Hybrid Sentiment Analyser for Arabic Tweets using R

Sarah Alhumoud, Tarfa Albuhairi and Wejdan Alohaideb
College of Computer and Information Science, Al-Imam Muhammad Ibn Saud Islamic University, Riyadh, Saudi Arabia

Keywords: Sentiment Analysis, Data Mining, Machine Learning, Supervised Approach, Hybrid Learning.

Abstract: Harvesting meaning out of massively increasing data could be of great value for organizations. Twitter is one of the biggest public and freely available data sources. This paper presents a Hybrid learning implementation to sentiment analysis combining lexicon and supervised approaches. Analysing Arabic, Saudi dialect Twitter tweets to extract sentiments toward a specific topic. This was done using a dataset consisting of 3000 tweets collected in three domains. The obtained results confirm the superiority of the hybrid learning approach over the supervised and unsupervised approaches.

1 INTRODUCTION

Making one's voice reachable instantaneously to the globe was a sort of a dream before. In some parts of the world where opinion freedom is in its lower indexes there are barley any channels to vent and express opinions. Now, with the emergence of social networks, opinions could be shared with the globe in real time. The power of one video on YouTube, a post on Facebook or a tweet on Twitter can change a country's fate and can be an effective means of marketing (Dinh et al., 2014). A study done by (Storck, 2011) has displayed the role of social media in the Egyptian revolution and how it helped the revolutionaries to achieve their goals. It has been known now by decision makers in political and economic fields that they need to hear the client's voice and immediate opinions and to study their behaviour for them to gain further insight and move wisely. A study by (Dinh et al., 2014) has shown the effect of viral marketing on social networks and how to find the optimal seeding for the aids in social networks. An example of a mining application to social networks seeking value and meaning is IBM customer analytic tool which can measures customer sentiment through surveying data in social networks to make a company more aware of its customers, who they are and what they want (IBM, 2011).

Social networks' applications offer means by which people build cyber social bonds based on their opinions, orientations, speciality, or hubbies and preferences. Considering the low freedom of expression in the Arab region by classical means, social networks have been an important ventilation mechanism. According to freedom house 2015 report, the Arab region is mostly ranked as low in civil liberties and political rights (Freedom House, 2015). In 2014, the number of Arabic internet users was 135 million, with more than 71 million of them are social networks active users (Mohammed Bin Rashid School of Government, 2014). Figure 1(a) shows the Arab usage of social networks in 2014. Moreover, a report was released in 2014 stating that the number of Arab Twitter active users is about 6 million while 2.4 million of them are from Saudi Arabia which makes it the highest Arab country in Twitter usage as shown in figure 1(b) (Arab Social Media Report, 2014). In Saudi Arabia, 60% of social networks' users were using Twitter in the year of 2013 (Reyaee and Ahmed, 2015). Having this extensive use of Twitter by the Saudi tweeters implies streaming a large amount of data that could be a rich and an invaluable source for analysis and study.

Sentiment Analysis (SA) is considered one of the text mining tasks (Han, Kamber and Pei, 2000) and one of the Natural Language Processing (NLP) concepts (Liu, 2012). SA is mainly the process of classifying text into two classes positive and negative to conclude the writer's orientation towards a certain topic or subject (Liu, 2012). After reviewing the literature, two approaches to implement SA have been concluded. The first is supervised approach or corpus-based approach(Abdulla et al., 2013). The second approach to SA is the unsupervised approach also called lexicon-based approach (Liu, 2012) and (Alhumoud et al., 2015).

Alhumoud, S., Albuhairi, T. and Alohaideb, W..
Hybrid Sentiment Analyser for Arabic Tweets using R.
In *Proceedings of the 7th International Joint Conference on Knowledge Discovery, Knowledge Engineering and Knowledge Management (IC3K 2015) - Volume 1: KDIR*, pages 417-424
ISBN: 978-989-758-158-8

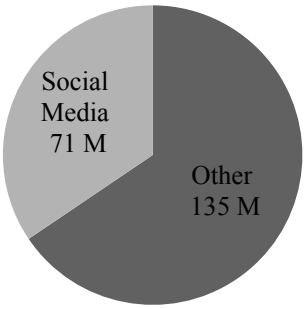

Figure 1(a): Arab internet usage in 2014.

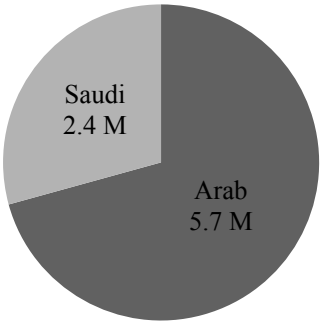

Figure 1(b): Arab Twitter usage in 2014.

One of the data mining tools that is used to implement SA is R programming language (R-project, 2015) R is an open source programming language that can perform data mining main tasks such as clustering, and classification (Kosorus, Honigl and Kung, 2011).

This research presents an implementation of SA using R programming language applying the hybrid learning approach and comparing it to the supervised and unsupervised learning approaches. The methodology and results of each approach will be discussed later in this paper.

This paper is organised as following: related work section where a list of similar work presented and discussed. Then, the methodology section that describes the method that was followed to implement the learning schemes. After that, the results and discussions' section, presenting results of each approach. Finally, the conclusion where the paper is wrapped up pointing out the marked findings.

2 RELATED WORK

Based on a survey that was published in 2014 by (Medhat, Hassan and Korashy, 2014) there are two main approaches for sentiment analysis. Those are machine learning and lexicon based approaches. Supervised learning (also known as corpus-based approach) falls under machine learning approach. There are several data mining algorithms that can be used to implement the supervised learning such as such as *Support Vector Machine* (SVM), *Naïve Bayes* (NB), *Decision Tree* (D-Tree) and *K-Nearest Neighbors* (KNN). The supervised learning consists of five sequenced stages based on multiple papers that implemented SA supervised learning such as (Shoukry and Rafea, 2012) and (Abdulla et al., 2013). These stages are: building the dataset, building the classifier (model), training the classifier, testing the classifier and finally using the classifier. Building the dataset is the first stage, creating labelled data and dividing it into training dataset and testing dataset. The second stage is building the classifier model where one of the data mining algorithms is chosen as a learning approach for the classifier. After that the classifier needs to be trained using the training dataset enabling it to classify new data. Once the classifier has been trained its accuracy is measured by testing it using the testing dataset. Where the classifier hide the labels and tries to classify the data based on the experience it gained in the training stage then it compares the classification result with the actual values. Finally, a new dataset with no labels is fed into the classifier to find its ability to classify based solely on previous learning.

Lexicon based approach, also known as unsupervised approach is a way where the classifier rely on number of lexicons helping it to classify the text. A lexicon is list of words or phrases, idioms or adjectives that hold a sentimental meaning. There are several ways to construct a lexicon, one of them is dictionary based. It starts with seeds of sentimental words then increasing the size of the dictionary by adding the synonym and antonyms (Medhat, Hassan and Korashy, 2014). Once a lexicon is completed it is ready for the classifier. An unsupervised classifier is an algorithm that performs matching between lexicon words and the words in the text. It usually treats words (in the lexicon and text) as numbers. Where positive words represented as positive integers and negative words as negative integers. Additionally, using certain mathematical processes it assigns the polarity of whole text as: +1, -1 or 0 for positive, negative or neutral, respectively (Shoukry and Rafea, 2012). There is no limitation for the number of lexicons that can be used for classification such as the work proposed by (Ibrahim, Abdou and Gheith, 2015) where they used two lexicons one holds sentimental words and the other for phrases and idioms.

The tools that are used for SA purposes are wide such as R programming language. The interest in R language for data mining is growing rapidly (KD Nuggets, 2013) (Magoulas and King, 2014). This increased interest in R is noticeable in both industry and academia. A key advantage of R programming language other than being an open source is data structures, visualization and packages. R provides a wide range of data structures that can represent data in understandable forms. Also, R provides the ability to transform results into graphs which makes it easier to understand. R has lots of packages that are specialized to different tasks and domains. (Jović, Brkićand and Bogunović, 2014). Authors (Kumar et al., 2011) performed clustering using R to implement a framework that provides a platform for developing data mining application. They used k-means and fuzzy k-means algorithms where fuzzy k-means got better results. While (Adamov, 2014) have used R language to pre-process http server logs for web usage mining. Moreover, the Authors (Nisa, Andrianto and Mardhiyyah, 2014) tried to solve the problem of forest fire by applying clustering function using R language.

R language is used to perform classification that solves problems of different domains. Work by (Keka and Hamiti, 2013) shows that they applied data mining to show electricity consumption over time. Electricity loads classification was based on two factors: seasons (summer and winter) and the usage intervals during the day. They applied unsupervised approach to achieve the loads classification. Another interesting work done by (Tsatsoulis and Hofmann, 2014) in that they tried to use data mining to classify albums of the rock musician Tom Waits into two classes. The first class is for albums that were produced before 1983 (period A) and second class for albums that were produced after 1983 (period B). They used word cloud package in R in order to help them to visualize the two classes. Their results using *RTextTools* package show that Maximum Entropy was able to classify the albums correctly with 95% of accuracy. A work presented by (Hosch, 2014) using R programming language and nine algorithms. The aim was to build a classifier that is able to classify open source projects into functional categories. Their experiments gave the classifier an accuracy of 81%.

R language has been used for SA purposes by many researchers as SA is one of the applications to classification (Ofek, 2015).The authors of (Fiaidhi et al., 2012) used TwitteR package to implement twitter client and used the *searchTwitter()* function to collect tweets from twitter. They applied an unsupervised approach. For the supervised approach authors (Horakova, 2015) have propose SA machine learning for Czech language. They hired *RTextTools* to build classifiers with five algorithms. They collected a dataset of 7000 users review about electronic devices. Also, a pre-processing model was applied. For filtering, they have used tagging, tokenization, stemming and lemmatization. The result of the experiments shows that SVM is among the best classifiers with 73% while NB is the worst with 51% accuracies respectively. Their results also show that all the classifiers have improved their accuracy when increasing the size of the dataset except for NB. They also, have experimented the classification on raw data and the results show a decrease in the accuracy. Table 1 compares between papers that used R to apply classification function, how they apply it, and to solve what problem.

3 METHODOLOGY

For the convenience of R language, the hybrid learning approach was implemented using R. R provided two packages that facilitates the implementation of SA, those are *RTextTools* and *tm* packages. *RTextTools* was designed to make machine learning accessible (Jurka et al., 2015). Whereas *tm* was designed to apply text mining techniques (Feinerer and Hornik, 2015). SA implementation was done using the SVM data mining algorithm.

In general, the hybrid learning approach contains five main stages: building the dataset, building the classifier (model), training the classifier, evaluating the classifier, and using the classifier to get overall sentiment of a new dataset as shown in Figure 2. The main difference between supervised and hybrid learning is in building the training dataset, this will be explained in the next section.

3.1 Building the Training Dataset

The training dataset is built from rows of single sentimental words and their labels. A labelled word tagged positive or negative according to the word sentiment. The training dataset contains 3690 sentimental words, 1370 words were positive and remaining 2320 words are negative. The training dataset contains 1000 MSA sentimental words are found in (Arabic MPQA subjective lexicon and Arabic opinion holder corpus, 2012) and 2690 Saudi dialect sentimental words were added manually with

Table 1: Comparison of classification approaches in related work.

Paper citation	Application	Data source	Dataset size	Pre-processing	Filtering	Results and approach
(Tsatsoulis and Hofmann, 2014)	Classify albums based on publishing periods by analysing song's lyrics.	Lyrics of rock musician Tom Waits	20 albums	▪ Converting to lowercase ▪ Removing punctuation	▪ Tokenizing ▪ Removing stop words	▪ Supervised ▪ Maximum Entropy 95%.
(Hosch, 2014)	Classify open source projects into functional categories.	Ohloh database	Not defined	-	▪ Tagging	▪ Supervised ▪ Used 8 algorithms ▪ Accuracy 81%
(Keka and Hamiti, 20113)	Electricity consumption over time.	Electrical working stations	2,976 observations	-	-	-
(Horakova, 2015)	SA for Czech language.	Czech portal heureka.cz	7,000 user reviews	Text pre-processing module.	▪ Tagging ▪ Tokenization ▪ Stemming ▪ Lemmatization	▪ Supervised ▪ Used 8 algorithms ▪ Best accuracy: ▪ Maximum Entropy ▪ Random Forests ▪ Increasing dataset size rise accuracy in all algorithms except NB. ▪ Effect of no filtering and pre-processing is decreasing accuracy.
(Fiaidhiet al., 2012)	SA on tweets	Twitter	1500 tweets	-	▪ TF-IDF	▪ Unsupervised

the help of 1000 sports tweets collected form Twitter.

In R language, the dataset need to be represented in a specific format such as Comma Separated Value (CSV). Create_matrix() function was used to generate the document term matrix. Several pre-processing options from the tm package are available, including stripping whitespace, removing sparse terms and setting number of minimum frequency of sentimental words in the whole document. In addition, word stemming, and stop word removal for several languages. However, currently those tools do not support Arabic language. Stop words were removed manually using simple algorithm with stop words' list. The hybrid learning approach could be implanted on languages other than Arabic, such as English. For the English language it is even more convenient and easy to use the full features of R SA packages.

In the supervised approach, each line contains one instance, a tweet. This instance has several words that do not affect the sentiment but causes confusion in the classifier, hence, decreasing the accuracy. If these words are removed from the instance, only sentimental words will remaining which is similar to the lexicon in the unsupervised approach. Using sentimental lexicon as a training dataset minimises the classifier confusion. Therefore, the hybrid learning incorporates the advantages of a data mining algorithm in supervised approach and lexicon based approach in unsupervised approach to better teach the classifier.

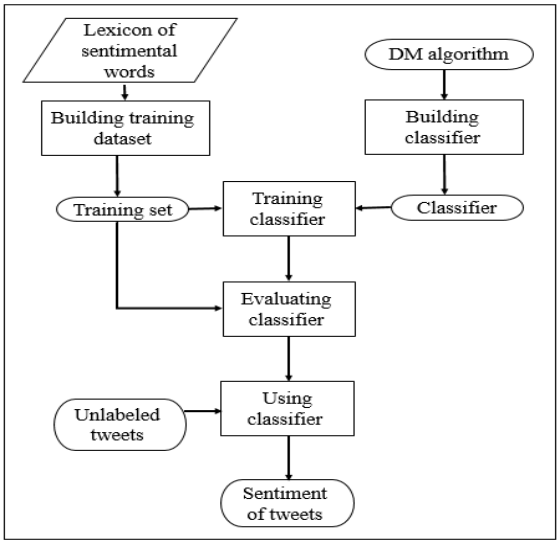

Figure 2: Classification stages of hybrid approach.

The difference between supervised and hybrid learning in building the training dataset reflected a remarkable difference in the results as well.

The hybrid learning also avoids the pre-processing cost associated with the supervised approach. Building a training dataset in the supervised learning requires several steps, first, collecting n instances, second pre-processing them with O(n) time complexity, third normalizing them with O(n) time complexity as well. While the hybrid learning approach does not require the second nor third steps.

3.2 Building the Classifier

Second step to the classification is building the classifier. SVM data mining algorithm was used to build the classifier. This algorithm works efficiently in text classification and it showed superior performance in previous related studies (Khasawneh et al., 2013), (Shoukry and Rafea, 2012) and (Abdulla et al., 2013).

3.3 Training the Classifier

Training the classifier is the third step in classification approaches. SVM classifier used linear kernel, unigram case and the minimum term frequency which is equal to one. Unigram means reading the dataset word by word, it is one of the n-gram cases which splits a string into n-grams with min and max gram.

3.4 Evaluating the Classifier

Evaluating the classifier is the fourth step in the classification which uses the training dataset as testing dataset and the produced classifier to expect the sentiment of each tweet. The result of the evaluation is measured by computing the precision, and recall. The precision and recall equations are presented below:

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

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

Where TP, FP, TN, and FN are true positive, false positive, true negative and false negative, respectively.

In precision and recall, the highest level of performance is equal to one and the lowest is zero (Sokolova et al., 2006). If precision and recall was more than 0.5, the classifier is accepted to classify new dataset.

3.5 using the Classifier

Fifth step comprises the classification of a new dataset by using the produced classifier to expect overall sentiment polarity of each tweet in the dataset. The tweets must be normalized before building the dataset. In addition, the dataset should be using the same format of the built container. Classification accuracy was measured by computing the number of correct classified tweets.

4 RESULTS AND DISCUSSIONS

This research presents the implementation of the hybrid learning approach compared to the supervised and unsupervised learning approaches. The same dataset is used for all three implementations. Both hybrid learning and supervised learning used SVM. Supervised classifier trained on sports domain using 1000 tweets. The unsupervised approach has two words' dictionaries, positive and negative. Positive words were 1370 words, while 2320 words are negative. These words are the same words used to train the hybrid classifier.

The hybrid learning classifier precision and recall were 0.960 and 0.970, respectively. Table 2 shows a comparison between the accuracy of the three classifiers hybrid learning, supervised and unsupervised, denoted by "Hyb", "Sup" and "Unsup" consecutively. These classifiers were classifying new dataset in different sizes.

Table 2: Accuracy of classifiers classifying sports domain.

Dataset Size	Hyb	Sup	Unsup
100	91.50%	90.00%	86.30%
250	90.60%	81.50%	80.70%
500	89.90%	80.80%	78.60%
1000	89.30%	84.30%	81.30%
Average	90.30%	84.20%	81.70%

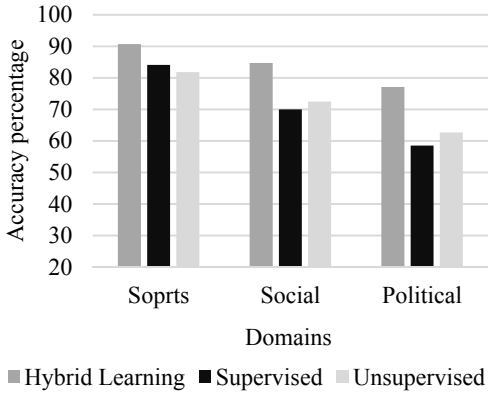

Figure 3: Accuracy for hybrid learning, supervised and unsupervised approaches in three different domains.

The results of classifying several datasets in sports domain proved superiority of the hybrid learning classifier over both supervised and unsupervised approaches scoring an improvement in

accuracy of 6% and 8% over supervised and unsupervised approaches respectively.

Figure 3 shows the comparison in accuracy of the three learning approaches in three domains, sports, social and political, each having 1000, 500 and 500 tweets respectively. The classifiers were trained using the sports domain. Testing the classifiers using new domains will exercise the learning scalability of the classifiers.

The proposed hybrid classifier scored an increase in accuracy of 7%, 15% and 19% in sports social and political domains respectively over the supervised approach.

Additionally, a similar enhancement was noticeable with the hybrid learning over the unsupervised approach scoring 9%, 12% and 14% in sports social and political domains respectively. The results show that when using the same SVM classifiers to classify new datasets in new domains, the hybrid learning approach achieve better accuracy than supervised and unsupervised. Proving a better scalability in different domains.

Moreover, increasing the training dataset size in hybrid learning is easier than the supervised approach as it does not require the pre-processing steps as in the supervised approach.

Enhancing the classifier needs some insight on the reasons that might lead to the incorrect classification. The following are examples of situations that may lead to classifier confusion or error:

- A tweet that has a negation word and this inverts the tweet polarity.
- A tweet having two opposite sentiments and could be classified to the wrong polarity.
- A tweet that is having an ambiguous or unknown sentiment to the classifier.
- The sentiment of a tweet is specified by a non-sentimental word that could not be added to the training set
- Words that have two different sentiments in two different domains

The hybrid learning approach's accuracy can be improved by increasing the size of the training dataset, and by using a words' stemmer.

5 CONCLUSIONS

This paper presented the new hybrid learning approach for Arabic SA in Twitter examining the classification of randomly collect tweets in three domains. The results confirm that the hybrid learning approach has better accuracy than both supervised and unsupervised approaches. The hybrid learning approach scored an enhancement in accuracy of 7%, 15% and 19% in sports, social and political domains over the supervised approach and a similar enhancement over the unsupervised approach. Moreover, the hybrid learning approach proves better scalability when applied to new different domains over the supervised and the unsupervised approaches. Another convenient feature of the hybrid approach is avoiding the overhead associated with the supervised pre-processing approach.

REFERENCES

Abdulla, N. Ahmed, N. Shehab, M. & Al-Ayyoub, M. (2013) *Arabic Sentiment Analysis: Lexicon-Based and Corpus-Based.* Proceedings of the IEEE Jordan Conference on Applied Electrical Engineering and Computing Technologies (AEECT). Amman, pp. 1- 6.

Adamov, A. (2014) *Data Mining and Analysis in Depth. Case Study of QAFQAZ University HTTP Server Log Analysis*. Proceedings of the IEEE 8th International Conference on Application of Information and Communication Technologies (AICT), Astana, pp.1-4.

Alhumoud, S. Altuwaijri, M. Albuhairi, T. & Alohaideb, W. (2015) *Survey on Arabic Sentiment Analysis in Twitter*. Proceedings of the International Conference on Computer Science and Information Technology (ICCSIT), Paris, pp. 364 – 368.

Arab Social Media. (2014) *Twitter in the Arab Region*, Dubai School of Government, [Online] Available from: https://shar.es/129INW. [Accessed: 16th June 2015].

Dinh, N. Huiyuan, Z. Nguyen, .T. & Thai, T. (2014) *Cost-Effective Viral Marketing for Time-Critical Campaigns in Large-Scale Social Networks*, IEEE/ACM Transactions on Networking, 22, pp.2001-2011

EthraArab. (2015) *قائمة Stop Words للهجة السعودية*, [Online] Available from: http://wp.me/p5VrNb-5o. [Accessed: 8th July 2015].

Feinerer, I. & Hornik, K. (2015) *Package 'tm'*. [Online] Available from: http://cran.r-project.org/web/packages/tm/tm.pdf. [Accessed: 17th June 2015].

Fiaidhi, J.Mohammed, O.Mohammed, S.Fong, S.& Kim, T. (2012) *Opinion Mining Over Twitterspace: Classifying Tweets Programmatically Using The R Approach*. Proceedings of the Seventh International Conference on Digital Information Management, Macau, pp. 313 - 319.

Freedom House, (2015) *Freedom in the World*, [Online] Available from: https://freedomhouse.org/report-

types/freedom-world#.VZcL2PkW6wN. [Accessed: 4th July 2015].

Han, J. Kamber, M. & Pei, J. (2000) *Data Mining: Concepts and Techniques*. Morgan Kaufmann.

Horakova, M. (2015) *Sentiment Analysis Tool Using Machine Learning*. Proceedings of the 2nd Global Conference on Computer Science, Software, Networks and Engineering. Turkey, pp. 192-204.

HOSCH, L. (2014) *Using Machine Learning To Classify Open Source Projects*. Department of Computer Science, Bachelor Thesis, University of Friedrich-Alexander. [Online] Available from: http://dirkriehle.com/uploads/byhand/theses/2014/hoesch_2014_arbeit.pdf. [Accessed: 16th June 2015].

IBM. (2011) *Customer Analytics Pay off*, IBM. [Online] Available from: http://www-01.ibm.com/software/analytics/rte/an/customer-analytics/. [Accessed: 1st September 2015].

Ibrahim, H. Abdou, S. & gheith, M. (2015) *Sentiment Analysis for Modern Standard Arabic and Colloquial*, International Journal on Natural Language Computing (4). [Online] Available from: arxiv.org/ftp/arxiv/papers/1505/1505.03105.pdf. [Accessed: 27th Aug 2015].

Jović, A. Brkić, K. & Bogunović, N. (2014) *An Overview Of Free Software Tools For General Data Mining*. Proceedings of the International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO), Opatija, pp. 1112 - 1117.

Jurka, T. Collingwood, L. Boydstun, A. Grossman, E. & Atteveldt, W. (2015*) Package 'RTextTools'*. [Online] Available from: http://cran.r-project.org/web/packages/RTextTools/RTextTools.pdf. [Accessed: 17th June 2015].

KD Nuggets. (2013) *Top Languages For Analytics, Data Mining, Data Science*, [Online] Available from: http://www.kdnuggets.com/2013/08/languages-for-analytics-data-mining-data-science.html. [Accessed: 16th June 2015].

Keka, I. & Hamiti, M. (2013) *Load Profile Analyses Using R Language*. Proceedings of the International Conference on Information Technology Interfaces (ITI), Cavtat, pp. 245 - 250.

Khasawneh, R. Wahsheh, H. Al Kabi M. & Aismadi, I. (2013) Sentiment Analysis of Arabic Social Media Content: A Comparative Study. Proceedings ofthe8th International Conference for Internet Technology and Secured Transactions (ICITST). London, pp.101-106.

Kosorus, H. Honigl, J. & Kung, J. (2011) *Using R, WEKA and RapidMiner in Time Series Analysis of Sensor Data for Structural Health Monitoring*. Proceedings of the 2nd International Workshop on Database and Expert Systems Applications, Toulouse, pp. 306 - 310.

Kumar, P. Ozisikyilmaz, B. Liao, W. Memik, G. & Choudhary A. (2011) *High Performance Data Mining Using R on Heterogeneous Platforms*. Proceedings of the 2011 IEEE International Symposium on Parallel and Distributed Processing Workshops and PHD Forum (IPDPSW), Shanghai, pp. 1720-1729.

Liu, B. (2012) *Sentiment Analysis and Opinion Mining*. Morgan & Claypool.

Magoulas, R. & King, J. (2014) *2013 Data Science Salary Survey Tools, Trends, What Pays (and What Doesn't) For Data Professionals*. O'Reilly, [Online] Available from: http://www.oreilly.com/data/free/files/strata survey.pdf. [Accessed: 16th June 2015].

Medhat, W, Hassan, A. Korashy H. (2014) *Sentiment analysis algorithms and applications: A survey*, Ain Shams Engineering Journal (5). P. 1093–1113. [Online] Available from: http://www.sciencedirect.com/science/article/pii/S2090447914000550.[Accessed: 27th Aug 2015].

Mohammed Bin Rashid School of Government. (2014) *The Arab World Online 2014: Trends in Internet and Mobile Usage in the Arab Region*, Mohammed Bin Rashid School of Government. [Online] Available from: http://www.mbrsg.ae/getattachment/ff70c2c5-0fce-405d-b23f-93c198d4ca44/The-Arab-World-Online-2014-Trends-in-Internet-and.aspxl. [Accessed: 2th July 2015].

Nisa, K. Andrianto, H. & Mardhiyyah, R. (2014) *Hotspot Clustering Using DBSCAN Algorithm and Shiny Web Framework*. Proceedings of the International Conference on Advanced Computer Science and Information Systems (ICACSIS), Jakarta, pp. 129-132.

NLP for Arabic. (2012) *Arabic MPQA Subjective Lexicon & Arabic Opinion Holder Corpus*. [Online] Available from: http://nlp4arabic.blogspot.com/2012/05/arabic-mpqa-subjective-lexicon-arabic.html. [Accessed: 17th June 2015].

Ofek, N. Rokach, L. Caragea, C. & Yen, J. (2015) *The Importance of Pronouns to Sentiment Analysis: Online Cancer Survivor Network Case Study*, Proceedings of the 24th International Conference on World Wide Web Companion, pp. 83-84.

Reyaee, S. & Ahmed, A. (2015) *Growth Pattern of Social Media Usage in Arab Gulf States: An Analytical Study*. Social Networking, 4, pp. 23-32.

R- project. (2015) *The R Project for Statistical Computing*, [Online] Available from: http://www.r-project.org/l. [Accessed: 4th July 2015].

Shoukry, A. & Rafea, A. (2012) *Sentence Level Arabic Sentiment Analysis*. Proceedings of the International Conference on Collaboration Technologies and Systems, Denver, USA, pp. 546 - 550.

Sokolova, M. Japkowicz, N. and Szpakowicz, S. (2006) *Beyond Accuracy, F-Score and ROC: A Family of Discriminant Measures For Performance Evaluation*. In Hutchison, D. Kanade, T. Kittler, J. Kleinberg, J.M. Mattern, F. Mitchell, J.C. Naor, M. Pandu Rangan, C. Steffen, B. Terzopoulos, D. Tygar, D. & Weikum, G. (Eds.). *AI 2006: Advances in Artificial Intelligence*. Lecture Notes in Computer Science (4304). Australia: Springer Berlin Heidelberg.

Stop Words (2014) *Stop-Words*[Online] Available From: Https://Code.Google.Com/P/Stop-Words/. [Accessed: 17th June 2015].

Storck, M. (2011) *The Role of Social Media in Political Mobilisation: a Case Study of the January 2011*

Egyptian Uprising. [Online] Available from: http://www.culturaldiplomacy.org/academy/content/pdf/participant-papers/2012-02-bifef/The_Role_of_Social_Media_in_Political_Mobilisation_-_Madeline_Storck.pdf/. [Accessed: 2th July 2015].

Tsatsoulis, C. & Hofmann, M. (2014) *Focusing On Maximum Entropy Classification of Lyrics byTom Waits*. Proceedings of the IEEE InternationalAdvance Computing Conference, Gurgaon, pp. 664 - 667.

Vinodhini G. & Chandrasekaran, RM. (2012) *Sentiment Analysis and Opinion Mining: A Survey*. International Journal of Advanced Research in Computer Science and Software Engineering. (2). P. 283- 292. [Online] Available from: http://www.dmi.unict.it/~faro/tesi/sentiment_analysis/SA2.pdf. [Accessed: 16th June 2015].