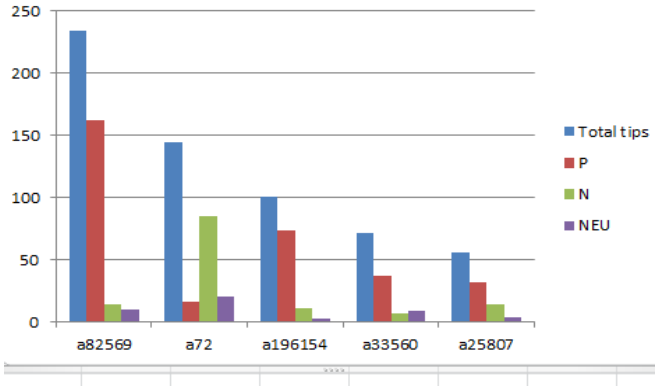

Fig : Top five most tipped venues.

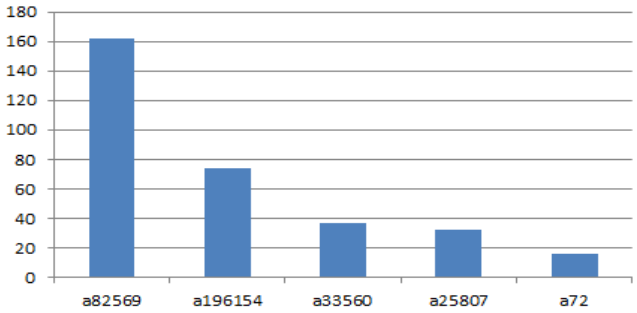

Fig: Top five popular venues based on positive tips.

Even the check-ins data of the venues in the New York city are analysed and the top five popular venues are predicted based on the check-in's data. The graph below depicts these results:

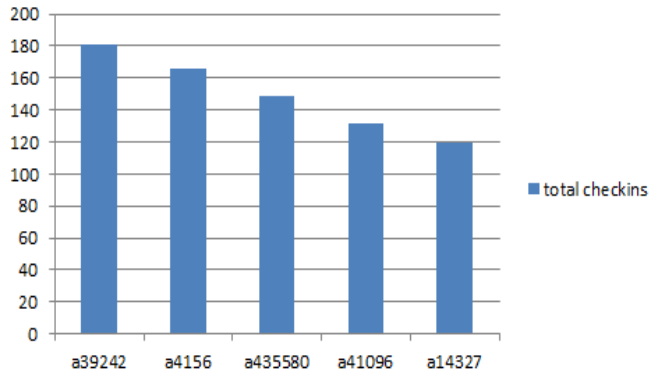

Fig : Top five popular venues based on check-in's

VI. CONCLUSION

Geo social networks can be beneficial for governments and other organisations as they provide facilities and security against different situations through proper management. Such networks may be of general benefit to citizens by providing recommender systems, traffic safety, health care, etc., what's more, business people to present new items in different territories by observing the geo-social information of a specific zone. However, these advantages must be accomplished with better investigation that utilizes a lot of information from various geo-interpersonal organizations. This is conceivable with trend setting innovation and better examination, and in addition a framework with high computational abilities. In our work we analysed the foursquare data which is a geo social networking site. The data collected is about the restaurants tips in New York City. For this data we have performed sentiment analysis on the tips and then predicted the top five popular venues in the city. Thus decision making can be done accordingly based on the results.

VII. FUTURE ENHANCEMENT

Geo social systems can be valuable for governments as they give offices and security against debacles through legitimate administration and lessening the dread of the spread of disease. Similarly, these networks may be of general benefits to the citizens by providing recommender systems, traffic safety, health care, etc., what's more, business people to present new items in different territories by checking the geo-social information of a specific zone. The data can be well analysed by making the image and content analysis more efficient. Future work will be focused on detecting if any spam tips are present and then developing a hybrid system to analyse different data from different sources.

REFERENCES

1. HERN, "Online Volunteer Map Philippine after Typhoon Haiyan",The Guardian, November 15,2013.
2. M. Haklay: "How extraordinary is resolved geographic information? Relative examination of Open Street Map and Ordnance Survey datasets," Condition and Planning .B, orchestrating and structure, vol. 37, no. 4, p. 682. 2010.
3. M. Zook et.al. "Volunteer Information and Crowdsourcing Disaster Relief: A Case Study on the Haitian Earthquake." "World Medical and Health Policy," Volume 2, No.2 ,Pp.7-33,2010
4. Hoot suite, "hootsuite.com".
5. 140kit, "https://github.com/WebEcologyProject/140kit."
6. Stefanidis, et.al "Choosing Environmental Information From Social Media Feeds," Geo Journal, pp. 1-20,2012
7. Z. Cheng, et.al, "Exploring Millions of Suspicious Imprints in Site-Sharing Services.", ICWSM, vol. 2011, pp. 81-88, 2011.
8. Crooks, A. Croitoru et.al. "# Earthquake: Twitter as Distributed Sensor System", Transactions in GIS, 17 (1), p.124-147, 2012.
9. Chow, J. Bao et.al "Towards Site-Based Social Network Services," second ACM SIGSPATIAL International Workshop on Location Based Social Networks,pp.31-38,ACM,2010
10. Papadimitriou, P.Symeonidis et.al "Geo-Social Recommendations", ACM Recommender Systems 2011.Workshop on Personalization in Mobile Applications,2011.
11. Y. Zheng, Y. Chen et.al. "GeoLife 2.0: A Location-Based Social Network Service," IEEE Tenth International Conference on Mobile Data Management: Systems, Services, and Middleware, 2009 (MDM '09), Pp. 357-358, 2009.
12. O'Connor, M. Krieger et.al "Tweet motiv: Explorative Search and Thematic Summary for Twitter," in ICWSM, 2010.
13. L. Ferrari, A. Rosi et.al. "Removing urban precedents from region based relational associations", In Proc. of the third ACMLBSN, 2011.
14. Xia, R. Schwartz et.al. A. Langdon, J. Ting and M. Naaman "City Beat: Real-time Social Media Visualization of Hyper locality Data" In Proc. the WWW Conference 2014.

Published By:
Blue Eyes Intelligence Engineering and Sciences Publication (BEIESP)

15. Paul, Anand et al. “Smart Buddy: Defining Human Behavior Using Big Data Analytics on the Social Internet of Things.” IEEE Wireless Communications 23.5 (2016): 68-74.
16. J. Bobadilla, F. Ortega et.al. “Recommender systems Survey”, Knowledge-Based Systems, 46: 109-132, 2013.
17. RA Becker, R. Caceres et.al. “A Town History: Use of Mobile Network Data for Urban Planning”, IEEE Pervasive Computing, 10 (4): 18-26, 2011.
18. M. Vasconcelos et al. Tips, Done’s and To-Dos: Uncovering User Profiles in Four Square. In Proc. of ACM WSDM, 2012.
19. Ratti, S. Williams et.al. “Compact Landscapes: Use of Location Data from Mobile Phones for Urban Analysis”, Environment and Planning B Planning and Design, 33 (5): 727, 2006.
20. Hefez et.al. “Tarsius: A Traffic-Sensitive Route Search System Under Uncertainty Conditions,” SIGSPATIAL'11, pp. 517-520, Chicago, 2011
21. R. Levin and Y. Kanza, “Tars: Traffic-discerning course look for”, Geo informatics, 18 (3): 461-500, 2014.
22. R. Lan, M.D. Adelfio et.al. “Spatio-common affliction following using news articles”, HealthGIS'14, pages 31-38, Dallas, TX, 2014.
23. R. Lan et.al. “The Image of Health: Card-based Collaborative Temporal disease tracing”, Health GIS'12, pages 27-35, Redondo Beach, CA, 2012.
24. L.A. Waller and C.A. Gotway, “Associated Spatial Statistics for public health data,” Book: John Wiley and Sons, Hoboken, New Jersey, Vol. 368. 2004.
25. Rathore, M. Mazhar et.al. “Nonstop medical Emergency Response System: Use of IoT and big data for public Health.”
26. Ahmad, Awais et.al. “ A gainful multi-dimensional gigantic data blend approach in machine-to-machine.”
27. Rathore, et al. “Use mixed and tunneled blended media acquires a fast gigantic data condition.” Multimedia Tools and Applications (2017): 1-26.