View metadata, citation and similar papers at core.ac.uk brought to you by CORE
provided by Academy Publication Online

ISSN 1798-4769
Journal of Language Teaching and Research, Vol. 8, No. 6, pp. 1041-1049, November 2017
DOI: http://dx.doi.org/10.17507/jltr.0806.03

Sentiment Analysis of Spanish Words of Arabic Origin Related to Islam: A Social Network Analysis

Mohamed M. Mostafa
GUST, Kuwait;
University of Malaga, Spain

Nicolás Roser Nebot
University of Malaga, Spain

***Abstract*—With the arrival of Muslims in 711 till their expulsion in the 1600s, Arabic language was present in Spain for more than eight centuries. Although social networks have become a valuable resource for mining sentiments, there is no previous research investigating the layman's sentiment towards Spanish words of Arabic etymology related to Islamic terminology. This study aim at analyzing Spanish words of Arabic origin related to Islam. A random sample of 4586 out of 45860 tweets was used to evaluate general sentiment towards some Spanish words of Arabic origin related to Islam. An expert-predefined Spanish lexicon of around 6800 seed adjectives was used to conduct the analysis. Results indicate a generally positive sentiment towards several Spanish words of Arabic etymology related to Islam. By implementing both a qualitative and quantitative methodology to analyze tweets' sentiments towards Spanish words of Arabic etymology, this research adds breadth and depth to the debate over Arabic linguistic influence on Spanish vocabulary.**

***Index Terms*—Spanish words of Arabic origin, sentiment analysis, Text mining, Twitter**

I. Introduction

In 711 a contingent of Muslims led by Tariq Ibn Ziyad crossed North Africa via Gibraltar to the Iberian peninsula. In a few years, Muslims ruled nearly all Iberia or modern day Spain and Portugal. Muslims stayed in Iberia from 711 till their expulsion in 1600s. Over more than eight centuries, Arabic became the language of culture and administration (Plann, 2009). Some authors argue that the greatest Arabic influence of Arabic on Spanish language is lexical, where around 8% of Spanish vocabulary may be traced to Arabic origin (Quintana & Mora, 2002). This includes probably thousands of modern geographic locations that still hold Arabic toponyms dating back to the Moors. Figure 1 shows only 17 of such toponyms in Spain. In fact, it is quite common to find places that contain the word "Guada-", meaning "river/valley" in Arabic. Examples include "Guadalajara" or "wadi al-hijarah/valley of stones", "Guadalaviar" or "wadi al-abyad/white valley" and "Guadalcazar" or "wadi al-qasr/valley of castle." (Please refer to Appendix for more details).

© 2017 ACADEMY PUBLICATION

Figure 1. Major Spanish Toponyms (cities of more than 20,000 inhabitants) of Arabic origin

Social media users represent are estimated to represent around 67 percent of the billion Internet active users (Eirinaki, Pisal & Singh, 2012). Thus, social networks play a major role in shaping public opinion's attitudes in areas as diverse as voting, buying products and stock markets prediction (Bai, 2011; Eirinaki, Pisal & Singh, 2012). Kim and Hovy (2004) define an opinion "as a statement in which the opinion holder makes a specific claim about a topic using a certain sentiment.". Online opinions expressed in social networks have been exploited in mining individuals' sentiments through text filtering and mining (Zhang, Zeng, Li, Wang & Zuo, 2009). Social networks' opinions are generally analyzed through natural language processing (NLP) sentiment analysis techniques. Sentiment analysis is also known as opinion mining, review mining, emotional polarity analysis (EPA), or appraisal extraction (Zagal, Tomuro & Shepitsen, 2012). Sentiment analysis can thus be regarded as a method to extract knowledge in order to find hidden patterns in an unstructured text or opinion expressed on blogs or tweets. In order to measure a sentiment score, the text's sentiment is usually compared to a lexicon/dictionary. This comparison determines the strength of the sentiment.

In fact, social network and online digital data created are expected to reach four zettabytes by the end of 2020 (Paper, Ugray & Johnson, 2014). It should be noted also that knowledge discovered through social networks are extremely useful since huge numbers of opinions expressed regarding certain topics are highly unlikely to be biased. This is probably why "social media applications, such as Facebook and Twitter, are increasingly being used by both large and small companies to gain business benefits" (He & Chen, 2014, p. 92). In a similar vein, Parise (2009, p. 2) argues that "social media tools will have a major impact on knowledge management". However, almost all online communications are *noisy* and as such pose considerable lexical and syntactic problems (Boiy & Moens, 2009; Derks, Fischer & Bos, 2008; Thelwall, Buckley, Paltoglou, Cai & Kappas, 2010; Pederson, 2001; Dave, Lawrence, & Pennock, 2003; Turney & Littman, 2003). Because of this problem, several online sentiment analyses techniques have been developed in languages as diverse as Arabic (Ahmed & Almas, 2005), Chinese (e.g., Xu, Liao & Li, 2008), English (e.g., Jansen, Zhang, Sobel & Chowdury, 2009), and multi-languages (Abbasi, Chen, & Salem, 2008).

Although several studies have investigated sentiment in different fields (Cai, Spangler, Chen & Zhang, 2010; Leong, Lee & Mak, 2012), no previous studies have focused solely on investigating sentiment of Islamic-related Spanish words of Arabic etymology as used on Twitter. In this study we aim to fill this research gap. We believe that by investigating such words polarity we add depth to the knowledge base on sentiment analysis and text mining. Through the use of both qualitative and quantitative methodologies, we also add breadth to the debate over Islamic-related words sentiment. Finally, by investigating solely Twitter texts, rather than traditional offline data, this research enriches the knowledge base of an under-represented area.

This paper is organized as follows. Next section provides a brief account of related work of major areas of sentiment analysis applications. Section three deals with the research method used to conduct the analysis. In this section issues related to research design, sampling and data analysis techniques are presented. In Section four the results of sentiment analysis are presented. Finally, Section five presents research implications, limitations and explores avenues for future research.

II. Literature Review

Sentiment analysis techniques have been used in diverse areas such as tracking sentiment trends in online discussion boards (e.g., Tong, 2001), differentiating between informative and emotional social media content (Denecke & Nejdi,

© 2017 ACADEMY PUBLICATION

2009), tracking political opinions (Thomas, Pang & Lee, 2006) mining suggestions from product reviews (Vishwanath & Aishwarya, 2011), determining consumers' dissatisfaction with online advertising campaigns (e.g., Qiu, He, Zhang, Shi, Bu & Chen, 2010), tracking emotions in emails (Mohammad, 2012), classifying consumers' positive and negative product reviews (Turney, 2002), detecting Internet hot spots (e.g., Li & Wu, 2010), and predicting stock market movements (Wong, Xia, Xu, Wu & Li, 2008). In this section we review the major five distinct sentiment analysis categories, namely political orientation analysis, consumers' product reviews, movie reviews, stock market predictions and patients' mood detection.

Several studies have investigated Twitters' sentiment analysis in election campaigns in European countries such as Sweden and Germany (Larsson & Moe, 2011; Tumasjan, Sprenger, Sandner & Welpe, 2011). In a similar vein, Williams and Gulati (2008) predicted electoral based on the total number of Facebook supporters. Malouf and Mullen (2008) investigated political ideological biases using social network analysis. Golbeck, Grimes and Rogers (2010) analyzed the US Congress Tweets and reported major reasons behind congressmen usage of Twitter. Similarly, Ekdale, Namkoong and Perlmutter (2010) investigated US political bloggers' behavior and argued that extrinsic motivation was the main motive behind blogging. Similar results were reported by Gil De Zuniga, Puig-I-Abril and Rojas (2009). Tweeting behavior during the Arab Spring was also investigated by several researchers (Papacharissi & Oliveira, 2012; Lim, 2012). Results showed that protesters used Twitter an alternative to the blocked access to the Internet. Park, Lim, Sams, Nam and Park (2011) analyzed Korean politicians' visitor boards' comments and found a gender gap in terms of positive and negative comments. Zappavigna (2011) analyzed tweets related to Obama's presidential elections victory in 2008. Other researchers investigating political sentiment analysis include Efron (2004), Thomas, Pang & Lee (2006) and Park, Kim & Barnett (2004)

Extracting sentimens from consumers' reviews has been extensively investigated (Blair-Goldensohn, Hannan, McDonald, Neylon, Reis & Reynar, 2008; Yi, Nasukawa, Bunescu & Niblack, 2003). A polarity system to analyze consumers' comparison comments was developed by Feldman, Fresko, Netzer and Ungar (2007). Hu and Liu (2004) analyzed consumers' sentiments related to several electronic products such as digital cameras and mobile phones. In a similar vein, Miyoshi and Nakagami (2007) analyzed electronic products consumer sentiments using adjective-noun pairs in a sentence. Zhang, Xu and Wan (2012) analyzed consumers' sentiments in Chinese language online texts. In fact, this study was an extension of a study reported by Ding, Liu and Yu (2008) and by liu (2010). Pekar & Ou (2008) evaluated 268 reviews of major hotels using attributes such as room service, facilities and food to analyze composite sentiments towards hotels. Na, Khoo and Wu (2005) classified sentiments related to products' reviews based on data mining techniques.

Na, Thet, and Khoo (2010) analyzed movie reviews by comparing textual characteristics across four different genres. Zhuang, Jing and Zhu (2006) summarized online texts movie reviews sentiments using machine learning techniques. Pang, Lee and Vaithyanathan (2002) classified online sentiments related to movie reviews using support vector machines (SVM). Wijaya & Bressan (2008) used a similar technique, while Na and Khoo (2008a) correctly segmented customers' reviews into relevant sections pertaining to different aspects of the movie. In another study, the same authors (Thet, Na & Khoo, 2008b) used computational linguistics to segment movie reviews comments.

Das and Chen (2001) analyzed sentiments related to Yahoo! Finance's discussion board. Gu, Konana, Liu, Rajagopalan and Ghosh (2006) predicted different stocks' future returns based on comments posted on the same platform. Bollen, Mao and Zeng (2011) predicted stock market movements using Twitter posts. Other studies investigated investors' sentiments related to factors such as air disasters (Kaplanski & Levy, 2010), earthquakes (Shan & Gong, 2012), and sports events (Chang, Chen, Chou & Lin, 2012).

Several studies have recently investigated the use of sentiment analysis in patients' mood detection. For example, Rodrigues et al. (2016) developed a Portuguese language sentiment analysis tool to detect emotional polarity among online cancer community in Brazil. Porter et al. (2013) applied sentiment analysis techniques on online cancer posts in order to detect changes in individuals' moods as a consequence of interaction with other patients within the same community. In a similar vein, Akay et al. (2015) investigated cancer patients and their families' sentiments as a reaction to the environment surrounding them. Twitter has been recognized also as a powerful gauge of several medical issues such as understanding of attitudes towards immunization (Love et al., 2013), forecasting affordable care markets (Wong et al., 2015), supporting decision-making in healthcare (Swain, 2016), predicting patients' feedback (Smith and Lee, 2012), analyzing drug reviews (Na et al. 2012) and investigating medical sentiments concerning patients' health status (Denecke and Deng, 2015).

III. METHOD

a. Twitter sampling

Launched on July 13, 2006, Twitter revolves around posting short updates of 140 maximum characters, which is approximately the size of a newspaper headline. It is estimated that there are around 500 million active twitters (Bliss, Klouman, Harris, Danforth & Dodds, 2012). Thelwall, Buckley and Paltoglou (2011) found that around 80 percent of Twitter users either disseminate information regarding their daily experiences or update their followers on what they actually doing. We selected Twitter for this study because it is the most large and popular microblog Web site. We used data representing a random set of Twitter posts from November 15, 2016, to December 15, 2016. The data comprised

© 2017 ACADEMY PUBLICATION

4586 out of 45860 tweets generated. All retweets and duplicated tweets were eliminated. Sample selection has been varied by day of the week and hours in the day in order to guarantee representativeness. The sample is comparable in size to similar research studies. For example, Qiu, He, Zhang, Shi, Bu, and Chen (2010) used a sample of 3783 opinion sentences.

b. Lexicon

Miao, Li, and Zeng (2010) argued that there are generally two widely used methods for sentiment orientation. The first one is known in the text mining community as the lexicon-based approach, while the second one is known as the corpus-based method. However, it should be noted that only few authors have used the corpus-based method in analyzing sentiment orientation. Both methods require either a pre-defined dictionary or a corpus of subjective words. In either method, the sentiment score is determined via a comparison between the sentence presented and an expert-defined entry in the dictionary. Several lexicons have been used in the literature such as the General Inquirer (Stone, Dunphy, Smith, & Ogilvie, 1966), the sentiment-based lexicon (Taboada, Brooke, Tofiloski, Voll, & Stede, 2011), the SentiWordNet (Baccianella, Esuli, & Sebastiani, 2010), the LIWC dictionary (Pennebaker, Mehl, & Niederhoffer, 2003), the Q-WordNet (Agerri & Garcia-Serrano, 2010) or the lexicon of subjectivity clues (Wiebe, Wilson, Bruce, Bell, & Martin, 2004). However, in this study we use the Spanish version of the Hu and Liu (2004) lexicon because it has been used successfully in similar applications (Miner, Delen, Elder, Fast, Hill, & Nisbet, 2012). The lexicon is similar in size to the Opinion Finder lexicon (Bollen, Mao, & Zeng, 2011).

IV. Results

Several libraries/software packages such as the twitteR, the maps, the plyr, the stringr and the ggplot2 libraries within the R version 3.3 environment were used to conduct the quantitative sentiment score. Figure 2 shows the distribution of positive sentiment scores obtained, while Figure 3 shows the distribution of negative sentiment scores. From the graphs, we recognize some asymmetry. For example, some words have high bars at the left end (positively skewed), while others show negative skeweness as measured by high tail on the left-hand side of the histogram. The visualization of the sentiment distribution in Figures 2 and 3 further underlines the fact that most words fall on the neutral point (0) or within the band of circa -1/+1. This result is in line with Lindgren (2012), who argued that the focus of sentiment analysis should be on either the positive or negative sentiments.

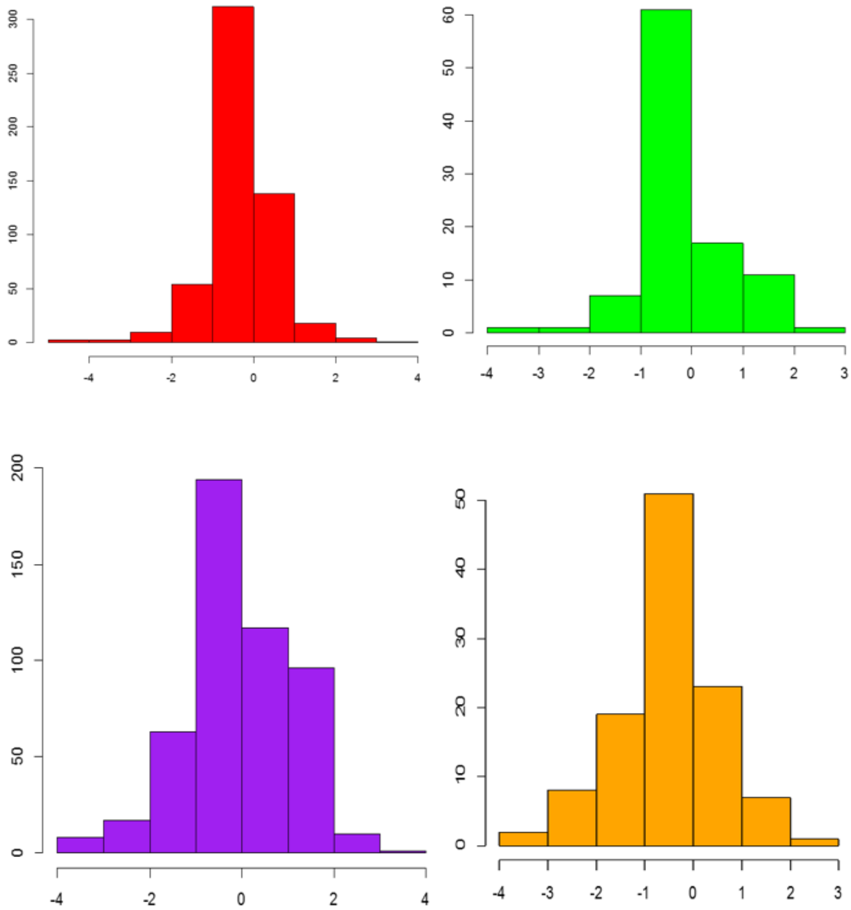

Figure 2. Major words with positive sentiment scores

© 2017 ACADEMY PUBLICATION

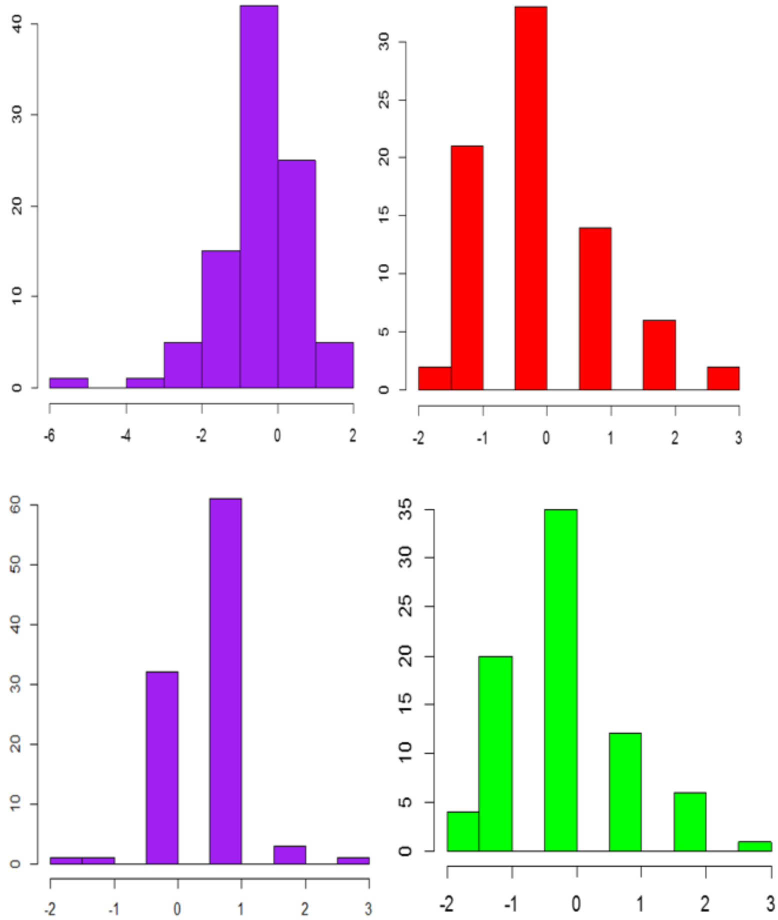

Figure 3. Major words with negative sentiment scores

Table 1 presents the division of random sample of tweets among each word used in the analysis. Following similar research only Spanish tweets were used to avoid text analysis complications (Thelwall, Buckley & Paltoglou, 2011). From Table 1 we see that most words have a positive sentiment score with the exception of four words (Islam Wahabita, Islamización, Islamismo, and Islamofobia. The overall mean score for the twenty Spanish words of Arabic etymology related to Islamic terminology was 0.2331.

TABLE 1.
SPANISH WORDS AND THEIR AVERAGE, MINIMUM AND MAXIMUM SENTIMENT SCORES

ID	Word	Tweets	Mean	Min	Max
1	ISLAM	1000	0.6710	-4	5
2	ISLÁMICA	541	0.1867	-5	4
3	ISLÁMICO	506	0.4348	-4	4
4	ISLAM CEUTÍ	415	0.5398	-3	4
5	ISLAM CHIÍ	398	0.0729	-3	3
6	ISLAM SUNÍ	287	0.6969	-2	3
7	ISLAM SUNNITA	236	0.6864	-2	5
8	ISLAMOFA	179	0.1117	-3	3
9	ISLAMICIDAD	132	0.2121	-2	2
10	ISLAMISTA	119	0.7479	-2	4
11	ISLAM WAHABITA	111	-0.0180	-4	3
12	ISLAMIZADO	99	0.6768	-2	3
13	ISLAMIZADORA	99	0.2929	-4	3
14	ISLAMIZAR	94	0.0106	-6	2
15	ISLAMO-ARABIZANTE	83	0.4458	-3	2
16	ISLAMÓFOBA	78	0.0897	-2	3
17	ISLAMIZACIÓN	78	-0.0128	-2	3
18	ISLAMÓFOGA	48	0.2917	-1	3
19	ISLAMISMO	42	-0.6429	-6	2
20	ISLAMOFOBIA	41	-0.3947	-4	1
	Total and grand mean	4586	**0.2331**		

V. IMPLICATIONS, LIMITATIONS AND FUTURE RESEARCH

© 2017 ACADEMY PUBLICATION

In this paper we analyzed sentiment polarity of more than 4500 social media tweets expressing sentiments towards twenty Spanish words of Arabic etymology related to Islamic terminology. Although any tweet is limited to 140 characters, millions of tweets posted on Twitter almost on a daily basis. Such tweets might provide an unbiased representation of individuals' sentiment towards a specific topic. Moreover, such tweets may be used by policy makers to gauge opinions regarding a specific issue. Ignoring such sentiments might put policy makers on the defensive and could also create significant image problems. The speed of social media might also render politicians efforts based on traditional media useless. However, it should be noted that while we conducted sentiment analysis to objectively classify individuals' opinions towards twenty Spanish words of Arabic etymology related to Islamic terminology, our analysis does not reveal the underlying reasons behind forming such opinions. Thus, future research might use topic recognition techniques in order to determine the most representative topics behind each sentiment, which allows us to gain comprehensive knowledge regarding the underlying causes of positive or negative sentiments. Although he lexicon-based approach we used in this study can detect basic sentiments, it sometimes fails to recognize subtle forms of linguistic expressions (Boiy and Moens, 2009). Finally, individuals' opinions might in fact be a manipulation of some online opinion makers posing as real individuals. This might distort sentiments of real individuals. Thus, future research should attempt to distinguish genuine sentiments from fake opinions.

APPENDIX. R CODES USED TO PRODUCE THE MAP IN FIGURE 1

```
require(maptools)
require(raster)
require(maps)
data(world.cities)
adm = getData("GADM", country= "Spain", level = 2)
mar = adm[adm$NAME_0 == "Spain" & adm$NAME_1 !="Islas Canarias",]
mar$coso = rep(1, length(mar$NAME_2))
plot(mar, bg = "grey80", axes= T)
plot(mar, lwd = 10, border= "cornsilk", add= T)
plot(mar, col= c("bisque","red")[mar$coso], add=T)
grid()
box()
Ciudades = world.cities[ world.cities$country.etc == "Spain" &
(world.cities$pop > 20000),]
Ciudades
points(x= Ciudades$long, y= Ciudades$lat, cex = 5 *
(Ciudades$pop/max(Ciudades$pop)), pch = 19, col = "indianred2" )
Ciudades2 = Ciudades[ Ciudades$name %in% c("Guadalajara", "Albacete", "Almeria", "Granada", "Madrid",
"Malaga","Murcia", "Algeciras", "Sevilla", "Valladolid", "Calatayud", "Zaragoza", "Benalmadena", "Cadiz", "Cordoba",
"Jaen", "Medina del Campo") , ]
Text (Ciudades2$long, Ciudades2$lat +  0.6*
(Ciudades2$pop/max (Ciudades2$pop)), Ciudades2$name, cex = 0.8)
```

REFERENCES

[1] Abbasi, H., Chen, H., & Salem, A. (2008). Sentiment analysis in multiple languages: Feature selection for opinion classification in Web forums. *ACM Transactions on Information and Systems*, 26, 1-34.

[2] Agerri, R., & Garcia-Serrano, A. (2010). Q-WordNet: Extracting polarity from WordNet senses. Paper Presented at the 7th Conference on International Language Resources and Evaluation. Malta (www.irec-conf.org).

[3] Ahmed, K., & Almas, Y. (2005). Visualising sentiments in financial texts? Proceedings of the 9th International Conference on Information Visualization, 363-368.

[4] Akay, A., Dragomir, A., Erlandsson, B. (2015). Network-based medeling and intelligent data mining of social media for improving care. *IEEE Journal of Biomedical Health Informatics*, 19, 210-218.

[5] Bai, X. (2011). Predicting consumer sentiments from online text. *Decision Support Systems*, 50, 732-742.

[6] Baccianella, S., Esuli, A., & Sebastiani, F. (2010). SentiWordNet 3.0: An enhanced lexical resource for sentiment analysis and opinion mining. *Proceedings of the 7th Conference on International Language Resources and Evaluation* (www.irec-conf.org).

[7] Blair-Goldensohn, S., Hannan, K., McDonald, R., Neylon, T., Reis, G., & Reynar, J. (2008). Building a sentiment summarizer for local service reviews. Paper presented at the WWW 2008 Workshop on NLP Challenges in the Information Explosion Era (NLPIX 2008), Beijing, April, 22.

[8] Bliss, C., Klouman, I., Harris, K., Danforth, C., & Dodds, P. (2012). Twitter reciprocal networks exhibit assortativity with respect to happiness. *Journal of Computational Science*, 3, 388-397.

[9] Boiy, E., & Moens, M. (2009). A machine learning approach to sentiment analysis in multilingual Web texts. *Information Retrieval*, 12, 526-558.

[10] Bollen, J., Mao, H., & Zeng, X. (2011). Twitter mood predicts the stock market. *Journal of Computational Science*, 2, 1-8.

[11] Cai, K., Spangler, S., Chen, Y., & Zhang, L. (2010). Leveraging sentiment analysis for topic detection. *Web Intelligence and Agent Systems*, 8, 291-302.
[12] Chang, S., Chen, S., Chou, R., & Lin, Y. (2012). Local sports sentiment and returns of locally headquartered stocks: A firm-level analysis. *Journal of Empirical Finance*, 19, 309-318.
[13] Dave, K., Lawrence, S., & Pennock, D. (2003). Mining the peanut gallery: Opinion extraction and semantic classification of product reviews. *Proceedings of the 12th International Conference on the World Wide Web (WWW-03),* ACM Press, New York, 519-528.
[14] Denecke, K., and Deng, Y. (2015). Sentiment analysis in medical settings: New opportunities and challenges. *Artificial Intelligence in Medicine*, 64, 17-27.
[15] Denecke, K., & Nejdi, W. (2009). How valuable is medical social media data? Content analysis of the medical web. *Information Sciences*, 179, 1870-1880.
[16] Derks, D., Fischer, A., & Bos, A. (2008). The role of emotion in computer-mediated communication: A review. *Computers in Human Behavior*, 24, 766-785.
[17] Ding, X., Liu, B., & Yu, P. (2008). A holistic lexicon-based approach to opinion mining. *Proceedings of the International Conference on Web Search and Web Data Mining*. CAN, 231-240.
[18] Efron, M. (2004). Cultural orientation: Classifying subjective documents by co-citation analysis. AAAI Fall Symposium on Style and Meaning in Language, Art, and Music.
[19] Eirinaki, M., Pisal, S., & Singh, J. (2012). Feature-based opinion mining and ranking. *Journal of Computer and System Sciences*, 78, 1175-1184.
[20] Ekdale, B., Namkoong, K., & Perlmutter, D. (2010). Why blog? (then and now): Exploring the motivation for blogging by popular American political bloggers. *New Media & Society*, 12, 217-234.
[21] Feldman, R., Fresko, M., Netzer, P., & Ungar, L. (2007). Extracting product comparisons from discussion boards. *Proceedings of the 7th IEEE International Conference on Data Mining (ICDM' 2007)*. Los Alamitos, CA, 469-474.
[22] Gil De Zuniga, H., Puig-I-Abril, E., & Rojas, H. (2009). Weblogs, traditional sources online and political participation: An assessment of how the internet is changing the political environment. *New Media & Society*, 11, 553-574.
[23] Golbeck, J., Grimes, J., & Rogers, A. (2010). Twitter use by the US Congress. *Journal of the American Society for Information Science and Technology*, 61, 1612-1621.
[24] Gu, B., Konana, P., Liu, A., Rajagopalan, B., & Ghosh, J. (2006). Predictive value of stock message board sentiments. McCombs Research Paper Series No. IROM-11-06.
[25] Havre, S., Hetzler, E., Whitney, P., & Nowell, L. (2002). ThemeRiver: Visualizing thematic changes in large document collections. *IEEE Transactions on Visualization and Computer Graphics*, 8, 9-20.
[26] He, W., & Chen, Y. (2014). Using blog mining as an analytical method to study the use of social media by small businesses. *Journal of Information Technology Case and Application Research*, 16, 91-104.
[27] Hu, M., & Liu, B. (2004). Mining and summarizing customer reviews. *Proceedings of the 10th International Conference on Knowledge Discovery and Data Mining (ACM SIGKDD 2004)*, 168-177.
[28] Kaplanski, G., & Levy, H. (2010). Sentiment and stock prices: The case of disasters. *Journal of Financial Economics*, 95, 174-201.
[29] Kim, S., & Hovy, E. (2004). Determining the sentiment of opinions. *Proceedings of the International Conference on Computational Linguistics (COLING 2004),* East Stroudsburg, PA, 1367.
[30] Larsson, A., & Moe, H. (2011). Studying political microblogging: Twitter users in the 2010 Swedish election campaign. *New Media & Society*, 14, 727-747.
[31] Leong, C., Lee, Y., & Mak, W. (2012). Mining sentiments in SMS texts for teaching evaluation. *Expert Systems with Applications*, 39, 2584-2589.
[32] Li, N., & Wu, D. (2010). Using text mining and sentiment analysis for online forums hotspot detection and forecast. *Decision Support System*, 48, 354-368.
[33] Lim, M. (2012). Clicks, cabs, and coffee houses: Social media and oppositional movements in Egypt, 2004-2011. *Journal of Communication*, 62, 231-248.
[34] Lindgren, S. (2012). "It took me about half an hour, but I did it!" Media circuits and affinity spaces around how-to videos on YouTube. *European Journal of Communication*, 27, 152-170.
[35] Liu, B. (2010). Sentiment analysis and subjectivity. Handbook of Natural Language Processing. Second Edition, (editors: N. Indurkhya and F. J. Damerau).
[36] Liu, Y., Huang, X., An, A., & Yu, X. (2007). ARSA: A sentiment-aware model for predicting sales performance using blogs. *Proceedings of the 30th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval,* New York, NY, 607-614.
[37] Love, B., Himelboim, I., Holton, A., & Stewart, K. (2013). Twitter as a source of vaccination information: Content drivers and what they are saying. *American Journal of Infection Control*, 41, 568-570.
[38] Malouf, R., & Mullen, T. (2008). Taking sides: User classification for informal online political discourse. *Internet Research*, 18, 177-190.
[39] Miao, Q., Li, Q., & Zeng, D. (2010). Fine-grained opinion mining by integrating multiple review sources. *Journal of the American Society for Information Science and Technology*, 61, 2288-2299.
[40] Miyoshi, T., & Nakagami, Y. (2007). Sentiment classification of customer reviews on electronic products. *Proceedings of the International Conference on Systems, Man and Cybernetics*, 2028-2033.
[41] Mohammad, S. (2012). From once upon a time to happily ever after: Tracking emotions in mail and books. *Decision Support Systems*, 53, 730-741.
[42] Na, J., Khoo, C., & Wu, P. (2005). Use of negation phrases in automatic sentiment classification of product reviews. *Library Collections, Acquisitions, & Technical Services*, 29, 180-191.

[43] Na, J., Kyaing, W., Khoo, C., Chang, Y., & Theng, Y. (2012). Sentiment classification of drug reviews using a rule-based linguistic approach. In H. Chen and G. Chowdhury (Eds.). *The outreach of digital libraries: A globalized resource network*. *Lecture Notes in Computer Science*, Vol. 7634, Springer: Berlin, 189-198.
[44] Na, J., Thet, T., & Khoo, C. (2010). Comparing sentiment expression in movie reviews from four online genres. *Online Information Review*, 34, 317-338.
[45] Pang, B., Lee, L., & Vaithyanathan, S. (2002). Thups up? Sentiment classification using machine learning techniques. *Proceedings of the 2002 Conference on Engineering Methods in Natural Language Processing*. Morristown, NJ, 79-86.
[46] Paper, D., Ugray, Z., & Johnson, J. (2014). Mega data value creation: A roadmap to success. *Journal of Information Technology Case and Application Research*, 16, 119-126.
[47] Parise, S. (2009). Social media networks: What do they mean for knowledge management? *Journal of Information Technology Case and Application Research*, 11, 1-11.
[48] Park, H., Kim, C., & Barnett, G. (2004). Socio-communicational structure among political actors on the Web. *New Media & Society*, 6, 403-423.
[49] Park, S., Lim, Y., Sams, S., Nam, S., & Park, H. (2011). Networked politics on cyworld: The text and sentiment of Korean political profiles. *Social Science Computer Review*, 29, 288-299.
[50] Pederson, T. (2001). A decision tree of bigrams is an accurate predictor of word sense. *Proceedings of the Second Annual Meeting of the North American Chapter of the Association for Computational Linguistics,* 79-86.
[51] Pekar, V., & Ou, S. (2008). Discovery of subjective evaluations of product features in hotel reviews. *Journal of Vacation Marketing*, 14, 145-155.
[52] Pennebaker, J., Mehl, M., & Niederhoffer, K. (2003). Psychological aspects of natural language use: Our words, our selves. *Annual Review of Psychology*, 54, 547-577.
[53] Plann, S. (2009). Arabic: Another "other Spanish language"? *International Journal of Multilingualism*, 6, 369-385.
[54] Portier, K., Greer, G., Rokach, L., et al. (2013). Understanding topics and sentiment in an online cancer survivor community. *Journal of the National Cancer Institute*, 47, 195-198.
[55] Qiu, G., He, X., Zhang, F., Shi, Y., Bu, J., & Chen, C. (2010). DASA: Dissatisfaction-oriented advertising based on sentiment analysis. *Expert Systems with Applications*, 37, 6182-6191.
[56] Quintana, Lucía & Juan Pablo Mora. (Oct. 2002). Enseñanza del acervo léxico árabe de la lengua española. XIII Congreso Internacional de ASELE. Retrieved from http://cvc.cervantes.es/ensenanza/biblioteca_ele/asele/asele_xiii.htm (retrieved 21, 02, 2017).
[57] Rodrigues, R., Dores, R., Camilo-Junior, C., & Rosa, T. (2016). SentiHealth-Cancer: A sentiment analysis tool to help detecting mood of patients in online social networks. *International Journal of Medical Informatics*, 85, 80-95.
[58] Shan, L., & Gong, S. (2012). Investor sentiment and stock returns: Wenchuan earthquake. *Finance Research Letters*, 9, 36-47.
[59] Smith, P., & Lee, M. (2012). Cross-discourse development of supervised sentiment analysis in the clinical domain. *Proceedings of the 3rd workshop in computational approaches to subjectivity and sentiment analysis*. Stroudsburg, PA, USA, 79-83.
[60] Stone, P., Dunphy, D., Smith, M., & Ogilvie, D. (1966). The general inquirer: A computer approach to content analysis. Cambridge, MA: The MIT Press.
[61] Swain, A. (2016). Mining big data to support decision making in healthcare. *Journal of Information Technology Case and Application Research,* 18, 141-154.
[62] Taboada, M., Brooke, J., Tofiloski, M., Voll, K., & Stede, M. (2011). Lexicon-based methods for sentiment analysis. *Computational Linguistics*, 37, 267-307.
[63] Thelwall, M., Buckley, K., Paltoglou, G., Cai, D. & Kappas, A. (2010). Sentiment strength detection in short informal text. *Journal of the American Society for Information Science and Technology*, 61, 2544-2558.
[64] Thelwall, M., Buckley, K., & Paltoglou, G. (2011). Sentiment in Twitter events. *Journal of the American Society for Information Science and Technology*, 62, 406-418.
[65] Thet, T., Na, J., & Khoo, C. (2008a). Sentiment classification of movie reviews using multiple perspectives. *Proceedings of the International Conference on Asian Digital Libraries (ICADL).* Springer Verlag, Berlin, 184-193.
[66] Thet, T., Na, J., & Khoo, C. (2008b). Aspects-based sentiment analysis of movie reviews on discussion boards. *Journal of Information Science*, 36, 823-848.
[67] Thomas, M., Pang, B., & Lee, L. (2006). Get out the vote: Determining support or opposition from congressional floor-debate transcripts. *Proceedings of the 2006 Conference on Empirical Methods in Natural Language Processing (EMNLP).*
[68] Tong, R. (2001). An operational system for detecting and tracking opinions in online discussions. Working Notes of the Workshop on Operational Text Classification, New Orleans, LA, 1-6.
[69] Tumasjan, A., Sprenger, T., Sandner, P., & Welpe, I. (2011). Election forecasts with Twitter: How 140 characters reflect the political landscape. *Social Science Computer Review*, 29, 402-418.
[70] Turney, P., & Littman, M. (2003). Measuring praise and criticism: Inference of semantic orientation from association. *ACM Transactions on Information Systems*, 21, 315-346.
[71] Vishwanath, J., & Aishwarya, S. (2011). User suggestions extraction from customer reviews. *International Journal on Computer Science and Engineering*, 3, 1203-1206.
[72] Wiebe, J., Wilson, TT., Bruce, R., Bell, M., & Martin, M. (2004). Learning subjective language. *Computational Linguistics*, 30, 277-308.
[73] Wijaya, P., & Bressan, S. (2008). A random walk on the red carpet: Rating movies with user reviews and Page Rank. *Proceedings of the 17th ACM Conference on Information and Knowledge Management*, ACM Press, New York, 951-960.
[74] Williams, C., & Gulati, G. (2008). What is a social network worth? Facebook and vote share in the 2008 presidential primaries. The Annual Meeting of the American Political Science Association. Boston, MA: APSA, 1-17.
[75] Wong, C., Sap, M., Schwartz, A. et al., (2015). Twitter sentiment predicts Affordable Care Act marketplace enrollment. *Journal of Medical Internet Research*, 17, e 51.

© 2017 ACADEMY PUBLICATION

[76] Wong, K., Xia, Y., Xu, R., Wu, M., & Li, W. (2008). Pattern-based opinion mining for stock market trend prediction. *International Journal of Computer processing of Languages*, 21, 347-361.
[77] Xu, D., Liao, S., & Li, Q. (2008). Combining empirical experimentation and modeling techniques: A design research approach for personalized mobile advertising applications. *Decision Support Systems*, 44, 710-724.
[78] Yi, J., Nasukawa, T., Bunescu, R., & Niblack, W. (2003). Sentiment analyzer: Extracting sentiments about a given topic using natural language-processing techniques. *Proceedings of the 3rd IEEE International Conference on Data Mining (ICDM' 2003).* Los Alamitos, CA, 427-434.
[79] Zagal, J., Tomuro, N., & Shepitsen, A. (2012). Natural language processing in game studies research: An overview. *Simulation & Gaming*, 43, 356-373.
[80] Zappavigna, M. (2011). Ambient affiliation: A linguistic perspective on Twitter. *New Media & Society*, 13, 788-806.
[81] Zhang, C., Zeng, D., Li, J., Wang, F., & Zuo, W. (2009). Sentiment analysis of Chinese documents: From sentence to document level. *Journal of the American Society for Information Science and Technology*, 60, 2474-2487.
[82] Zhang, W., Xu, H., & Wan, W. (2012). Weakness finder: Find product weakness from Chinese reviews by using aspects based sentiment analysis. *Expert Systems with Applications*, 39, 10283-10291.
[83] Zhuang, L., Jing, F., & Zhu, X. (2006). Movie review mining and summarization. *Proceedings of the 15th ACM Conference on Information and Knowledge Management*. New York, NY, 43-50.

Mohamed M. Mostafa is a doctoral student in Spanish/Arabic Translation at the University of Malaga, Spain. He has also earned an MA in Translation Studies from the University of Portsmouth, UK, an MS in Applied Statistics from the University of Northern Colorado, USA, an MSc in Functional Neuroimaging from Brunel University, UK, an MSc in Social Science Data Analysis from Essex University, UK, an MBA and a BSc at Port Said/Suez Canal University, Egypt. He was employed at universities in the USA, Portugal, Egypt, Cyprus, Turkey, France, Jordan, United Arab Emirates, Bahrain and Kuwait. His current research interests include data mining, social networks analysis, artificial intelligence applications and translation studies.

Nicolás Roser Nebot obtained a doctorate degree from the Autonomous University of Madrid in 1997. His thesis title is "Politics and Religion: the Islamic Concept". He has worked as a professor of Arabic language and translation at the University of Malaga since 1991, where he becomes Senior Lecturer in 2001. He specializes in the field of Specialized Translation Arabic/Spanish/Arabic. He was granted an award of Excellence in 2001 for this PhD thesis from the Autonomous University of Madrid. He is an expert in the topic of the political theory of Islam, the didactics of Arabic as a foreign language, and the translation of the authoritative Islamic texts as well as in the translation of manuscripts and classical Arabic texts. He has supervised, and is currently supervising, several doctoral theses on subjects related to his research.

© 2017 ACADEMY PUBLICATION