Methodology for sentiment analysis in twitter posts about mobile learning

DOI: 10.46932/sfjdv2n2-125

Received in: March 1st, 2021
Accepted in: May 30th, 2021

Hernán Gil-Ramírez
Ph.D. in Educational Processes in Virtual Spaces (University of Salamanca, Spain)
Technological University of Pereira
Carrera 27 #10-02, Pereira (Risaralda, Colombia)
E-mail: hegil@utp.edu.co

Rosa María Guilleumas-García
Ph.D. in Educational Processes in Virtual Spaces. (University of Salamanca, Spain)
Technological University of Pereira
Carrera 27 #10-02, Pereira (Risaralda, Colombia)
E-mail: roguiga@utp.edu.co

ABSTRACT
Analysis of social networks has become of great interest to researchers from different areas, including educators, due to Twitter’s growing importance as a space for discussion and dissemination of knowledge and opinions. This reality demands the development of analysis processes that allow to know the topics of interest in the network, the positive or negative feelings in relation to those topics and who the network influencers are. Those objectives guided this research work and in order to achieve them, we developed a methodological proposal for sentiment analysis of tweets. This article describes the process followed, which involved 1) detecting the structure of the communication network, 2) calculating the general metrics, 3) representing the communication network, 4) identifying and analyzing the clusters, 5) calculating their metrics as well as those of the individual nodes and 6) establishing the polarity of the posts published in the network. This paper also describes the methodology followed to identify trends and topics of interest in the hashtags and web domains included in the tweets. The proposal for analysis presented here is intended to help researchers interested in the field of social networks, to understand the complex interactions that take place in these environments and the way in which information is disseminated, valued and converted into topics of interest thanks to the network users’ actions.

Keywords: Social Network Analysis, Sentiment Analysis, Influence in Social Networks, Twitter.

1 INTRODUCTION
Advances in mobile communications and broadband create a space to develop proposals for mobile learning (M-learning). Although the studies carried out by specialized educational organizations can offer some insights about how to implement such proposals, their publications are infrequent. Statistical studies can also help find responses, but they reach the public with a long delay. A third option

---

1 This paper was produced within the research project titled Sentiment Analysis of Twitter posts to assess the convenience of implementing mobile learning in the classroom (Code 4-19-6). The research was financed by the Technological University of Pereira (Colombia) through the Vice-chancellor of Research, Innovation and Extension.
that allows overcoming the limitations just mentioned are research studies supported on the monitoring of opinions in Twitter to assess users’ posts about mobile learning. Identification of the emotions (positive or negative) generated over time by this educational proposal and detection of predominant topics of interest in the tweets and their promoters may also enlighten educators on the suitability of implementing educational mobile learning proposals.

Our research was focused on learning with mobile technology and we sought to address the analysis of the trends in the topics of interest, the polarity, and the potentially most influential users in the publications about mobile learning on Twitter, by using a combination of Social Network Analysis, Text Mining and Sentiment Analysis. We used R language to perform the different calculations, and Gephi for the graphical representation of the networks.

The methodological proposal we present in this article was developed for the analysis of communication networks about education on Twitter but may well be used in fields such as politics, health, natural disasters, social and migratory movements, or the finance.

2 DATA IMPORT AND SYSTEMATIZING

To import the data, we created a grid with the following 18 tweet fields: created_at, status_id, screen_name, text, source, reply_to_screen_name, is_retweet, favorite_count, retweet_count, hashtags, urls_expanded_url, mentions_screen_name, lang, retweet_count, retweet_screen_name, description, followers_count, favourites_count.

We used Twitter’s API Rest to import a total of 27,668 tweets in English. The tweets, imported on a weekly basis during the period going from September 28 to December 28, 2019, included the terms “mlearning OR m-learning OR mobile + learning”. We employed R Language to import the data and Excel to store them.

3 DATA ANALYSIS

In this section, we will present the procedures that allowed us to tackle the analysis of the data pertinent to each of the objectives proposed for the research.

3.1 DETECTING THE STRUCTURE OF THE COMMUNICATION NETWORK

In order to detect the structure of the communication network that was created by the flow of tweets with the terms “mlearning OR m-learning OR mobile + learning”, we used social network analysis techniques (Social Network Analysis). The structure emerged from the connections between the users who generated or forwarded the tweets (source nodes) and those mentioned in them or receiving responses
(target nodes). This structure was dependent on the topics addressed and the users leading the conversation (Smith, Rainie, Shniderman and Himelboin, 2014). The process of analysis included three levels: general, cluster, and individual.

### 3.1.1 General Level

At the general level, we identified the general metrics and created the communication network graph.

* Diameter, density, reciprocity, order and size were some of the metrics calculated at this level.*

* Diameter is the longest geodetic distance between the actors connected in the network and allows to understand the dimensions of the network. The density of the network is given by the proportion of existing connections in relation to all possible connections and is therefore calculated by dividing the number of actual connections by those that could potentially exist (Ognyanova, 2016).*

* Density is expressed as a percentage. In case that the value obtained is 100%, it must be understood that the number of connections present is equal to the number of possible connections. Such would be a perfectly dense network: all potential connections have come to exist. In directed networks such as Twitter, every node has the possibility to establish a connection with another, but the happening of such an event does not automatically imply the existence of a relationship between the two connected nodes, an acceptance of the relationship or a response from the target node. On this characteristic lies the interest of reciprocity. In a directed graph, this metric indicates the proportion of pairs of nodes (A, B) that are in a reciprocated relationship measured against all possible ones.*

* Reciprocity, a numerical constant between zero and one, reveals the level of interconnection within directed networks (Ognyanova, 2016).*

* Order and size of the network (Barbera, 2017) allow to identify the composition of the network (G). The order of the network is given by the set of users (nodes) that constitute it (V); and the size represents the set of existing relationships between the nodes (E). Thus, it is possible to express a network as: G = (V, E).*

The graph of a communication network offers a snapshot of the imported data and allows the observation and intuitive evaluation of the network metrics.

In this graph, users were represented by a dot. Since Twitter is a directed network, the relationship between each pair of users (source and target) was represented by an arrow showing the direction of the connection (from source to target). In case of reciprocity between a pair of users, a bidirectional arrow was employed, i.e., an arrow with heads at both ends.
To facilitate the display of both the graph and the existing relationships between users and clusters, graphic representation of social networks should take initially a general approach (i.e., present the unprocessed graph); at a second stage, a filter can be employed to eliminate the less connected nodes and reveal the clusters (Hawksey, 2011).

3.1.2 Cluster Level

Once analyzed the network from a general approach, we deepened the analysis to get a better understanding of the way information circulated within the network. The number of relationships established among users within the network allowed to identify several clusters. To analyze them, the following steps were followed: Identification of clusters, calculation of cluster metrics and graphic representation of the clusters.

The existing clusters in the communication network were identified using the Newman-Girvan method (Despalatovic, Vojkovic and Vukicevic, 2014), which allowed to place in independent groups those users more connected to each other than to other users of the network. "Community structure detection algorithms try to find dense subgraphs in directed or undirected graphs, by optimizing some criteria, and usually using heuristics." (The igraph core team, 2020, p.1).

The analysis at this level included the study of the following metrics: the number of clusters in the network, as well as the number of users and relationships per group; Transitivity (also called the clustering coefficient), a measure of the probability that the nodes adjacent to a node were connected (to it); and Modularity, “High modularity for a partitioning reflects dense connections within communities and sparse connections across communities” (Ognyanova, 2016).

The clusters’ graphic representation was the end point of the analysis at this level and facilitated understanding not only the previously calculated metrics, but also the relationships between the different elements of the network.

3.1.3 Individual Level

Once finished the study of the clusters, it was still possible to deepen the analysis one further level and explore who were the individuals that formed the network and how they acted in it. For this purpose, it was necessary to calculate the individual measurements or properties of the nodes and identify the patterns of action that existed in the network.

The communication network created from the flow of tweets in which the terms mlearning OR m-learning OR mobile + learning were included, was directed. Among the measures that reveal the importance of the nodes, we calculated Degree and Betweenness Centrality:
Degree was calculated to determine the number of relationships adjacent to each node. This parameter is considered a measure of direct influence. The degree can be of two different types: in degree and out degree. In-degree reflects the number of incoming relationships, or, in other words, the connections received by a node. It is considered a sign of popularity; Out-Degree reflects the number of outgoing connections, that is, those that originate in the node itself. It is considered a measure of productivity (Hanneman and Riddle, 2005).

Betweenness Centrality. This metric allows to determine the ability of users (nodes) to connect other users (nodes). It reflects the frequency with which a node appears in the shortest path in which two nodes of the network connect. Thus, it becomes the most efficient path to connect a pair of nodes not directly connected to each other in the network (Hanneman and Riddle, 2005). Users (nodes) with high values of Betweenness Centrality are in a position of power, because they control the flow of information in the network, and, in case of disappearing, can disconnect whole sectors of it (Barbera, 2017).

Network analysis is a way of looking at the world from the study of the form and structure of relationships (Smith, 2013). This approach focuses on the relationships established by individuals rather than on the individuals themselves. The analysis of these connections revealed the patterns of the communication network studied (Smith, 2013), where a few users had many connections, while the great majority had very few.

The nodes of the communications network were classified in five different categories according to their In-Degree and Out-Degree, based on what was proposed in Gil y Guilleumas (2019): pattern 1 included users who received no connections, but established them with other network users; pattern 2 grouped those who received connections from other users, but did not themselves initiate any; pattern 3 was applied to those who received a lower number of connections than what they initiated; pattern 4 grouped users who received connections in a greater number than what they initiated; and, finally, pattern 5 was reserved for users who received and established the same number of connections, as long as this number was not zero.

3.2 ESTABLISHING POSITIVE OR NEGATIVE POLARITY OF POSTS

After studying the particularities of the network, at the general, cluster and individual level, the focus turned to the content of the messages published.

An interesting option within this type of analysis consists in studying the feelings expressed in the tweets in relation to mobile learning. Sentiment Analysis is a field of linguistics and computer science that seeks to automatically determine the feeling expressed in a non-structured text as is the case of a
Tweet. Sentiments can be characterized as a positive or negative evaluation (Taboada, 2016); they refer to beliefs or personal judgements not based on proof or certainty, and often represent the emotional state of whom publishes the text or their point of view on a particular topic (Chong, 2016).

Although tweet sentiment analysis can be done using machine learning, that option requires prior computer training. We opted instead for using semantic analysis, which employs a sentiment lexicon. The process requires a previous processing of the text to eliminate stop words and carry out a linguistic normalization by stemming or lemmatization. Once this step has been completed, the existence of the words of the text in the lexicon is verified and the polarity value, either positive or negative, is assigned to the text under scrutiny (Liu, 2012). One way to analyze the feeling of a text is by considering it as the sum of the feelings of each of the words that make it up (Amat, 2017). This option offers a good balance between complexity and results.

In sentiment lexicons, words are associated with a feeling. For example, the BING lexicon categorizes words into positive and negative categories; the AFINN lexicon assigns words a score that varies between -5 and +5, understanding that negative scores indicate negative feelings, and positive scores indicate positive feelings; and the NRC lexicon classifies words in a binary way ("yes" / "no") into categories of positive, negative, anger, anticipation, dislike, fear, joy, sadness, surprise, and trust (Silge and Robinson, 2019).

To analyze the polarity of feelings (positive or negative) in relation to mobile learning, it is necessary to use a dictionary or lexicon appropriate to the language of the data to be analyzed. Since our data were in English, we used the Bing lexicon (Liu and Hu, 2004) and a complementary lexicon that responded to the context studied (the topics on mobile learning), built from the research corpus and previous tweets about education, as well as some words not found in the Bing lexicon.

The analysis began with the cleaning (debugging) of the text of the tweets, a process which included the following stages: converting the text to lower case; suppressing punctuation marks, numbers, control characters, extra whitespace and stop words, RT, and URLs. These elements were eliminated because they do not provide valuable information and, on the contrary, difficult text analysis.

Next, the words of the tweets were reduced to their root (Stemming) and the search terms (mlearning, mobile + learning and m-learning) were also eliminated, to avoid distortion in relation to the true participation of other terms within the corpus of data.

With the text resulting from this process and using both the Bing and the complementary lexicon, the score of each of the imported tweets was calculated. First, to get the positive score, we compared the
words of each tweet from the corpus with the positive list. Next, the same procedure was applied, but this time comparing the words of the tweet to the negative list.

The final score was calculated by subtracting the negative score from the positive one (Score = positive score - negative score) (Liu and Hu, 2004). If Score > 0, it was interpreted that the tweet expressed in general, “a positive opinion”. If Score < 0, it meant that the tweet expressed in general, “a negative opinion”. If Score = 0, then the tweet was considered to express a “neutral opinion”. The appearance of an opinion word in a sentence did not necessarily imply that the tweet were positive or negative. From the previous calculations, it was possible to establish the users’ emotional state, or point of view on mobile learning.

Other procedure that helped obtain an overview of the sentiments expressed by users was the representation of the frequencies of tweets’ sentiments. It was also useful to represent the texts associated with the identified feelings in a word cloud and a dendrogram. These graphic representations made it easier to visualize the relationships identified among the words.

Word clouds helped to understand the text of the tweets and facilitated the visual identification of the most used words. Dendrograms allowed us to display words in a tree diagram organized in clusters of categories, which were then divided in other subcategories.

Real life events can affect users’ opinions about the topics they talk about. For that reason, our research also studied the overtime evolution of users’ feelings in relation to mobile learning with the purpose of establishing relationships and understanding the changes that took place throughout the period studied.

To refine tweets sentiment analysis, we examined the frequency with which ‘no’ or other negative words (adverbs of negation) preceded sentiment words, as these elements change the sentiment expressed, often into its opposite.

Additionally, some complementary analyses were carried out, such as calculating the frequency of the words used in the tweets with positive and negative sentiment, and the analysis of the different words used according to the sentiment expressed.

Up to this point of the analysis, we had considered words as individual and independent units. However, language is created from combinations of words. With this in mind, we decided to explore the text by n-grams. An n-gram is a sequence of n consecutive words. With the purpose to explore the most recurrent word associations in the corpus of data, we divided the text into bigrams and identified and counted the occurrences of the most frequent. The procedure enabled us also to analyze the relationships between the words using a graph of networks.
Finally, taking as a reference Chong (2016), the tweet texts’ polarity was associated with the users who posted them, in order to determine their emotional state or disposition towards the subject studied. This information is particularly significant in the case of users who may be experts in the topics studied. In any case, the results of the sentiment analysis were evaluated by human experts on the subject.

3.3 DETERMINING THE PREDOMINANT TOPICS OF INTEREST

The objectives of our research took us next to identifying the predominant topics of interest in the communication network created by the flow of tweets which included the terms used to import the data. For this purpose, we employed Text Mining techniques to identify the trends in the hashtags and web domains included in the tweets.

3.3.1 Hashtag trends

Hashtags can be considered as a widely used resource to promote or spread a topic of interest. It is a common occurrence these days that the media, such as television or radio, identify current events or hot topics from the most used hashtags, which are then called trending topics. In addition, hashtags make it easier for users to locate or retrieve tweets related to a trend at a certain time. Therefore, the identification of the most used hashtags constituted the first step in detecting trends in the topics of interest of the network.

In addition to identifying the hashtags, we considered it necessary to analyze the content associated to them. Our purpose was to find out what was really being said, and the topics or opinions of collective interest that were being spread through this medium.

To do this, initially, we made a word cloud out of the hashtags used in the communication network. Once the most used were identified, we calculated their individual frequency, i.e., how many times each one was used in the corpus being analyzed.

This analysis could be further deepened by assessing the extent to which the use of a certain hashtag was creating a trend because many users employed it during extended periods of time or if, on the contrary, there were just a few users behind it.

With this purpose, we calculated each hashtag’ trend coefficient (Gil y Guilleumas, 2019), a value which considered not only its frequency of publication, but also the number of unique users who had included the hashtag in their tweets and the number of days during which it had been posted in relation to the total. This combination of variables can help minimize the effect, on a trend, of an individual user or a small group of users who may employ a hashtag in bulk.
3.3.2 Trends in web domains

The identification of the predominant web domains in the imported tweets was used as a basis to determine which sites to connect to, or which software to use.

To calculate the trends in the web domains, we calculated their trend coefficient (Gil y Guilleumas, 2019), considering not only the number of times they had been referenced, but also the number of users who had employed them and how many days had been present in the tweets posted. The use of this array of variables intended to reduce the possibility of a single user or a reduced number of them turning a web domain into trend by massively posting it.

In this context, although the identification of the trends of the websites is a first step in understanding the dynamics of the network, it is necessary to progress in the identification of the potential uses of these websites as support for the researched object.

3.4 IDENTIFYING POTENTIAL INFLUENCERS

Furthering the detailed analysis of the network, we continued with the identification of the potential influencers in the communication network created by the flow of tweets in which the terms "mlearning OR m-learning OR mobile + learning" were included. This identification was particularly useful for our research and we consider it of great interest for anyone forming their personal learning network around a specific topic and needing to identify potential influencers in that field of study.

To calculate the potential influence of users, it is not enough to consider their followers. Aspects such as popularity, number of tweets published, interactions with other users (Mentions, Responses), and Retweets (RT) of their tweets are aspects of great import.

The potential influence of the users that made up the communication network was calculated using the potential influence coefficient (Gil y Guilleumas, 2019), which employed the following variables: Outreach capacity (Followers); Popularity (In-Degree); Interconnection capacity (Betweenness Centrality); Posting volume (tweets published); Relationships (mentions and responses); and Retweets.

To conclude, it should be noted that Twitter has acquired great importance as a space for discussion and dissemination of knowledge and opinions on various topics, what requires analysis processes to know the topics of interest in the network, the positive or negative feelings in relation to those issues and who the network influencers are. The analysis of social networks has therefore become a line of work of great interest for researchers from different areas, among others the educational.

The methodology for analysis presented in this document aims to help researchers and, anyone interested in the field of social networks, to understand the complex interactions that take place in these
environments and the way in which information is disseminated, valued and converted into topics of interest thanks to the actions of the users that configure the network.
REFERENCES


Gil H., y Guilleumas R.M. (2019). Twitter en la detección de tendencias en TIC para su uso en educación, editorial UTP


The igraph core team (2020). Functions to deal with the result of network community detection, [Online]. Available: https://igraph.org/r/doc/communities.html