Repositório ISCTE-IUL

Data extraction and preparation to perform a sentiment analysis using open source tools

The example of a Facebook fashion brand page

António Teixeira

Instituto Universitário de Lisboa (ISCTE-IUL), Lisboa, Portugal
Av. das Forças Armadas, Lisboa
1649-026 Lisboa, Portugal
toze_teixeira@sapo.pt

Raul M. S. Laureano

Instituto Universitário de Lisboa (ISCTE-IUL), Lisboa, Portugal
BRU-IUL and ISTAR-IUL, Lisboa, Portugal
Av. das Forças Armadas, Lisboa
1649-026 Lisboa, Portugal
raul.laureano@iscte.pt

Abstract - **Social Media is a subject that is being very discussed in the present. The increasing availability of internet and the growth of Social Media platforms drove organizations attention to Opinion Mining and Sentiment Analysis. One of the most popular Social Media Platform is Facebook. In this Social Media Platform users/consumers can express their feelings with comments or with emotion buttons. With the functionalities of Facebook, they can criticize, praise, suggest or expect. The user's interactions with a brand page as posts, likes, shares or comments are getting more relevance in present days. The analysed data can give to decision-makers new approaches to run their business, understand their brand value against the competitors or even understand better what are the evaluations of customers or potential customers about the brand products or services. The purpose of this paper is to explain how to extract and prepare data collected in Facebook using open-source tools to perform a Sentiment Analysis.**

Keywords – Sentiment analysis, Social media, Facebook, Data extraction, Open source.

I. INTRODUCTION

Social Media is the new trend, a trend that combines with marketing and Information Technology (IT). This phenomenon is not recent, but with the constant advances in IT, people communication with these tools is increasing. The amount of data produced and shared in the Social Media Platforms increased massively and the way to collect this type of data changed in the present days. Sharing online content is an integral part of modern life, this is somewhat related with social transmissions which are driven in part by the increase of information likely to be shared and the willingness to share. This statement suggests why contents that evokes more emotions are shared more often. [1]

This new way of living the Web gave people another way to disseminate what are their believes, opinions or even thoughts in Social Media Platforms. This opening to their feelings and believes can influence political, economic and social behaviours. An opinion is simply a positive or negative sentiment, view, attitude, emotion, or appraisal about an entity or an aspect of the entity [2]. This massive opinion sharing, led that the decision makers had to be more cautious and worried to analyse all these opinions, especially with the more negatives that can harm the image of brands or products. Recent literature points out that the term Sentiment Analysis was perhaps first used by Nasukawa and Yi [3]. Sentiment Analysis, also called Opinion Mining, is the process of collecting opinions, emotions or even attitudes from a text analysis [4]. The purpose is to understand and determine a communication through the contextual polarity in a text or reaction in a Social Media platform [5].

In the early stage of Sentiment Analysis, the methods to analyse were totally focused on paper or analysing what we call word to mouth. With the increase of social networks, the media starts evolving the public, and the public opinion started to gain expression.

At the end of last century, Porter said that the "change in the way office functions can be performed, is one of the most important types of technological trends occurring today for many firms, though few are devoting substantial resources to it."[[6]. In a study about Social Media usage as a business activity "86% of marketers said that Social Media was important to their businesses" still only 26% of companies can measure social activity (ROI) [7]. The results of these type of analysis (as Sentiment Analysis) can be very useful for decision makers and be used in different departments of the organization [8].

Retail uses Sentiment Analysis to gain a meaningful understanding of the feelings of customers and different areas of retail can take advantage of it. For instance [9]:

- Customer Service: Sentiment Analysis can give useful information about the preferences of current and new consumers. Knowing opinions, likes, dislikes the company can improve strategies by creating or improving methods to approach to their consumers in better way to create a positive sentiment and erase the negative ones from them. As is totally online, is totally accessible 24/7 we can be aware always about what is happening and what are being said about the brand, and so the brand can participate at the moment in the conversation to soften the sentiment if it is negative.

- Revitalize Brand: The feelings about the brand can be quantified through Sentiment Analysis. We can quantify the sentiments with the set of information we gather from the own brand, products and services provided, i.e.

we can track all spikes in the comments. This analysis can be used for better marketing strategies or even for the responsible of the products verify the acceptance or not about them.

- About Competition: Why not to understand what people feels about your competitors? This analysis can be useful to measure and understand what consumers feels about the competition. Even can try predict trends and develop strategies to gain advantage over them, or even benchmark the brand process against the competitor.

- Gain Business Intelligence: The sets of information gathered via Sentiment Analysis can be provided and added to Business Intelligence. Using this information in a correct manner, can be precious to organizations and can help to measure impact decisions to gain advantages and leverage the business. Being in real-time, can determine if the identification of the feeling is positive, negative and act at the moment. Even the stakeholders can have reports or notifications if or why the sentiments or feelings are positive or negative.

Facebook is one of the best places where we can find manifestation of feelings. Nowadays any event raises several reactions and comments. Facebook is also the most popular Social Media Platform and have more advantages over the other platforms, such as Twitter or Instagram. Another good point, is related with the character limitation (which do not have) and is worldwide used. In this way, we will have more content to analyse and it will be more accurate. Facebook has already almost 2 billion users active [10] and the communication between consumers and organization can be made instantly.

Informing consumers about the company or products we can make posts with text, numbers, pictures and videos. These pieces of information can be liked, have reactions, shared or commented. Globally we can easily create communications to inform consumers, as the posts we share at the company page can be totally public. If a person follows the company page, will get information about all the brand page activity. From a corporate point of view, Facebook should be used to create stakeholder communities and therefore it will be easier to manage "the structural and flow properties in a relationship" [11]. Table I shows the terminology used in Facebook for user activities [12].

TABLE I. FACEBOOK TERMINOLOGY

User	Activity	Facebook Term
1	Status update	Post, Status; Status post
2	Agree with status update of user 1	Like; Love, Haha; Wow
2	Disagree with status update of user 1	Sad; Angry
2	Make a comment	Comment
2	Reply a comment	Reply
2	Share User 1 status update	Share

Social Media are changing the way how consumers search products. In a recent past, magazines were viewed and read; and now they are being substituted by blogs or even for Social Media Platform of each brand. Today, the products are being promoted in the Social Media Platforms and the received feedback its being given by comments, shares or even likes. What would be an expert opinion before, now it is given by thousands of interactions in a post made in a platform. The elitist industry of fashion was changed by Social Media and now is a consumer-friendly industry.

One study was made using Facebook 611 posts published by the page fans for qualitative analysis. Researchers concluded: "Facebook brand pages support the Social Media marketing opportunities and goals for building brand awareness, gathering insights and knowledge for future steps, community involvement and engaging in open and honest dialog, as presented in the related work section. Marketing practitioners could use the topic-category frequency of occurrence as a measure for successful Social Media marketing utilization over time." [13]. Moreover, the main topics that fans wrote in the brand page are: product (52%), sales (13%), brand (8%) and competitor (4%). Regarding the category of the posts, that the fans commented in the brand page, the main are: suggestions & requests (28%), affect expression (28%), sharing (27%), information inquiry (16%), complaints & criticism (4%), gratitude (4%) and praise (1%). Also, 4% of the comments include a competitor reference [13].

So, the collection and analysis of the Facebook data can help organizations to have a better understanding about products, customers, competitors and any aspect of the environment needed to support executives and managers in making strategic decisions for an organization. This is called Competitive intelligence [13]. But to gain this Competitive Intelligence we need to understand the business, define the goals and understand if the data we collect are useful or if the data collected need to be prepared, to be useful when delivered to decision makers [15].

Demonstrating that open source tools can help small organizations with small IT and Marketing teams is one of our goals. Contrary to what one might think, using open source tools can help small teams and drive them to achieve the objective of collect, prepare and present useful data. It is critical to realize that this work is critical otherwise it could be delivered useless data. To illustrate, *Parfois* was chosen because it is a recognized and international brand, with almost 800 stores worldwide in almost 60 countries. Having the same product, same posts in different country's such as Portugal, France, Belgium, Hungary, Qatar, Panama, Georgia, Romania, Italy, Saudi Arabia.

II. SENTIMENT ANALYSIS: THE PROCESS

Most of the products to make Sentiment Analysis use Natural Language Processing techniques as well as a number of complex rules-based algorithms, in order to have an exhaustive and detailed set of information about the feelings that had been expressed in words in Social Media Platforms. The sentiments can be classified as: positive, negative or neutral. To get these sentiments, the process uses text analytics, linguistic and accepted language processing so it can verify and analyse only the subjective information.

The whole process is complex and comprises five steps [16]:

1. Data collection: In this step, we collect the data from Social Media Platforms. The data is in an unstructured form.

2. Data preparation: Here we clean all non-textual contents from the data we collected.

3. Sentiment detection: Then, we examine the opinions and thoughts collected. Only the opinions, beliefs, views are maintained.

4. Sentiment classification: In this step, we classify the sentiments gathered in the data collected as positive, negative, neutral.

5. Presentation of output: Here the main goal is converting the information in "readable way".

As already mentioned, the focus of this paper is to demonstrate how to make the first two steps, using only open-source tools.

A. Data Collection

The first step for this process involved a research how to extract data from Facebook. We find Facebook Graph API. This API is a low-level HTTP API that allows us to perform various tasks, being one of them the extraction of data from Facebook platform. This API also provides all core functionalities of Facebook application using Rest web services. For the goal of this paper, we used the Feed connection of the Page object. The Feed represents the list of all Post objects, which contains details like post date, likes, comments, reactions, products advertised, and shares.

Still, there are some concerns with API like the total number that is called. When the API is called a big number of consecutive time, gets block. For example, if we want get the total number of posts of a page, it will be block at same time. So, the workaround its extract the data with time intervals.

Along with the API we used also a Java client called RestFB. This API is a very minimal and simply way to fetch information from Facebook and even publish into the application new items. With this client, the access to the core functions of the Facebook API is simplified and even we can use it, to automatize the extraction of the data. For the automation, we started a Java project where we started to develop an application using RestFB API that in turn uses Facebook Graph API. Using RestFB, anyone can make a custom implementation, as we made one to adapt for our research purpose.

To know what we did, we will insert here the principal lines of code that we used to extract data from Facebook.

In this line, is the API Call using the version 2.6 of Facebook API:

FacebookClient facebookClient= new DefaultFacebookClient(Constants.MY_ACCESS_TOKEN, Version.VERSION_2_6);

Calling the feed of the page we want, in this case, *Parfois* Facebook page (official page of *Parfois*) with date range parameters. As we already wrote, due the limits of number of calls of the Facebook API we separated the data collection by years. So, by example if we want all posts from year 2016 we manage that configurations like this:

*SimpleDateFormat dateFormatRange = **new** SimpleDateFormat("dd/MM/yyyy");*
Date *initDate* = dateFormatRange .parse("01/01/2016");
Date endDate = dateFormatRange.parse("01/01/2017");
Connection<Post> feed = facebookClient.fetchConnection("Parfois/feed", Post.class, Parameter.with("since", initDate), Parameter.with("until", initDate),Parameter.with("limit", 1));

After the post's fetch, we start iterating the Feed to collect all posts from the range of date we parametrized before:

for (List<Post> feedConnectionPage : feed) {
for (Post post : feedConnectionPage) {
Number of shares of the post
numberShares = getShares(post.getId(),facebookClient);
Product advertised in the post
product = searchProduct(post.getMessage());
Date of creatiin of the post
postDate = getCorrectDate(post.getCreatedTime());

In the next line of code, we are fetching the reactions of the post:

reactionsConn = facebookClient.fetchConnection(post.getId()+"/reactions", Reactions.ReactionItem.class);

It was created a function that returns a list with the emotions of the post and as input it receive the fetch result:

emotionsList = getEmotions(reactionsConn);

Detailing the function, we can see that we are iterating twice, so we can get all information about the reactions of the post:

for (List<ReactionItem> reactionsPost : reactionsConn) {
for (ReactionItem reactionItem : reactionsPost) {
reaction = **new** PRFReactions();
reaction.setReactionId(reactionItem.getId());
reaction.setName(reactionItem.getName());
reaction.setType(reactionItem.getType());
reactionsList.add(reaction);
countReactions(reactionItem.getType());
} }

Now it is the time to fetch the comments made in the post. Like before first we fetch the comments of the post and we call a function to retrieve all comments information's and store it:

Connection<Comment> comments = facebookClient.fetchConnection(post.getId()+"/comments", Comment.class, Parameter.with("fields","from,message,id"));
commentsList = getComments(comments);

And now time to store all the post information collected:

```
numberPosts++;
postInformation = new PostInformation(post.getId(),
post.getMessage(),numberShares,
numLikes, numWow, numLove, numHaha,
numAngry, numSad, commentsList, emotionsList,
product, postDate, Constants.PAGE);
facebookPosts.put(numberPosts, postInformation);
}}
```

All this information is stored in a structure defined by us, in this case:

HashMap<Integer,PostInformation> facebookPosts

We took the option to save all the data into the relational data base so we can access information via query or extract it to a file. We made a function where the input is the structure we defined and then we insert into the data base the information collected:

```
facebookPosts.forEach( (k,v) -> setDataToInsert(v));
facebookPosts.forEach( (k,v) -> setCommentsToInsert(v));
facebookPosts.forEach( (k,v) -> setReactionsToInsert(v));
```

As it could be seen, the posts are being made with a structure, so we had to "personalize" the code of the extraction.

B. Data Preparation

To get the first set of results from the automation process of extract the data, we had to adapt the code to the brand posts.

The first set of data extracted had the timeline from 1ˢᵗ of January of 2017 and 15ᵗʰ February of 2017. This set contained 46 posts, 217 comments and 15994 reactions. In the output, we got in our findings, we had some issues. The next figures were in the comments.
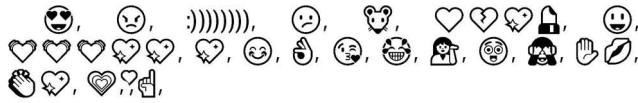
About these kinds of figures, we must have two concerns: or we clean them from the comments, or we use them too for Sentiment Analysis. In literature, it refers as emoticons and gives some examples (Table II) [17]. Emoticons are facial expressions pictorially represented using punctuation and letters; they express the user's mood. In the present we have emojis too. Emojis are used much like emoticons and exist in various genres, including facial expressions, common objects, places and types of weather, and animals [17].

TABLE II. PART OF THE DICITIONARY OF EMOTICONS AND EMOGIS

Emoticon	Polarity
:-) :) :o) :] :3 :c)	Positive
😎, ♡, ♡	Extremely-Positive
:D C:	Extremely-Positive
:-(:(:c :[	Negative
D8 D; D= DX v.v	Extremely-Negative
☹, ♡	Extremely-Negative
:|	Neutral

Others authors refer to emoticons in their studies. One study concluded that the simplest to detect the way polarity (i.e., positive and negative affect) of a message is based on the emoticons it contains. Emoticons have become popular in recent years, and are now included in all write messages in the social networks. Emoticons are primarily face-based and represent happy or sad feelings, although a wide range of non-facial variations exist: for instance, <3 represents a heart and expresses love or affection, :(is a sad face, :) is a happy face [19].

For our purposes, we can filter the emojis or substitute them with the correspondent emotion. We need to decide which one is the best solution to accomplish our goals. There are some API's we can use to run this tasks and that we can insert in our own API. For gain usage of comments made/having emoticons we must apply other parameters in text tools, when mining comments. Even for irony detection. Is called a Noisy-text Analysis **Erro! A origem da referência não foi encontrada.**.

Another issue we had is the different languages used in the comments. We had Spanish, Italian, Portuguese, French, English and Arabic. One of the options will be the attempt of filtering the comments by language and insert into the separated databases. With this process, we could have cleaner data when running the text mining algorithm.

Following our path in the open source, we will use Google Translation API to detect the language used in the comment, as we show in the next lines:

```
Translate translate = createTranslateService();
List<Detection> detections =
translate.detect(ImmutableList.of(textToCheck));
 for (Detection detection : detections) {
    //TODO: Check how many different languages
 }
```

After this we can separate the comments by language and then insert them in his own comments table. As future work, we must create different tables for comments in each language.

At last, other issue would be cleaning the emoticons and separate the real comments from the ones which are only mentions of friends in comments. For future work, we are investigating if using Stanford Named Entity Recognizer (NER) will help us to achieve our goal that is discarding or just put apart the comments that are mentions to friends.

III. THE RELATIONAL DATABASE

The data collection is being made with a customized tool, that uses Facebook graph API, and a Java API who invokes Facebook API. The data is being stored at a relational data base, and its being analysed if the actual structure of the data base is correct and if is adequate to our purposes.

All the data extracted is stored in a relational database in the next three tables:

- Facebook_Posts: Contains the general data about the post.

- Facebook_Comments: Contains all data about the comments.

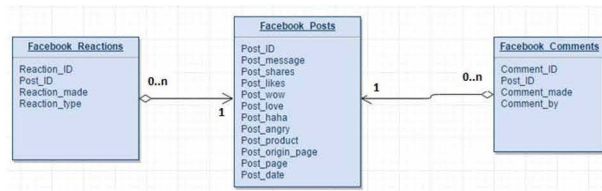- Facebook_Reactions: Includes data collected about the reactions.

Figure 1: Relational database

The general data we store in the relational data base is:
i) Post identification (ID): The unique identifier of the post;
ii) Post message: The message that was in the post;
iii) Number of shares: Number of times the post was shared;
iv) Number of total reactions types: Total number of reactions, and total number of reactions type;
v) Post product: Product announced in the post;
vi) Post page: Page where the post was made;
vii) Post date: Date the post was made;
viii) Person: Name of the person who made the reaction or the comment.
ix) Comments: Comments made in the post.

This structure was the structure we choose to store the data extracted. This data will be separated to better understanding in this first phase. At the moment, we iterate each post, we automatically generate two more sets of data, with the comments and other with the reactions:

emotionsList = getEmotions(reactionsConn);
commentsList = getComments(comments);

As we had separated the comments and the reactions we can make a first approach to Sentiment Analysis because we have the reaction of the posts in one separated table.

IV. FUTURE WORK

A. Data collection and preparation

At the present day, we are investigating if a NoSQL database will be needed. We are searching for the best/appropriate open source tools so we can make a full open source process. Regarding this topic, we need to ensure the tools are reliable and give us the best and more accurate results possible. The cleaning/discarding useless comments is another topic being discussed at the present so we can have a better and correct Sentiment Analysis.

B. Measures

We need to be able to measure and filter the results of Sentiment Analysis, so we must look at "quality metrics". In quality metrics, we include opinions, feelings, satisfaction ratings, comments, replies as well as the quality of engagement over time. Table III shows some examples of measures we can use in our result set and some of the analyses can be pertinent for business and decision makers [20].

TABLE III. OBJECTIVES, REQUIREMENTS & OUTCOMES IN SOCIAL MEDIA ANALYTICS

Common Use-Cases for Social Media Analytics	Required Business Insight	Enabling Social Media Analytics Techniques	Pertinent Social Media Performance Metrics
Social Media Audience Segmentation	Which segments to target for acquisition, growth or retention? Who are the advocates and influences for brand or product?	Social Network Analysis	Active Advocates Advocate Influence
Social Media Information Discovery	What are the new or emerging business relevant topics or themes? Are new communities of influence emerging?	Natural Language Processing Complex Event Processing	Topic Trends Sentiment Ratio
Social Media Exposure & Impact	What are the brand perceptions among constituents? How does brand compare against competitors? Which Social Media channels are being used for discussion?	Social Network Analysis Natural Language Processing	Conversation Reach Velocity Share of Voice Audience Engagement
Social Media Behavior Inferences	What is the relationship among business relevant topics and issues? What are the causes for expressed intent (buy, churn etc.)?	Natural Language Processing Clustering Data Mining	Interests or Preferences (Theme) Correlations Topic Affinity Matrices

Key Performance Indicators (KPI) must be easily identified by all organization. The efforts in the early stage of definition and elaboration of the measures can help all organization to have a strong Sentiment Analysis process [19]. KPIs should be set that relate to data returned on campaigns or products posts, campaign sentiment, brand sentiment, mentions, individually and in combination with other sources as for example de e-commerce website or loyalty programs.

C. Sentiment Analysis tools

Table IV presents some of the software's that are used for Sentiment Analysis, describing the main functionalities and characteristics of each one.

TABLE IV. FUNCTIONALITIES OF SENTIMENT ANALYSIS OPEN SOURCE TOOLS

Tool	Functionalities and characteristics
Knime	Data integration from files, data bases; Data Transformation; Data Mining; Social Network Analysis; Image processing; Text Processing and Text Mining.
Hive (Cloudera Virtual Machine CDH)	Ad-hoc queries support; MapReduce usage for data processing; Text mining; Document indexation; MapReduce algorithm usage; Predictive models;
Pentaho Community Edition	Data integration; Connection with Hadoop NoSQL and analytical data bases; MapReduce usage to reduce development cycles; Preparation and modulation for non-structured data; Data integration with parallel processing
R	Comprehensive statistical analysis; Is a programming language; Leading tool for statistics and machine learning; Integrate with other languages; Interact with many data sources; One of the most used tool in scientific environment.

V. CONCLUSION

The motivation of this research was creating a personalized tool to extract and prepare the data from Facebook of a fashion brand page. After this process, the ultimate motivation is use a Sentiment Analysis tool to get the sentiments from Facebook comments towards the brand. Another objective is to collect the reactions created by some posts about some products and with the help of other teams understand why products with high expectation of being a best seller were not. With marketing team is to understand if people who ask information about prices, buy the product and compare with competitors.

This small data already collected and prepared gave us the visibility of what is a minimum data set result. Now with a small set of results we need to insure to have a bigger set, so we can perform more accurately the opinion mining. Due the way people express themselves, many posts are discarded. So, having a huge number of posts and consequently comments and reactions will give us a better and accurate result.

Gave us too the chance to observe how can be difficult to make Sentiment Analysis in an international brand page, where we have a lot of comments in different language. We need to research and test open source tools so we can find and choose the best tool that serves our purpose.

After understanding what information, we want to collect and personalize the extraction software it is very easy to use it. So, the advice we give is to first understand the business to personalize and collect the correct data. After the collecting process, for small IT department and low resources can use any open source data base as MongoDB, MySQL, PostgreSQL or extract the data directly to a file. Combining all this tools can be very helpful for small organizations to start giving the first steps in the world of Sentiment Analysis. We show that it is not necessary high developed IT skills to extract data from Social Media Platforms. However, it is very important to define the aims of the analysis, i.e., identify the answers for the questions we are looking for. Only with a successful business understanding, it is possible to succeed in a Sentiment Analysis and thus contribute to the brand competitive advantages.

Sentiment Analysis can help gain insight into so much about how consumers feel about a brand, about brands campaigns, products and services. There are so many tools that can help IT departments or marketing departments. But as always it depends on business, time, and the most important in majority of company's; budget and resources. As is important to think about the right steps in order take actions, thinking about how we can improve, personalize or maintain our own tools for Sentiment Analysis can be very important too.

ACKNOWLEDGMENT

REFERENCES

[1] J. Berger, "Arousal Increases Social Transmission of Information", Psychological Science, vol. 22, no. 7, pp. 891–893. 2011.

[2] B. Liu, Web Data Mining: Exploring Hyperlinks, Contents, and Usage Data, Springer, 2006.

[3] T. Nasukawa, and J. Yi, Sentiment Analysis: Capturing Favorability using Natural Language Processing. In Proc. of the 2nd international conference on Knowledge capture, October 23–25, pp. 70–77, 2003.

[4] B. Liu, Sentiment Analysis and Subjectivity. Handbook of Natural Language Processing 2, p. 627-666, 2010.

[5] E. Younis, "Sentiment Analysis and Text Mining for Social Media Microblogs using Open Source Tools: An Empirical Study", In Proc. Of International Journal of Computer Applications. vol. 112, no. 5, pp. 44-48. 2015.

[6] M. Porter, "Technology and Compertitive Advantage", Journal of Business Strategy, vol. 5, no. 3, pp. 60 – 78. 1985. doi: 10.1108/eb039075

[7] M. Stelzner,.Social Media Marketing Industry Report. How Marketers use Social Media to Grow. Social Media Examiner, 2014.

[8] P. Melville, V. Sindhwani, and R. Lawrence. Social Media Analytics: Channeling the Power of the Blogosphere for Marketing Insight. In *Proc.of the WIN*. 2009.

[9] W. He, J. Shen, X. Tian, Y. Li, V. Akula, G. Yan, and R. Tao, "Gaining Competitive Intelligence from Social Media Data", Industrial Management & Data Systems, vol. 115, no. 9, pp. 1622-1636, 2015. Doi:10.1108/IMDS-03-2015-0098.

[10] J. Edwards, 'Facebook Inc.' Actually Has 2.2 Billion Users Now — Roughly One Third of the Entire Population of Earth, Business Insider. 2014 (retrieved from http://www.businessinsider.com/facebook-inc-has-22-billion-users-2014-7).

[11] J. Kiezmann, K. Hermkens, I. McCarthy, and B. Silvestre, "Social Media? Get Serious! Understanding the Functional Building Blocks of Social Media", Business Horizons , vol. 54, no. 3, pp. 241–251, 2011. doi: 10.1016/j.bushor.2011.01.005.

[12] S. White. Facebook Terminology. Local Blitz Marketing. 2016 (retrieved from https://localblitzmarketing.com/facebook-terminology/).

[13] I. P. Cvijikj, and F. Michahelles "Understanding Social Media Marketing:A Case Study on Topics, Categories and Sentiment on a Facebook Brand Page", In Proc. MindTrek'11, September 28-30, 2011. pag. 177-182. doi:10.1145/2181037.2181066.

[14] L. Dey, S. Haque, A. Khurdiya, and G. Shroff, "Acquiring Competitive Intelligence from Social Media", In Proc. of the 2011 Joint Workshop on Multilingual OCR and Analytics for Noisy Unstructured Text Data, September, 2011.

[15] E. Gallinucci, M. Golfarelli, and S. Rizzi, "Advanced Topic Modeling for Social Business Intelligence", Information Systems, vol.53, pp. 87-106, 2015. doi: 10.1016/j.is.2015.04.005.

[16] M. Godsay, "The Process of Sentiment Analysis: A Study", International Journal of Computer Applications, vol. 126, no.7, pp. 26-30, 2015.

[17] A. Agarwal, B. Xie, I. Vovsha, O.Rambow, and R. Passonneau, "Sentiment Analysis of Twitter Data in Workshop on Language in Social Media", In Proc. of LSM 2011 - Workshop on Languages in Social Media, 23 June, pp. 30-38, 2011.

[18] https://en.wikipedia.org/wiki/Emoji

[19] Dongi Saothawan, Consumer Social Media Analytics in retail - Prerequisites, business value and application for Social Media Marketing,September 2014, d0i: 10.13140/2.1.3174.2406

[20] U. Ruhi, "Social Media Analytics as a Business Intelligence Practice: Current Landscape & Future Prospects", Journal of Internet Social Networking & Virtual Communities, vol. 2014, pp. 1-12, 2014. doi: 10.5171/2014.920553.

[21] J. Lovett, Social Media Metrics Secrets. Indianapolis: Wiley Publishing, Inc., 20