https://doi.org/10.17576/apjitm-2019-0801-07

http://www.ftsm.ukm.my/apjitm
Asia-Pacific Journal of Information Technology and Multimedia
Jurnal Teknologi Maklumat dan Multimedia Asia-Pasifik
Vol. 7 No. 2, June 2019): 79 - 96
e-ISSN: 2289-2192

ASPECT-BASED SENTIMENT ANALYSIS METHODS IN RECENT YEARS

ZOHREH MADHOUSHI
ABDUL RAZAK HAMDAN
SUHAILA ZAINUDIN

ABSTRACT

Sentiment Analysis (SA) is the computational treatment of opinions, sentiments and subjectivity of text. Aspect-based Sentiment Analysis (ABSA) is a specific SA that aims to extract most important aspects of an entity and predict the polarity of each aspect from the text. A review of the recent state-of-the-art in ABSA, shows the remarkable growing in finding both aspect, and the corresponding sentiment. Current methods are categorized based on their proposed algorithms and models. For each discussed study, aspect extraction method and sentiment prediction method, the dataset, domain and the reported performance is included. The main goal of this work is to review ABSA techniques with brief details. The main contributions of this paper consist of the refined categorizations of a great number of recent articles, comparing them and the illustration of the recent trend of research in the ABSA.

Keywords: Aspect-based Sentiment Analysis, aspect extraction, sentiment prediction

INTRODUCTION

"Capturing public opinion about social events, political movements, company strategies, marketing campaigns, and product preferences is garnering increasing interest from the scientific community for the exciting open challenges, and from the business world for the remarkable marketing fallouts and for possible financial market prediction" (Cambria et al. 2013).

This is the reason that a new field of opinion mining or Sentiment Analysis (SA) has emerged. Liu (2012) define SA as "*the field of study that analyses people's opinions, sentiments, evaluations, appraisals, attitudes, and emotions towards entities such as products, services, organizations, individuals, issues, events, topics, and their attributes."* SA is categorized into three levels; document, sentence, and phrase/aspect level (Liu 2012; Medhat et al. 2014). Document-level and sentence-level sentiments cannot provide sufficient information that is important for decision making. We can obtain such information with aspect level SA. If a reviewer gives a feedback on a particular product, he usually comments on some aspects of that product. This does not mean that the reviewer likes or dislikes that product totally. Although his overall opinion on the product can be positive or negative, the reviewer usually writes both positive and negative about different aspects of the product. This idea leads to ABSA, which was named as feature-based opinion mining by Hu and Liu (2004). The basic task in ABSA is to extract and summarize opinions that people express on entities and aspects of those entities.

A book chapter of Liu (2012) is on ABSA specifically. It describes ABSA, its methods and many sub-problems that arise from the main problem. For example, dealing with implicit and explicit aspect and sentiment. Another recent survey by Schouten & Frasincar (2016) is similar to this paper. Since this field is growing so fast, this current review covers more recent developments compare to the previous work. A new categorization of the ABSA methods is also included. A shortcoming of previous surveys is that they explain each task of ABSA separately without considering the fact that ABSA is a process and tasks need to be presented sequentially. Currently, most of the work uses accuracy, precision and recall performance evaluation, but some less common metrics are used as well. These methods are Ranking Loss, Mean Absolute Error (MAE), and Mean Squared Error (MSE), Normalized Discounted Cumulative Gain (nDCG) (Schouten & Frasincar 2016) and Rand Index. We are not covering the explanation about these measurements in this review.

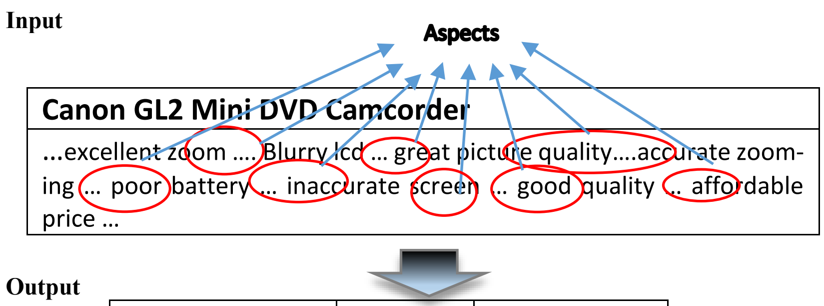
The main goal of this work is to review ABSA techniques with brief details. The main contributions of this paper consist of the refined categorizations of a great number of recent articles, comparing them and identification of the recent trend of research in this field. This paper is organized as follows: in the next section we present a basic definition about ABSA and discuss its tasks. Section 2 includes the review methodology. Section 3 presents recent ABSA techniques and their related articles. Section 4 includes result and discussion, and finally the conclusion and future work are tackled in Section 5.

DEFINITION

Liu (2012) define that an opinion (or sentiment) as a quintuple, '*(e, a, o, h, t)'*. In this quintuple '*e*' is the name of an entity, 'a' is an aspect of '*e*', '*o*' is the orientation of the sentiment about aspect '*a*' of entity '*e*', '*h*' is the sentiment holder, and '*t*' is the time when the sentiment is expressed by '*h*'. The sentiment orientation '*o*' can be positive, negative or neutral, or more detailed with different intensity levels. Given a set of text D, objective of ABSA is to discover all sentiment quintuples '*(e, a, o, h, t)* 'in D, or at least '*a*' and 'o' where '*e*' is fixed.

Based on the current review (Table 1, 2, 3, 4, 5) most of the recent works divide the ABSA process (showed in Figure 1) in to two tasks of aspect extraction and polarity/rating estimation, and some others try to join both tasks in one phase. There are also works such as by Lek & Poo (2013) and Pavlopoulos & Androutsopoulos (2014) that divide the ABSA process in three different tasks of aspect extraction, aspect- sentiment pair prediction and polarity/rating estimation.

Consider the two phase process. In the first task they extract aspects. The goal of this task is to extract aspects of the reviewed item and to group synonyms of aspects, as different people may use different words or phrases to refer to the same aspect, for example; display, screen, LCD. The second task is polarity/rating estimation. This task aims to determine the sentiment on the aspect. Whether it is positive, negative or neutral or has a numerical range (usually ranges from 1 to 5) (Liu 2012). The definition is illustrated in Figure 1 inspired by Moghaddam & Ester (2013).



Aspect	Rating	Polarity
Zoom	5	Positive
Price	4	Positive
Battery	2	Negative
screen	1	Negative
...	...	...

FIGURE 1. ABSA system input and output

RESEARCH METHODOLOGY

Since polarity/rating estimation in ABSA is usually performed as a consecutive task of aspect extraction, in this study we use the aspect extraction categorizations for grouping state-of-the-art techniques.

The articles presented in this review are summarized in five separated table (Table 1, 2, 3, 4, 5). Organization of all tables are the same and as follow: Different works in the literature is presented in the first column. Dataset of each work is presented in the second column. Domains that each work covers are presented in column three. If author present the result of the ABSA tasks in the work, it is shown in column four.

ABSA METHODS

Zhang & Liu (2014) classified aspect extraction methods into three categories of language rule, sequential and topic model. We add two other categories of Deep learning and hybrid methods.

Liu (2012) has categorized sentiment estimation methods into two groups of supervised and lexicon based methods. Since there are work with one, two or three different tasks for ABSA, we use aspect extraction method to categorize the methods and we will explain the rest of the methods in each category and name it the aspect polarity/rating estimation methods. Figure 1 shows ABSA methods and technique.

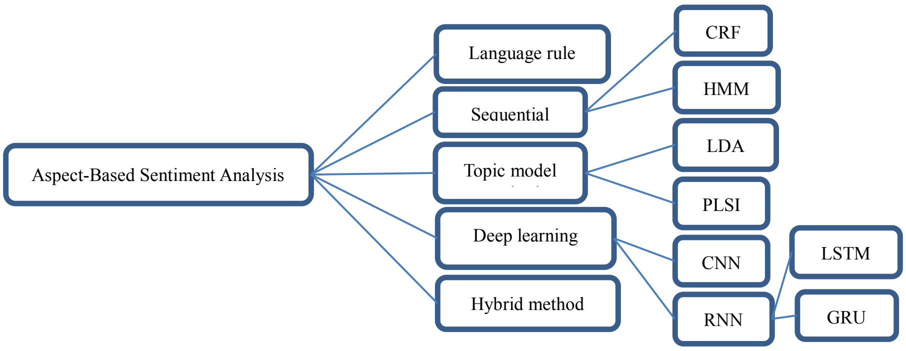


FIGURE 2. ABSA method and techniques

LANGUAGE RULE MODEL METHOD

Most of the early works on ABSA usually apply a set of filters on high-frequency noun phrases to identify aspects. An aspect can be expressed by a noun, adjective, verb or adverb. Usually in reviews, people talk about relevant aspects frequently which give the idea that aspects should be frequent nouns. But of course, not all of these frequent nouns are aspects. Therefore, different filtering techniques are applied on frequent nouns to filter out non-aspects. Most of these models try to find the most frequent nouns and noun phrases of the reviews in dataset, ordered by decreasing sentence frequency in the first step. Which means how many sentences contain the noun or noun phrase. Hu & Liu (2004) first determines all frequent noun phrases from full text reviews as candidate aspects. Then two pruning methods are applied to remove those candidate aspects with meaningless string, based on association rule mining, and those which are subsets of others (redundant). This is a reasonably effective and popular baseline (Liu 2012). Marrese-Taylor et al. (2014) improves the algorithm to estimate the orientation of sentence for compound aspects. In Twitter context, Lek & Poo (2013) takes all nouns, abbreviations, @mentions, or hashtags as candidate aspects. Closest adjective, verb, adverb, or hashtag in the right and left of each aspect considered as sentiment word. Poria et al. (2014) uses implicit aspect corpus. For each implicit aspect, synonyms and antonyms were obtained from WordNet and Semantics extracted from SenticNet. Then aspect parser is built based on several rules. Lizhen et al. (2014) uses dependency parser without extracting aspects and creates 6 tuple feature vector including feature, sentiment words, number of over-modifier of sentiment word, average score of general modifiers of sentiment word, number of negation words and the punctuation of the sentence. A new feature weighting algorithm is presented that improves TF (Term Frequency) (Luhn 1957) and TF-IDF (Term Frequency-Inverse Document Frequency (Sparck Jones 1972).

Zhang et al. (2014) present three methods, noun phrase extraction, Named Entity Recognition and a combination of both for aspect extraction. The semantic-based approach in Liu et al. (2014) is similar to Hu & Liu (2004). Moreover, they estimate personalized aspect polarity estimation for each individual user from his/her review. Their model removes irrelevant pairs of aspect-sentiment if they are not similar to any of the pre-defined aspects and then group relevant pairs into their corresponding aspect. A pair is grouped into an aspect if the semantic similarity between the noun of the pair and the pre-defined aspect word is above the specific threshold. WordNet Similarity is used to compute the semantic similarity between words.

The idea in Yaakub et al. (2012) is that customers usually gave comments on products based on some pre-known aspects. To extract the aspects from reviews; each frequent noun is mapped with an aspect ontology. If no useful relation between candidate aspect and aspect ontology is found, the candidate aspect is ignored. Then they extract the nearest sentiment word to group it with the aspect. The result is collected to create a data warehouse along with structured information. Finally, the author creates many data cubes to calculate and analyse the orientation of some groups of customers for products in certain levels based on the ontology that they have on a scale of -3 to +3 from strongly negative to strongly positive. More recent overview of the ontological approach can be found in Abu Latiffi and Yaakub (2018). de Albornoz et al. (2011) uses some predefined aspects and WordNet similarity to find other features. Then it computes a single score for each feature, based on the polarity of the sentences that is computed using three different classifiers.

Aspect-sentiment relationship is an advantage that some recent works use to extract new aspects and sentiments. The intuition behind these works is that most of the time sentiments are about aspects and sentiments are easier to find (Liu 2012). Therefore their relationship can be used for identifying new aspects (and sentiments). Zhuang et al. (2006) created a list of predefined aspect and also uses dependency parser to find aspects. It used GI (General Inquirer) and WordNet lexicons to complete the list of sentiment word. Then it identifies some templates of dependency relation in training data, and identify valid aspect-sentiment pairs in test data that follows one of those templates extracted from the training data. Wu et al. (2009) extends classical dependency parsing to phrase level. This parser is used to extract noun and verb phrases as aspect candidates. A normal dependency parser detects dependency of individual words only, while a phrase dependency parser detects dependency of phrases. Du & Tan (2009) utilize another type of relations between aspects and sentiments. They first consider all noun phrases as aspects and all adjectives as sentiments and then build a graph based on aspects and sentiments co-occurrence in reviews. A graph clustering algorithm (information reinforcement) is applied to find aspects highly related to sentiments.

Using syntactic structures of the aspects with exact matching is another simple approach for aspect extraction which cannot handle similar syntactic structures and therefore fails to generalize for new data. Liu et al. (2005) and Moghaddam & Ester (2010) employ this method. Approaches based on tree kernels are proposed to address this limitation. The kernel-based methods simply evaluate the similarity between two trees via computing a kernel function between them. Wu et al. (2009) use phrase dependency parsing. They proposed a new tree kernel function to model the phrase dependency trees. To extract the aspect and the related sentiment, the kernel is computed through training instances, then support vector machine (SVM) is used for classification. Peng et al. (2010) define several tree kernels for aspect-sentiment pair extraction and sentiment classification. They consider the aspect-sentiment pair extraction as an aspect-base sentiment classification task that treats the correct pairs as positive examples and incorrect pairs as negative examples.

Hai et al. (2011) utilized the co-occurrence matrix of aspect-sentiment for mining a set of rules for extraction of new pairs. In a similar work by Qiu et al. (2011) the dependency idea is improved to develop double-propagation method for extracting both sentiment words and aspects simultaneously. Liu et al. (2012) also utilized the relation between sentiment word and aspect for aspect-sentiment pair extraction. However, they formulated the sentiment relation identification between aspects and sentiment words as a word alignment task. They perform monolingual word alignment. The associations between aspects and sentiment words are measured by translation probabilities, which can capture sentiment relations between sentiment words and aspects more correctly than language rules or patterns. Lal & Asnani (2014) can be seen as an advanced extension of Hu &Liu (2004) method. It is designed specifically to identify aspects that mention implicitly in review sentences. Secondly, the approach distinguished sentiment words and aspect words with predefined rules. For example, sentiment words can only occur in the rule antecedents, while rule consequents must be aspects. Thirdly association rules are made directly from the Latent Semantic Analysis (LSA) matrix of sentiments and aspects. Landauer et al. (2013) explain that LSA is a mathematical and statistical approach for text vector representation. The idea is that semantic information can be derived from a word-document co-occurrence matrix and words and documents can be represented as points in a (high-dimensional) Euclidean space and dimensionality reduction is a crucial part of it. Table 1 shows language rule model techniques for ABSA.

TABLE 1. Language rule methods

Author	Dataset	Domain	Result (average)
Hu and Liu (2004)	FBS	Digital camera Cellular phone Mp3 player DVD player	Precious Recall Accuracy 0.693 0.642 0.842
Liu et al. (2005)	FBS	Digital camera Cellular phone Mp3 player DVD player	precision: 88.9/79.1 recall: 90.2 / 82.4
Zhuang et al. (2006)	IMDB	movie review	Precious Recall Accuracy 0.483 0.585 0.529
Blair-Goldensohn et al. (2008)	From Tripadvisor.com, maps.google.com,	Restaurant and Hotel	Static aspect classification results: Precious Recall F-score 85.2 66 74
Ding et al. (2008)	FBS	digital cameras, DVD player, MP3 player, cellular phones, router, anti-virus software	Precious Recall F-score 1.91 0.90 0.90
Wu et al. (2009)	FBS	Diaper, Cell Phone, Digital Camera, DVD Player, and MP3 Player	Precious Recall F-score 47.1 44.7 45.8
Du and Tan (2009)	From www.ctrip.com	Hotel reviews (in Chinese)	Precision of aspect sentiment pair: 78.90

Moghaddam and Ester (2010)	Create a dataset from Epinions.com	Camcorder Cellular Phone Digital Camera DVD Player MP3 Player	Aspect extraction Precious Recall 80 87 Polarity prediction Ranking Loss 0.49
de Albornoz et al. (2011)	Creates 1000 reviews from booking.com http://nil.fdi.ucm.es/index.ph p?q= node /456	hotel	F-score Logistic 70 LibSVM 69 FT 67

SEQUENTIAL MODEL METHOD

Aspect extraction can be seen as a special case of the general information extraction problem. The most prominent methods for information extraction are based on sequential learning (or sequential labelling). The current state-of-the-art sequential learning methods are HMM and CRF. These methods infer a function from labelled (supervised) training data to apply to unlabelled data. HMM is a generative probabilistic model with two dependency assumptions. One the hidden variable at time t, namely y_t, depends only on the previous hidden variable y_{t-1} (Markov assumption). Second the observable variable at time t, namely x_t, depends only on the hidden variable y_t at that time. The parameters are then learned by maximizing the joint probability distribution p (x, y). CRF is a discriminative probabilistic model that can come in many different forms. The form that most closely resembles the HMM is known as a linear-chain CRF. The parameters of a CRF are learned by maximizing the conditional probability distribution p(y|x).

Jin et al. (2009) propose a hybrid approach integrating POS information with the lexicalization technique under the HMM framework. In this model the current tag is related with the previous tag and also correlates with previous observations (word token and part of speech). Shariaty & Moghaddam (2011) proposes CRF-based models for fine grain sentiment analysis. For a more domain independent extraction, Jakob & Gurevych (2010) trained a CRF model on multi-domain review sentences. A list of domain independent features are such as tokens, POS tags, etc. is used. Another CRF-based mode is presented in Choi & Cardie (2010). In this work, a set of sequential rules are mined using a sequential rule mining technique considering class labels, dependency and word distance. The skip-tree CRF models are proposed by Li et al. (2010) to detect product aspects and sentiments. Many features are used such as Word features including word's token, lemma, part of speech, negation, superlative, comparative, dictionaries features including WordNet and SentiWordNet (Esuli & Sebastiani 2006), sentence feature including the number of negative and positive words identified by SentiWordNet, syntactic features from dependency tree such as parent word and its polarity, edge features such as conjunction words and syntactic relationship. Kiritchenko et al. (2014) presents a new sequence tagger for aspect terms extraction and supervised classifiers for aspect category detection.

The major limitation of the Language rule methods that we described in previous section is that they require us to tune various parameters manually. Therefore, the models cannot be generalized for unseen dataset. Sequential method overcome the limitations of language rule methods by learning the parameters from the data automatically. Although the supervised methods can achieve reasonable effectiveness, but the weakness of these models is that they require labelled data for training. Labelled data are not usually available. Building enough labelled data is very expensive and needs much human effort. Therefore, it is desired to develop a model that works with unlabelled data or less labelled data. Additionally, a wide range of product and service reviews is being written in many languages on the internet every day. Therefore, supervised, language-dependent and domain specific models are not practical. Table 2 shows sequential model method for ABSA.

TABLE 2. Sequential method

Author	Dataset	Domain	Result (average)
Jin et al. (2009)	FBS (camera), Amazon. com	camera	aspect extraction F1: 78.8% - 82.7% sentiment sentence extraction F1: 84.81% - 88.52% sentiment classification F1: 70.59% - 77.15%
Li et al. (2010)	-------	Movie, product	Precision Recall F-Score 82.6 76.2 79.3
Jakob & Gurevych (2010)	--------	----------	Precious Recall F-score 65.26 69.46 67.30
Choi & Cardie (2010)	The MPQA corpus	----------	Precision Recall F-Score 48.0 87.8 62.0
Shariaty & Moghad-dam (2011)	The authors Create 200 reviews from Epinion.com	Camcorder	Precision Recall F-Score 90 50 75
Ki-ritchenko et al. (2014)	SemEval-2014 dataset	Restaurant Laptop	Aspect extraction Restaurant: Precision Recall F-Score 84.41 76.37 80.19 Laptop: Precision Recall F-Score 78.77 60.70 68.57 Polarity prediction Restaurant: Acc. 70.49 Laptop: Acc. 80.16
Qiu et al. (2011)	FBS	Digital camera, DVD player, mp3 player, cell phone	Precious Recall F-score 0.88 0.83 0.86
Yu et al. (2011)	Reviews were crawled from cnet.com, viewpoints.com, reevoo.com and gsma-rena.com	Camera, Laptop, MP3 player, Phone	F 1 : 70.6 - 76.0
Liu et al. (2012)	http://irchina.org.cn/coae2008.html http://si-faka.cs.uiuc.edu/~wang296/Data/in-dex.html FBS	Camera, Car, Lap-top, Phone, Hotel, MP3, Restaurant	
Yaakub et al. (2012)	FBS	Digital camera, Cel-lular phone, Mp3 player, DVD player	Precious Recall F-score 0.890 0.872 0.880

Lek and Poo (2013)	- Stanford Twitter Sentiment (STS) - Sanders Twitter Corpus (STC) - Telecommunication Company - Dataset (TCD1/TCD2)	Twitter	TCD1 Precious Recall F-score 68.3 71.5 69.8 TCD2 Precious Recall F-score 68.9 68.9 68.9 STS Precious Recall F-score 93.7 44.0 59.9 STC Precious Recall F-score 87.5 46.8 61.0
Lal & Asnani (2014)	FBS	Cell phone	Aspects extraction: Precious Recall 67.0 83.6 Using LSA for aspect reduction (best result): Precious Recall 71.42 66.67 Rules Mined: Precious Recall 53.57 71.8
Zhang et al. (2014)	SamEval 2014 dataset	Laptop Restaurant	Aspect extraction: Laptop Restaurant F-score F-score 65.88 78.24 Aspect sentiment polarity estimation: Precious Recall F-score 65.26 69.46 67.30
Lizhen et al. (2014)	----------	Digital product	Precious Recall 67 83
Marrese-Taylor et al. (2014)	FBS	Tourism	Precious Recall F-score 72 76 73
Poria et al. (2014)	FBS http://alt.qcri.org/semeval2014/task4/index.php?id=data- and- tools	Digital camera, DVD player, mp3 player, cell phone Restaurant, laptop	
Liu et al. (2014)	TripAdvisor Yelp	Hotel Restaurant	

TOPIC MODEL METHOD

Current unsupervised models are topic models which are based on Probabilistic Latent Semantic Indexing (PLSI) and Latent Dirichlet Allocation (LDA). Latent Dirichlet Allocation (LDA) model is a generative probabilistic model for collections of discrete data such as text (Blei et al. 2003). The basic idea is that documents are represented as mixtures of latent topics, and topics are associated with a distribution of the words of the vocabulary.

Lu et al. (2009) first creates sentiment phrases (head term, modifier), and then models are learned to generate sentiment phrases only and not all the words of a review. They cluster the head terms using PLSI to extract aspects. The polarity of a head term is considered as the polarity of the corresponding short comment. Brody & Elhadad (2010) apply LDA model on reviews to extract topics as aspects. Moghaddam & Ester (2011) introduce Interdependent Latent Dirichlet Allocation (ILDA) model that utilize the assumption of interdependency between aspects and sentiments. In Chen et al. (2014), prior knowledge is learned automatically from a large amount of review data available on the review websites. Then this prior knowledge used by a topic model to find more relevant aspects. Some recent works try to jointly extract aspects and their polarity in a single phase (Moghaddam & Ester 2011; Sauper & Barzilay 2014). Bagheri et al. (2014) proposed a model that can extract aspects automatically using the structure of reviewed sentences. The model captures multiword aspects, and relaxes the bag-of-words (Salton 1989) assumption from topic modelling.

In contrast to sequential models, there is no need for manually labelled data. In addition, topic models perform both aspect extraction and grouping at the same time in an unsupervised manner. The limitation of topic models is that they normally need a large volume of unlabelled data to be trained accurately. Also considering the output we expect, they are very much complicated. Zhang and Liu (2014) states that these models only able to find some general aspects, and has difficulty in finding detailed aspects. The author explain that topic models are too statistics centric and come with its limitations. The author believe it could be beneficial if we can move more toward natural language and knowledge centric for a more balanced approach.

The summarization of all above methods is shown in Table 3. The table shows topic model method for ABSA. Results are presented in column 6. Some models use Rand Index to evaluate the result.

TABLE 3. Topic model methods

Author	Dataset	Domain	Result (average)
Chen et al. (2014)	http://www.cs.cornell.edu/zhychen/ from Amazon.com	Electrical and digital products in 36 domains	Average Precious at 5: 90 Average Precious at 10: 85
Bagheri et al. (2014)	FBS	2 digital cameras, DVD player, MP3 player, cellular phones	Rand Index 85.18
Lu et al. (2009)	The authors create a data set by collecting feedback comments for 28 eBay sellers with high feedback scores.	unknown	Aspect rating prediction: Correlation (Kendal's Tau): 0.11-0.49 Ranking Loss (AVG of 3): 0.15-0.63 Evaluation of representative phrases: Precision: 0.26-0.59 Recall: 0.29-0.63
Moghaddam &Ester (2011)	They built a crawler to extract reviews from Epinions.com Not available	Camcorder Cellular Phone Digital Camera DVD Player Mp3 Player	Aspect extraction (Average Rand Index) 0.83 Polarity prediction (Average Rand Index) 0.73

DEEP LEARNING METHOD

Most of the above works that use machine learning techniques for ABSA use classifiers with manually engineered features. Deep learning automatically learns latent features as distributed vectors and have recently been shown to outperform many machine learning methods on similar tasks. Because in ABSA we may have more than one aspect in each review and consequently more than one class for each review, a simple supervised model cannot classify each review to different classes. To overcome this challenge a growing body of literature has developed several deep learning architecture and presented new methods. We continue our literature on these architectures and methods.

A number of studies have used Attention mechanism to let the model learn representation with attention on a specific part of text. Luong et al. (2015) states that the concept of "attention" in training neural networks, allows models to learn alignments between different modalities. In this line of work, usually word2vec representation of aspect is used by authors for aspect representation. But there are works arguing that in addition to sentence, aspect representation need to be learned separately. Ma et al. (2018) propose the interactive attention networks that interactively learn attentions in both sentences and aspects, and generate the representations for each one separately. Baziotis et al. (2017) use Bidirectional LSTM for both aspect and sentence. Then, concatenates the hidden layers and adds attention on top of it.

There are works that combine classical models with deep learning. Liu et al. (2015) regarded the task as BIO sequence labelling problem. They propose a general class of discriminative models based on recurrent neural networks (RNNs) and word embedding. They also combine linguistic features like part-of- speech (POS) tags and chunk information directly with the output layer of RNN, and learn their related weights in training. Jebbara & Cimiano (2016) use BIO sequence labelling with RNN for aspect extraction. For sentiment polarity prediction, a bidirectional GRU processes the input sentence as a concatenation of a sequence of word vectors, sentic vectors from SenticNet lexicon, POS tag vectors and distance embedding vectors. Wang et al.(2016) integrates CRF with Recursive Neural Network and developed a Recursive Neural Conditional Random Fields model. They also append some linguistic features such as POS tags and sentiment lexicon to the hidden vector of each word. Du et al. (2016) integrate CNN with LDA.

Tang et al. (2015) presents TC-LSTM, TD-LSTM models which considers the aspect information during training. The basic idea in TD-LSTM is to model the preceding and following contexts surrounding the aspect string, so that contexts in both directions could be used as feature representations for sentiment classification. The difference in TC-LSTM is that the input at each position is the concatenation of word embedding and aspect vector, while in TD-LSTM the input at each position only includes the embedding of current word.

In Ruder et al. (2016a), Word embedding are fed into a sentence-level bidirectional LSTM. Final states of forward and backward LSTM hidden layer are concatenated together and with the aspect embedding and fed into a bidirectional review-level LSTM. At every time step, the output of the forward and backward LSTM is concatenated and fed into a classifier, which outputs a probability distribution over sentiments. Ruder et al. (2016b) take aspect extraction as a multi-label classification problem, and output probabilities over aspects classes using CNN. The same model was used by Xu et al. (2016) on a different dataset. For the sentiment towards an aspect, they concatenate an aspect vector with every word embedding and apply a convolution over it. For aspect extraction, Poria et al. (2016) uses a 7-layer CNN to tag each word in the review data as aspect and non-aspect. They have also used many features such as POS tags and two different word embeddings; Google embedding and Amazon embeddingto combine them with CNN. Dhanush et al. (2016) proposes a model made of separate models for aspect extraction and sentiment classification. The first model extracts aspects by tagging aspects in a sentence using Recurrent Neural Network (RNN) and the second model classify sentences using Convolution Neural Network (CNN).

TABLE 4. Deep learning methods

Article	Dataset	Domain	Result (best Average)
Wang et al. (2014)	SemEval 2014 Task 4 dataset	Laptop Restaurant	Accuracy Restaurant: 84.0 Accuracy Laptop: 68.9
Tang et al. (2015)	Twitter conversation by (Dong et al., 2014)	Twitter	TD-LSTM Accuracy: 0.708 F-score: 0.690 ------------------ TC-LSTM Accuracy: 0.715 F-score: 0.695
Liu et al. (2015)	SemEval 2014 Task 4 dataset	Laptop Restaurant	Best F1-score: Restaurant: 78.00 Laptop: 81.56
Saeidi et al. (2016)	Dataset is created by the author based on the QA on city neighbourhoods	urban neighbourhoods	Best results: Aspect F1-score: 0.697 Sentiment Accuracy: 0.875
Dhanush et al. (2016)	SemEval-2014 task 4 dataset	Laptop Restaurant	Precision Recall F1score 0.886 0.824 0.854
Ruder et al. (2016a)	SemEval-2016 task (Pontiki et al., 2016)	Restaurant Laptop	F-score: Restaurants:: 85.3 Laptops: 80.1
Ruder et al. (2016b)	SemEval-2016 Task 5	Restaurant Laptop	F-score: Restaurants: 68.108 Laptops: 45.863
Jebbara & Cimiano, (2016)	Task 2 of the ESWC-2016 Challenge on Semantic Sentiment Analysis dataset	Restaurant Laptop Hotel	Accuracy: 0.811
Poria et al., (2016)	FBS (Hu &Liu, 2004) SemEval 2014 task 5	Multi-domain	FBS: Precision: 86.18 Recall: 90.19 Laptop F1: 82 .32 Restaurant F1: 87 .17
Xu et al., (2016)	Yelp Dataset	Restaurant Computer	Restaurant Accuracy: 68.34 Computer Accuracy: 76.90
Baziotis et al., (2017)	SemEval-2017 Task 4	Twitter	F1-score: 0.82

HYBRID METHOD

Every categorization has its exceptions, and the categorization used in this review is no different. In this category, the works falls in more than one of the above categories or cannot be categorized in any of above categories. Blair-Goldensohn et al. (2008) uses a hybrid method to extract aspects, a MaxEnt classifier to find the frequent aspects where aspects are pre-defined for labelled data, and a rule based method that uses frequency information same as Hu and Liu (2004) and syntactic patterns to find the non-frequent aspects. A model in Zhao et al. (2010) captures aspect-specific sentiment words that are used most commonly with their corresponding aspects, and general sentiment words that are shared across aspects. Sauper &

Barzilay (2014) combine topic modelling with HMM, where the HMM models the sequence of words with types (aspect word, sentiment word, or background word). Popescu & Etzioni (2007) is a modification of (Hu & Liu 2004) which removes those nouns that their Pointwise Mutual Information (PMI) with product class is low. Then, these possible aspects fed into a Naïve Bayes classifier to output a set of explicit aspects. It uses NLP parser to determine syntactic dependencies of words in each sentence and then generates a set of syntactic rules for extracting sentiment phrases. Then, uses some external knowledge (in the form of word lists) to remove those potential sentiment phrases without positive or negative polarity. Then, an unsupervised classification technique is applied on the extracted sentiment phrases to classify them as positive or negative. There are also some recent hybrid deep learning models such as Wang et al. (2016) which integrate RNN with CRF and Du et al. (2016) which integrate CNN with LDA.

TABLE 5. Hybrid methods

Author	Dataset	Domain	Result (Average)
Wang et al. (2016)	SemEval 2014 dataset	Laptop Restaurant	Restaurant Aspect F1-score: 84.93 Restaurant Sentiment F1-score: 84.14 Laptop Aspect F- score: 78.42 Laptop Sentiment F- score: 79.44
Du et al. (2016)	http://jmcauley.ucsd.edu/data/amazon/links.html	Electronics, Movies, TV, CDs and Vinyl, Clothing, Shoes, Jewellery	Accuracy: Electronics: 92.08% Movies and TV: 92.05% CDs and Vinyl: 94.38%, Clothing Shoes and Jewellery: 93.22%
Zhao et al. (2010)	restaurant review data set used in (Ganuet al., 2009; Brody & Elhadad, 2010) and a hotel review data set used in (Baccianella et al., 2009)	Restaurant Hotel	Aspect extraction: Average Restaurant F score: 0.705 nDCG@10 Restaurant: 0.897 nDCG@10 Hotel: 0.789 Polarity prediction: Average Restaurant F score: 0.726 average P@5 Restaurant: 0.825
Popescu & Etzioni (2007)	FBS	Digital, camera, Cellular phone Mp3 player DVD player	Precision Recall 0.94 0.76
Blair-Goldensohn et al. (2008)	From Tripadvisor.com, maps.google.com,	Restaurant and Hotel	Static aspect classification results: Precious Recall F-score 85.2 66 74
Sauper & Barzilay (2014)	Yelp restaurant reviews by our previous system (Saup er et al., 2010), dictated patient summaries at the Pediatric Environmental Health Clinic (PEHC) at Children's Hospital Boston	Restaurant, medical	Aspect cluster prediction precision recall 74.3 86.3 Sentiment classification accuracy: 82.5

RESULT AND DISCUSSION

One of the important characteristic of ABSA system is the ability of such a system to work in different domain (Cambria et al. 2013; Liu 2012). Among all aforementioned methods for aspect extraction, sequential models, because of their supervised nature, are not suitable to work in different domain compare to language models and topic models. Meaning that the method must be more unsupervised than supervised in nature to make the system as domain independent as possible. Comparing these methods, a great body of literature use language rule models for ABSA tasks that are employing machine learning to model language.

So far, the research community has mainly focused on opinions about electronics products, hotels, and restaurants. The performance results still need improvements on different domains and datasets. Thus, the problems remain to be highly challenging. Results shows that if one focus on specific domain, then high accuracies can be attained. But when the one works on several domains, the situations gets considerably harder. Without considering the method or techniques, another ongoing problem in most of the articles is un-ability of the systems to extract implicit aspects and predict the polarity of implicit sentiments. Also it is clear in Table 1, 2, 3 that evaluation methods are different, so it is difficult to compare the results. Benchmark initiatives like Pontiki et al. (2015) and Pontiki et al. 2014), that provide a testing environment are a good example of how this can be achieved. We also can see a move from traditional word-based models, towards semantically rich ABSA models with usage of word embedding and deep learning models.

CONCLUSION

In this review recent methods for two well-known tasks of ABSA, aspect extraction and sentiment estimation, is reviewed in the literature, and categorized. The limitation of each work is discussed. Also different domains and datasets reviewed in each work. From this review it is clear that this field of study still has a long way to go. Particularly in terms of how its effectiveness is defined and measured. In some works, a presented method is able to perform aspect extraction and sentiment estimation jointly, while in some other works separated models for each tasks are provided. There are also works that just perform one task, mostly sentiment estimation.

The review shows most of the articles because of their supervised nature cannot work in different domains. Another issue is that most of the classical works cannot extract implicit aspects and detect the polarity/rating of implicit sentiments. Also with lacking a conventional evaluation measurement the results are not easily comparable. A larger number of successful recent works use modern deep learning models for both tasks. But considering that deep learning models that used in the literature are supervised and domain dependent consequently. Future research should focuse on developing unsupervised deep learning models. These models do not need feature engineering but they needs large data to perform well.

ACKNOWLEDGEMENT

This research was supported by Universiti Kebangsaan Malaysia under grant code FRGS/1/2013/ICT07/UKM/02/3.

REFERENCES

Abu Latiffi, M. I. & Yaakub, M. R. 2018. Sentiment Analysis: An Enhancement of Ontological-Based Using Hybrid Machine Learning Techniques. *Asia-Pacific Journal of Information Technology and Multimedia.* 7(2):61-69.

Bagheri, A., Saraee, M. & De Jong, F. 2014. ADM-LDA: An Aspect Detection Model Based on Topic Modelling Using the Structure of Review Sentences. *Journal of Information Science.* 40(5): 621-636.

Baziotis, C., Pelekis, N. & Doulkeridis, C. 2017. DataStories at SemEval-2017 Task 6: Siamese LSTM with Attention for Humorous Text Comparison. pp. 390-395. 10.18653/v1/S17-2065.

Blair-Goldensohn, S., Hannan, K., Mcdonald, R., Neylon, T., Reis, G. A. & Reynar, J. 2008. Building a Sentiment Summarizer for Local Service Reviews. *Proceedings of WWW Workshop on NLP in the Information Explosion Era*. pp. 1-10

Blei, D. M., Ng, A. Y. & Jordan, M. I. 2003. Latent Dirichlet Allocation. *The Journal of Machine Learning Research*. 3:993-1022.

Brody, S. & Elhadad, N. 2010. An Unsupervised Aspect-Sentiment Model for Online Reviews. *Human Language Technologies: The 2010 Annual Conference of the North American Chapter of the Association for Computational Linguisticn.* 804-812.

Cambria, E., Schuller, B., Xia, Y. & Havasi, C. 2013. New Avenues in Opinion Mining and Sentiment Analysis,"in *IEEE Intelligent Systems*, 28(2): 15-21. doi: 10.1109/MIS.2013.30

Chen, Z., Mukherjee, A. & Liu, B. 2014. Aspect Extraction with Automated Prior Knowledge Learning. *Proceedings of ACL*, pp. 347-358.

Choi, Y. and Cardie, C. 2010. Hierarchical sequential learning for extracting opinions and their attributes. In Proceedings of the ACL 2010 Conference Short Papers (ACLShort '10). Association for Computational Linguistics, Stroudsburg, PA, USA, 269-274.

de Albornoz J.C., Plaza L., Gervás P., Díaz A. (2011) A Joint Model of Feature Mining and Sentiment Analysis for Product Review Rating. In: Clough P. et al. (eds) Advances in Information Retrieval. ECIR 2011. Lecture Notes in Computer Science, vol 6611. Springer, Berlin, Heidelberg.

Dhanush, D., Thakur, A. K. & Diwakar, N. P. 2016. Aspect-Based Sentiment Summarization with Deep Neural Networks. *International Journal of Engineering Research and Technology*, pp. 371-375.

Ding, X., Liu, B. & Yu, P. S. 2008. A Holistic Lexicon-Based Approach to Opinion Mining. *Proceedings of the 2008 International Conference on Web Search and Data Mining*, pp. 231-240.

Du, H., Xu, X., Cheng, X., Wu, D., Liu, Y. & Yu, Z. 2016. Aspect-Specific Sentimental Word Embedding for Sentiment Analysis of Online Reviews. *Proceedings of the 25th International Conference Companion on World Wide Web*, pp. 29-30.

Du, W. & Tan, S. 2009. An Iterative Reinforcement Approach for Fine-Grained Opinion Mining. *Proceedings of Human Language Technologies: The 2009 Annual Conference of the North American Chapter of the Association for Computational Linguistics*, pp. 486-493.

Esuli, A. & Sebastiani, F. 2006. Sentiwordnet: A Publicly Available Lexical Resource for Opinion Mining. *In Proceedings of the 5th Conference on Language Resources and Evaluation (LREC'06)*, pp. 417-422.

Hai, Z., Chang, K. & Kim, J.-J. 2011. Implicit feature identification via co-occurrence association rule mining. In Proceedings of the 12th international conference on

Computational linguistics and intelligent text processing - Volume Part I (CICLing'11), Alexander Gelbukh (Ed.), Vol. Part I. Springer-Verlag, Berlin, Heidelberg, pp. 393-404.

Hu, M. & Liu, B. 2004. Mining and summarizing customer reviews. In Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining (KDD '04). ACM, New York, NY, USA, pp. 168-177. DOI=http://dx.doi.org/10.1145/1014052.1014073.

Jakob, N. & Gurevych, I. 2010. Using anaphora resolution to improve opinion target identification in movie reviews. In Proceedings of the ACL 2010 Conference Short Papers (ACLShort '10). Association for Computational Linguistics, Stroudsburg, PA, USA, pp. 263-268.

Jebbara, S. & Cimiano, P. 2016 Aspect-based sentiment analysis using a two-step neural network architecture. *Semantic Web Evaluation Challenge*, pp. 153-167.

Jin, W., Ho, H. H. & Srihari, R. K. 2009. OpinionMiner: a novel machine learning system for web opinion mining and extraction. In Proceedings of the 15th ACM SIGKDD international conference on Knowledge discovery and data mining (KDD '09). ACM, New York, NY, USA, pp. 1195-1204. DOI: https://doi.org/10.1145/1557019.1557148.

Kiritchenko, S., Zhu, X., Cherry, C. & Mohammad, S. 2014. Nrc-Canada-2014: Detecting Aspects and Sentiment in Customer Reviews. *Proceedings of the 8th International Workshop on Semantic Evaluation (SemEval 2014)*. pp. 437-442.

Lal, M. & Asnani, K. 2014. Aspect Extraction & Segmentation in Opinion Mining. *International Journal Of Engineering And Computer Science* 3(5): 5873-5878.

Landauer, T. K., Mcnamara, D. S., Dennis, S. & Kintsch, W. 2013. *Handbook of Latent Semantic Analysis*. New York. Taylor & Francis.

Lek, H. H. & Poo, D. C. 2013. Aspect-Based Twitter Sentiment Classification. *Tools with Artificial Intelligence (ICTAI), 2013 IEEE 25th International Conference on*, pp. 366-373.

Li, F., Han, C., Huang, M., Zhu, X., Xia, Y.-J., Zhang, S. & Yu, H. 2010. Structure-Aware Review Mining and Summarization. *Proceedings of the 23rd International Conference on Computational Linguistics*, pp. 653-661.

Liu, B. 2012. Sentiment Analysis and Opinion Mining. *Synthesis Lectures on Human Language Technologies.* 5(1): 1-167.

Liu, B., Hu, M. & Cheng, J. 2005. Opinion observer: analyzing and comparing opinions on the Web. In Proceedings of the 14th international conference on World Wide Web (WWW '05). ACM, New York, NY, USA, pp. 342-351. DOI: https://doi.org/10.1145/1060745.1060797.

Liu, K., Xu, L. & Zhao, J. 2012. Opinion Target Extraction Using Word-Based Translation Model. *Proceedings of the 2012 Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning*, pp. 1346-1356.

Liu, P., Qian, X. & Meng, H. 2014. Seeing Stars from Reviews by a Semantic-based Approach with MapReduce Implementation. ASE BigData/SocialInformatics/PASSAT/BioMedCom 2014 Conference, Harvard University, December 14-16, 2014, pp. 1-5.

Liu, P., Joty, S. & Meng, H. 2015. Fine-Grained Opinion Mining with Recurrent Neural Networks and Word Embeddings. *Conference on Empirical Methods in Natural Language Processing (EMNLP 2015)*, pp. 1433-1443.

Lizhen, L., Wei, S., Hanshi, W., Chuchu, L. & Jingli, L. 2014. A Novel Feature-Based Method for Sentiment Analysis of Chinese Product Reviews. *China Communications,* 11(3): 154-164.

Lu, Y., Zhai, C. & Sundaresan, N. 2009. Rated aspect summarization of short comments. In Proceedings of the 18th international conference on World wide web (WWW '09). ACM, New York, NY, USA, pp. 131-140. DOI: https://doi.org/10.1145/1526709.1526728.

Luhn, H. P. 1957. A Statistical Approach to Mechanized Encoding and Searching of Literary Information. *IBM Journal of Research and Development*, 1(4): 309-317.

Luong, M.-T., Pham, H. and Manning, C. D. (2015) 'Effective Approaches to Attention-based Neural Machine Translation', *Emnlp, (September),* pp. 11. doi: 10.18653/v1/D15-1166.

Ma, Y., Peng, H. & Cambria, E. 2018. Targeted Aspect-Based Sentiment Analysis via Embedding Commonsense Knowledge into an Attentive LSTM. *The Thirty-Second AAAI Conference on Artificial Intelligence*. pp. 5876–5883.

Marrese-Taylor, E., Velásquez, J. D. & Bravo-Marquez, F. 2014. A Novel Deterministic Approach for Aspect-Based Opinion Mining in Tourism Products Reviews. *Expert Systems with Applications.* 41(17): 7764-7775.

Medhat, W., Hassan, A. & Korashy, H. 2014. Sentiment Analysis Algorithms and Applications: A Survey. *Ain Shams Engineering Journal*. 5(4): 1093-1113.

Moghaddam, S. & Ester, M. 2010. Opinion Digger: An Unsupervised Opinion Miner from Unstructured Product Reviews. *Proceedings of the 19th ACM international conference on Information and knowledge management*. pp. 1825-1828.

Moghaddam, S. & Ester, M. 2011. ILDA: Interdependent LDA Model for Learning Latent Aspects and Their Ratings from Online Product Reviews. *Proceedings of the 34th international ACM SIGIR conference on Research and development in Information Retrieval*, pp. 665-674.

Moghaddam, S. & Ester, M. 2013. Opinion Mining in Online Reviews: Recent Trends. *Tutorial at WWW2013*. [31 December 2018]

Pavlopoulos, J. & Androutsopoulos, I. 2014. Multi-Granular Aspect Aggregation in Aspect-Based Sentiment Analysis. *Proceedings of EACL*, pp. 78-87.

Peng, J., Chunxia, Z., Hongping, F., Zhendong, N. & Qing, Y. 2010. An Approach Based on Tree Kernels for Opinion Mining of Online Product Reviews. In Proceedings of the 2010 IEEE International Conference on Data Mining (ICDM '10). IEEE Computer Society, Washington, DC, USA, pp. 256-265. DOI=http://dx.doi.org/10.1109/ICDM.2010.104.

Pontiki, M., Galanis, D., Pavlopoulos, J., Papageorgiou, H., Androutsopoulos, I. & Manandhar, S. 2014. Semeval-2014 Task 4: Aspect Based Sentiment Analysis. *Proceedings of the 8th International Workshop on Semantic Evaluation (SemEval 2014)*, pp. 27-35.

Pontiki, M., Galanis, D., Papageorgiou, H., Manandhar, S. & Androutsopoulos, I. 2015. Semeval-2015 Task 12: Aspect Based Sentiment Analysis. *Proceedings of the 9th International Workshop on Semantic Evaluation (SemEval 2015), Association for Computational Linguistics, Denver, Colorado*, pp. 486-495.

Popescu, A.-M. & Etzioni, O. 2007. Extracting product features and opinions from reviews. In Proceedings of the conference on Human Language Technology and Empirical Methods in Natural Language Processing (HLT '05). Association for Computational Linguistics, Stroudsburg, PA, USA, pp. 339-346. DOI: https://doi.org/10.3115/1220575.1220618

Poria, S., Cambria, E., Ku, L.-W., Gui, C. & Gelbukh, A. 2014. A Rule-Based Approach to Aspect Extraction from Product Reviews. *SocialNLP 2014* 28, pp. 1-10.

Poria, S., Cambria, E., & Gelbukh, A. 2016. Aspect Extraction for Opinion Mining with a Deep Convolutional Neural Network. *Knowledge-Based System*. 108:42-49.DOI: http://dx.doi.org/10.1016/j.knosys.2016.06.009

Qiu, G., Liu, B., Bu, J. & Chen, C. 2011. Opinion Word Expansion and Target Extraction through Double Propagation. *Computational Linguistics* 37(1): 9-27.

Ruder, S., Ghaffari, P. & Breslin, J.G.(2016b). INSIGHT-1 at SemEval-2016 Task 5: Deep Learning for Multilingual Aspect-based Sentiment Analysis. *Proceedings of SemEval-2016,* pp. 330–336. *Available at: http://www.aclweb.org/anthology/S16-1053* [11 November 2018].

Ruder, S., Ghaffari, P., & Breslin, J. G. (2016a). A Hierarchical Model of Reviews for Aspect-based Sentiment Analysis, *In Proceedings of the Conference on Empirical Methods on Natural Language Processing (EMNLP 2016)*, pp. 1-7.

Saeidi, M., Bouchard, G., Liakata, M. and Riedel, S., 2016. Sentihood: Targeted aspect based sentiment analysis dataset for urban neighbourhoods. *arXiv preprint arXiv:1610.03771*. [11 November 2018].

Salton, G. 1989. Automatic Text Processing: The Transformation, Analysis, and Retrieval Of. *Reading: Addison-Wesley*.

Sauper, C. & Barzilay, R. 2014. Automatic Aggregation by Joint Modeling of Aspects and Values. *arXiv preprint arXiv:1401.6422* . pp. 89-127.

Schouten, K. & Frasincar, F. 2016. Survey on Aspect-Level Sentiment Analysis. *IEEE Transactions on Knowledge and Data Engineering,* 28(3): 813-830.

Shariaty, S. & Moghaddam, S. 2011. Fine-Grained Opinion Mining Using Conditional Random Fields. *Data Mining Workshops (ICDMW), 2011 IEEE 11th International Conference on*, pp. 109-114.

Sparck Jones, K. 1972. A Statistical Interpretation of Term Specificity and Its Application in Retrieval. *Journal of Documentation* 28(1): 11-21.

Tang, D., Qin, B. & Liu, T. 2015. Deep learning for sentiment analysis: Successful approaches and future challenges. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery* 5(6): 292–303.

Wang, Y., Huang, M., Zhao, L., and Zhu, X. 2014. Attention-based LSTM for Aspect-level Sentiment Classification. *Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing,* pp. 606–615.

Wang, W., Pan, S. J., Dahlmeier, D. & Xiao, X. 2016. Recursive Neural Conditional Random Fields for Aspect-Based Sentiment Analysis. *arXiv preprint arXiv:1603.06679*, pp. 616-626.

Wu, Y., Zhang, Q., Huang, X. & Wu, L. 2009. Phrase Dependency Parsing for Opinion Mining. *Proceedings of the 2009 Conference on Empirical Methods in Natural Language Processing.* 3: pp. 1533-1541.

Xu, J., Chen, D., Qiu, X. and Huang, X., 2016. Cached long short-term memory neural networks for document-level sentiment classification. *arXiv preprint arXiv:1610.04989,* pp. 1-10.

Yaakub, M. R., Li, Y., Algarni, A. & Peng, B. 2012. Integration of Opinion into Customer Analysis Model. *Web Intelligence and Intelligent Agent Technology (WI-IAT), 2012 IEEE/WIC/ACM International Conferences on*, pp. 164-168.

Yu, J., Zha, Z.-J., Wang, M. & Chua, T.-S. 2011. Aspect Ranking: Identifying Important Product Aspects from Online Consumer Reviews. *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies-Volume 1*, pp. 1496-1505.

Zhang, F., Zhang, Z. & Lan, M. 2014. Ecnu: A Combination Method and Multiple Features for Aspect Extraction and Sentiment Polarity Classification. *Proceedings of the 8th International Workshop on Semantic Evaluation (SemEval 2014)*, pp. 252-258.

Zhang L., Liu B. (2014) Aspect and Entity Extraction for Opinion Mining. In: Chu W. (eds) Data Mining and Knowledge Discovery for Big Data. Studies in Big Data, vol 1. Springer, Berlin, Heidelberg, pp. 1-49.

Zhao, W. X., Jiang, J., Yan, H. & Li, X. 2010. Jointly Modeling Aspects and Opinions with a Maxent-Lda Hybrid. *Proceedings of the 2010 Conference on Empirical Methods in Natural Language Processing*, pp. 56-65.

Zhuang, L., Jing, F. & Zhu, X.-Y. 2006. Movie review mining and summarization. In Proceedings of the 15th ACM international conference on Information and knowledge management (CIKM '06). ACM, New York, NY, USA, pp. 43-50. DOI=http://dx.doi.org/10.1145/1183614.1183625

Zohreh Madhoushi
Abdul Razak Hamdan
Suhaila Zainudin (corresponding author)
Center for Artificial Intelligence Technology
Faculty of Information Science and Technology
Universiti Kebangsaan Malaysia
43600 Bangi, Selangor Darul Ehsan, Malaysia
zohrier@gmail.com, arh@ukm.edu.my, suhaila.zainudin@ukm.edu.my

Received: 6 October 2018
Accepted: 29 April 2019
Published: 30 June 2019