International Journal of Computing and Digital Systems
ISSN (2210-142X)
Int. J. Com. Dig. Sys. 7, No.3 (May-2018)
http://dx.doi.org/10.12785/ijcds/070304

# Comparative Study on Efficiency of Using Supervised Learning Techniques for Target-Dependent Sentiment Polarity Classification in Social Media

**Shadi Abudalfa[1]and Moataz Ahmed[2]**

[1]*Information & Computer Science Department / Collage of Computer Science and Engineering,*
[2]*King Fahd University of Petroleum and Minerals, Dhahran, Kingdom of Saudi Arabia*

*Received 24 Dec. 2017, Revised 29 Jan. 2017, Accepted 13 Feb. 2017, Published 1 May 2018*

**Abstract:** Classifying polarity of sentiments expressed in micro-blogs, such as tweets, is an active research area nowadays. The research direction has been focusing on classifying sentiments towards specific targets, i.e., topics, in the micro-blog. A more recent direction currently addresses the problem of detecting the target then identifying the sentiment toward it. While the former direction is referred to as target-dependent sentiment classification, the latter direction is referred to as open domain targeted sentiment classification. Many approaches have been proposed in the literature for automatic sentiment classification. Most of these approaches use supervised learning techniques that exploit only labeled data for training their proposed models. This paper presents an invited extension to a recent survey published by the authors. In this paper, we compile and present the accuracy reported by researchers with respect to the application of different techniques when applied to the same dataset. Our study presents comparisons between different techniques with regard to both the target-dependent and the open domain targeted sentiment classification. The study identifies some gaps to be addressed in future research. For instance, it shows that performance of both target-dependent and open domain targeted sentiment classification is still limited, and further future research could be promising.

**Keywords:** Text Mining, Social Opinions, Sentiment Analysis, Polarity Classification, Supervised Learning, Target-Dependent

## 1. INTRODUCTION

Nobody can deny effect of social media on societies nowadays. Anyone can login and share his sentiment (opinion) freely which leads to increase popularity and effect of social media. A lot of researches are done for classifying these sentiments automatically. Sentiment classification plays an important role in many applications of natural language processing (NLP) [1][2]. It is also one of the active research areas in text mining which have gained much attention nowadays [3][4]. Our research is focused specifically on sentiment polarity classification. The main goal of this research area is identifying opinions [5] and classifying polarities [6].

State-of-the-art techniques for sentiment identification deal with three different levels of input size: document, sentence, or word. The interested level of our research is the sentence level and especially a short sentence namely micro-blog in social media. Different online tools are manifested nowadays for opinion mining of micro-blogs. Typically, the input to such tools is a short sentence that is gathered from the social media by querying about a specified target (what the opinion is about). The output is the opinion polarity that is inferred from the input text and expressed in one of three options: positive, negative, or neutral.

In this paper, we compile and present the accuracy reported by researchers with respect to the application of different supervised learning techniques when applied to the same dataset. We summarize previous related works and find the gaps for suggesting some future works.

This work is an extension to a recent survey by the authors that has just shown that most of such supervised learning techniques have been applied on target independent sentiment classification and little ones are used with target-dependent approach [7]. Moreover, they use different datasets as a result we cannot make comparisons between these related works. This gap encouraged us to find researches on target-dependent sentiment analysis that using the same dataset.

The rest of the paper is organized as follows: Section 2 introduces theoretical background and describes evaluation metrics. Section 3 presents researches for target-dependent sentiment classification applied to the

*E-mail:shadi_abudalfa@hotmail.com, g201305390@kfupm.edu.sa,moataz@kfupm.edu.sa*

same dataset. Section 4 previews researches for open domain targeted sentiment classification applied to the same dataset. Finally, Section 5 concludes the paper and presents suggestions for future work.

## 2. BACKGROUND

Most of developed tools for sentiment analysis are based on target independent strategy for identifying sentiment in micro-blogs. Thus, these applications may fail to assign correct sentiment to a micro-blog that includes more than one topic (target). For example, consider this micro-blog: "Windows is much better than iOS!" A target independent system would always classify this micro-blog as positive since it contains only positive words (much better). However, if "iOS" is a target of interest, a target dependent system would classify this micro-blog as negative. Otherwise, it would be classified as positive if the target is "Windows".

A more challengeable scenario deals with detecting the name entities (topics or targets) in the micro-blog and identifying sentiments toward them. Referring to the above example, the system will detect firstly words "Windows" and "iOS" as topics and then identify opinions toward them as discussed previously. Such scenarios are referred to as open domain targeted sentiment classification which helps in detecting opinions for many related topics such as identifying opinions for a company along with its products and facilities. Next two subsections describe theoretical background for achieving open domain targeted sentiment classification.

### A. Name Entity Recognition

Name entity recognition [8] is a basic task in natural language processing (NLP). The task of name entity recognition and classification identifies named entities (such as name of person or organization) in readable text (such as micro-blog). The output of this operation is a categorization tag that describes each named entity.

Open domain targeted sentiment analysis uses name entity recognition [9] to identify all named entities in the micro-blog. The next phase in open domain targeted sentiment classification is based on determining which name entity represents the targeted topic in the selected micro-blog. After finding the targeted name entity, we can follow the same steps that are used in target-dependent sentiment classification to complete processes of open domain targeted sentiment classification.

### B. Sequence Labeling

Since entity recognition deals with entities (elements) in the input sentence (such as micro-blog), the research direction is shifted from sentence level into word (token) level. As a result, we need to deal with a sequence of words (tokens) that form each micro-blog. The most famous method that is used for classifying sequence of tokens is called sequence labeling. Sequence labeling [10] is used broadly in NLP for classifying each token instead of classifying the whole micro-blog.

In open domain targeted sentiment analysis, each micro-blog is represented as a sentence of tokens. Then sequence labeling identifies all words that are related to names and classify them as persons, organizations, etc. The typical way to set this up as a sequence labeling problem is called BIO tagging. Each token is labeled as "B" (beginning) tag if it is the first word in a named entity, or it is labeled as "I" (inside) tag if it is a subsequent token in a named entity, otherwise the word will be tagged as "O" (outside) tag. We can use other encoding strategy with sequence labeling but BIO tagging is the most famous one and it is used also with open domain targeted sentiment classification.

There are three approaches can be used for developing sequence labeling in open domain targeted sentiment classification. The first one converts the problem into a traditional classifying method by using BIO encoding, then we can use any classifier such as SVM [11]. The second approach uses deep learning with neural network for building the model of open domain targeted sentiment classification. The third approach uses hidden Markov models such as HMM [12] and CRF [13] for building the model of open domain targeted sentiment classification.

### C. Evaluation Metrics

Empirical results obtained from experiments provide a good way to evaluate sentiment classification systems. In this section, we describe measures that are used to assess solutions for target-dependent and open domain targeted sentiment classification. These measures are used against labeled tweets that are collected from Twitter.

#### a) Accuracy

The accuracy is the ratio of all samples which are classified correctly. We can simply calculate it by using the following formula:

$$Accuracy = \frac{Number\_of\_Correclty\_Classified\_Samples}{Number\_of\_All\_Samples} \times 100\% \quad (1)$$

#### b) Precision

Precision is the ratio of samples which are correctly classified as positive to all samples classified as positive.

#### c) Recall

Recall (which also known as sensitivity or true positive rate) is the ratio of samples which are classified correctly as positive to all positive samples.

#### d) F1-score

The F1-score (also known as F-score or F-measure) is the harmonic mean of precision and recall, and its best value is 1 while the worst score is 0. It is calculated as:

$$F1-score = 2 \times \frac{\mathrm{Pr}ecision \times \mathrm{Re}call}{\mathrm{Pr}ecision + \mathrm{Re}call} \quad (2)$$

The F1-score is basically used with binary classification and there are different modifications [14] to

use it with multiclass classification such as the macro-average F1-score, and the micro-average F1 score. The macro-average F1-score is straight forward. It is calculated by taking the average of the precision and recall of the system on different sets (each set is generated by using binary classifier with selected two classes). In micro-average F1-score, we firstly calculate the individual true positives, true negatives, false positives, and false negatives of each different set. Then use the sum of these values to find the micro-average precision and the micro-average recall. Finally the micro-average F1-score will be the harmonic mean of the micro-average precision and the micro-average recall.

We can use macro-average method for studying how the system performs across overall sets of data. From the other side, micro-average method can be used when dataset varies in size to come up with a specific decision.

*e) Acc-all*

This metric is used specifically with open domain targeted sentiment classification. It measures the accuracy of the entire named entity span along with the sentiment span. It primarily measures the correctness of O labels.

*f) Acc-Bsent*

This metric is used specifically with open domain targeted sentiment analysis. It measures accuracy of identifying the start of a named entity (B-labels) along with the sentiment expressed towards it. Thus is focuses only on the beginning of named entities.

## 3. TARGET-DEPENDENT SENTIMENT CLASSIFICATION

In this section, we present some researches that employ supervised learning techniques for target-dependent sentiment classification. Authors of these researches use the same dataset and try to add improvement with respect to classification accuracy and F1-score over previous researches in the state of the art. Next subsection describes the used dataset followed by results that are reported in these researches. We conclude the section by analyzing the results.

### *A. Describing Dataset*

The dataset that is used in this direction is collected by authors of [15] and has been utilized by many other researchers[1] such as [16]. The dataset consists of 6248 tweets for training and 692 tweets for testing. The distribution of sentiment polarities of micro-blog (in both training and testing data) is 25% are positive tweets, 25% are negative tweets, and the rest 50% are neutral tweets.

### *B. Results and Analysis*

In this subsection, we present performance of using different supervised learning teachings for target-dependent sentiment classification. Table I describes all compared models in this work. All compared supervised learning models are reported in [17] except SSWE which is proposed by [18] and reported by [16] as comparable work. Table II presents a summary of the best achieved results by using supervised learning models that are described in Table I. The reported results show classification accuracy and macro-average F1-score.

TABLE I. DESCRIBING COMPARED METHODS FOR TARGET-DEPENDENT SENTIMENT CLASSIFICATION

| Method | Description |
|---|---|
| SSWE | Sentiment-specific word embedding model [18]. |
| SVM-indep | SVM classifier uses only target-independent features [19]. |
| SVM-dep | SVM classifier uses target-independent features concatenated with target-dependent features of [19]. |
| RecursiveNN | Standard recursive neural network with target-dependent dependency tree [15]. |
| AdaRNN-w/oE | Adaptive recursive neural network (RNN) [15]. |
| AdaRNN-w/E | Adaptive recursive neural network (RNN) [15]. |
| AdaRNN-comb | Adaptive recursive neural network (RNN) [15]. |
| Target-dep | SVM classifier uses rich target-independent and target-dependent features [16]. |
| Target-dep+ | SVM classifier uses rich target-independent, target-dependent, and sentiment lexicon features [16]. |
| LSTM | Long short-term memory model (recurrent neural network) uses Glove vector. It classifies target-dependent sentiment based on target independent strategy [17]. |
| TD-LSTM | Target-Dependent LSTM [17]. |
| TC-LSTM | Target-Connection LSTM [17]. |

As shown in the Table II, each research provides better results in comparison with previous ones. The best achieved F1 score is 69.9% while the best classification accuracy is 71.5%. We can notice that the best accuracy and F1 score are reported by different researches. This means that the improvement is not significant between these two researches. We can notice also that each additive improvement is very small in comparison with previous ones. Moreover, the best achieved result is still limited (did not exceed 71.5%). Thus, it is obvious that more work might be done as future work for improving classification accuracy and macro-average F1-score.

TABLE II. BEST ACCURACY AND F1-SCORES ACHIEVED FOR TARGET-DEPENDENT SENTIMENT CLASSIFICATION

| Method | Acc | Macro-F1 |
|---|---|---|
| SSWE | 62.4 | 60.5 |
| SVM-indep | 62.7 | 60.2 |
| SVM-dep | 63.4 | 63.3 |
| RecursiveNN | 63.0 | 62.8 |
| AdaRNN-w/oE | 64.9 | 64.4 |
| AdaRNN-w/E | 65.8 | 65.5 |
| AdaRNN-comb | 66.3 | 65.9 |
| Target-dep | 69.7 | 68.0 |
| Target-dep+ | 71.1 | **69.9** |
| LSTM | 66.5 | 64.7 |
| TD-LSTM | 70.8 | 69.0 |
| TC-LSTM | **71.5** | 69.5 |

[1]https://github.com/duytinvo/ijcai2015

## 4. Open Domain Targeted Sentiment Classification

In this section, we present some researches that employ supervised learning techniques for open domain targeted sentiment classification. Authors of these researches use the same dataset and try to add improvement with respect to precision, recall and F1-score over previous researches in the state of the art. Next subsection describes the used dataset followed by results that are reported in these researches. We conclude the section by analyzing the results.

### A. *Describing Dataset*

Experimental works in this research direction are conducted by using corpus (dataset) that is collected originally by authors of [20] which is available publically[2]. This corpus is used by other research works such as [21] and [22]. Thus, using the same corpus gives a possibility to make real comparisons with previous related works. The corpus includes both English and Spanish tweets where each word (token) is located in a separated line. Table III shows statistics of the corpus as illustrated in paper [21].

TABLE III. Dataset for Open Domain Sentiment Analysis

| Domain | #Sent | #Entities | #+ | #- | #0 |
|---|---|---|---|---|---|
| English | 2,350 | 3,288 | 707 | 275 | 2,306 |
| Spanish | 5,145 | 6,658 | 1,555 | 1,007 | 4,096 |

### B. *Results and Analysis*

Table IV shows a summary of best results that are achieved in the state of the art for open domain targeted sentiment classification. These results are reported in [21] and [22] for making comparisons with previous works. Since the open domain targeted sentiment classification includes two tasks (entity recognition and sentiment classification), the results show performance of these two tasks individually.

We cannot use directly metric of classification accuracy since open domain targeted sentiment classification consists of two tasks. Authors of [20] reported their results by using acc-all and acc-Bsent, but these metrics are not used be the other researches. Thus, the reported results in this table include only precision, recall, and F1-score for making accurate comparisons between the successive researches.

We can notice that the reported result by using Spanish tweets is better than result of using English tweets. This means that there is still an effort needed to improve accuracy of classifying English tweets. Moreover, the overall best result is still limited. Thus, we need to do more research in this direction for improving performance of open domain targeted sentiment classification.

It is interesting to clarify that the first research [20] proposed three models and next research [21] mimics these three models for proving efficiency of their proposed approach. While authors of [22] proposed a new model and they did not mimic the three former models. This provides a gab that may be filled in future work by applying the model of [22] to mimic the three basic models of open domain targeted sentiment classification.

It is clear also that research work [21] provides the highest precision when using collapsed model. This means that this approach returned substantially more relevant results than irrelevant ones. Thus, it important to give this model more attention in future works.

## 5. Conclusion and Future Work

In this work, we presented a comparative study concerning the different target-dependent sentiment classification techniques using different supervised learning approaches. Performance of the different techniques is compared based on the same dataset. The comparison considered techniques from two research directions: target-dependent sentiment classification and open domain targeted sentiment classification. Findings could be used for improving performance of sentiment classification techniques in the future.

Based on our observations, we can say that the best achieved results are still limited. The accuracy did not exceed 71.5% for target-dependent sentiment classification. Meanwhile, the resulted F1-score of open domain targeted sentiment classification did not overtake 44.13%. Clearly, there is a room for improvement. For example, employing different feature reduction schemes might improve performance. Likewise, using more pre-trained word embeddings, such as fastText [23], might result in better accuracy. Similarly, employing different kernels within SVM could improve accuracy.

Moreover, our survey reveals that, to the best of our knowledge, all techniques use only labeled data. These techniques suffer from the scarcity of data. One problem here is that available labeled data suffer from what is called the "annotation" problem. This problem arises when the judgment of human annotating the micro-blog is not that accurate [24][25]. Another problem is two-fold: first, it is not easy to assess the accuracy of such techniques on different datasets; second, and more important, the applicability of such techniques is limited to those datasets with labeled data, which are not in real life. This introduces the needs for semi- and un-supervised techniques.

Future work should investigate applicability and suitability of using different unsupervised and semi-supervised learning techniques for both target-dependent and open domain targeted sentiment analysis. Moreover,

[2]http://www.m-mitchell.com/code/index.html

merging more than one method of machine learning should be investigated as well.

## ACKNOWLEDGMENT

The authors wish to acknowledge King Fahd University of Petroleum and Minerals (KFUPM) for providing the facilities to carry out this research.

TABLE IV. BEST PRECISION, RECALL, AND F1-SCORES ACHIEVED FOR OPEN DOMAIN TARGETED SENTIMENT CLASSIFICATION

| Model | English | | | | | | Spanish | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Entity Recognition | | | Sentiment Analysis | | | Entity Recognition | | | Sentiment Analysis | | |
| | P. | R. | F1 | P. | R. | F1 | P. | R. | F1 | P. | R. | F1 |
| CRF_Pipeline [20] | 65.74 | 47.59 | 55.18 | 46.8 | 33.87 | 39.27 | 71.29 | 58.26 | 64.11 | 43.8 | 35.8 | 39.4 |
| CRF_Collapsed [20] | 54 | 42.69 | 47.66 | 38.4 | 30.38 | 33.9 | 62.2 | 52.08 | 56.66 | 39.39 | 32.96 | 35.87 |
| CRF_Joint [20] | 59.45 | 43.78 | 50.32 | 41.77 | 30.8 | 35.38 | 66.05 | 52.55 | 58.51 | 41.54 | 33.05 | 36.79 |
| Neural Net_Pipeline [21] | 60.69 | 51.63 | 55.67 | 43.71 | 37.12 | 40.06 | 70.77 | 62 | 65.76 | 46.55 | 40.57 | 43.04 |
| Neural Net_ Collapsed [21] | 64.16 | 44.98 | 52.58 | **48.35** | 32.84 | 38.36 | **73.51** | 53.3 | 61.71 | **49.85** | 34.53 | 40 |
| Neural Net_Joint [21] | 61.47 | 49.28 | 54.59 | 44.62 | 35.84 | 39.67 | 71.32 | 61.11 | 65.74 | 46.67 | 39.99 | 43.02 |
| Sentiment Scope(SS) [22] | 63.18 | 51.67 | 56.83 | 44.57 | 36.48 | 40.11 | 71.49 | 61.92 | 66.36 | 46.06 | 39.89 | 42.75 |
| SS (+word emb) [22] | **66.35** | **56.59** | **61.08** | 47.3 | **40.36** | **43.55** | 73.13 | **64.34** | **68.45** | 47.14 | **41.48** | **44.13** |
| SS (+POS tags) [22] | 65.14 | 55.32 | 59.83 | 45.96 | 39.04 | 42.21 | 71.55 | 62.72 | 66.84 | 45.92 | 40.25 | 42.89 |
| SS (semi) [22] | 63.93 | 54.53 | 58.85 | 44.49 | 37.93 | 40.94 | 70.17 | 64.15 | 67.02 | 44.12 | 40.34 | 42.14 |

## REFERENCES

[1] B. Pang, and L. Lee , "Opinion mining and sentiment analysis," *Foundations and Trends in Information Retrieval*, vol. 2, no. 1-2, pp. 1–135, 2008.

[2] A. Shoufan, and S. Al-Ameri "Natural Language Processing for Dialectical Arabic: A Survey," *Proc. the Second Workshop on Arabic Natural Language Processing*, pp. 36–48, Jul. 2015.

[3] S. Salloum, M. Al-Emran, and K. Shaalan "Mining Social Media Text: Extracting Knowledge from Facebook," *International Journal of Computing and Digital Systems*, vol. 6, no.2, Mar. 2017.

[4] O. Kotevska, and A. Lbath, "Sentiment Analysis of Social Sensors for Local Services Improvement," *International Journal of Computing and Digital Systems*, vol. 6, no.4, July. 2017.

[5] M. Tsytsarau, and T. Palpanas, "Survey on mining subjective data on the web," *Data Mining and Knowledge Discovery*, vol. 24, no. 3, pp. 478- 514, 2012.

[6] B. Pang, L. Lee, and S. Vaithyanathan, "Thumbs up? Sentiment classification using machine learning techniques," *Proc. the 2002 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pp. 79–86, 2002.

[7] S. Abudalfa, and M. Ahmed, "Survey on Target Dependent Sentiment Analysis of Micro-Blogs in Social Media," *Proc. the 9th IEEE GCC Conference & Exhibition*, pp. 7-10, May. 2017.

[8] E. Sang and F. Meulder, "Introduction to the CoNLL-2003 Shared Task: Language-Independent Named Entity Recognition," *Proc. the 7th Conference on Natural Language Learning*, vol. 4, pp. 142-147, 2003.

[9] L. Ratinov and D. Roth, "Design Challenges and Misconceptions in Named Entity Recognition," *Proc. the 13th Conference on Computational Natural Language Learning*, pp. 147-155, 2009.

[10] N. Nguyen and Y. Guo, "Comparisons of Sequence Labeling Algorithms and Extensions," *Proc. the 24th International Conference on Machine Learning*, pp. 681–688, 2007.

[11] C. Scholkopf, J. C. Burges, and A. J. Smola, *Advances in Kernel Methods: Support Vector Learning*, MIT Press, 1999.

[12] M. Stamp, "A Revealing Introduction to Hidden Markov Models," Department of Computer Science, San Jose State University, 2018.

[13] J. Lafferty, A. McCallum, and F. Pereira "Conditional random fields: Probabilistic models for segmenting and labeling sequence data," *Proc. 18th International Conference on Machine Learning*, pp. 282–289, 2001.

[14] S. Parambath, N. Usunier, and Y. Grandvalet, "Optimizing F-measures by cost- sensitive classification," *Proc. Neural Information Processing Systems (NIPS)*, pp. 2123–2131, 2014.

[15] L. Dong, F. Wei, C. Tan, D. Tang, M. Zhou, and K. Xu, "Adaptive Recursive Neural Network for Target-dependent Twitter Sentiment Classification," *Proc. the 52nd Annual Meeting of the Association for Computational Linguistics (Short Papers)*, pp. 49–54, Jun. 2014.

[16] D. Vo, and Y. Zhang, "Target-Dependent Twitter Sentiment Classification with Rich Automatic Features," *Proc. the Twenty-Fourth International Joint Conference on Artificial Intelligence (IJCAI)*, pp. 1347- 1353, 2015.

[17] D. Tang, B. Qin, X. Feng, and T. Liu, "Effective LSTMs for Target-Dependent Sentiment Classification," *Proc. the 26th International Conference on Computational Linguistics: Technical Papers*, pp. 3298–3307, Dec. 2016.

[18] D. Tang, F. Wei, N. Yang, M. Zhou, T. Liu, and B. Qin, "Learning sentiment-specific word embedding for Twitter sentiment classification," *Proc. the 52nd Annual Meeting of the Association for Computational Linguistics*, pp. 1555–1565, 2014.

[19] L. Jiang, M. Yu, M. Zhou, X. Liu, and T. Zhao, "Target-dependent Twitter Sentiment Classification," *Proc. the 49th Annual Meeting of the Association for Computational Linguistics*, pp. 151–160, Jun. 2011.

[20] M. Mitchell, J. Aguilar, T. Wilson, and B. Durme, "Open Domain Targeted Sentiment," *Proc. the 2013 Conference on Empirical Methods in Natural Language Processing*, pp. 1643-1654, Oct. 2013.

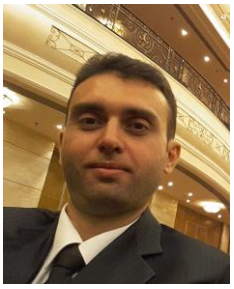
[21] M. Zhang, Y. Zhang, and D. Vo, "Neural Networks for Open Domain Targeted Sentiment," *Proc. the 2015 Conference on Empirical Methods in Natural Language Processing*, pp. 612–621, Sep. 2015.

[22] H. Li and W. Lu, "Learning Latent Sentiment Scopes for Entity-Level Sentiment Analysis," *Proc. the Thirty-First AAAI Conference on Artificial Intelligence (AAAI-17)*, pp. 3482- 3489, 2017.

[23] P. Bojanowski, E. Grave, A. Joulin, T. Mikolov, "Enriching Word Vectors with Subword Information," *arXiv, preprint, arXiv:1607.04606*, 2017.

[24] K.Veselovska, J. Hajic, Jr. and J Sindlerova, "Creating Annotated Resources for Polarity Classification in Czech," *Proc. KONVENS 2012 (PATHOS 2012 workshop)*, Sep. 2012.

[25] A. El-ghobashy, G. Attiya, and H. Kelash, "A Proposed Framework for Arabic Semantic Annotation Tool," *International Journal of Computing and Digital Systems*, vol. 3, no. 1, pp. 47-53, 2014.



**Shadi Abudalfa** received the BSc and MSc Degrees both in Computer Engineering from the Islamic University of Gaza (IUG), Palestine in 2003 and 2010 respectively. He is a lecturer at the University Collage of Applied Sciences (UCAS), Palestine. He is currently a PhD candidate in Computer Science and Engineering at King Fahd University of Petroleum and Minerals (KFUPM), Kingdom of Saudi Arabia. From July 2003 to August 2004, he worked as a research assistant at Projects and Research Lab in IUG. From February 2004 to August 2004, he worked as a teaching assistant at Faculty of Engineering in IUG. Abudalfa is a member of IEEE and his current research interests include artificial intelligence, data mining, data clustering, machine learning, and sentiment analysis.

**Moataz Ahmed** received his PhD in computer science from George Mason University in 1997. Dr. Ahmed is currently a faculty member with the Information and Computer Science Department, King Fahd University of Petroleum and Minerals, Kingdom of Saudi Arabia. He also serves as an Adjunct/Guest Professor in a number of universities in the US and Italy. During his career, he worked as a software architect in several software houses. His research interest includes soft computing-based software engineering, especially, soft- ware testing, software reuse, and cost estimation; and software metrics and quality models. He has supervised a number of theses and published a number of scientific papers in refereed journals and conferences in these areas.