

A Twitter Sentiment Analysis Model for Measuring Security and Educational Challenges: A Case Study in Saudi Arabia

¹Hassan Abdullah Alqarni, ²Yahya AlMurtadha and ³Abdelrahman Osman Elfaki

¹Department of Education and Psychology,
College of Education and Arts, University of Tabuk, Tabuk, KSA

²Department of Computer Science,

³Department of Information Technology,
Faculty of Computing and Information Technology, University of Tabuk, Tabuk, KSA

Article history

Received: 9-01-2018

Revised: 08-03-2018

Accepted: 14-03-2018

Corresponding Authors:

Yahya AlMurtadha
Department of Computer
Science, Faculty of Computing
and Information Technology,
University of Tabuk, Tabuk,
KSA
Email: y.murtadha@ut.edu.sa
and

Abdelrahman Osman Elfaki
Department of Information
Technology, Faculty of
Computing and Information
Technology, University of
Tabuk, Tabuk, KSA
Email: a.elfaki@ut.edu.sa

Abstract: Ensuring the good psychological health of the community is one of the highest priorities in modern societies. Therefore, having a sense of the community's rhythm and mood is a very important factor in understanding what challenges it may be facing. Psychological challenges differ from one society to another. Hence every community has its own specific psychological scales. In the context of present-day Saudi Arabia and many other countries, the measuring of educational and security challenges is critical as it can enable decision-makers to avoid anticipated risks. Traditional psychological scales in the form of questionnaires are time consuming to administer and analyze, especially where data need to be collected from a massive sample distributed over a wide geographical area. Such scales are impractical and ineffective in terms of providing critical results especially in today's rapidly changing environment. Therefore this research proposes an approach to identify and measure the educational and security challenges facing Saudi society through Twitter sentiment analysis. The psychological measurement standards of three key categories of education and security challenges were identified and broken down into selected keywords that best identified these challenges. Then Arabic tweets that contained those keywords were analyzed in order to develop a model that could classify new tweets into one of the three types of challenges. The proposed model was better able to predict tweets belonging to the cross-cultural and ethics of dialogue and rules of difference challenges than those for the dominant negative social values challenge. The results of this research show that the sentiment analysis of tweets could provide a faster and cheaper alternative to the use of traditional psychological scales.

Keywords: Security and Educational Challenges, Arabic Tweets, Sentiment Analysis

Introduction

Psychological rigidity has become one of the hallmarks of the health community (Rappaport and Seidman, 2000). Problems and issues in social and psychological domains are defined as challenges. Classical social and psychological researchers deal with studying a problem in specific ambit by applying a psychological test that is suitable for that particular

problem. According to Urbina and Anastasi (1997), a psychological test is "an objective and standardized measure of behavior of an individual's performance of tasks that have usually been prescribed beforehand". The most common type of psychological test is a paper-and-pencil test consisting of a series of questions that need to be answered by a respondent for sake of the measurement. The answers to these questions produce a test score. Psychometrics has

been recognized as the approved term for the science behind psychological testing (Mellenbergh, 1989). In this study, we introduce Twitter as psychological test medium. Specifically, we use Twitter to measure the extent of some security and educational challenges in Kingdom of Saudi Arabia.

A number of security and educational challenges are threatening educational security. Each nation has its own educational ideas that stem from its values and its stylistic and philosophical constants, which shape its identity and character. Educational security protects nations through the maintenance of those educational values (Ali, 2013). The security and educational challenges can be categorized into three main types: (1) Cross-cultural, (2) ethics of dialogue and rules of difference and (3) dominant negative social values. The cross-cultural challenge has been defined as following a civilizational style to apply some beliefs and believing that those beliefs represent the absolute truth (Arif, 1995). The ethics of dialogue and rules of difference challenge has been defined as a sort of communication between individuals that esteems those variations with others which allows for true listening in a safe environment that offers possibilities for the transformation of self-awareness in each individual" (Ashki, 2006). Dominant social values have been defined as those values that the majority of people in a society support at a particular time (Prilleltensky, 1989). If these values are negative, they are called dominant negative social values. An example of a dominant negative social value in Saudi Arabia at the moment could be a value such as "men and women are considered not to be equal."

The objective of sentiment analysis is to determine the attitude of a speaker, writer, or other subject with respect to some topic (Cambria and Hussain, 2012). Recently, social media has grown exponentially and now contains a huge amount of human activity, making it a fertile source for sentiment analysis (Pfeffer, 2014). Twitter was created to be a fast communication medium for people from all walks of life and is considered one of the most important social networking sites on the World Wide Web with a global impact. The use of Twitter sentiment analysis to understand and predict emotions, opinions, or moods has been proven theoretically and demonstrated practically (Maynard and Bontcheva, 2016). Sánchez *et al.* (2017) highlighted the value of Twitter as a new data source for the social sciences. Twitter represents an abundant source for many human studies due to its quantitative and qualitative content, not least because more than 140 million active users publish over 400 million tweets every day (Li *et al.*, 2012). Moreover, in the Middle East, Twitter is becoming a prominent player in

socio-political events such as the Arab Spring (MacEachren *et al.*, 2011). Vosoughi *et al.* (2015) discussed the importance of Twitter in the analysis and study of human behavior and refer to the fact that less than 10% of Twitter accounts are private.

The importance of measuring educational and security challenges together with the availability of a huge amount of data on human activities online, encouraged us to measure educational and security challenges by using social media and more precisely by using Twitter.

Twitter is currently the most popular social networking program in Saudi Arabia. In 2015, according to the Alyaum website (2015), 500,000 tweets were made per day in Saudi Arabia and no doubt this figure has increased since then. In this study, we introduce a novel approach to measure psychological hardness by using tweets that have been initiated in Saudi Arabia. The research question we seek to address is: What is the relation between educational and security challenges and timing in Saudi Arabia? It should be noted that in this study we target only Arabic tweets. The methodology section provides details about the data collection and analysis procedures.

The traditional method of measuring educational and security challenges is to conduct a survey by using a questionnaire that is filled in directly by targeted persons at a specific time or at a time of their own choosing. Hence, the answers are dependent on the respondents' mood at the time they complete the questionnaire. According to Choi (2011), educational and security challenges are influenced by variable factors, making them vulnerable to change. For instance, in a psychological test to measure educational and security challenges that was developed by Mikhamar (1996), the first question is "Whatever the obstacles, I can achieve my goals" and the respondent is asked to choose one answer from "Always," "Sometimes" or "Never." A respondent could choose "Always" but at another time he/she might choose "Sometimes." Consequently, measuring psychological hardness at one point in time, regardless of the sample type or size, may not provide a clear picture of the phenomenon under study or accurate results.

Related Works

In the current decade, Twitter sentiment analysis has attracted many researchers and practitioners due to its promising results. Usually, sentiment analysis is applied in the business domain to understand customer behavior or feedback (Neuendorf, 2017). Numerous research and practical papers have proposed solutions based sentiment analysis for different business domains and recently sentiment analysis has been used to

provide solutions in the social sciences as well. In this study, we use sentiment analysis as a research tool for the field of psychology. Therefore, we focus our discussion of the related works on a few examples that have proposed sentiment analysis as a solution for problems in a number of different fields in order to show that sentiment analysis is suitable, applicable and successful in a number of different fields.

For instance, Bollen *et al.* (2011) used the Profile of Mood States (POMS) to model public mood and emotion by using Twitter sentiment analysis and they concluded that social, economic and political events do have a direct impact on public mood. As an example of using Twitter sentiment analysis in political science, Al-Khalifa (2012) used graph analysis to understand the underlying social structure of Saudi political activities on Twitter. Hend used NodeXL to visualize the shape of political networks for different hash tags and to ascertain the power of the relationships within these networks. In a similar vein, Ullmann (2017) used tweets to understand the effect of Twitter during the Egyptian revolution. In a different strand of research, Alabbas *et al.* (2017) used Twitter sentiment analysis to detect high-risk floods. On the other hand and on a more personal level, Desai *et al.* (2012) used Twitter sentiment analysis to check the usefulness of tweets that were issued during Kidney Week and their results showed that Twitter is an effective medium through which to improve medical awareness. Schwartz (2016) analyzed tweets to predict individual well-being by modifying the Satisfaction With Life (SWL) Positive Emotions, Engagement, Relationships, Meaning and Accomplishment (PERMA) psychological scales to make them suitable for Twitter sentiment analysis.

Many standard models and tools have been generated for the purpose performing Twitter sentiment analysis on English tweets, but unfortunately Arabic sentiment analysis lacks of mature models and tools. However, some notable efforts have been made in this direction. For instance, Alabbas *et al.* (2017), mentioned above, used classifier techniques to analyze short informal (colloquial) Arabic text. As examples of these researches are Aldayel and Azmi (2016), El Ballouli *et al.* (2017). To address the lack of data sources for this area of research, Refae and Rieser (2014) created an Arabic Twitter corpus for subjectivity and sentiment analysis that has been published by the European Language Resources Association (ELRA). The publication of this corpus by the ELRA indicates that the use of Arabic tweets in sentiment analysis is an attractive area of research. An enormous amount of research has also be undertaken on Twitter as data source for the sake of improving the automated understanding of the Arabic language (see the review by Boudad *et al.* (2017)).

However, to the best of our knowledge, our research is the first to use sentiment analysis of Arabic tweets to measure psychological behavior.

Development of a Sentiment Analysis Model for Measuring Security and Educational Challenges

The measurement of security and educational challenges through the analysis of tweets is both a theoretical and a practical challenge. From the theoretical perspective, many psychological tests can be used to measure security and educational challenges. However, for the purpose of this study we needed a psychological test that had been developed for the Middle East context so that we could modify it to make it suitable for Arabic tweet analysis. On the other hand, the work presented us with a practical challenge due to the lack of Arabic sentiment analysis tools. Currently, researchers and practitioners in the Arabic sentiment analysis domain are using a standard corpus that has been developed for commercial goals and which is thus not suitable for our purposes. To overcome this limitation, we developed a novel technique for Arabic sentiment analysis. Hence, our developed model makes two novel contributions: Offering a way to measure security and educational challenges by tweets and providing a new technique for Arabic sentiment analysis. In the following, we provide details of the developed model.

First, some suitable psychological tests were selected for each security and educational challenge. Due to the importance of measuring these challenges numerous research studies have used different psychological tests, such as Kobasa (1979) and Maddi and Kobasa (1984). This importance drew the attention of researchers from academic and community institutions to study the psychological rigidity of different sectors of society. It is a fact of life that psychological factors vary from place to place and that social or religious factors affect psychological factors as well. Hence, researchers have developed scales for psychological hardiness to match their area of interest. Therefore, we felt that it would be appropriate to use a general psychological scale to measure security and educational challenges that was developed to match the nature of Saudi Arabian society. One of the most well-known psychological hardiness scales for Arab countries is that proposed in Mikhamar (1996), which has been used in many PhD theses and research papers. The other source of psychological scales for measuring security and educational challenges in the Middle East is Arif (1995). Therefore, in this research, we used the psychological scales in Mikhamar (1996) and in Arif (1995). These psychological scales are in the form of a traditional questionnaire that consists of direct

questions targeted at a specific population and are evaluated by statistical methods.

We modified these psychological scales by combining related questions from both to configure specific categories. For instance, in the works of Mikhamar (1996) and in Arif (1995) we found that a total of five questions are used to measure “negative thinking,” so we replaced these five questions with the category “negative thinking”. Then we identified a number of keywords to represent “negative thinking”. In brief, in our modification of the psychological scales, we (1) chose a suitable psychological scale for each challenge; (2) grouped the related questions under a specific category; (3) Unification of Polarity. All categories should be appeared as positive or negative feelings; and (4) Representing each category by keywords. Note that this four-step process was evaluated and approved by a group of experts.

Table 1 shows a sample of the keywords and their associated challenge. The categories and keywords have been translated into English for the sake of clarity (the original Arabic words are in parentheses).

In order to develop our proposed model, we followed five steps:

Step 1

By using assistance of experts we identified a group of keywords that represent the security and educational challenges that the proposed model is expected to discover.

Step 2

We used RapidMiner Studio using the search Twitter operator (set for the Twitter streaming API) to search Twitter for Arabic tweets made within boundaries of Saudi Arabia. The resulting tweets were generated as records and each record consisted of {name, to-user, source (device), text, geo-location latitude, geo-location longitude, retweet count}. We removed retweets (-rt) and links (-http) to reduce noise and duplications. The output of this step was a dataset in which each tweet was represented as a record. In the traditional questionnaire

process, this dataset equates to the data obtained from filled-in questionnaires.

Step 3

We organized and cleaned the dataset by (1) tokenizing the dataset, i.e., extracting only the words from tweet texts; (2) removing Arabic stop words, such as “من” meaning “from”; and (3) stemming the remaining words to their original roots according to Arabic grammar. The output of this step was a dataset of clean tweet records containing only stem words on which the proposed classification model would be tested

Step 4

We developed the proposed classification model based on deep learning (Bengio *et al.*, 2013) as a multi-layer feed-forward artificial neural network that was trained with stochastic gradient descent using a back-propagation algorithm. The aim of this step was to build a model that could predict whether the new tweets were positive, negative or neutral educational and security challenges. The result was cumulative and based on the majority. As our aim was to identify security and educational challenges for the whole community, the output result reflected the community situation. A negative result showed the existence of a risk to security and educational values and vice versa for the positive result. The output of this step was a novel classification model.

Step 5

In this step, we validated the developed model by employing 10 k-fold cross-validation (Kohavi, 1995) to test the prediction accuracy of the developed model on the dataset. Cross-validation split the data into k subprocesses. Each time one subprocess was used as a validation set the remaining k-1 subprocesses were used as the training set. The results obtained from the k experiments were averaged to get a single validation value. The output of this step was a validated classification model.

Table 1: Sample of keywords and associated challenges

Category	Keywords
Challenge 1: Cross-cultural	
Negative thinking	death – failure – depression – grief – poverty
Laziness and poor pursuit of livelihood	Luck – fortune – chance
Limited scientific knowledge	Ignorance – I do not understand I don’t know
Challenge 2: Ethics of dialogue and rules of difference	
Inability to express own opinion	Hesitant – awkward – stuttering
Lack of self-confidence	I’m weak – I’m stupid – I cry
Lack of communication skills	Don’t speak to me, go away, go out
Challenge 3: Dominant negative social values	
Negative trend toward women	Women are incomplete religion – women are lacking in intelligence – women do not understand
Negative trend toward the elderly	Old – old-fashioned – gray-haired
Negative relationship between friends	My benefit – what is your benefit to me? – what is your value for me?

Table 2: Summary of the process for developing the sentiment analysis model

Step	Process	Output
1.	Prepare psychological scale	Keywords represent a new psychological scale to measure security and educational challenges by using Twitter sentiment analysis
2.	Get dataset	Dataset: Group of Arabic tweets
3.	Organize and clean the dataset	Cleaned dataset
4.	Develop the classification model	Classification model
5.	Validate the classification model	Validated classification model

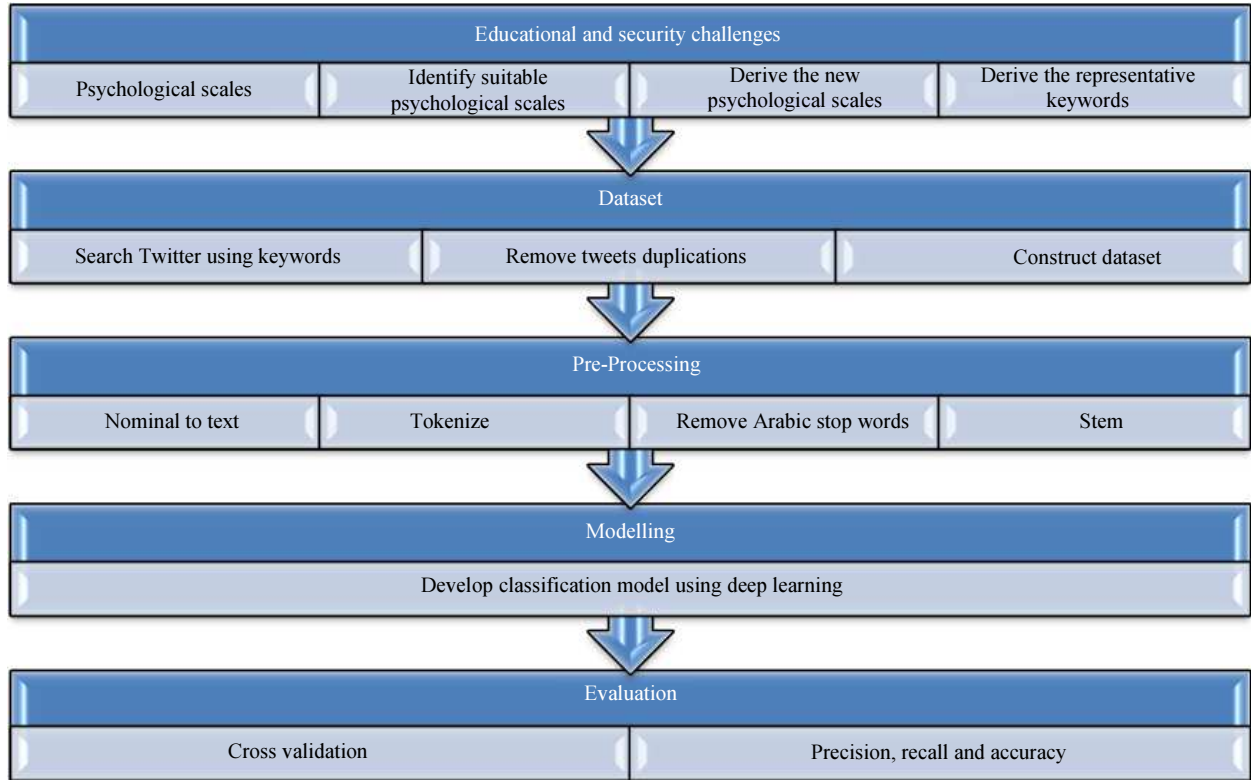


Fig. 1: Life cycle of proposed sentiment analysis model

Table 2 summarizes the process followed in developing our proposed sentiment analysis model and Figure 1 illustrates the life cycle of the proposed model.

Results

The results were obtained by developing the classification model through applying a deep learning operator using Rapid miner. In line with Al-Rubaiee *et al.* (2016) and Khan *et al.* (2014), precision, recall and accuracy were used to test the performance of the developed model. Precision calculates the ratio of relevant tweets in the middle of the true positive and false positive tweets belong to a particular category; recall calculates the ratio of retrieved relevant tweets over the total number of relevant tweets; and accuracy:

$$Precision = TP / (TP + FP) \quad (1)$$

$$Recall = TP / (TP + FN) \quad (2)$$

$$Accuracy = TP + TN / (TP + TN + FP + FN) \quad (3)$$

Where:

TP = True positive

FP = False positive

TN = True negative

FN = False negative

Table 3 shows the precision and Table 4 the recall results for Challenge 1, the cross-cultural challenge with an accuracy of 64.06+/-5.56%. From the tables, the highest predicted precision (73.48%) and recall (75.91%) were for negative tweets.

Table 3: Challenge 1: Precision

Challenge 1: Cross-cultural	True positive	True negative	True neutral
Class recall	58.20%	75.91%	40.78%

Table 4: Challenge 1: Recall

Challenge 1: Cross-cultural	Class precision (%)
Predicted positive	57.89
Predicted negative	73.48
Predicted neutral	45.16

Table 5: Challenge 2: Precision

Challenge 2: Ethics of dialogue and rules of difference	Class precision (%)
Predicted positive	67.52
Predicted neutral	45.31
Predicted negative	61.26

Table 6: Challenge 2: Recall

Challenge 2: Ethics of dialogue and rules of difference	True positive	True neutral	True negative
Class recall	75.60%	30.21%	63.55%

Table 7: Challenge 3: Precision

Challenge 3: Dominant negative social values	Class precision (%)
Predicted negative	65.78
Predicted neutral	45.12
Predicted positive	63.37

Table 8: Challenge 3: Recall

Challenge 3: Dominant negative social values	True negative	True neutral	True positive
Class recall	57.48%	34.26%	78.64%

Table 5 shows the precision and Table 6 the recall results for Challenge 2 (ethics of dialogue and rules of difference). From the tables, the highest predicted precision (67.52%) and recall (75.60%) calculated are for the positive tweets. According to equation (3), the accuracy is 61.99+/-6.23%.

Table 7 gives the precision and Table 8 the recall for Challenge 2 (dominant negative social values). According to equation (3), the accuracy is 61.46+/-6.67%. From the tables, the highest predicted precision (65.78%) is for the negative tweets while the highest recall (78.64%) is for the positive tweets.

The precision results indicate that the developed model can best predict the tweets that reflect Challenge 2, while the recall results indicate that it can better retrieve the positive tweets associated with Challenge 3. As the attitude of the public changes with frequent changes in the mood and in response to events, this implies that the keywords for the challenges especially Challenge 3 (dominant negative social values) might be need to be reviewed in order to develop an adaptive method that can look for keywords that change based on modifications in the mood of the public at large.

The results of this research can be used in the following: The assembly of a mechanism that acts as an

early warning system about the public mood; the identification of possible educational and security threats and the protection of the community as part of a strategy of prevention rather than cure.

Conclusion

The impact of Twitter as the most famous micro-blogging site in Saudi Arabia is a phenomenon that has attracted researchers interested in studying societal behavior. In this research, we introduced a novel way to measure educational and security challenges by applying sentiment analysis to Arabic tweets. The methodology consisted of the following steps: (1) Identifying the keywords needed to discover security and educational challenges in Twitter content; (2) using Twitter API streaming to construct a dataset of Twitter users for a specific period of time; (3) using data mining algorithms to develop a classification model based on the identified keywords to classify tweets; and (4) evaluating the performance of the developed model.

This aim of this research was to develop a classification model that could determine whether the current tweet stream contained tweets representing three major security and educational challenges, namely, the

(1) cross-cultural, (2) ethics of dialogue and rules of difference and (3) dominant negative social values challenges. The results showed that the model was able to predict Challenge 1 and Challenge 2 tweets more effectively than Challenge 3 tweets.

In addition, this research provided a faster and cheaper alternative method for two traditional psychological scales. As future work, detailed scientific studies should be undertaken to compare our developed model with traditional statistical methods.

Acknowledgment

This study was funded by the Deanship of Scientific Research, University of Tabuk, Tabuk, Saudi Arabia, under grant no. S-1437-314.

Authors Contributions

Hassan Abdullah Alqarni: Designed the research and experiments plan, organized the study, findings interpretation and contributed to the writing of the manuscript.

Yahya AlMurtadha: Research design, designing coding and executing the experiments, data analysis, findings interpretation and contributed to the writing of the manuscript.

Abdelrahman Osman Elfaki: Research design, participated in all experiments, participated in the data-analysis and contributed to the writing of the manuscript.

Ethics

This article is original and unpublished elsewhere. The corresponding authors confirm that all of the authors have read and approved the manuscript and there are no ethical issues involved.

References

- Alabbas, W., H.M. al-Khateeb, A. Mansour, G. Epiphaniou and I. Frommholz, 2017. Classification of colloquial Arabic tweets in real-time to detect high-risk floods. Proceedings of the International Conference on Social Media, Wearable and Web Analytics, Jun. 19-20, IEEE Xplore Press, London, UK, pp: 1-8. DOI: 10.1109/SOCIALMEDIA.2017.8057358
- Aldayel, H.K. and A.M. Azmi, 2016. Arabic tweets sentiment analysis-a hybrid scheme. J. Inform. Sci., 42: 782-797. DOI: 10.1177/0165551515610513
- Ali, A.S., 2013. Educational security and its role in maintaining identity and achieving comprehensive security (Arabic Paper). Proceedings of the Scientific Symposium on Security and the Role of Universities in Strengthening it, (RUS' 13), Riyadh, KSA.
- Al-Khalifa, H.S., 2012. A first step towards understanding Saudi political activities on Twitter. Int. J. Web Inform. Syst., 8: 390-400. DOI: 10.1108/17440081211282883
- AL-Rubaiee, H., R. Qiu, K. Alomar and D. Li, 2016. Sentiment analysis of Arabic tweets in e-learning. J. Comput. Sci., 12: 553-563. DOI: 10.3844/jcssp.2016.553.563
- Alyaum website, 2015. Retrieved in 10 Oct 2017. <http://www.alyaum.com/article/4099587>
- Arif, N.M., 1995. A study of the biography of the term and the significance of the concept. International Institute of Islamic Thought (IIIT), USA.
- Ashki, M.S., 2006. Islamic approaches and principles of dialogue. Fellowship Papers, Salam Institute.
- Bengio, Y., A. Courville and P. Vincent, 2013. Representation learning: A review and new perspectives. IEEE Trans. Patt. Anal. Mach. Intell., 35: 1798-1828. DOI: 10.1109/TPAMI.2013.50
- Bollen, J., H. Mao and A. Pepe, 2011. Modeling public mood and emotion: Twitter sentiment and socio-economic phenomena. Proceedings of the 5th International AAAI Conference on Weblogs and Social Media, Jul. 17-21, Barcelona, Spain, pp: 450-453.
- Boudad, N., R. Faizi, R. Oulad Haj Thami and R. Chiheb, 2017. Sentiment analysis in Arabic: A review of the literature, Ain Shams Eng. J. DOI: 10.1016/j.asej.2017.04.007
- Cambria, E. and H. Hussain, 2012. Sentic Computing: Techniques, Tools and Applications. 1st Edn., Springer, Paises Bajos, ISBN-10: 9400750706, pp: 153.
- Choi, M., 2011. Employees' attitudes toward organizational change: A literature review. Hum. Resour. Manage., 50: 479-500. DOI: 10.1002/hrm.20434
- Desai, T., A. Shariff, A. Shariff, M. Kats and X. Fang *et al.*, 2012. Tweeting the meeting: An in-depth analysis of twitter activity at kidney week 2011. PLoS ONE, 7: e40253-e40253. DOI: 10.1371/journal.pone.0040253
- El Ballouli, R., W. El-Hajj, A. Ghandour, S. Elbassuoni and H. Hajj *et al.*, 2017. Proceedings of the 3rd Arabic Natural Language Processing Workshop, Co-located with EACL 2017). Valencia, Spain.
- Khan, F.H. S. Bashir and U. Qamar, 2014. TOM: Twitter opinion mining framework using hybrid classification scheme. Decision Support Syst., 57: 245-257. DOI: 10.1016/j.dss.2013.09.004
- Kobasa, S.C., 1979. Stressful life events, personality and health: An inquiry into hardiness. J. Personality Soc. Psychol., 37: 1-11. DOI: 10.1037/0022-3514.37.1.1

- Kohavi, R., 1995. A study of cross-validation and bootstrap for accuracy estimation and model selection. Proceedings of the 14th International Joint Conference on Artificial Intelligence, Aug. 20-25, Morgan Kaufmann, San Mateo, CA, pp: 1137-1143.
- Li, R., K.H. Lei, R. Khadiwala and K.C. Chang, 2012. TEDAS: A twitter-based event detection and analysis system. Proceedings of the IEEE 28th International Conference on Data Engineering, Apr. 1-5, IEEE Xplore Press, Washington, DC, USA, pp: 1273-1276. DOI: 10.1109/ICDE.2012.125
- MacEachren, A., A. Jaiswal, A. Robinson, S. Pezanowski and A. Savelyev *et al.*, 2011. SensePlace2: GeoTwitter analytics support for situational awareness. Proceedings of the IEEE Conference on Visual Analytics Science and Technology, Oct. 23-28, IEEE Xplore Press, Providence, RI, USA, pp: 181-190. DOI: 10.1109/VAST.2011.6102456
- Maddi, S.R. and S.C. Kobasa, 1984. The Hardy Executive. 1st Edn., Dow Jones-Irwin, Homewood, ISBN-10: 0870943812, pp: 131.
- Maynard, D. and K. Bontcheva, 2016. Challenges of evaluating sentiment analysis tools on social media. Proceedings of the Language Resources and Evaluation Conference, (REC' 16), Portoroz, Slovenia.
- Mellenbergh, G.J., 1989. Item bias and item response theory. *Int. J. Educ. Res.*, 13: 127-143. DOI: 10.1016/0883-0355(89)90002-5
- Mikhmar, I. M., 1996. Arental rejection and its relation to the psychological hardness of university students. *J. Psychol. Stud.*, 6: 275-299.
- Neuendorf, K.A., 2017. The Content Analysis Guidebook. 2nd Edn., SAGE Publications, Inc.
- Pfeffer, J., 2014. Social media for large studies of behavior. *Science*, 346: 1063-1064. DOI: 10.1126/science.346.6213.1063
- Prilleltensky, I., 1989. Psychology and the status quo. *Am. Psychol.*, 44: 795-802. DOI: 10.1037/0003-066X.44.5.795
- Rappaport, J. and E. Seidman, 2000. Handbook of Community Psychology. 1st Edn., Springer, New York, ISBN-10: 0306461609, pp: 1011
- Refaee, E. and V. Rieser, 2014. An Arabic twitter corpus for subjectivity and sentiment analysis. Proceedings of the 9th International Conference on Language Resources and Evaluation, (LRE' 14), Reykjavik, Iceland.
- Sánchez, C.R., M. Craglia and A.K. Bregt, 2017. New data sources for social indicators: the case study of contacting politicians by Twitter. *Int. J. Digital Earth*, 10: 829-845. DOI: 10.1080/17538947.2016.1259361
- Schwartz, H.A., 2016. Predicting individual well-being through the language of social media. *Pacific Symp. Biocomput.*, 21: 516-527. PMID: 26776214
- Ullmann, S., 2017. Conceptualising force in the context of the Arab Revolutions: A comparative analysis of international mass media reports and Twitter posts. *Discourse Commun.*, 11: 160-178. DOI: 10.1177/1750481317691859
- Urbina, S. and A. Anastasi, 1997. Psychological Testing. 7th Edn., Prentice Hall, Upper Saddle River, NJ, ISBN-13: 9780023030857.
- Vosoughi, S., H. Zhou and D. Roy, 2015. Enhanced twitter sentiment classification using contextual information. Proceedings of the 6th Workshop on Computational Approaches to Subjectivity, Sentiment and Social Media Analysis, Sept. 17-17, Lisboa, Portugal, pp: 16-24. DOI: 10.18653/v1/W15-2904