

Customer Intelligence Analytics on Social Networks

Brano MARKIĆ*, Sanja BIJAKŠIĆ and Arnela BEVANDA

University of Mostar, Bosnia and Herzegovina

Discovering needs, habits and consumer behavior is the primary task of marketing analytics. It is necessary to integrate marketing and analytical skills with IT skills. Such knowledge integration allows access to data (structured and unstructured), their analysis and finding out information about the opinions, attitudes, needs and behavior of customers. In the paper, the hypothesis is set that software tools can collect data (messages) from social networks, analyze the content of messages and get to know the attitudes of customers about a product, service, tourist destination, with the ultimate goal of improving customer relations. Experimental results are based on the analysis of the content of social network Facebook by using the package and function 'R language'. This language showed a satisfactory application and development power in analysis of textual data on social networks for marketing analytics.

Keywords: consumer behavior, text mining, R language, word cloud, unstructured data, consumer analytics, consumer intelligence.

JEL Classification: M31

1. Introduction

Collecting data about customers, their profiles, opinions and attitudes, as well as development and implementation of various methods of data analysis are the study subject of marketing business function, qualitative and quantitative research methods, especially methods of data and text mining. A marketing analyst usually collects (or has access to) huge amounts of data that has to be analyzed. The qualitative marketing analysis endeavor, based on textual data, to learn about the behavior of customers and service users, their satisfaction or complaints on price, product quality, service, tourist destination or other object of analysis (Grigsby M., 2015; Chapman and McDonnell, 2015). Social networks are objective, unbiased sources of information for marketing analysis because they reflect the actual, real and genuine customer reviews. Therefore, the social network for marketing analysis is an indispensable data source.

Facebook is one of the most used social network in Bosnia and Herzegovina. Therefore, access to potential customers by informing them about promotions, prices, changing assortment is "simplest" via Facebook. Marketing analytics comprises of business processes and information technology (IT) that enable marketers to measure performance of marketing programs. Marketing analytics include a few subsets among

* Corresponding Author:

Prof. Brano Markić, PhD, Faculty of Economics, University of Mostar, Bosnia and Herzegovina

Article History:

Received 15 July 2016 | Accepted 27 July 2016 | Available Online 3 August 2016

Cite Reference:

Markić, B., Bijakšić, S. and Bevanda, A., 2016. Customer Intelligence Analytics on Social Networks. *Expert Journal of Marketing*, 4(1), pp. 24-30

them and the most important ones are: behavioral analytics, business intelligence, customer intelligence analytics, predictive analytic, social analytics, real time analytics, sale analytics (Hemann and Burbary, 2015). Customer intelligence analytics is based on analyzing data available about customers, their behavior and monitoring whether or not their customers actually like them. It uses data on social networks as customer feedback to find how the company is perceived and what can be done to improve this perception.

2. The Content Analysis of Social Networks

The Facebook social network is an accessible online communication service used for business purposes. It allows you to connect with users. So, for example, any company representative can easily create invitations for an event and send them to friends or potential customers. Using these practices, companies can save money, because these campaigns cost almost nothing and are made very quickly. Simply, invitations sent via social networks save any company money and time.

Marketing and promotion are particularly valuable for social networks and vice versa. Measuring the success of promotional activities and marketing decisions is carried out by analyzing the content on social networks. For the promotion of products or tourist destinations on the social network Facebook user can use one, two or all three facilities: personal profile, group and fans page.

The idea of Facebook social network is based on a personal profile and communication among friends but the profile cannot be used for commercial purposes. For the promotion activities customers can access the group. In the group are linked profiles with some common characteristics so they can interchange information on familiar topics (e.g. from a same professional area as is marketing analytics).

The third facility designed for business purposes is a fan page. In it the company can promote a product, service, tourist destination. Friends (fans) read messages, get informed about products, events, services, prices. Commitment is to shape the page so it is interesting to the users and they want to follow it and become engaged. It becomes a channel for a two-way communication and a social network channel for the promotion of ideas, products, and services. The company may have a fan page and analyze how many potential buyers monitor and analyze the page but also what potential buyers want. Companies that constantly innovate existing products or create new products and services need to communicate with customers, but the customers also have a need to monitor their news, ask them questions and seek answers. Figure 1 exemplifies a brand page on Facebook:



Figure 1. Brand Page on Facebook social network

Facebook becomes a channel of communication with the market. Companies that often, on a daily basis, inform customers, publish sweepstakes, undertaken action, price changes, provide new products or services have at disposal a social network as an indispensable channel of communication with consumers and it's environment. The number of fans becomes the main measure of the relevance of the page as well as the product or service. The obligation of a brand's management is to recognize and exploit the potential of information technology.

On social networks there are huge amounts of data that are often do not even analyzed. There are many reasons for it but the two most significant. The first is the lack of time and Insufficient level of training and knowledge management. These reasons may be called subjective. Other reasons can be attributed to complex of the data analysis on social networks in terms of building the appropriate software tools and recognition of algorithms that enable analysis.

2.1. Data Mining Methods for Content Analysis of Social Networks

Data prepared for analytical purposes today are most commonly found stored in a data warehouse. Namely, formatted data bases contain large quantities of detailed data that cannot be directly used for analytical purposes, therefore they are aggregated into dimensional model of data warehouse. These data can be reached and analyzed by various data mining algorithms in order to extract regularities or laws hidden in a dimensional model. However, marketing nowadays, cannot be satisfied with such formatted data in the database or data warehouse. The best example are social networks as virtual space for exchange of opinions, feelings, requirements, motives, ideas, views about companies, products, services, events, and destinations. The data on social networks are in a text form, and its layout is not previously determined (it is not formatted as a database or data warehouse). Business requires real-time data and information (Jeffery, 2010).

The user (customer) forms a website, a blog and becomes a part of social networks: (Twitter, Facebook, YouTube, Instagram, LinkedIn and many others) and by using these forms of web technologies, one 'leaves' useful data for marketing analysts about events, personal impressions on products, services, quality or product defects, their needs and desires.

In the process of text or data mining, data is first prepared in a form (dataset) that can then be analyzed using some algorithms. The steps in the process of data preparing and the implementation of text mining algorithm are illustrated Figure 2.

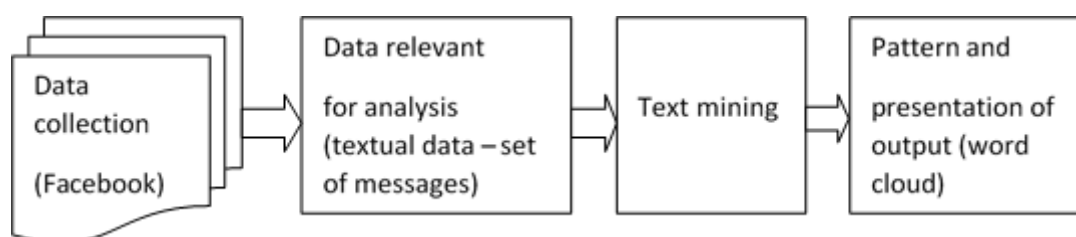


Figure 2. Steps in customer intelligence analytics on social networks

In Figure 2, there is a conceptualization of the overall concept of content analysis of social networks into five stages and begins with collecting data on social networks and ends with the presentation of data (usually the visual form).

After collecting the data, it follows their cleaning, prior to analysis by a data mining algorithm. Preparation of the text is the process of eliminating stop words (e.g. - ; , : ?) or words that are not relevant to the analysis.

In analyzing the data, only the messages that contain the object of analysis will be retained in the dataset. The purpose is to translate the documents into simpler form so that they are suitable for parsing.

Text mining algorithms retrieve data and perform their processing. This can be, for example, a classification into particular groups, where Naive Bayes algorithm or maximum entropy is applicable. The research results can be presented in the form of a word cloud. It is a method of data visualization when high frequency words are written in the biggest font and reflect the most common terms that appear in the text.

The programming language R contains appropriate functions to find Facebook page and analyze responses and reactions of users of social networks. For the marketing analytics are interesting all the social networks and blogs, where potential customers leave their opinions, views or express satisfaction products, events, services, tourist destinations.

3. Research Methodology and Experimental Results

Each research dynamics is based on data. Data collection, their selection and display in an appropriate format acceptable to analyze is a complex and demanding process in terms of time. In this paper we focus on the details of Medjugorje tourist destination on the social network Facebook.

3.1. Data Collection

Access to data, their collection, selection and display in a form acceptable for marketing analytics provides R language, its packages and functions. A precondition for data collection from Facebook is to access URL address: <https://developers.facebook.com>, registration and filling required information on Facebook developer page (form). This is followed by the installation packages devtools and Rfacebook from github:

```
>install.packages("devtools")
>library(devtools)
>install_github("Rfacebook", "pablobarbera", subdir="Rfacebook").
```

After installing these packages we need to connect our R session with our test application and authenticate it to our Facebook Profile for Data Mining. The package Rfacebook offers a very easy function for that. All we need is to copy *app id* and *app secret* from our *app settings* on the Facebook developer page:

```
>require("Rfacebook")
> fb_oauth <- fbOAuth(app_id=" app id from our app settings ", app_secret =" app secret oour app settings ", extended_permissions = TRUE).
```

Sequential diagram in the following figure shows the registration steps of brand analysis application on Facebook and launching R session.

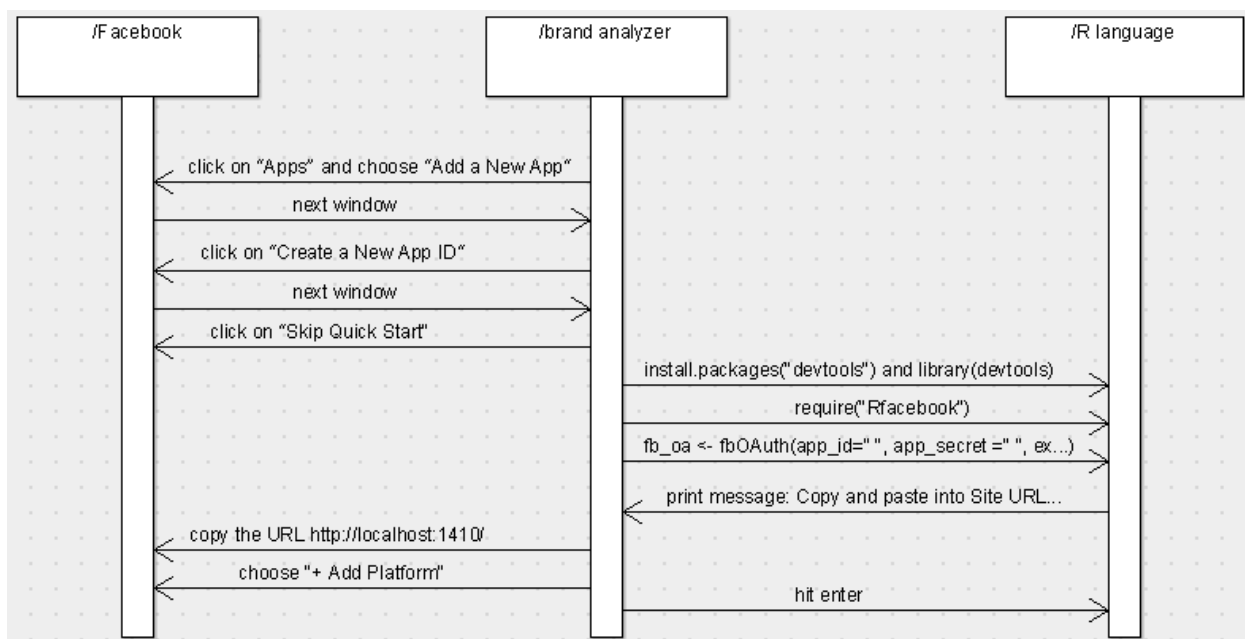


Figure 3. Sequence diagram of Facebook connecting and starting the R session

A very useful step is saving of our variable *fb_oauth* as `save(fb_oauth, file="fb_oauth")`.

The variable *fb_oauth* is available to use in an easy manner next time. After accessing and connecting to the server of Facebook social network, the collection of necessary data follows according. Our analysis focuses on Medjugorje as one of the most important tourist destinations in Bosnia and Herzegovina.

3.2. Data Analysis

Customer intelligence analytics tends to be part of the overall culture based on the data and directs to the data recorded at the social networks. Our marketing analytics concentrates in learning how the Facebook page “*Medjugorje*” has become popular, and what the most frequent words in messages of Medjugorje are. The collection of data starts with the R function *getPage()*:

```
>pageMedjugorje <- getPage("Medjugorje", token= b_oauth, n = 2000)
```

This function returns a data frame with information about all its posts. We were requested 2000 posts and the API returns 1702 of them:

25 posts 50 posts 75 posts 100 posts.....1670 posts 1695 posts 1702 posts >. The first four posts are:

```
> pageMedjugorje[1:,:]
  from_id  from_name  message
1 169173713621 Medjugorje affidiamoci a Maria, preghiera da dire con fede..
  created_time  type  link
1 2016-04-23T08:38:42+0000 link 1 http://www.amicidilazzaro.it/index.php/atto-di-affidamento-
a-maria-santissima-giovanni-paolo-ii
  Id  likes_count  comments_count  shares_count
1 169173713621_10154121964663622 92 3 8
```

The total number of likes, comments and shares for the post Medjugorje is:

```
> colSums(Filter(is.numeric, pageMedjugorje[,]) )
  likes_count  comments_count  shares_count
  400798 12025 57217
```

There is a huge amount of likes 400798 what automatically indicates that the Medjugorje is nice and appropriate tourist destination.

Our focus are messages and presentation of the most frequent words in the form of word cloud. For illustration will be selected only four messages from pageMedjugorje[] data frame:

```
> pageMedjugorje[1:2,3]
[1] "affidiamoci a Maria, preghiera da dire con fede.."
[2] "alcune ricerche scientifiche sulle apparizioni di Medjugorje"
```

Now, we first select all messages from data frame *pageMedjugorje[]*, create corpus, load necessary packages (twitterR; tm; Gentry, 2013), transform *fb_Corpus* to text and prepare the text by eliminating stop words (e.g. - ; , : ?) or words that are not relevant to the analysis:

```
> fb_wcloud=pageMedjugorje[,3]
> fb_Corpus<-Corpus(VectorSource(fb_wcloud))
>library(twitterR) >library(tm)
>fbCorpus<- tm_map(fb_Corpus, PlainTextDocument) >fbCorpus <- tm_map(fb_Corpus,
stripWhitespace)
>fbCorpus <- tm_map(fb_Corpus, removePunctuation) >library(wordcloud)
The next step is to create a term matrix that contains frequencies of terms for learning. In our example,
the number of terms is 11706 and the number of documents 1702.
>fb_tdm<- TermDocumentMatrix(fbCorpus, control = list(minWordLength = 5))
>fb_tdm
<TermDocumentMatrix (terms: 11706, documents: 1702)>
```

The terms (attributes) are selected based on the frequency in the documents (Bijakšić, Markić and Bevanda, 2013). Attributes that exceed a certain threshold will consist a list of index of terms. To view the results of processing, and these will form the terms that appear most frequently in messages of Facebook page of Medjugorje as a destination, it is necessary to load the package *wordcloud()* in a R session. It follows forming the documents term matrix and sorting the words in a descending order. The rows of the matrix dtm are terms and the columns are the documents (messages):

```
>dtm<- TermDocumentMatrix(fbCorpus)
>m_dtm<- as.matrix(dtm) >m.S<- sort(rowSums(m_dtm),decreasing=TRUE)
>m<- data.frame(word = names(m.S),freq=m.S)
```

In the end, only two commands of R language are enough to display word cloud terms in color and these terms reflect the views, perceptions and opinions about Medjugorje:

- Facebook Developers, 2016. *Facebook Developers* [online] Available at: <https://developers.facebook.com> [Accessed on March 3, 2016]
- Gentry J., 2013. *Twitter client for R*, [online] Available at: <http://cran.r-project.org/web/packages/twitterR/vignettes/twitterR.pdf> [Accessed on March 3, 2016]
- Grigsby M., 2015. *Marketing Analytics. A practical guide to real marketing science*. London: Kogan Page.
- Hemann C. and Burbary K., 2015. *Digital Marketing Analytics: Making sense of Consumer Data in Digital World*. Indianapolis: Que Biz-Tech series.
- Jeffery M., 2010. *Data-driven Marketing. The 15 Metrics Everyone in Marketing Should Know*. Chicago: Kellogg School of Management.
- Rfacebook', 2016. *Package Rfacebook'* [online] Available at: <https://cran.r-project.org/web/packages/Rfacebook/Rfacebook.pdf> [Accessed on March 3, 2016]

