International Academic Institute for Science and Technology

International Academic Journal of Science and Engineering
Vol. 2, No. 2, 2015, pp. 164-174
ISSN 2454-3896

International Academic Journal of Science and Engineering

www.iaiest.com

Increase Business Intelligence Based on opinions mining in the Persian Reviews

Hamed Azizi Vamerzani[a] , Maryam Khademi[b]

[a] *Department of Computer Engineering, Islamic Azad University South Tehran Branch, No. 209, North Iranshahr St., Tehran, 11365-4435, Iran.*

[b] *Department of Applied Mathematics, Islamic Azad University South Tehran Branch, No. 209, North Iranshahr St., Tehran, 11365-4435, Iran.*

Abstract

By improving electronic business and managing system of customer's request, a huge load of Textual data has produced direct and indirect by users. Customers of one specific product and even managers need to observe comprehensive and effective information of obtained opinions in order to make the best decision in minimum time. Research shows that 81% of internet users have searched on the website before buying a product or their accessories. Unfortunately many of these researches are allocated to English language and just few projects have been done in Persian. In this paper, the provided framework has the ability of predicting the polarity of a text by user's textual review and extracting the critical Features and then by means of an improved SVM model, user's opinion about a product and its Features has been classified. The trained model can predict the polarity with a high degree of accuracy.

Keywords: Sentiment Analysis, Opinion Mining, Persian Language, Text Mining, Natural Language Processing

Introduction:

Human beings love to express their opinions. This car is better than that one. This political candidate is dishonest. That restaurant's food is delectable. And the World Wide Web has provided countless forums for collecting these opinions and presenting them to the world. But they also present a challenge as with other textual information, there is now so much of it that humans have difficulty navigating the sea of data. To address this issue, the field of opinion mining and sentiment analysis has arisen to provide automatic and semi-automatic methods for taking expressions of opinion in text and providing useful analysis and summarization for users.

Sentiment analysis or opinion mining is the computational study of people's opinions, appraisals, attitudes, and emotions toward entities, individuals, issues, events, topics and their attributes. The task is technically challenging and practically very useful. For example, businesses always want to find public or consumer opinions about their products and services. Potential customers also want to know the opinions of existing users before they use a service or purchase a product (Liu and Zhang, 2012).

Nevertheless, finding and reading numerous potential reviews is a time-consuming and tedious task. Therefore, designing tools to automatically extract people's opinion from their online comments has attracted an increasing attention in recent years.

Two major pieces of information available in each online review are its textual content and the numerical rating for each post from other Users which respectively indicate the aspects of customer concerns and the customer sentiment (Chung and Tseng, 2012).

Persian is an Indo-European language, spoken and written primarily in Iran, Afghanistan, and a part of Tajikistan. The amount of information in Persian language on the internet has increased in different forms. As the style of writing in Persian language is not firmly defined on the web, there are too many web pages in Persian with completely different writing styles for the same words. Different writing styles produce many challenges about Persian text processing such as, informal and colloquial words, declensional suffixes, multiple types of writing for a word and word spacing.

In this study, we study the main challenges of Persian language, and experiment our model on a Persian Cell-Phones review dataset. The Roadmap to the remaining part of the paper is as follows: Section 2 describes the related work; Section 3 explains the methodology proposed Section 4 presents the implementation details that have been undertaken and the section 5 concludes the paper with the future works explained.

RELATED WORKS

Sentiment classification for product reviews has recently attracted much attention from the natural language processing community. Researchers have investigated the effectiveness of different machine learning algorithms using various features to develop sentiment classifiers on languages like English or Chinese.

Presently, there are very few researches conducted on sentiment analysis for other languages like Persian. This is due to the fact that lexical resources and data collections for languages other than English are very limited and it is very time and labour-consuming to generate these resources. Some example works for sentiment analysis on languages other than English are as follows.

Syed et al. (Syed et al., 2010) proposed new approach based on the identification and extraction of SentiUnits from the given text in Urdu language, using shallow parsing. SentiUnits are the expressions, which contain the sentiment information in a sentence.

Gural Vural et al. (Vural et al., 2013) present a framework for unsupervised sentiment analysis in Turkish text documents. They customize the SentiStrength sentiment analysis library by translating its lexicon to Turkish. Then apply their framework to the problem of classifying the polarity of movie reviews.

Our focus on Persian language and as we Sayed earlier, there are very rare works for Persian sentiment analysis particular in text review analysis but some of the best works in Persian language proposed as follows.

Bagheri et al. (Bagheri and Saraee, 2014) Introduced sentiment classification using different feature selection methods for online customer reviews in Persian language. This paper considers the problem of sentiment classification using different feature selection methods for online customer reviews in Persian language. Three of the challenges of Persian text are using of a wide variety of declensional suffixes, different word spacing and many informal or colloquial words. They study these challenges by proposing a model for sentiment classification of Persian review documents. The proposed model is based on stemming and feature selection and is employed Naive Bayes algorithm for classification.

Shams et al. (Shams et al., 2012) Proposed new unsupervised LDA-based sentiment analysis and they called LDASA. They Generate Persian Clues, at the first phase, an automatic translation approach used to translate the existing English clues to Persian. Next, iterative refinement approach is used to correct the erroneous clues resulted from previous step. Then, topic-based polar sets are achieved from these clues and finally, each document is categorized into its related polarity using a classification algorithm. They experiment their model on Persian documents about hotels, cellphones and digital cameras and reached to 76% accuracy on average.

PERSIAN TEXTS CHALLENGES IN SENTIMENT ANALYSIS

Persian language grammar and writing is extremely complex and this complexity makes analysis harder than other languages, some Persian specific features which are different from other language such as English are proposed in below (Shamsfard, 2011).

Persian text is formal and informal writing but has same meaning, Table 1 show some examples of formal and informal (colloquial) writing.

Table 1. Example of Formal and informal word in Persian text

Formal Text	Informal Text	English Translate
گرونه	گران	expensive
نشون نمیده	نمایش نمی‌دهد	Don’t Show
باتریش	باتری آن	Battery

Another problem is different type of space. Sometimes space may occur between two words and sometimes within a word. More ever pseudo-space (short space) pseudo-space may be used as intra-word space to separate words part (Shamsfard, 2011). This problem affects correct word tokenizing in Persian

text. Thus, there are many words which can be written with space, short space, or without any spaces. For example " می خواهم، میخواهم ، میخواهم and Mikhaham" are all forms of 'I want. '
Adding space within a word usually occurs in multi-part words in which the first part ends with a multi-form letter as in " فیلم برداری or فیلمبرداری or Filmbardari " means 'video recording'.
Persian is a derivational and generative language in which many new words may be built by concatenating words and affixes. So the possibility of facing a new word that is not available in the lexicon is high (Shamsfard, 2011).
There are words in Persian that have same written forms with or without different pronunciations. Those with different pronunciations can be disambiguated with short vowels with are usually not written special in colloquial writing. For example "اشکال " if it is pronounced as "eshkal", then the word means 'problem' and if it is pronounced as "ashkal", then it means 'shapes' (Shamsfard, 2011).
Usually Persian-language users write English words between Persian sentence especially in their review about technology and science for example "ساپورت نمی‌کند OTG این گوشی " "This phone Support OTG Cable".
In Persian language like other languages some of the words or feelings in the news and due to the existence of dialogue, define their polarity Hot early as negative features in the mobile phone but in some entity like Steam Iron is a positive feature .
As well as many Persian speakers in reviews are good proverbs, irony they use For example, the handset is like a sturdy rock.
If a system works just based on a Sentiment dictionary, it will get into trouble during recognition of statements without Sentiment words. For example, "the phone has worked for a week".

PROPOSED METHOD

İn this paper, a framework has designed by means of that various text which has written by Persian-language users in popular websites like Digikala, that has the ability of forecasting the polarity of the statements based on an under supervision model with a high degree of accuracy and displaying the information according to the characteristics which user has commented about. Different parts of this framework will be explained later. The proposed method diagram is illustrated in Fig. 1.

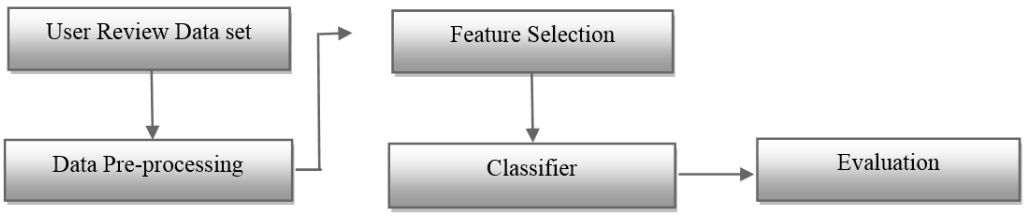


Figure 1. Proposed method diagram

User Review Data set

Two sets of data have used for the purpose of construction and implementation. The first one consists of 1000 users' review comments which divided into two positive and negative sets of 500.

Using these sets, our model has been trained and in next step, is capable to predict the desired result by means of a collection of test data which includes 100 review comments. (The data has been collected from an internet sales website).

Data Pre-Processing

Preprocessing of text in Persian language is of a great importance because of its specific challenges and complexities. There will be definitely no correct answer if the process is not done properly on text.

First step – text is tokenized: In other word, an input statement is separated into its components in order to carry out data processing operations properly.

Second step – remove stop words: "stop words" are some words that despite frequently repeated in the text, are semantically little importance. At first look, they appear to be the relevance and define words. However, most of verbs, auxiliary verbs, nouns, adverbs and adjectives are also stop words. In most applications of the text, deleting the words, significantly improves the processing results and result in calculation time reduction and increasing speed increasing. For this reason, these words are often deleted in preprocessing phase. İn this implementation, a list of stop words collected more than 1,500 entries

Third step – stemming: At this stage, for the purpose of assimilate the various forms of a word, integration and subsequent processing, words should be stemmed. Stemming is the process of converting words to their stem. As it was mentioned before, during writing a review, Persian-language users may have type their sentences in the form of speech or formal which show in Table 2.

Based on need and repeat essential words, such stemming has been created to be able to achieve the stem of both groups.

Table 2. Example of formal and informal words in Persian language

English Stem	Stem	Formal word	Informal word (colloquial)
Battery	باتری	باتریهای، باتری‌ها	باتریشم، باتریش، باتریم
Expensive	گران	گرانی	گرونه، گرونتر

It is noteworthy to mention that the stem of this section is not the exact word stem used in linguistics but a representative for words that are semantically and syntactically in the same field. This procedure in text processing is very important; because it makes a machine with two words of the same family but apparently different seems like two words with distinct stem.

Forth step - To explore some of the characteristics of a compound (two-parts) words, one needs to use 1-gram method. This case was investigating the repeated compound words. Several features were extracted from it. There are some examples of 2-gram method which extracted a word "آنتن-دهی" (Coverage) and "تراکم_پیکسل" (Pixel density).

Fifth step – word matrix-text: Vector space model is a model that is used in many classification algorithms. Here, too, the same model has been used to create vector words that each element of this vector shows the number of repetitions of the word, especially in a context.

$$doc_{i_{tf}} = [tf_1, tf_2, \text{K}, tf_n] \quad (1)$$

In vector 1, n is the number of unique words in the dictionary words and is the number of repetitions of the word i in user reviews. In general texts often contain words that increase the ability to text

classification but are frequently used in the literature. So it is better to introduce these words to the system with less weight. One way for weighting words as means to figure out this problem is utilization of IDF weighting. In this trend, the words that have come in a wide range of literature review have given less weight. The method is that the repetitions of words are multiplied in the phrase where here and demonstrate the total number of text user reviews and the number of repetitions of the word i in the literature, respectively. Hence the text word vector display with tf-idf can be shown like that is in Equation 3.

$$doc_{i_{tf}} = \left[tf_1 \times log\left({|D| \big/ df_1} \right), tf_2 \times log\left({|D| \big/ df_2} \right), \text{K}, tf_n \times log\left({|D| \big/ df_n} \right) \right] \quad (2)$$

$$c_i = \frac{1}{|c_i|} \sum_{j=1}^{P} d_j \quad (3)$$

In this paper, before performing the operation on the vector text, normalization is done on them as means to obtain a unit vector. Hence, each element of the vector has a value in the range [0, 1].

Feature Selection

Selecting the characteristics that here are the words is one of the most important points in the categories of literature. Opinions expressed by Users, are a set of positive and negative comments about various aspects of the issue. For creating a rating system for any features of a mobile phone, eight separate trees are designed that each child node, it leads to an important feature of a mobile phone. For example, "ژیرسکوپ" meaning "gyroscope" and "دو_سیمکارت" meaning "dual SIM" are each a feature that related to the "facilities" of a mobile phone. Each child nodes may also have children that are synonymous with that feature. For example, "acceleration sensor" is the child node of gyroscope. In this way, specific features will be extracted out of users' texts and according to the labels that each user's text already has, training step will start.

Classification

A supervised learning algorithm needs some of data to be labeled as training data and some of them as test data. These two datasets must be separate to prevent from false results in evaluating the performances of the methods. Therefore multiple runs of the experiments are usually needed with different datasets at each turn (Farhoodi and Yari, 2010).

In this research, we used Fast Large Margin SVM with L2 Logistic Regression Solver, This method applies a fast margin learner based on the linear support vector learning scheme proposed by RE Fan, et al (Fan et al., 2008). Although the result is similar to those delivered by classical SVM or logistic regression implementations, this method has many advantages than traditional SVM like this linear classifier is able to work on data set with millions of examples and attributes and supports two popular binary linear classifiers ،Logistic Regression and linear SVM. Given a set of instance-label pairs (xi, yi), i = 1, . . . , l, xi ∈ Rn, yi ∈ (Liu et al., 2014), both methods solve the following unconstrained optimization problem with different loss functions ξ(w; xi, yi):

$$\min_{w} = \frac{1}{2} w^T w + c \sum_{i=1}^{l} \xi(w, x_i, y_i) \quad (4)$$

Where C > 0 is a penalty parameter. For SVM, the two common loss functions are max (1 −yiwT xi,0) and max(1−yiwT xi,0).

In cross validation phase we use Repeated random sub-sampling validation .This method randomly splits the dataset into training and validation data. For each such split, the model is fit to the training data, and predictive accuracy is assessed using the validation data. The results are then averaged over the splits .

The advantage of this method (over k-fold cross validation) is that the proportion of the training/validation split is not dependent on the number of iterations (folds).

Evaluation

We use three evaluation measurements that are commonly used in sentiment analysis to evaluate the effectiveness of our proposed method, i.e., Precision, Recall and F-score. These evaluation measurements for each corpus are calculated by the following formulas:

$$precision = \frac{TP}{TP + FP} \quad (5)$$

$$Recall = \frac{TP}{TP + FN} \quad (6)$$

$$F - score = \frac{2 \times (Precision \times Recall)}{Precision + Recall} \quad (7)$$

Where, TP (True-Positive) is the number of sentences classified correctly as positive, FP(False-Positive) is the number of sentences that were incorrectly classified as positive, TN(true-negative) is the number of sentences correctly classified as negative and FN(False-Negative) is the number of sentences that are incorrectly classified as negative. In Table 3, the evaluation results of the proposed systems performance are presented.

Table 3. Performance Result

Methods	Training Time (s)	Recall	precision	F-score
Baghery et al (Naïve Bayes)	-	85.47	84.57	84.78
Fast Large Margin (SVM)	5.3	87.42	93.03	90.15

And finally, as the display model predictions and viewing features that users have been commenting about it. An example is given in Table 4 and Fig. 2.

Table 4. two example of positive and negative user review

ID	**User Comment**	**English Translate**
1	اصلا این گوشیو نخرید توی بازی ها سریع داغ میکنه و به نظرم وزن گوشیم خیلی بالاست و خیلی	Do not buy this phone, Phone gets hot fast when playing games And I think that weight is too

	از مواقع حسگر ضربان قلبش کار نمیکنه	high. And much of the time, heart rate sensor does not work.
2	گوشی فوق العاده ای هست من یکساله این گوشی رو دارم و واقعا راضی ام باتریش هم در حد متوسط هست وکیفیت صفحه نمایشش عالیه هنگ هم نمیکنه در کل با این قیمت واقعا توپه	The phone is great I bought this phone for a year. I'm really satisfied. Battery is moderate. Display Quality Awesome Never hanging , and with this price is Great.

xampleSet (10 examples, 5 special attributes, 8 regular attributes) Filter (10 / 10 examples): all

id	prediction...	confidence(negative)	confidence(positive)	text	ارزش	امکانات	باتری	دوربین	طراحی	نمایش	کارایی	گارانتی
1	negative	0.579	0.421	غرید توی بازی ها سریع داغ میکنه و به نظرم وزن گوشیم خیلی بالاست و خیلی از مواقع حسگر ضربان قلبش کار نمیکنه	0	0.082	0	0	0.096	0	0.158	0
2	positive	0.189	0.811	اقعا راضی ام باتریش هم در حد متوسط هست وکیفیت صفحه نمایشش عالیه هنگ هم نمیکنه در کل با این قیمت واقعا توپ	0.127	0	0.094	0	0	0.140	0.104	0
3	positive	0.355	0.645	سلام من دیروز این مدلشو از بازار خریدم گوشی عالی هستش و سرعت فوق العاده ای داره	0.126	0	0	0	0.125	0	0.052	0
4	positive	0.354	0.646	ضمن گوشی خیلی خوش دستی هم هست امکاناتشم نسبت به قیمتش خیلی خوبه.دوربینشم که حرف نداره.رندگارو خیلی گ	0.058	0	0	0.043	0	0	0	0
5	positive	0.355	0.645	ب نظر من گوشی زیبایی هستش و به نسبت قیمتش امکاناتش خوبه	0.168	0	0	0	0	0	0	0
6	positive	0.288	0.712	و چند تا نرم افزار و بازی براتون بریزن و همون جا نصب کنید.چون اگه اون نرم افزار ها رو براتون نریزن ممکن هست	0	0	0	0	0	0	0.187	0
7	negative	0.598	0.402	اه این گوشی را خریدم تایک ماه پیش خوب بود ولی یه روز که زدم تو شارژ دیگه روشن نشد وبعد شروع کرد به داغ شدن	0	0	0	0	0	0	0.160	0
8	negative	0.614	0.386	ن اسند جی510 دارم اصلا ازش راضی نیستم بیشتر دوستانم که هواوی دارن ازش راضی نیستن حالا به خودتون مربوطه	0	0	0	0	0	0	0	0
9	negative	0.632	0.368	قته وسرعتش بشدت پایین میاد و بعضی برنامه ها رو به درستی اجرا نمیکنه. با اینکه برنامه سنگین هم نصب نکردم.اص	0	0	0	0	0	0	0.189	0
10	negative	0.656	0.344	. روز استفاده متوجه این عیب ها در گوشی شدم لطفا راهنمایی کنید :1-خصاویر زمینه خود گوشی ر g6 سلام من دو ماهه	0.067	0.029	0.125	0	0.033	0.037	0	0.060

Figure 2. Prediction Result by Rapidminer

Fig. 2 shows that the result of forecasting 10 comments about a particular model of a mobile phone. After extraction of features that users have commented on, the below chart is illustrated

After obtaining the polarity of statements and the features that user has commented about, it is time for the results visualization. So a customer who is going to buy the product or even a mobile phone manufacturer can find advantages and disadvantages of that product in the shortest time with respect to the other users' opinions Fig. 3.

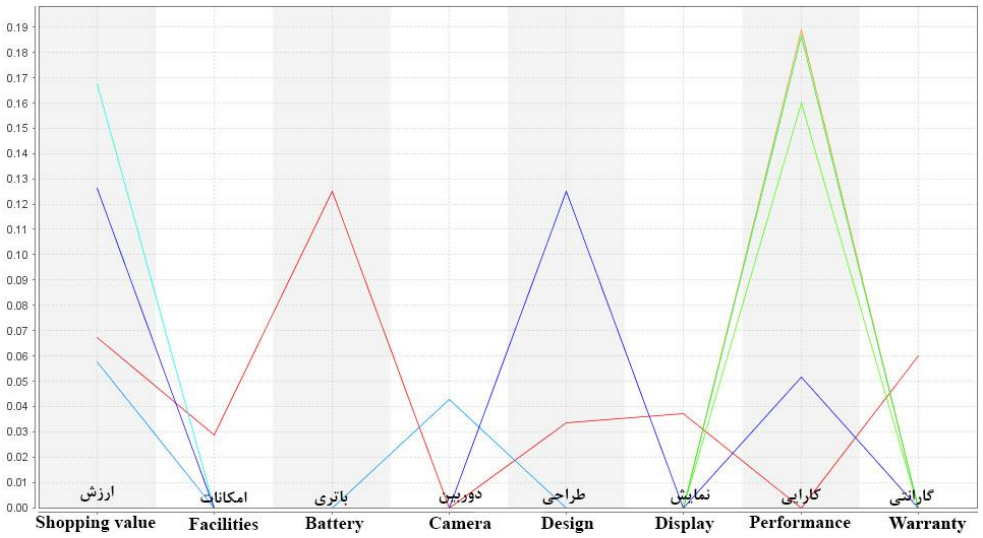


Figure 3. Visualization features Expressed by Users

CONCLUSIONS AND FUTURE WORKS

This model does not have any affiliation to a particular product or application and it can be easily implemented in any environment where customers say their opinions about products or services. Especially at customer request management system, many managers are seeking comprehensive and reliable reporting for clients and their feedback and naturally, in a large organization that offers services or products to customers that textual data is a very high volume production, by using this applied model can reach to very efficient and accurate reports by means of the data that customers provide for organizations in a fraction of a second.

Because of the special challenges that exist in Persian literature analysis, there are lots of things to do in order to identify the polarity of the mental sentences. And also, the large volume of text information and development tools, text information can be stored in the No-SQL databases like Hadoop and results can be discussed

References:

ADAMIC, L. 2014. Social Network Analysis [Online]. Coursera. Available: https://class.coursera.org/sna-004 2014].

ANJERANI, M. & MOEINI, A. Selecting influential nodes for detected communities in real-world social networks. Electrical Engineering (ICEE), 2011 19th Iranian Conference on, 2011. IEEE, 1-6.

BAGHERI, A. & SARAEE, M. 2014. Persian Sentiment Analyzer: A Framework based on a Novel Feature Selection Method. arXiv preprint arXiv:1412.8079.

CHALKIADAKIS, G., ELKIND, E. & WOOLDRIDGE, M. 2011. Basic Concepts. Computational aspects of cooperative game theory Synthesis Lectures on Artificial Intelligence and Machine Learning.

CHEN, W., WANG, C. & WANG, Y. Scalable influence maximization for prevalent viral marketing in large-scale social networks. Proceedings of the 16th ACM SIGKDD international conference on Knowledge discovery and data mining, 2010. ACM, 1029-1038.

CHUNG, W. & TSENG, T.-L. 2012. Discovering business intelligence from online product reviews: A rule-induction framework. Expert Systems with Applications, 39, 11870-11879.

CLAUSET, A., NEWMAN, M. E. & MOORE, C. 2004. Finding community structure in very large networks. Physical review E, 70, 066111.

FAN, R.-E., CHANG, K.-W., HSIEH, C.-J., WANG, X.-R. & LIN, C.-J. 2008. LIBLINEAR: A library for large linear classification. The Journal of Machine Learning Research, 9, 1871-1874.

FARHOODI, M. & YARI, A. Applying machine learning algorithms for automatic Persian text classification. Advanced Information Management and Service (IMS), 2010 6th International Conference on, Nov. 30 2010-Dec. 2 2010 2010. 318-323.

GLEISER, P. M. & DANON, L. 2003. Community structure in jazz. Community structure in jazz, 6, 565-573.

GOLDENBERG, J., LIBAI, B. & MULLER, E. 2001. Using complex systems analysis to advance marketing theory development: Modeling heterogeneity effects on new product growth through stochastic cellular automata. Academy of Marketing Science Review, 9, 1-18.

GOYAL, A., BONCHI, F., LAKSHMANAN, L. V. & VENKATASUBRAMANIAN, S. 2010. Approximation analysis of influence spread in social networks. arXiv preprint arXiv:1008.2005.

GRANOVETTER, M. 1978. Threshold models of collective behavior. American journal of sociology, 1420-1443.

ISELLA, L., STEHLÉ, J., BARRAT, A., CATTUTO, C., PINTON, J. F. & VAN DEN BROECK, W. 2011. What's in a crowd? Analysis of face-to-face behavioral networks. Journal of theoretical biology, 271, 166-180.

JACKSON, M. O. 2008. Representing and Measuring Networks. Social and economic networks. Princeton University Press Princeton.

KEMPE, D., KLEINBERG, J. & TARDOS, É. Maximizing the spread of influence through a social network. Proceedings of the ninth ACM SIGKDD international conference on Knowledge discovery and data mining, 2003. ACM, 137-146.

KEMPE, D., KLEINBERG, J. & TARDOS, É. 2005. Influential nodes in a diffusion model for social networks. Automata, languages and programming. Springer.

LESKOVEC, J., KRAUSE, A., GUESTRIN, C., FALOUTSOS, C., VANBRIESEN, J. & GLANCE, N. Cost-effective outbreak detection in networks. Proceedings of the 13th ACM SIGKDD international conference on Knowledge discovery and data mining, 2007. ACM, 420-429.

LESKOVEC, J. & MCAULEY, J. J. Learning to discover social circles in ego networks. Advances in neural information processing systems, 2012. 539-547.

LIU, B. & ZHANG, L. 2012. A Survey of Opinion Mining and Sentiment Analysis. In: AGGARWAL, C. C. & ZHAI, C. (eds.) Mining Text Data. Springer US.

LIU, L., SONG, W., WANG, H., LI, C. & JINGLI, L. 2014. A novel feature-based method for sentiment analysis of Chinese product reviews. Communications, China, 11, 154-164.

MICHALAK, T. P., AADITHYA, K. V., SZCZEPANSKI, P. L., RAVINDRAN, B. & JENNINGS, N. R. 2013. Efficient computation of the Shapley value for game-theoretic network centrality. Journal of Artificial Intelligence Research, 607-650.

NARAYANAM, R. & NARAHARI, Y. 2010. A shapley value-based approach to discover influential nodes in social networks. IEEE Transactions on Automation Science and Engineering, 1-18.

PANDIT, S., YANG, Y. & CHAWLA, N. V. Maximizing information spread through influence structures in social networks. Data Mining Workshops (ICDMW), 2012 IEEE 12th International Conference on, 2012. IEEE, 258-265.

SHAKMATT. 2015. NodeXL Sample Files [Online]. Available: http://nodexl.codeplex.com/wikipage?title=NodeXL%20Teaching%20Resources.

SHAMS, M., SHAKERY, A. & FAILI, H. A non-parametric LDA-based induction method for sentiment analysis. Artificial Intelligence and Signal Processing (AISP), 2012 16th CSI International Symposium on, 2-3 May 2012 2012. 216-221.

SHAMSFARD, M. 2011. Challenges and open problems in Persian text processing. Proceedings of LTC, 11.

SHAPLEY, L. S. 1952. A value for n-person games. DTIC Document.

SYED, A., ASLAM, M. & MARTINEZ-ENRIQUEZ, A. 2010. Lexicon Based Sentiment Analysis of Urdu Text Using SentiUnits. In: SIDOROV, G., HERNÁNDEZ AGUIRRE, A. & REYES GARCÍA, C. (eds.) Advances in Artificial Intelligence. Springer Berlin Heidelberg.

VURAL, A. G., CAMBAZOGLU, B. B., SENKUL, P. & TOKGOZ, Z. O. 2013. A Framework for Sentiment Analysis in Turkish: Application to Polarity Detection of Movie Reviews in Turkish. In: GELENBE, E. & LENT, R. (eds.) Computer and Information Sciences III. Springer London.

ZACHARY, W. W. 1977. An information flow model for conflict and fission in small groups. Journal of anthropological research, 452-473.

ZHOU, S., YUE, K., FANG, Q., ZHU, Y. & LIU, W. An efficient algorithm for influence maximization under linear threshold model. Control and Decision Conference (2014 CCDC), The 26th Chinese, 2014. IEEE, 5352-5357.