AN EVALUATION OF SENTIMENT ANALYSIS ON SMART ENTERTAINMENT AND DEVICES REVIEWS

Donia Gamal, Marco Alfonse, El-Sayed M. El-Horbaty, Abdel-Badeeh M.Salem

Abstract: *Ever expanding utilization of the internet and online activities such as booking, blogging, e-commerce and conferencing, leads us to analyze very large quantities of structured data and unstructured data through Sentiment Analysis (SA). SA refers to the application of Natural Language Processing (NLP), computational linguistics, and data mining to classify whether the review is positive or negative. SA of this customer-generated data is extremely valuable to get a clearer perspective of public opinion and mood. In this paper, we evaluate the most popular Machine Learning (ML) algorithms such as Artificial Neural Networks (ANN), Support Vector Machines (SVM), and Nave Bayes (NB), which are utilized for SA on different user reviews datasets such as movie reviews, product reviews, and smart electronics devices over the last five years. The results show that using the ANN classifier along with the unigram as a feature extractor accomplishes a high accuracy 90.3%.*

Keywords: *Natural Language Processing Text Mining; Machine Learning; Opinion Mining (OM);Movie Reviews; Sentiment Analysis; Sentiment Classification (SC)*

ACM Classification Keywords: *I.2.7 Natural Language Processing, I.2.6 Learning*

Introduction

In recent years, the web is the fundamental birthplace of data. There are various internet business sites accessible where individuals examine on various issues of items. Users can share their experiments and their point of view of the public by using online websites like Amazon, blogs, IMDB, Yelp, and e-commerce. All of that enlarges the text datasets day by day due to the large collection of information in the form of electronic document. This data can be partitioned into two principles domains: facts and opinions, while facts concentrate on the objective information transmission, the opinions express the sentiment [Raghuvanshi and Patil, 2016].

Sentiment analysis is a computational study of opinions, attitudes, sentiments, and thoughts expressed in texts towards an entity. Sentiment analysis (also known as opinion mining, opinion extraction, reviews analysis or attitude analysis) is the task of identifying, extracting and classifying sentiments concerning various topics, expressed in a written text. Sentiment analysis helps in achieving different objectives such as observing public mood with respect to the political movement, market, the measurement of customer loyalty, movie sales prediction and much more.

Sentiment analysis aims to mine the written reviews of customers for a specific product by classifying the reviews into positive, negative or neutral opinions. It can be performed by using the ML approach, the Lexicon Based (LB) approach and the hybrid approach [Medhat et al., 2014].

Sentiment analysis has six main sub-tasks, which are sentiment classification, Sentiment Lexicon Generation (SLG), Sentiment Quantification (SQ), Opinion Extraction (OE), Feature-Based Summary (FBS), and Opinion Spam (OS). SC concerns with classifying a part of the text based on sentiment. The sentiment might be a judgment, attitude, mood or assessment of an object such as a film, hotel, book, smart device, product, etc., which can be in the form of three levels, which are a document level, sentence level or feature/aspect level. SLG is the task of marking words with a sentiment polarity to produce a sentiment lexicon. SQ is the task of estimating the prevalence of different sentiments for a given set of texts. OE is the task of extracting all opinions of the entities in user reviews, and categorizes them. FBS concerns with constructing a summary of the features. Features are product attributes, components and other aspects of the product [Buche et al., 2013]. OS is the task of detecting the spam content in data, such as fake and untruthful reviews and comments. In this paper, we concern with SC on different user reviews datasets.

The rest of the paper is organized as follows: section 2 demonstrates the data preparation, the background information related to the comparative study is presented in section 3, the evaluation of sentiment classification approaches is shown in section 4, the results of various ML algorithms with different feature extractors used for SC are discussed in section 5 and finally section 6 contains the challenges in sentiment analysis, conclusion, and future work

Data Preparation

Data procurement and data pre-processing are the most basic tasks required for SA. Online texts comprise usually lots of noise and uninformative parts, for example, HTML tags, Java scripts, hashtags, and advertisements. In addition, on the word level, many words in the text do not have an impact on the general orientation of it. Pre-processing the data considers removing the noisy redundant data and arranging the cleaned text for SA. Any data of high quality lead the analytical process to a better result in reducing the noise of the text improves the performance of the classifier and speeds up the classification process. There are some points that might help to have the data properly pre-processed, which are:

- Eliminating the most common stop words from being included in the process of Feature Extraction (FE).
- Stemming text, as the reviews are generally used with informal language. It is necessary to bring words into their original form, and losing out on potential features.
- Correct the spellings, since internet users usually use informal language, there are often wrong spellings in Text.
- Map the emoticons opinions; there are many emoticons that are used frequently in user reviews. Changing emoticons to positive or negative

opinions and eliminate emoticons that are ambiguous, unclear or unrelated to sentiments is a necessary task.

- Negation handling, it is a difficult process in SA as it inverses the polarity. Negation expresses by sarcasm and implicit sentences, which do not contain any negative words.

Literature Review

The idea of SA was at first shown by [Pang et al., 2003]. Few algorithms were utilized for SA, for example, Maximum Entropy (ME), Nave Bayes (NB) and SVM to accomplish SC. These algorithms are normally utilized for topic classification. The authors gathered movie reviews from IMDb.com. They explored different avenues regarding different FE methods where SVM yielded the most noteworthy accuracy 82.9% with unigrams features.

Lin and colleagues [Lin et al., 2012] used a Joint Sentiment-Topic (JST) model and a re-parameterized version of JST called Reverse-JST. While the greater part of the current ways to deal with SC favor supervised learning, both JST and reverse-JST models target sentiment and topic detection simultaneously in a weakly supervised fashion. The JST gives accuracy 71.2%, while reverse-JST gives accuracy 70.2%.

Smeureanu and Bucur [Smeureanu and Bucur, 2012] proposed an algorithm to classify sentiments of the users' reviews in a movie dataset based on NB taxonomy. The authors tested its performance on the movie review dataset, which gives accuracy 79.94%.

Mudinas and colleagues [Mudinas et al., 2012] introduced a new concept-level SA called pSenti that seamlessly integrates LB and ML approaches. In contrast with pure LBsystems, it achieves a better accuracy in SC with 82.30%.

Kalaivani and Shunmuganathan [Kalaivani and Shunmuganathan, 2012] compared three supervised ML algorithms, which are K-Nearest Neighbor (KNN), NB, and SVM for SC of the movie reviews. The results demonstrate that the SVM approach gives greater precision than NB and KNN approaches. The SVM approach gives over 80% precision.

Socher and colleagues [Socher et al., 2013] proposed a Recursive Neural Tensor Networks (RNTN) trained on the new tree-bank. The RNTN approach for a single sentence sentiment detection enhances the accuracy from 80% to 85.4%.

Liu and colleagues [Liu et al., 2013] introduced a fundamental framework for SA on large datasets utilizing NB with the Hadoop framework. The results exhibit that the NB classifier could scale up enough. The authors utilized a dataset with size exceeds 400K, the average accuracy remains below 82% under any circumstances.

Moraes and colleagues [Moraes et al., 2013] focused on analysing the behaviour of SVM and ANN regarding different ratios of positive and negative reviews. The results show that ANN can be a promising approach when the task includes sentiment learning, however, the SVM has a tendency to be steadier than ANN to manage noise terms in an unbalanced data context. The ANN has a better accuracy than SVM with 90.3% on cameras dataset of 1000 terms.

Basari and colleagues [Basari et al., 2013] utilized the SVM to recognize the patterns and analyse the data that are used for classification. They faced issues like tackling the double optimization. It was handled by using a hybrid Practical Swarm Optimization (PSO). The results indicate the improvement change of accuracy level from 71.87% to 77%.

Deng and colleagues [Deng et al., 2014] applied SVM combined with Importance of a Term in a Document (ITD) in order to extract features on various datasets. Their approach certainly beats BM25 on two of three datasets while the distinction is in distinctive on the small Cornell movie review dataset. The accuracy of the combined method is 87.44% on the Stanford movie review data set, which is greater than the BM25, which is 87.10%.

Tripathy and colleagues [Tripathy et al., 2015] aimed to apply the advances in deep learning, including more intuitive model architectures to the SC problem.

The authors performed a few tests with approaches that have customarily been utilized for SA, like SVM/A ne neural networks. Their proposed architecture, Recursive Neural Network and Recurrent Neural Network (RNN+RecNN), is able to accomplish accuracy of 83.88% without any handcrafted features at all.

Tripathi and Naganna [Tripathi and Naganna, 2015] presented a comparative study of various classification algorithms in combination with various FE methods. The results clearly show that the linear SVM gives more accuracy than NB with 84.75%. As well, the authors demonstrated that the accuracy increments for the bigrams which are in contrast with the results for [Pang et al., 2003].

Bhadane and colleagues [Bhadane et al., 2015] implemented a group of algorithms for aspect classification of product review using SVM mixed with domain specific lexicons. The results demonstrate that they have accomplished around 78% accuracy.

Tang and colleagues [Tang and Liu., 2015] presented Neural Network (NN) models for SC for document level, which are Convolution-Gated Recurrent Neural Network and Long Short Term Memory - Gated Recurrent Neural Network (Conv-GRNN and LSTM-GRNN). This approach encodes semantics of sentences and their relations in document representation, and is effectively trained end-to-end with supervised SC objectives. The results demonstrate that their approaches achieve accuracy on all these datasets with 63.7% on Yelp dataset.

Tripathy and colleagues [Tripathy et al., 2016] attempted to classify movie reviews utilizing numerous supervised ML algorithms, such as NB, ME, Stochastic Gradient Descent (SGD) and SVM. The authors applied n-gram approach on IMDb dataset. The NB acquired an accuracy of 86.23%, ME acquired an accuracy of 88.48%, SVM acquired an accuracy of 86.97% and SGD acquired accuracy of 85.11%.

Ashok and colleagues [Ashok et al., 2016] proposed two approaches in FE, where a stream of Cornell-movie review are pre-processed and classified. Then the authors applied various ML algorithms for SC, which are NB, RF, SVM, and

ME. For the first approach, the classification accuracy of ME is 77.17%. For the second approach, RF achieved accuracy 70.5%.

Povoda and colleagues [Povoda et al., 2016] applied two ML algorithms, which are NB and KNN for SC on two different datasets; the movie reviews and hotel reviews. The results show that the NB yielded better results for the movie reviews with 82.43% accuracy and beats KNN approach, which yielded 69.81%. However, for the hotel reviews, the accuracies are much lower and both the classifiers yielded similar results.

Wawre and Deshmukh [Wawre and Deshmukh, 2016] compared two supervised ML algorithms (NB and SVM) for SA on a movie review. The NB gives accuracy 65.57% comparing with SVM with accuracy 45.71% in SA of text.

Hegde and Seema [Hegde and Seema, 2017] proposed Incremental Decision Tree Classification (IDTC) that utilizes the iterative technique to classify product reviews. The results demonstrate that the SVM is much better compared to NB. The NB gives 78.44% and SVM gives 80.34 %, while IDTC gives 83.5%.

Evaluation of Sentiment Classification

In this section, we exhibit results produced by various ML algorithms with different FE methods and tested on different datasets, as movie reviews, and smart devices reviews (Electronics, GPS, etc.). Many researchers have focused on the use of traditional classifiers and others use ensembles of multiple classifiers to improve the accuracy of classification. The following tables present the comparison between different ML algorithms that are used for SA during the period 2012 to 2017.Table 1 focuses on movie reviews, while table 2 concerns with books and smart devices reviews.

Table 1. Evaluation of sentiment analysis using machine learning algorithms on movie reviews

Authors	Dataset	The Feature Extraction / Selection Method	The Classifier	Accuracy
[Smeureanu and Bucur ,2012]	Movie Reviews	Term Frequency	NB	79.94%.
[Mudinas, et al., 2012]	Movie Reviews	N/A	pSenti	82.30%
[Kalaivani and Shunmuganathan, 2012]	Movie Reviews	N-gram	NB	68.80%
			SVM	81.71%
			KNN	65.44%
[Socher et al., 2013]	The Stanford Sentiment Treebank (Based on Movie Reviews)	N-gram	RNTN	85.40%
[Liu et al., 2013]	Movie Reviews	Map Reduce	NB	82%
[Moraes et al., 2013]	Movie Reviews	Unigram	ANN	86%
			SVM	85.20%
			NB	72.50%
[Basari et al., 2013]	Movie Reviews	N-gram	SVM	72.20%
			SVM-PSO	76.20%
[Deng et al., 2014]	Cornell Movie Reviews	BM25	SVM	88.70%
		ITD		88.50%
	The Stanford Movie Reviews	BM25		87.10%
		ITD		88.00%
[Tripathy et al., 2015]	The Stanford Sentiment Treebank (Based on Movie Reviews)	No Handcrafted Features	RNN + RecNN	83.88%
[Tripathi and Naganna, 2015]	Movie Reviews	Term Occurrence	NB	70%
			SVM	75.25%
		Term Frequency	NB	68.50%
			SVM	84%
		Binary Term Occurrence	NB	70%
			SVM	76.50%

		TF-IDF	NB	67.50%
			SVM	84.75%
[Tang and Liu, 2015]	Yelp 2013	Document Representation	Bi-Gated-NN	63.70%
	Yelp 2014			65.50%
	Yelp 2015			66%
	IMDB			42.50%
[Tripathy et al., 2016]	Movie Reviews	Unigram	NB	83.65%
			ME	88.48%
			SVM	86.98%
			SGD	85.12%
		Bigram	NB	84.06%
			ME	83.23%
			SVM	83.87%
			SGD	62.36%
		Trigram	NB	70.53%
			ME	71.38%
			SVM	70.20%
			SGD	58.41%
		Unigram + Bigram	NB	86%
			ME	88.42%
			SVM	88.88%
			SGD	83.36%
		Bigram + Trigram	NB	83.83%
			ME	82.95%
			SVM	83.64%
			SGD	58.74%
		Unigram + Bigram + Trigram	NB	86.23%
			ME	83.36%
			SVM	88.94%
			SGD	83.34%
[Ashok et al., 2016]	Movie Reviews	Average Vector	NB	64.17%
			SVM (Gaussian kernel)	55.67%
			SVM (Linear kernel)	64.33%
			ME	51.33%
			RF	70.50%
		Bag-of-centroids	NB	65.17%
			SVM (Gaussian kernel)	68.33%
			SVM (Linear kernel)	73.17%
			ME	77.17%

			RF	51%
[Povoda et al., 2016]	Movie Reviews	Chi-squared	NB	82.43%
			SVM	69.81%
[Wawre and Deshmukh, 2016]	Movie Reviews	N/A	NB	65.57%
			SVM	45.71%

Table 2. ***Evaluation of sentiment analysis using machine learning algorithms on books and smart devices reviews***

Authors	Dataset	The Feature Extraction / Selection Method	The Classifier	Accuracy
[Lin et al., 2012]	Books Reviews	N-gram	Baseline	60.60%
			JST	70.50%
			Reverse JST	69.50%
	DVD Reviews		Baseline	59.20%
			JST	69.50%
			Reverse JST	66.40%
	Electronics Reviews		Baseline	58.60%
			JST	72.60%
			Reverse JST	72.80%
[Moraes et al., 2013]	GPS Reviews	Unigram	ANN	87.30%
			SVM	84.50%
			NB	65.10%
	Books Reviews		ANN	81.80%
			SVM	80.90%
			NB	76.20%
	Cameras Reviews		ANN	90.30%
			SVM	89.60%
			NB	81.80%
[Deng et al., 2014]	Amazon Product Reviews	BM25	SVM	87.70%
		ITD		88.70%
[Bhadane et al., 2015]	Product Reviews	N-gram	SVM	78%

[Hegde and Seema, 2017]	Product Reviews	Term Frequency	NB	79%
			SVM	82%
			IDTC	88.50%
		Binary term	NB	70%
			SVM	75%
			IDTC	75%
		TF-IDF	NB	67.50%
			SVM	70.50%
			IDTC	78.50%

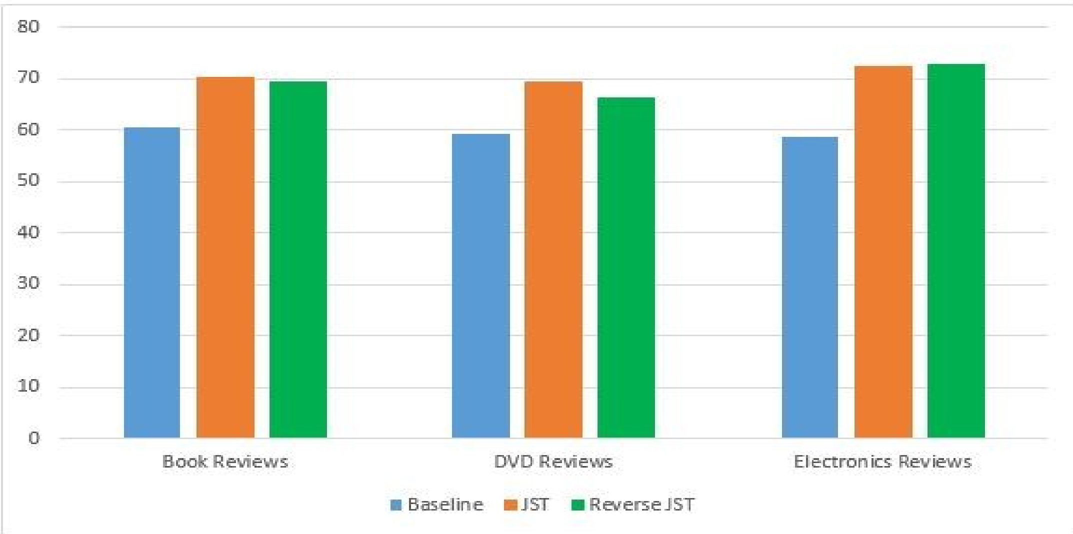


Figure1. Sentiment Classification Results using N-gram

Figure 1 shows the results of applying n-gram with baseline, JST, and reverse JST on different datasets, which are book reviews, DVD reviews and electronics reviews. The JST gives acceptable accuracies on these different datasets. The reverse JST gives the highest accuracy on electronics reviews and the baseline gives the lowest accuracy.

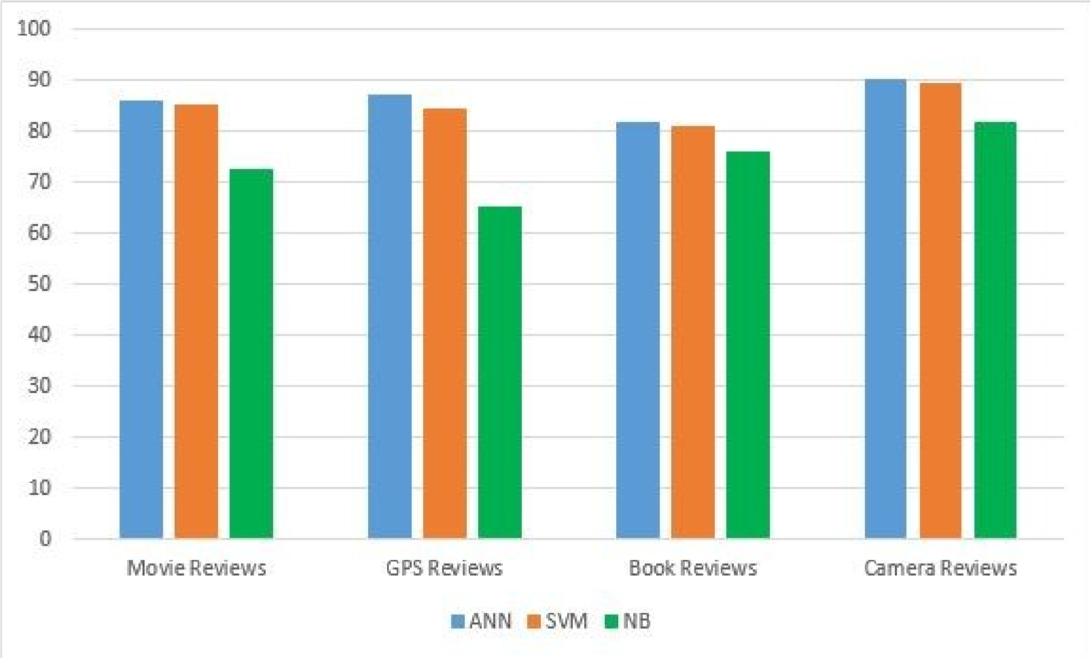


Figure 2: Sentiment Classification Results using Unigram

Figure 2 shows the results of applying the unigram with ANN, SVM, and NB on various datasets, which are movie reviews, GPS reviews, book reviews and camera reviews. Applying the ANN on the Camera reviews dataset gives the greatest accuracy among all the accuracies shown.

These results are to be illustrated in the next section.

Discussion

From the experiments above it is observed that the FE techniques indeed have an impact on the performance of a classifier. The most common feature extractors used are; n-gram, term frequency and Term Frequency-Inverse Document Frequency (TF-IDF). The most well-known ML algorithms used are SVM and NB, which accomplish high accuracy for classifying sentiment when combining different features. Some researchers used a hybrid approach to improve the accuracy of SC. The hybrid Practical Swarm Optimization classifier with the n-gram as a feature extractor gives a promising result, which is better than using a single SVM classifier with n-gram.

For the movie reviews dataset, the ANN with unigram, the SVM with (BM25, ITD, or unigram), and the ME with unigram give accuracies greater than 85%, the SVM with BM25 gives the highest accuracy among these algorithms. Some researchers combine multiple methods of FE to obtain a better accuracy. The integration of unigram and bigram with the SVM and the combination of unigram, bigram, and trigram with the SVM give accuracies greater than using SVM with a single FE (unigram, bigram, or trigram).

For the products reviews and smart devices datasets, the ANN with unigram, the SVM with (BM25 or ITD), and the ITDC with the term frequency give accuracies greater than 85%, the ANN with unigram gives the highest accuracy among these algorithms.

In addition, it is noticed that the removal of the low information gain features, also known as stop words, effects on the performance of a classifier, since the low information gain features are words that were not demonstrative of the sentiment and thus were not pertinent in deciding the polarity of a post.

Conclusions and Future work

This paper demonstrates a comprehensive, state-of-the-art review of the research work done in SC using the most common ML algorithms on different users' reviews datasets such as movie reviews and product reviews during the last five years.

We deduced from this comparative study that the most commonly used algorithms for sentiment classification on movie reviews and smart electronic devices are SVM and NB and the most commonly used feature extractor is N-gram. Among the discussed algorithms, a portion of ML algorithms is not exploited comprehensively such as; Naive Bayes Multinomial (NBM), Complement Naive Bayes (CNB), Sequential Minimal Optimization (SMO), Filtered Classifier (FC), Radial Basis Function Neural Network (RBFNN) and Logistic Model Tree (LMT).

The OM incorporates several challenges, which makes researchers concentrate on this important topic like:

- Opinions can contain numerous abbreviations, idiomatic expressions, orthographic mistakes, ironic sentences, or colloquial expressions.
- SA classification is domain dependent. Applying these algorithms to other domains requires adaptation.
- Time impact, opinions may change over time due to product improvement.
- Enhancing the accuracy of sentiment classification.

The future work is to utilize a hybrid supervised ML algorithm with different FE methods for SC to enhance the accuracy of SA classification.

Bibliography

[Ashok et al., 2016] Ashok, M., S. Rajanna, P. V. Joshi, S. Kamath, A personalized recommender system using machine learning based sentiment analysis over social data, Proceedings of IEEE Students’ Conference on Electrical, Electronics and Computer Science (SCEECS), pp. 1–6, 2016

[Basari et al., 2013] Basari, A. S. H., B. Hussin, I. G. P. Ananta, J. Zeniarja, Opinion mining of movie review using hybrid method of support vector machine and particle swarm optimization, International Journal of Procedia Engineering, vol.53, pp.453–462, 2013

[Bhadane et al., 2015] Bhadane, C., H. Dalal, H. Doshi, Sentiment analysis: Measuring opinions, International Conference on Advanced Computing Technologies and Applications (ICACTA), Procedia Computer Science, vol.45, pp.808–814, 2015.

[Buche et al., 2013]. Buche, A., D. Chandak, A. Zadgaonkar, Opinion mining and analysis: a survey, International Journal of Natural Language Computing (IJNLC), vol.2, no.3, pp.39–48, 2013.

[Deng et al., 2014] Deng, Z.-H., K.-H. Luo, H.-L. Yu., A study of supervised term weighting scheme for sentiment analysis, International Journal of Expert Systems with Applications, vol.41, no.7, pp.3506–3513, 2014.

[Hegde and Seema, 2017] Hegde, R., S. Seema., Aspect based feature extraction and sentiment classification of review data sets using incremental machine learning algorithm, Proceedings of Third International Conference on Advances in Electrical, Electronics, Information, Communication, and Bio-Informatics (AEEICB), IEEE, pp.122–125, 2017.

[Kalaivani and Shunmuganathan, 2012] Kalaivani, P., K. L. Shunmuganathan, Sentiment classification of movie reviews by supervised machine learning approaches, Indian Journal of Computer Science and Engineering, vol.4, no.4, pp.285–292, 2012.

[Lin et al., 2012] Lin, C., Y. He, R. Everson, S. Ruger, Weakly supervised joint sentiment-topic detection from text, International Journal of IEEE Transactions on Knowledge and Data engineering, vol.24, no.6, pp.1134–1145, 2012.

[Liu et al., 2013] Liu, B., E. Blasch, Y. Chen, D. Shen, G. Chen., Scalable sentiment classification for big data analysis using naive bayes classifier, Proceedings of the IEEE International Conference on Big Data, pp. 99–104, 2013.

[Medhat et al., 2014] Medhat, W., A. Hassan, H. Korashy, Sentiment analysis algorithms and applications: A survey, International Journal of Ain Shams Engineering Journal, vol.5, no.4, pp. 1093–1113, 2014.

[Moraes et al., 2013] Moraes, R., J. F. Valiati, W. P. G. Neto, Document-level sentiment classification: An empirical comparison between SVM and ANN, International Journal of Expert Systems with Applications, vol.40, no.2, pp.621–633, 2013.

[Mudinas, et al., 2012] Mudinas, A., D. Zhang, M. Levene, Combining lexicon and learning based approaches for concept-level sentiment analysis, Proceedings of the First International Workshop on Issues of Sentiment Discovery and Opinion Mining, ACM 5, 2012.

[Pang et al., 2003] Pang, B., L. Lee, S. Vaithyanathan, Thumbs up?: sentiment classification using machine learning techniques, Proceedings of the Association for Computational Linguistics (ACL-02) conference on Empirical methods in natural language processing, vol.10, pp.79–86, 2003.

[Povoda et al., 2016] Povoda, L., R. Burget, M. K. Dutta, Sentiment analysis based on support vector machine and big data, Proceedings of The 39th International Conference on Telecommunications and Signal Processing (TSP), IEEE, pp.543–545, 2016.

[Raghuvanshi and Patil, 2016] Raghuvanshi, N., J. M. Patil, A naive bays classifier to classify movie, International Journal of Advanced Research in Computer Engineering & Technology (IJARCET), vol.5 no.4, pp.1014–1017, 2016.

[Smeureanu and Bucur, 2012] Smeureanu, I., C. Bucur, Applying supervised opinion mining techniques on online user reviews, International Journal of Informatica economica, vol. 16, no.2, pp.81–91, 2012.

[Socher et al., 2013] Socher, R., A. Perelygin, J. Y. Wu, J. Chuang, C. D. Manning, A. Y. Ng, C. Potts, Recursive deep models for semantic compositionality over, a sentiment treebank, Proceedings of the conference on Empirical Methods in Natural Language Processing (EMNLP), pp.1631–1642, 2013.

[Tang and Liu, 2015] Tang, B. Q., Duyu, T. Liu, Document modeling with gated recurrent neural network for sentiment classification, Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP), pp.1422–1432, 2015.

[Tripathi and Naganna, 2015] Tripathi, G., S. Naganna, Feature selection and classification approach for sentiment analysis, International Journal of Machine Learning and Applications, vol. 2, pp.1–16, 2015.

[Tripathy et al., 2015] Tripathy, A., A. Agrawal, S. K. Rath, Classication of sentimental reviews using machine learning techniques, International Journal of Procedia Computer Science, vol.57, pp.821–829, 2015.

[Tripathy et al., 2016] Tripathy, A., A. Agrawal, S. K. Rath., Classification of sentiment reviews using n-gram machine learning approach, International Journal of Expert Systems with Applications, vol.57, pp.117–126, 2016.

[Wawre and Deshmukh, 2016] Wawre, S. V., S. N. Deshmukh, Sentiment classification using machine learning algorithms, International Journal of Science and Research (IJSR), vol.5, no.4, pp. 819-821, 2016.

Authors' Information

Donia Gamal *is a demonstrator at the Faculty of Computer and Information Sciences, Ain Shams University, Cairo, Egypt; She got her B.Sc in computer science with very good with honors in 2013.*

E-mail: *donia.gamaleldin@cis.asu.edu.eg*.

Her research interests: *Sentiment Analysis, Machine Learning, and Artificial Intelligence.*

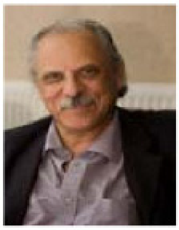
Marco Alfonse *is a Lecturer at the Faculty of Computer and Information Sciences, Ain Shams University, Cairo, Egypt. He got Ph.D. of Computer Science since August 2014, University of Ain Shams.*

His research interests: *Semantic Web, Ontological Engineering, Machine Learning, Medical Informatics, and Artificial Intelligence. He has 34 publications in refereed international journals and conferences..*

El-Sayed M. El-Horbaty: *He received his Ph.D. in Computer science from London University, U.K (1985)., his M.Sc. (1978) and B.Sc (1974) in Mathematics From Ain Shams University, Egypt.*

His work experience *includes 44 years as an in Egypt (Ain Shams University), Qatar(Qatar University) and Emirates (Emirates University, Ajman University and ADU University). He worked as*

Deputy Dean of the faculty of IT, Ajman University (2002-2008). He is working as a Vice Dean of the faculty of Computer & Information Sciences, Ain Shams University (2010-2017). Prof. El-Horbaty is current areas of research are parallel and distributed computing, combinatorial optimization, image processing, cloud computing, e-health and mobile cloud computing. His work appeared in journals such as Parallel Computing, International Journal of Mobile Network Design and Innovation, International Journal of bio-Medical Informatics and e-health, and International journal of Computers and Applications (IJCA), Applied Mathematics and Computation, and International Review on Computers and software. Also he has been involved in more than 26 conferences.



Abdel-Badeeh M Salem *is a Professor of Computer Science since 1989 at Ain Shams University, Egypt.*

His research includes *intelligent computing, knowledge-based systems, biomedical informatics, and intelligent e-learning. He has published around 250 papers in refereed journals and conferences. He has been involved in more than 400 Conferences and workshops as a Keynote Speaker, Scientific Program Committee, Organizer and Session Chair. He is a member of many national and international informatics associations.*