Survey on Deep Learning Approaches for Aspect Level Opinion Mining

AHMED R. ABAS [1], IBRAHIM EL-HENAWY [2], HOSSAM MOHAMED [3,*], AND AMR ABDELLATIF [4]

[1] Department of Computer Science, Faculty of Computer and Informatics, Zagazig University, Zagazig 44519, Egypt
[2] Department of Computer Science, Faculty of Computer and Informatics, Zagazig University, Zagazig 44519, Egypt
[3] Department of Computer Science, Faculty of Computer and Informatics, Zagazig University, Zagazig 44519, Egypt
[4] Department of Computer Science, Faculty of Computer and Informatics, Zagazig University, Zagazig 44519, Egypt

* Correspondence: h.hawash.research@gmail.com

Abstract

The the task of Aspect-based opinion mining (AbOM) is an emeraging research area, where aspects are mined, the corresponding opinion are scrutinized and sentiments are continuously changed, is gaining increased attention with growing feedback of clients and community across various social media streams. The gigantic improvements of deep learning (DL) techniques in natural language processing (NLP) tasks motivated research community to introduce a novel DL models and for AbSA, each investigate a diverse research points from different perspective, that cope with imminent problems and composite circumstances of AbOM. Consequently, in this survey paper, we concentrate on the limitations of the current studies and challenges relevant to mining of various aspects and their pertinent opinion, interrelationship delineations among different aspects, interactions, dependencies and contextual-semantic associations among various entities for enhanced opinion precision, and estimation of the automaticity of opinion polrity development. A laborious investigation of the later advancement is discussed depending on their contribution in the direction of spotlighting and alleviating the shortcomings related to Aspect Extraction (AE), AbOM, opinion progression (OP). The reported performance for each scrutinized study of Aspect Extraction and Aspect opinion Analysis is also given, revealing the numeriacal evaluation of the presented approach. Future research trends are introduce and delibrated by critically analysing the existing recent approaches, that will be supportive for researchers and advantageous for refining aspect based opinion classification.

Keywords: Sentiment Analysis, Opinion mining, Deepl Learning

1.Introduction

The areas Natural Language Processing (NLP) treat **opinion** mining (OM), also named as syntiment analysis, is an active research area to display emotions and to automatically discover the **opinion**s expressed within the text [1]. The object of OM is usually a product or service that is of keen interest among people, that they care to put a **opinion**

Received: April 15, 2020 Revised: June 20, 2020 Accepted: July 05, 2020

DOI: 10.5281/zenodo.3943130

towards it. Traditionally, OM has been considered as opinion polarity that whether someone has expressed positive, negative or neutral **opinion** about an event [2].

Since last decade, researchers are putting efforts to capture, quantify and measure dynamic public **opinion**s through different methods, tools and techniques, and thus allowing OM as one of the rapidly growing research areas [3]. OM applications have been widely spread to nearly every domain, like social events, consumer products and services, healthcare, political elections and financial services. Influential groups and business organizations, like, Google, Microsoft, SAP and SAS, have designed their own in-house capabilities that support them in decision making and assist them in developing better business applications to track and predict evolving market trends. From studies, OM has been generally categorized at three levels. Document-level [4], sentence-level [5] and aspectlevel OM [6] to classify whether a whole document, sentence (subjective or objective) and an aspect expresses a **opinion**, i.e., positive, negative or neutral. The Aspectbased **opinion** mining s (AbOM) helps to understand the problem of OM better comparatively, because it directly focuses on **opinion**s rather than language structure.

Where, an aspect is related to an entity, and the basic concept of an aspect is not just limited to judgement but also extends towards thoughts, point of views, ways of thinking, perspectives, an underlying theme or social influence towards an occurrence. Hence, AbOM provides a great opportunity to analyse opinions (public) over time across different contents present on media [7]. AbOM can be categorized by three main processing phases, i.e., Aspect Extraction (AE), Aspect opinion mining (AOM) and opinion progression (OP). The first phase deals with the extraction of aspects, which can be explicit aspects [8], implicit aspects [9], aspect terms [10], entities [11] and Opinion Target Expressions (OTE) [12]. The second phase classifies opinion polarity for a predefined aspect, target or entity [13]. This phase also formulates interactions, dependencies and contextualsemantic relationships between different data objects, e.g., aspect, entity, target, multi-word target, opinion word, for achieving improved opinion classification accuracy [14], [15]. The expressed opinion can be classified into ternary, or fine-gained opinion values [2], [16]. The third phase concerns with the dynamicity of peoples' opinion towards aspects (events) over a period of time. Social characteristics and self-experience are considered as the leading causes for OP[17].

1.1 Focus of this Survey

The field of AbOM is not a straight road, it has suffered many diverse changes and touched many new eras to ponder over. Researchers have been working hard to resolve multi-faceted challenges containing many issues. They have come up with thorough solutions of many complicated challenges through different machinelearning techniques, mostly deep-learning techniques, that represent their critical idea in the field. They have provided pictorial representations and numerical modeling for handling complex scenarios through different attention mechanisms and neural-memory networks. For detailed knowledge on deep-learning techniques, attention mechanisms and memory networks, we refer our readers to these surveys [18], [29], [30], [31]. Therefore, a comprehensive survey is the need of time to summarize the most recent developments in the field of AbOM. The focus of existing surveys is limited to the technical details or specific phases of AbOM. The critical issues and key challenges of AE, AOM, SE have not stated and summarized precisely. Those surveys have also become out-dated because of the exponential achievements and innovations in recent years. To fill this gap, we introduce a comprehensive survey related to **AbOM**. Our survey presents the systematic, detailed and thorough study about existing issues (challenges) of AE, **AOM** and SE, and their assembled solutions in most recent years. The survey also introduces new suggestions that induce some serious thoughts and considerations to improve present solutions, which would be helpful for future research directions; along with considering how OP is an significant and exciting research problem for many applications.

Based on the present study, the fundamental prospects are discovered and discussed that are essential for future advancements and developments, as they will provide an immense-boost and motivation for researchers to explore and devise new approaches that will overcome the critical issues and major challenges of **AbOM**.

DOI: 10.5281/zenodo.3943130

1.2 Organization of this Survey

The survey has been started with the brief introduction and paramount significance of **AbOM**. The rest of the survey is organized as follows: Section II provides the definitions of **opinion** with respect to aspect, and lists down the major issues and challenges related to AE, **AOM** and SE. Section III and IV discusses the major issues of AE and **AOM**, and concisely describes the recent solutions for these issues. Section V discusses the dynamics of SE. Section VI highlights the future research directions. Section VII concludes this survey

2. Key solutions

This section presents **AbOM** definitions, and outlines the major issues and sub-issues of AE, **AOM** and SE. This section also illustrates three main processing steps of **AbOM**, i.e., AE, **AOM** and SE, through framework for the ease and understandability of the reader.

2.1 Definitions (AbOM)

In **AbOM**, **opinion** is a subjective consciousness of human beings towards an aspect (objective existence). People get affected by the ups and downs of life, at any time and place, which results in the change of their **opinion** towards a specific aspect. This change depicts human behavior flexibility, decision autonomy and creative thinking. Pang and Lee [19] defined OM as: "A **opinion** is basically an opinion that a person expresses towards an aspect, entity, person, event, feature, object, or a certain target." Now days, researchers mostly use the term 'aspect' instead of 'feature' for specific application. Liu [5] defined the above terms like: 'an opinion is defined as a quintuple (ei, aij, sijkl, hk, tl), where ei is an entity, object or person, aij is the jth aspect of ei, sijkl represents **opinion** that an opinion holder hk expresses at time t for aij ∈ ei. sijkl could be positive, negative or neutral. Both ei and aij collectively become the opinion target'.

2.2 The Core Issues and Challenges of AbOM

In the light of above definitions, the big issues and major challenges facing by **AbOM** are basically the elements which constitute these definitions. These major issues and subissues that are discussed in this survey are listed below:

A. How to perform the process of AE effectively; why till now this remains a big challenge?

   a. i. How to extract explicit aspects, OTE, implicit aspects and aspects with neutral **opinion**? ii. How to perform cross-domain and crosslingual AE?

   b. iii. How to map relationships between different data objects for improved AE?

B. How to conduct **AOM**; how to achieve a **opinion**alcalculation model that performs an in-depth analysis of all the emotional aspects; why this is still a big challenge and hot research area?

   a. i. How to perform OM at aspect (target), entity and multi-word-target level?

   b. ii. How multi-task learning enhances the **opinion** prediction accuracy?

   c. iii. How the interactions, dependencies and contextual-semantic relationships between data objects contribute towards improved **opinion** classification?

C. How to measure the change of **opinion** value with time; why there is not any outstanding achievement which makes OP an open issue?

DOI: 10.5281/zenodo.3943130

a. How to identify factors and track obvious deficiencies in continuously-changing**opinion**al characteristics?

b. How to predict OP over Social Data?

These three issues are the core of AbOM and excellence could be achieved by resolving each issue at granular-level. We are goint to adresss these three main phases (issues) of AbOM, i.e., AE, AOM and SE, with an initial necessary preprocessing step that applies the preprocessing filters, e.g., tokenization, transforming cases, stemming and filtering of stop words, on the corpus for eliminating any needless information, e.g., special characters, stops words, repeated words, etc. [32]. The following step is to transform the data into word embeddings, e.g., word2vec [33], Glove [34], ELMo [35], Bert [36], and then compute positional information of words, accordingly [29].

3. Aspect extraction (AE)

AE deals with detection and discovery of explicit aspects, implicit aspects, entities, aspect categories and OTE. AE also tackles the relation distribution between different aspects in order to identify consistent, coherent and similar aspects from the dataset that help in improving the overall representation of extracted aspects.

3.1 Extraction of Aspects

3.1.1 Explicit Aspects and OTE Extraction

The foremost crucial issue towards AE is; the extraction of explicit aspects, entities, aspect categories and OTE upon which a opinion is being expressed. Where, an explicit aspect means that it is present in the text, e.g., "the camera of this mobile phone is great" is an example of an explicit aspect, i.e., camera, and explicit entity, i.e., mobile phone. Aspect category combines entity, and aspect of that entity into a pair, e.g., "This mobile looks great but expensive", in this example "MOBILE#APPERANCE" and "MOBILE#PRICE" are two mentioned categories, and the OTE in this example is the "mobile" because user is targeting the mobile's attribute. The sequence-labeling techniques, e.g., Conditional Random Field (CRF), have been proved effective for extracting different kind of aspects through neuralcontextual word embeddings used by deep-learning mechanism [37], [38]. Many researchers have integrated CRF and Recurrent Neural Network (RNN) to extract explicit aspects and opinion words through learning sequential features based on likelihood and backpropagation mechanism [8], [10]. For instance, Giannakopoulos et al. [39] utilized BiLSTM and CRF to handle the extraction and sequential-labeling of explicit aspects and OTE, simultaneously, through continuous word representations. They achieved high precision score and performed AE in both supervised and unsupervised manner. Rana and Cheah [40] also extracted explicit aspects and OTE by generating sequential-pattern rules through CRF on the basis of direct and indirect correlation between aspect and opinion words. They focused on learning pattern rules from users ' review, instead of manually designing them, to mitigate the effect of language constraints and grammatical rules. Jebbara and Cimiano [41] introduced RNN supplemented with character-level embeddings to extract OTE after scrutinizing its characteristics through a sequence-labeling system, i.e., CRF. The inclusion of character embeddings helped to rectify the robustness of unseen words and spelling errors. CRF has also been applied for grouping semantic, syntactic and lexical aspects according to their respective ambiguities [42], [43]. For instance, Ma et al. [44] discovered aspects through a hierarchical multi-layer Bidirectional Gated Recurrent Unit (BiGRU) that highlighted character features and high-level semantic features by capturing long-range dependencies between aspects and targets. The resulted features helped in estimating opinion labels through sequence-labeling CRF.

Furthermore, researchers have achieved improved aspect representations by utilizing neural-memory operations and attention encoders through memory interactions and attention mechanisms , respectively [38], [45], [46]. These two phenomena handle the interactions between aspects and opinion words that soothe the extraction procedure, and also help to achieve the multi-task learning environment [47]. For instances, Li and Lam [45] introduced a memory-interaction mechanism with two Long Short Term Memory (LSTM) for extracting aspect and opinion words through a multi-task-learning environment. Their memory operations were based on the positional information of aspects and

DOI: 10.5281/zenodo.3943130

opinion words, and global score of opinion terms. They also added the constraints of opinional sentences to facilitate the extraction process through another LSTM. Wang et al. [48] intoduced multilayer neural attentions with tensor operators for handling dual-interaction between aspects and opinion words. They guided attentions through input word embeddings and the prototype vectors of aspects and opinion words, in order to measure the attention score for each input word, where words having high-attention scores were considered as an aspect or opinion word. Further, Li et al. [49] deployed trimmed-history-attention mechanism to encode useful history-aware-aspect representation into sequentialaspect representations obtained from two-stacked LSTMs. The engagement of two-stacked RNNs and afterwards an auxiliary-opinion-based-word-recognition module helped in refining the boundaries of extracted OTE. Angelidis and Lapata [50] intoduced a weakly-supervised-neuralattention encoder that averaged word vectors in each segment to identify aspect words. The encoder assumed that every aspect present in the data contained a small set of seed words that could be used to identify segments containing aspects through a multi-task objective.

Though, as discussed previously; RNNs (LSTMs) have been mostly used for carrying out the AE procedure. But, to deal with the semantic and syntactic irregularities, mostly present in the user-generated social media text, Convolutional Neural Network (CNN) is proved extra effective than RNN [51]. Because CNN extracts local and position-invariant aspects whereas RNN captures longrange-semantic dependencies based on sequential information for classification [52]. For instance, Gu et al. [53] introduced cascaded-CNN structure for mapping the local semantic of each aspect through convolutional and maxpooling operations, which associated each aspect to its corresponding sentence for producing aspect category labels. Xu et al. [54] adjusted aspects of seen words along with its semantics for modeling label dependencies through a less complicated neural architecture. They utilized Dual-Embeddings CNN, i.e., cross-domain embedding and domain-specific embedding to encode the information of each input word and to achieve the AE procedure with less supervision, but they did not perform the semantics of conjunction words. Furthermore, Pham and Le [55] captured extra semantics in the text through three different representations of word embeddings, i.e., Word2Vec, Glove embeddings and one-hot character vectors. Each embedding was passed to an independent CNN channel and a non-linear activation function to produce global-sentence representation, followed by a single NN to carry out aspect-category-label-prediction task.

3.1.2 Implicit Aspects Extraction

Implicit and explicit aspects should be given same significance considering their relevance to the customers' reviews. Where, an implicit aspect means that it is not clearly stated in the text, e.g., "I could not use the camera because of very low charging" is an example of implicit aspect, i.e., battery, of an entity, i.e., camera. Unfortunately, explicit aspects have earned additional attention of researchers relative to implicit aspects [56]. Because, the identification of implicit aspects is a challenging task as they often do not contain any name or clue words. But with the immense achievement of NNs like LSTM that are excellent in learning implicit knowledge from data through gated mechanism because they work similar to human brain and encode the signifcance of each word present in the text. For instance, Sentic-LSTM (an extension of LSTM) introduced by Ma et al. [9] offered a solution for explicitly combining the implicit and explicit knowledge. The Sentic-LSTM adopted a sequence-encoder and a self-attention mechanism to calculate and incorporate affective-common-sense knowledge into a deep-neural-sequential model for handling the significance of multi-instances target.

Other studies are also classified based on supervised, unsupervised, hybrid and semi-supervised learning for categorizing implicit aspects through dependency parsing, association rule mining, semantic ontology, classification, clustering and rule-based approach [57]. SemEval 2015 involved the segregation of implicit attributes for entities from different domains through a unified representation model. They mapped different features (as nodes) in a heterogeneous network, i.e., linguistic features, from label aspects of the source domain to unlabeled aspects of the target domain, with the help of a mathematicalformalization framework and theoretic guarantees of convergence. Shu et al. [72] claimed that lifelong-machinelearning-CRF could be significantly improved for extracting aspects from different domains by retaining knowledge from previous domains, to facilitate future learning. Akhtar et al. [64] worked on an efficient Particle Swarm Optimization technique that automatically discovered the most relevant aspects, irrespective of the domain adaptation. Commonly, the research focuses on English language [73], but other languages also gained interest in recent years [11], [74]. The use of machine translation for multilingual-OTE extraction requires

DOI: 10.5281/zenodo.3943130

NLP resources to perform word alignment between POS tags and dependency features [75]. Similarly, the CRF system using different features such as POS, bigrams and lemmas obtained good results for Spanish and English [76]. Agerri and Rigau [43] also modelled multi-lingual-OTE extraction as a sequencelabeling task. Their language-independent model obtained different features clusters, i.e., local, shallow and independent features, from diverse data sources based on semantic distribution. Jebbara and Cimiano [77] used zero-shot-cross-lingual approach with multi-lingual word embeddings for predicting OTE. They used common vector space to transfer learning model from source language to target language [12].

Implicit aspect indicators were also identified through a probabilistic-graphical framework (CRF) for extracting implicit aspects [58]. Likewise, Chatteri et al. [59] used AspectFrameNet, where they considered aspects as Frame elements, for identifying implicit-aspect and explicit-aspect patterns as a sequence-labeling task (CRF). According to defined pattern rules, the AspectFrameNet updated itself with aspects' patterns for every next iteration. Gaillat et al. [60] classified both explicit and implicit aspects by utilizing pre-defined aspect categories in financial domain. The semantic relatedness of every input word was calculated with aspect-class label based on their occurrence, which was further provided to machine-learning algorithm for final classification. However, Poria et al. [61] extracted implicit aspect through Sentic Latent Dirichlet Algorithm (SLDA) that integrated common-sense reasoning in computation of word distributions. The clusters were formed by SLDA after capturing semantic association between words and multi-word expressions. The words having highest probability among clusters were considered as aspect terms. Despite of discussed solutions, implicit-aspect extraction still remains an issue which could be handled with improved feature engineering and large datasets.

3.1.3 AE with Neutral opinion

Every aspect plays an integral role for improving opinion classification accuracy. But, researchers mainly focused on aspects with positive and negative opinions, and aspects with neutral or linguistic opinions have been neglected [62]. At the same time, some researchers also performed proper encoding of text for detecting aspects with positive, negative and neutral opinions, after correspondingly performing end-to-end optimization through a appropriate NN [63-64], [51]. For instance, Wu et al. [65] came up with a chunk-level-extraction method for extracting neutral and implicit aspects. The method contained both rule-based and supervised-learning approaches to produce rational predictions and to achieve higher-level aspect representations through a deep NN architecture. Furthermore, Li et al. [11] investigated the opinion scope of neutral aspects through opinion-scope graphs that captured the boundless longdistance dependencies between aspects and entities.

3.1.4 Cross-Domain and Cross-Lingual AE

The limitations and characteristics of domain application contribute towards domain dependence, which is a big challenge in AE [66]. Inductive-transfer learning could be used to extract common aspects (linguistic) from two domains, only if they share a common feature space and same distribution characteristics [67], [68]. The use of pretraining and fine-tuning across different domains may cause inconsistency issues because different domains have different marginal-probability distribution and feature spaces [69]. Although, cross-domain-transfer learning have reported feasible solutions, but they make the regular labeling of data, expensive or nearly impossible [6], [62], [68], [67]. On the contrary, transductive learning has been proved effective for handling the issue of classifying unlabeled data [70]. For instance, Marcacini et al. [71] utilized transductive learning to incorporate knowledge from different domains through a unified representation model. They mapped different features (as nodes) in a heterogeneous network, i.e., linguistic features, from label aspects of the source domain to unlabeled aspects of the target domain, with the help of a mathematicalformalization framework and theoretic guarantees of convergence. Shu et al. [72] claimed that lifelong-machinelearning-CRF could be significantly improved for extracting aspects from different domains by retaining knowledge from previous domains, to facilitate future learning. Akhtar et al. [64] worked on an efficient Particle Swarm Optimization technique that automatically discovered the most relevant aspects, irrespective of the domain adaptation. Commonly, the research focuses on English language [73], but other languages also gained interest in recent years [11], [74]. The engagementment of machine translation for multilingual-OTE extraction requires NLP resources to perform word alignment between POS tags and dependency features [75]. Similarly, the CRF system using different features such as POS, bigrams and lemmas obtained good results for Spanish and English

DOI: 10.5281/zenodo.3943130

[76]. Agerri and Rigau [43] also modelled multi-lingual-OTE extraction as a sequencelabeling task. Their language-independent model obtained different features clusters, i.e., local, shallow and independent features, from diverse data sources based on semantic distribution. Jebbara and Cimiano [77] used zero-shot-cross-lingual approach with multi-lingual word embeddings for predicting OTE. They used common vector space to transfer learning model from source language to target language.

3.2 Interrelationship Delineations Among Aspects

The explicit representations of the extracted data objects are essential to precisely map relationships between particular entities and aspects. These relationships include co-occurrence relations [78], dependency relations [79], dependency-context information [37] and other relations [80], [81], [82], which could be mapped through hierarchical representation of aspects based on tree structure, opinion ontology or conceptual ontology. All these mapping representations help to understand the relations between different data objects and present product information in a structured way [83], [84], [85].

3.2.1 Co-occurrence Relation

The co-occurrence relations help to predict complete, coherent and consistent knowledge between distribution of words, e.g., beef, mutton, fish (aspect terms), into one aspect “meat” or it can be “food”. Many existing models do not incorporate the prior knowledge, e.g., consideration of topic distribution in each document, and encoding of word co-occurrence statistics to preserve topic consistency, for AE procedure, and thus they often deduced aspects of poor quality [78], [86]. To handle the issue, Souchen et al. [87] adopted a supervised approach for computing cooccurrence frequencies between annotated categories and lemmas. They also appended grammatical dependencies between words for extra correct aspect- category (implicit and explicit) detection. Furthermore, they also considered word co-occurrence frequencies for computing direct and indirect relations between words through association-rule mining in an unsupervised manner. They assigned activation value to each word through spreading-activation process, where the word having higher activation value was considered as an aspect category. He et al. [46] mapped the co-occurrence between words according to their context for discovering coherent aspects in the neural-embedding space. They successfully understated non-aspect words through attention mechanism and thus maintained the uniqueness of word embeddings. Li and Lam [45] also captured the co-occurrence patterns between aspects and opinion words through memory interactions. The neural-memory operators were designed to achieve interactions between data objects and to generate task-level information summary.

The co-occurrence relation between words could also be captured through association-rule-mining techniques [103], aspects-clustering [104], co-occurrence matrix [105] and graph-based methods [106]. Where, co-occurrence matrix focuses on relations between sentences and aspects with the intention of finding implicit aspects as well, and graph-based methods utilize probabilistic models to define edges’ weights as co-occurrence relation between words. Sometimes, the identification of co-occurrence relation and word-frequency information leads to the extraction of similar aspects from sentences, but in actual scenario these aspects should be considered differently. For example, “beautiful design” and “good looks” should be two different aspects of mobile phone, although both come under the category of “appearance”. To address this issue, Luo et al. [97] addressed the overlapping problem of potential aspect terms through lexico-syntactic examination of input data. They found unique aspect terms at subsumed relations in the subgraph through synsets in WordNet. Their method covered both prominent and distinct aspects adequately.

3.2.2 Semantic and Dependency Relation (Hierarchical and Ontology Structure)

Dependency patterns based on semantic and syntactic sequences help to discover the missing synonymous aspects and also distinguish the non-aspect words [82], [73]. These dependency relations could be achieved either through a hierarchical structure or an ontology tree. The hierarchical structure helps to achieve the distributed representation learning by performing the relational classification between two aspects (nodes) where nodes can be adjacent nodes or parent and child nodes [8]. The tree dependency information of sentences introduceed by Ye et al. [99] captured the syntactic and linguistic aspects through convolutional-stacked NN. Further, they incorporated an

DOI: 10.5281/zenodo.3943130

inference layer to achieve final tag scores. Furthermore, Yin et al. [37] encoded neural embeddings with syntactic and contextual information to explicitly learn the semantic embedding paths as grammatical relations in dependency tree through recurrent NN, and utilized sequence-labeling CRF for improved AE. Similarly, Luo et al. [10] also structured the sequence-labeling syntactic-dependency tree based on bidirectional-gated mechanism through BiLSTM for AE.

The exploitation of ontology tree and ontology feature as a knowledge-base help to define the relationships between different domain-concepts. These axioms can derive implicit stated information to generate aspects and review summaries for improved extraction performance, and also encode domain-knowledge into an ontology tree to reduce the need of training data [107]. For instance, Schouten et al. [100] encoded domain-knowledge into an ontology tree enriched with target and opinion lexicons for handling the task of aspect detection as binary-classification. The association of text to its domain-concept led to strong indication for aspect detection. Konjengbam et al. [102] formed a hierarchical-tree structure that included aspects, sub-aspects, their relations and the semantic relationship of various aspects. They created two ontology trees based on semantic relationship and semantic-similarity relationship to incorporate every product related information after exploring all the aspects and their related review-snippet. Besides hierarchical structure and ontology tree, heuristic patterns based on semantic-similarity also produce a reasonable approximate solution towards AE in a reasonable time [108]. For instance, Asghar et al. [101] composed a hybrid-integrated framework with an extended set of heuristic patterns, where noun, nounphrases (multi-words) and verbs were considered as candidate terms from the review data, to extract aspects those were grouped by the semantic-similarity measure.

4. Aspect Opinion Mining (AOM)

AOM, the second phase of AbOM, assigns the opinion score to each extracted aspect, target, or entity. It also includes the interactions, dependencies and contextual semantic relationships between different data objects to enhance the opinion classification accuracy. The AOM provides additional detailed information than general OM , because it predicts the opinion polarity of the given aspects, targets and entities present in the text. Initially, AOM was handled by manually designing features, such as n-grams, opinion lexicons or dependency information [109], [110]. But, with the progression of deep-learning mechanisms, AOM approaches have been projecting for automatically learning of aspects and their opinions, providing best solutions to many problems in AbOM [111], [112].

4.1.1 Aspect (Target)-level OM

Aspect (target)-level OM classifies opinions by learning separate representation for each data object. NN models get clear representation of words, i.e., adding extra information at the word-level, by measuring and identifying the semantic and syntactic relatedness between different data objects [51], [113]. The implementation of BiLSTM with character-level embeddings have performed clear representation of OTE and non-vocabulary words, and thus contributed in the acquisition of good opinion information [74]. For instance, Ma et al. [44] learned and pointed out non-vocabulary words through a hierarchical multi-layer BiGRU. The model produced target labels appropriately and thus made good prediction for opinion labels. Pham et al. [114] combined feed-forward NN, containing representation learning techniques of word embeddings, with compositional vector model for capturing semantic-information and richer-knowledge representations. The designed NN utilized existing aspectrating and higher-aspect representation layer to produce compositional opinions. Ghasedi and Huang [115] adopted a CrowdDeep autoencoder coupled with probabilistic approach for efficiently reconstructing and reducing the noise of corrupted-text data. The approach correctly categorized opinions by estimating the underlying information of text data. It also did not encounter any memory exhaustion problem during handling insufficient training parameters in very large datasets because of utilizing text data as a source of information. Besides NN, Asghar et al. [101] composed a hybrid-integrated framework to group extracted aspects according to their semantic-similarity. The opinion scores were assigned to aspects through a opinion scoring technique based on lexicon-based and corpusbased concepts. Besides sequential models like LSTM and GRU, CNN is also proved effective for achieving targeted opinion classification accuracy [73], [116] along with many other NLP tasks such as relation classification [56] and information retrieval [117]. For instance, Xue and Li [118] developed a gated-convolutional network, containing novel gated Tanh-ReLU units (GTRU), that utilized aspect embeddings for selecting n-gram features on each receptive field to perform aspect-category **OM** and aspect-term

DOI: 10.5281/zenodo.3943130

OM. The GTRU units via ReLU activation function, on the top of the convolutional layers, made the network appropriate for parallel computing that contributed towards reduced training time.

Additionally, Zhao et al. [119] introduced CNNbased-weakly-supervised-deep embeddings for learning high-level representations that reflected the general opinion distribution of sentences from a large number of weakly-labeled sentences. They utilized review ratings, as weak labels, and labelled sentences for training and finetuning deep NN, respectively. In addition, Pham and Le [55] used a multiple-CNN model, where every CNN was initialized with a different word embedding, i.e., Word2Vec, Glove and one-hot character embeddings, for capturing semantics in the text. Each CNN module was trained with the same non-linear activation function, simultaneously. The final vector representations were produced through a single NN for classifying opinions. Li et al. [120] utilized CNN layer for extracting salient features from the transformed word representations originated from a bidirectional RNN. They adopted a proximity strategy that correctly located the opinion indicators by scaling the input of CNN layer with positional relevance between word and target. The technique demonstrated its capability by handling both general and target-sensitive opinion. Although, RNN and CNN both significantly process out relevant information from one step to another through gated mechanisms and convolutional procedures for improved opinion classification. But with the longerinput sequences, sometimes it becomes difficult for a NN to keep the context of hidden vectors. Here, attention mechanism helps to mitigate long-sequence issue by encoding the most relevant information of the input sequences. It combines with neural word embeddings to capture the key parts of the sentence, which helps to explicitly combine the explicit and implicit knowledge and produces good opinion information [9], [63], [113]. For instance, He et al. [121] utilized attention-based LSTM that was trained on aspect-level data for capturing domainspecific opinion words. They hypothesized that domainknowledge could be fully exploited to achieve additional knowledge from documents, that were from similar domain, by using two transfer-learning methods, i.e., pretraining [122] and multi-task learning [56], [123], as these two tasks are highly related semantically.

Besides, Hu et al. [124] utilized aggregated-opinions and negationspecific words as attentions for capturing the semantic components from data to form sentence representation. Next, they incorporated the hierarchical structure of word, sentence and document representations into one-targetclass label for final opinion prediction. He et al. [125] adopted an autoencoder trained with a neural-attentionbased opinion classifier for learning aspect embeddings those were semantically related (meaningful). Further, they exploited a local-attention mechanism to create a weighted combination of aspect embeddings for a given target by averaging the target and context information. The adopted autoencoder provided good target representation with a high accuracy on the predicted opinion. Angelidis and Lapata [50] utilized a hierarchical-attention-based-neural architecture that adopted a greedy algorithm for discarding redundant opinions. The neural architecture required little knowledge and less supervision for opinion classification, and for generating well-balanced opinion-based summary from high-ranked opinions with minimal human intervention. Another category of network architectures are memory networks that provide an explicit context representation for every input word in the sequence [126]. Memory networks, like attention mechanisms, also mitigate the issue of learning long-range dependencies in sequential data, and also draw relationships between different data objects for improved opinion performance [14], [127]. For instance, Tang et al. [128] used an external memory which contained multiple-layers, where each layer acted as an attention mechanism, for learning significance of each context word. They considered aspect terms as a query and utilized contextual information for continuous text representation. The computed features of last layer were considered for opinion classification.

Furthermore, Zhu and Qian [129] performed the interactions between two memories, i.e., deep memory network and auxiliary memory, for learning the context of each input word and for explicitly generating connections between aspects and terms. They considered each memory as attention, and the result of auxiliary memory fed to the main memory for final opinion classification.

4.1.2 Entity-level OM

DOI: 10.5281/zenodo.3943130

Entity **OM** is considered as a subdomain of AOM. opinionscope graphs were introduced to jointly determine the named entities and their respective opinions based on assumption that each entity is surrounded by an unbounded window that is responsible for capturing its opinion polarity, which could not be achieved by fixed window of bounded size [11], [130]. But sometimes, reviews contain entities those share same aspects and it becomes difficult to capture opinion dependencies between aspects and entities. For example, "Bought Johnsons baby lotion, texture was very smooth, but bit expensive. Then I tried Mother Care but texture was bit oily". This example contains two aspects, i.e., price and texture, that are shared by two entities, i.e., Johnsons and Mother Care. Addressing the problem, Yang et al. [13] modelled the context, entity and aspect memory, simultaneously. Where, context memory utilized three layers, i.e., interaction, position and LSTM layer, for performing element-wise multiplication and concatenation of entity, aspect and context information. The other two memories, i.e., entity and aspect memory, effectively updated context memory in an iterative manner for predicting the opinion polarity towards each pair, i.e., aspect-entity pair, present in the data.

4.1.3 Multi-word Target OM

Reviews on social data mostly contain aspects with more than one word, e.g., hot-dog, mango pudding, Chinatown etc. Single-attention-based methods or merely LSTMs are not appropriate for multi-word expressions, because they performed aspect-mapping by taking the average of word vectors that worked well for single word targets, but could not fully capture the semantics of multi-word targets [121], [111]. In this way, they also might disturb the integrity of each attended word [63]. Therefore, studies have been designed for modeling the contribution (context) of each word (multi-word) present in the sentence through attention mechanisms and RNN architectures [131], [132]. For instance, Chen et al. [133] utilized multiple attentions (memory) that were non-linearly combined with the help of BiLSTM for handling complicated multi-word expressions. Every target in a sentence possessed its own appropriate memory and the weight of memory slices were defined according to the position of target in a particular sentence. However, Fan et al. [134] introduced a convolutional-based memory network incorporated with an attention mechanism for explicitly modeling both words and multi-words information in the sentence. The model mapped low-dimensional embedding space, extracted sequential aspects, and captured long-distance dependencies by storing the context information into a fixed-size window through BiGRU for opinion classification.

4.1.4 Multi-Task Learning OM

Traditional OM approaches have separately handled the task of ternary and fine-grained classification learning. But with the passage of time, multitask learning has been potentially validating in various domains through different data-learning techniques [56], [123], [135]. These techniques elegantly access resources which are developed for similar tasks to learn hidden representation of different tasks and to promote multi-task learning environment through NNs [136], [47]. For instance, Balikas et al. [137] performed ternary and finegrained opinion classification to jointly learn multiple independent tasks through BiLSTM. The tasks were correlated by sharing information, which was expected to improve the overall generalization ability. Each task was assigned to an independent softmax layer for final classification. Besides, Li et al. [138] adopted a multi-task deep memory network for effectively modeling the interactions between subtasks and targets through sharing a common semantic-space. The semantic-similarity helped to simultaneously learn the targets and their respective opinion polarities. Furthermore, Yin et al. [139] modelled a hierarchical-attention architecture, as machine comprehension problem, with multi-task framework for building different representations at both word-level and sentence-level for AOM. These representations interacted with aspect-questions, where pseudo question-answer pairs were constructed by utilizing some aspect-related keywords and ratings, for learning aspect-aware-document representation through input-encoders and iterativeattention modules.

4.2 Interactions and Contextual opinion Information for Improved AOM

4.2.1 Interaction between Data Objects

DOI: 10.5281/zenodo.3943130

Interactions and dependencies between data objects achieve refined and precise aspects' representation, and opinion classification. But some existing works, did not consider the neighboring aspects and inter-aspect relationships or dependencies in a sentence [113], [140]. Single NNs also lack the representation learning capability for capturing relevance between aspects and terms, and for drawing relationships between different aspects, and between aspects and word terms [126], [128]. However, the utilization of attention mechanisms and deep memory networks have been proving effective for drawing relationships between different data objects by focusing on the corrected aspect terms and by capturing long-distance, inter-aspect, clause-level and structural dependencies from the sentences. For instance, Tay et al. [14] adopted circular convolution and circular correlation of vectors for modeling relationship between aspects and words through an attention layer. The attention calculated the probabilistic weights of aspect-word fusion by utilizing an associativememory operator. These learned attention scores and all the hidden representations of the LSTM layer were then passed to the softmax for final opinion classification. The model was efficient and inexpensive, as it did not append any extra tasks on LSTM for learning word-aspect relationship. Liu et al. [141] also considered the correlations between words and aspects through a content-attention mechanism. They first captured the significant information about a given aspect from a global perspective through a sentence-level-content-attention mechanism. Next, to correlate aspects and words, and to embed the sequential words into customized memories, they used a context-attention mechanism. Both attentions mutually considered the whole meaning of the sentence conveyed by each word (aspect) for final opinion classification.

Moreover, Tay et al. [127] adopted memory networks that utilized parameterized-neural-tensor compositions and holographic compositions for modeling richer interactions between aspects and words. The rich-dyadic interactions skilled with complex-valued computations helped to achieve improved opinion classification performance. Zhu and Qian [129] introduced a novel deep memory network with an auxiliary memory for handling the relation between aspects and terms, based on the semantic relatedness, through an attention mechanism. The results of auxiliary memory were forwarded to main memory for capturing the context of each word for opinion classification. Gu et al. [142] discussed that aspects and their neighboring words should be given more reputation than other long-distant words. They introduced a position-awarebidirectional-attention network for mutually modeling the relation between aspect terms and sentence. They initially generated aspect-positional embeddings for each word in the corresponding sentence. And then concatenated the word embeddings and position embeddings to get the final hidden contextual representations of the aspects and input words to compute the attention weights for classification prediction.

Furthermore, Majumder et al. [143] also introduced a memory network architecture to model the connection between inter-aspects and its neighboring aspects. They used GRU for promoting contextual information of words, which was further utilized by an attention mechanism for achieving aspect-aware-sentence representation. The final precise representations of the aspects and opinion classification was achieved after continuously iterating through the memory network. However, Wang and Lu [144] captured and modelled the structural dependencies between targets and opinions, and between opinional words, through a segmentationattention-based LSTM for selecting opinion words those have most impact on classification. However, another challenge is to handle sentences that contain multiple aspects, as it includes two sub-challenges, first; to connect an aspect to the clause that contains its opinion information, and second; to conclude opinion for aspects that do not hold enough contextual or semantic information. Because, sometimes, sentences with conjunctions, e.g., but, not, also, though, however, only etc., are appropriate to predict the opinions for aspects, e.g., "Food is great, but I doubt about the freshness of meat". In this example, the positive opinion of "food" and presence of conjunction "but" ultimately determine the negative opinion for "meat". Another example, "The grocery list was short; I remember it contained only 5 to 7 items". In this example, the aspect "items" does not provide any opinion unless we consider it with aspect "grocery list". Hence, the negative opinion of "grocery list" cause "items" to have similar opinion. To handle these challenges, Hazarika et al. [145] focused on all aspects and the particular words of the sentence for generating aspectsentence representations through an attention-based LSTM. Further, they captured inter-aspect dependencies by sequentially modelling aspect-sentence representations through another LSTM for concluding opinion for aspects that did not hold enough contextual information. However, Wang

DOI: 10.5281/zenodo.3943130

et al. [15] handled word-level and clauselevel information in the sentence through a hierarchicalaspect-specific-attention network for improved opinion classification. They initially implemented elementarydiscourse units to segment a sentence into different nonoverlapping clauses. Next, they used a BiLSTM and a word-level attention for encoding clauses' information and for capturing the significance of each word present in the clause, respectively. Further, they utilized another BiLSTM and a clause-level attention for computing the attention weights of each clause and representing their relative significance. Moreover, Lin et al. [146] adopted deep memory network for modeling inter-aspect-semantic dependencies between different aspects through semantic attention and contextual learning to finalize opinions of all aspects. The attention exploited the information about every explicit aspect present in memory through semantic parsing. And the context-moment-learning module provided the complete context (background) of the particular aspects by learning the semantic distribution of the whole sentence.

4.2.2 Contextual-Semantic opinion Information

The impressive enhancement in attention mechanisms have considerably improved the abilities of NNs, and OM of the text is more becoming the contextual analysis of the text. Because the contextual scenario behind the text data derive meaningful actions for making things interesting and up-to-date. Therefore, a crucial step is the detection of the opinion context about the given aspect(target) for correct opinion prediction. Many researches have used attention mechanisms for achieving semantic relatedness between targets and context words to capture the target information and to compute context representation for opinion classification [128], [154] Besides, deep memory networks [129], LSTM [144], GRU [143], and positional information of aspects [142] has also been adopted to promote contextual-semantic opinion information of words, along with attention mechanisms. For instance, Fan et al. [156] used a BiLSTM via multigrained-attention network to capture relationships between similar context words, possessing similar/different opinion polarities. The network considered the temporal interaction between words by consistently linking and fusing information between aspect terms and context words, with no information loss. The network assigned attention scores to context words with respect to their aspect terms and targets, and vice versa. Attention mechanism also studied the attention difference of aspects sharing same context words. Furthermore, Yang et al. [151] used a multi-view co-attention network for modeling general semantics of opinions, targets and context words. First, they learned improved feature representations of single inputs, POS and word-position features through low-dimensional embeddings. Next, they exploited three independent LSTMs and attention to obtain the hidden states of opinion, target and context words, and to capture the significant opinion information from different representation subspaces at different positions. Further, Ma et al. [155] utilized LSTM and positionattention mechanism for computing the explicit positioncontextual information between the aspect and its context words for processing multi-aspects and their opinions within the sentence simultaneously.

But sometimes, simply relying on attention model cannot solve the issue of contextual-semantic-opinion information because performance degrades when the opinion of a context word is sensitive to the given target (word). Therefore, Ma et al. [157] got inspired by the cognitive characteristic of humans and focused on further effective representation of each context word. They observed that only using word embedding as external memory is not sufficient, hence, they extracted location, POS and opinion features through memory network to enrich the word representation of each context word. Each layer of model was designed to compute attention scores between aspects' representation and context words. Finally, aspect representations of last layer used through a ReLU function for final opinion classification. Overall, model showed better performance in distributing attention scores between words and in ignoring words without opinion. Yang et al. [140] modelled long-sequences between words, after considering different word locations in a relatively even manner, for performing targetdependent-opinion classification through a BiLSTM with additional attention layer(s) on top. One attention layer considered the context and intent significance of current input and target for computing attention score of each word through a dot-product mechanism. The other attention layer computed attention scores through a bilinear term. Further, Wang et al. [153] introduced six alternative target-sensitive-recurrent-attention-memory networks for capturing interaction between the aspects (targets) and their contextual opinions. The targetsensitive memory networks were capable of handling both general and target-sensitive

DOI: 10.5281/zenodo.3943130

opinions. However, Qiu et al. [158] utilized a sequence-labeling technique CRF for predicting rating of reviews by considering aspects with their related context. The core component of the model was a probabilistic discriminative model, a variant of CRF, called SentiCRF, that built the list of term pairs where each element of the pair was acted as the context for their counterpart. This contextual information helped to assign appropriate opinion scores to the term pairs.

5 opinion progression (OP)

SE, the third phase of AbOM, deals with the opinion's dynamicity with time. OP captures researchers' attention for last two decades, as it depicts the social diversity that formulates the overall agreement in opinion change [159]. The major two issues of OP are: the recognition of factors that opt people to change their opinion, e.g., cognitive attitude, adopting the majority sentiment, social behavior and influence of heterogeneous confidence, and the prediction of OP over social media.

5.1 Recognition of Factors in OP

5.1.1 Cognitive (Informational) Attitude

Cognitive attitude refers to the thoughts, perceptions and understanding of a certain aspect. It is basically an opinion that depicts the general knowledge of a person towards a certain aspect (object, event, entity, target or person), which compel him/her to change his opinion towards certain event [160]. For example, the opinion outcome on the mega events such as World Cup involved much dynamicity with the passage of time [161]. Keeping this factor in mind, Chi et al. [17] provided a theoretical framework that associated the affective-evaluation attitude with the cognitive-evaluation attitude to evaluate SE. They studied the attitude and predicted the decision making psychology of people towards an event over a period. They assessed that the differences in the peoples' attitudes concerning pre-event and post-event are due to the information insights and understanding towards the particular aspects, that make them revise their opinions with passing time. They clearly demonstrated that the context of a certain aspect (event) is the basic reason for decision making tendency of humans.

5.1.2 Adopting the Majority Opinion

Another issue involved in OP is the preference of majority opinions. People instead of showing their own opinion towards the aspect, adopt the neighbor's and relative's opinion and avoid the linkage to the minor community holding different opinion attitude. Fu and Wang [162] found that OP is basically based on majority-adaptation and minority-avoidance rule. The underlying network of their study was a majority-rule model in which people preferred to follow crowd to update their opinion and broke the link with the community holding a dissimilar (minor) opinion. They revealed that an increasing tendency of linking neighbors reduced the number of opinions. Their investigation clearly demonstrated that like-minded people holding the same opinion belonged to the singlecommunity in the network.

5.1.3 Social Behavior

Social behavior arises when the individuals interact with the community and its surrounding environment. During social interactions, the change of opinion is determined by the social impacts like harmonization and divergence. Since individuals get influenced from their neighbors and make a decision, but, the probability still holds that an individual instead of following its community, devises its own opinion [163], [164]. Chen et al. [165] also stated that social policies, relationships and norms have great impact on social evolution. They introduced ed three different social-acquaintance networks named kinship-priority-acquaintance network (KPAN), independence-priority-acquaintance network (IPAN), and hybrid-acquaintance network (HAN) based on the variations between the social culture and policies. They found that KPAN always achieved fragmentations in opinion due to the fact that individuals mostly believe in their relatives and neighbors for adopting opinions. In IPAN, OP was based on the western values that encouraged independent decision about opinions, but with consent discussion. HAN took into account both networks, i.e., KPAN and IPAN, and depicted that opinion reach a consensus on a very large scale. The findings of these

DOI: 10.5281/zenodo.3943130

networks facilitated in computing and predicting OP under diverse associated networks, and also helped in forming rational cultural policies for public guidance.

5.1.4 Heterogeneity of Limited Self-assurance and Influence (HCCI)

Sometimes, the fact of heterogeneity-confidence distribution over the network is ignored, in which people cannot communicate with each other due to different geographical locations or statuses in the society, but yet hold the same opinion level. To address the issue, Liang et al. [166] examined the HCCI in discrete-time scenarios, where all agents updated their opinion level in a synchronous manner after averaging their neighbor's opinions. In the network, every cluster kept the value of opinion for a finite time before promoting consensus of opinions, but this was not always true. Furthermore, to find the weighted and signed relationships between individual nodes in a heterogeneous network environment, Pengyi et al. [165] explored dynamic opinion patterns within heterogeneous relationships. After conducting a series of simulations they found that at a stable state the opinion level depends on the degree of social communication between networks. They observed that when harmoniousness parameter increases within the networks, opinions transit from the bipolarization to the consensus phase.

5.2 Forecasting OP over Social Data

Predicting the changing social behavior of individuals on real-time applications is a complex task. Nguyen et al. [167] presented a statistical-machine-learning model for analysing the collective opinion level. They further transformed their model for large structured networks; containing many individuals and tweets. They extracted aspects from history and then utilized them for predicting the opinion change in future. Besides, Guerra et al. [168] tracked the emotional reactions of social-media users during continual changing events on twitter, such as. breaking news, political talk etc. They found that users are more driven by the positive thoughts instead of negative thoughts, and users preferred to present their extreme thoughts instead of average thoughts towards the events. By taking into account these extreme thoughts, they introduced a feature-representation approach that was capable of discovering new aspects by capturing frequent changes on opinion level caused by the real world events. Lumin et al. [169] presented a hierarchical multidimensional model that first captured the nonredundant opinion patterns towards the event, and then discovered the closed opinion patterns through a clustering algorithm. The algorithm also maintained a opinion vector that hold reasons behind changed opinions. The model focused on the user-level opinion dynamicity and conducted a study on real dataset, i.e., "Japanese Earthquake" in March 2011. Charltron et al. [170] observed the dynamic communicability of evolving networks to recognize the top communicators on twitter. First, they computed the initial opinion of the communities and then detected correlation between the loss of users through a community-detection algorithm. They monitored stable connected twitter communities over a time-scale of months, and found that the people who have highest communicability-broadcast directories generally show positive opinion as compared to ordinary users. They also found that every community maintain a stable opinion level and remain connected by sending messages, but when a community shows a temporary large deviation from its usual opinion, then some communities hold all its users and some communities lost a relatively small number of users.

6. Discussion and Future Trends

To achieve excellent performance from NLP and specifically AbOM requires a lot of effort. Aspect extraction and opinion score determination are the core challenges of AbOM, which could not be handled through singlegeneral solution. Instead, researchers should consider many sub-issues and sub-challenges for resolving the major challenges. These sub-challenges would develop an effective tool for AbOM and increase the opinion classification performance (to some extent) at aspect-level. Some of them are discussed in this section.

6.1 Data Pruning and Cleansing

Mostly during the preprocessing phase, punctuation marks, common words, special characters, slangs and pronouns are removed from the text. To our perception, all of them make little contribution in analyzing opinion polarity from opinionated text. But the stop words and the redundant terms should be eliminated through an effective

DOI: 10.5281/zenodo.3943130

pattern-based approach. In preprocessing, a multiple domain-specific lexicon would also be an interesting development that will produce distance metrics, integrate fuzzy membership of unclassified reviews and perform automatic domainlabeling. It could also be utilized with AE techniques to enrich text with significant information (lexical).

6.2 Cross-Domain-Transfer Learning

Supervised approaches are convenient for dealing with labeled data, but relying on large training datasets make the task less efficient. Generally, the efficiency of AE process compromises during handling unlabeled datasets, and might fail when applying it on a new domain, e.g., a model trained on restaurant data is not applicable on hotel data. No doubt, it is hard to grasp that there would be a general model that will perform reasonably good on all domains. However, model representations could be enhanced by pre-training and fine-tuning process, if accompanied by an efficient self-attention mechanism to carry out cross-domain-transfer learning. Because, selfattention network has the capability to deal with complex data by relating input words according to their context and positional information. This challenge, we believe, will be a new era in the field of AbOM, but it requires proper advancements for text representation, e.g., statistical and linguistic features, and label propagation with effective transductive learning.

6.3 Contextual-Semantic Relationships

Semantic information is compulsory to associate words with sentences according to their contextual information, which could be helpful in the extraction of explicit and implicit aspects, neutral opinion expressions and complex text, e.g., negation words (shifting words), from the data. Based on semantic information, relationships between different data objects could also be mapped to improve opinion classification accuracy. Additionally, the inclusion of contextual-semantic information, e.g., ngrams, dependency and syntactic relations, and likeness between sentences, would also be advantageous for upgradation of knowledge extraction process and for handling complex scenarios, e.g., “We went to a Pakistani restaurant and they offered us 5 to 6 choices of cocktails”, here, “choices of cocktails” should be considered as a positive opinion but normally taken as neutral. Another example is “Try the hand ripped bread”, in this, the word “try” itself is not opinionally charged, but it carries opinional meaning when considered in the right context.

6.4 Aspect Summarization

Aspect summarization is entity-centric; it aims to produce brief summaries which are relevant to particular aspects (entities) and their opinions. These summaries could be created by context-dependent words, negation words, similar opinion words etc., but the challenge is to introduce a solution that will perform improved personalized summarization and give clear explanation on recommended aspects from the application. However, the actual presentation totally depends on the requirements and needs of the particular application.

6.5 Predicting Dynamicity of opinion

Some factors for opinion transition might be little ‘trust in government’ and some ‘political issues’. Therefore, social context should be improved to track opinion at user, group and multi-group level. This will help to understand the behavior and complex-posting patterns of social media users. The dominant patterns will analyze the OP in extra detail, and will also achieve different relationships between separated groups to produce more informative-social contexts. Most of the time, prediction model computes the evolving opinions of the existing users, but to validate the effects of new users on real data is a challenging research task in itself.

6.6 Multimodal OM

Lastly, as social media is not just overwhelmed by the text data, it also contains images, videos, emoticons, stickers etc., which depict user behaviors and attitudes towards the scenario. Therefore, it would be interesting to perform unification on such multimodal data by devising new approaches for extracting semantics of visual features.

DOI: 10.5281/zenodo.3943130

This fusion of cross-modal features helps to achieve extra comprehensive aspect analysis from the user-generated content over social media. Also, cognitive techniques could also be adopted for reading and studying comprehensive skills and behaviors of human beings through machine intelligence imitation. To have extra sympathetic on the topic, we refer our readers to [171]. However, researchers should also ponder upon other issues, like, preference of authors, emotional (emotional characteristics) analysis and computation etc. Apart from essential elements, AbOM deals with many other challenges which are not only limited to aspects' quality and flexibility. For instance, Robustness; to cope with informal writing style on user-generated content, e.g., spelling mistakes, grammatical errors, non-dictionary words, slang words, special characters, emoticons etc., for expressing opinion. Scalability; efficiency of the system as the problem grows and also as the system grows. Learning rate decay; the learning time of system should be reduced over time. Other major problems are overfitting reduction and improved regularization over predicting large NNs. One added is max-pooling; to make assumptions about different aspect features. Finally, an approach should always learn to stop when performance starts to degrade. The field of AbOM is at the never-ending stage where researchers are contributing day-by-day to satisfy it at several levels.

7. Conclusion

The survey has explored a very significant research domain in the field of NLP, i.e., AbOM a subfield of opinion analysis that has accomplished many excellent researches. The survey has stated the thorough overview of the recent progress in AbOM by depicting the state-of-the-art deep-learning techniques fashioned in locating the target, which can be an entity, or it can be an aspect related to a target or an entity, their relationships, their respective opinions and opinion progression dynamicity. An issue-based categorization of the recent solutions is presented each contributing towards improving the process of aspect extraction, aspect opinion analysis or opinion progression. Key challenges are included as future research directions, which should be considered to achieve excellence in opinion classification at aspect-level. Our survey can act as a foundation for researchers who want to know the recent progress of the field and help them formulate general strategies which would be applicable to most of the scenarios.

Funding: "This research received no external funding"

Conflicts of Interest: "The authors declare no conflict of interest."

References

1. E. Cambria, S. Poria, A. Gelbukh, and M. Thelwall, "Sentiment Analysis is a Big Suitcase," IEEE Intelligent Systems, vol. 32, no. 6, pp. 74-80, 2017.
2. B. Liu, Sentiment Analysis and Opinion Mining (Synthesis Lectures on Human Language Technologies). Morgan & Claypool Publishers, 2012.
3. K. Dave, S. Lawrence, and D.M. Pennock, "Mining the Peanut Gallery: Opinion Extraction and Semantic Classification of Product Reviews," Proc. 12th Int'l Conf. World Wide Web, pp. 519- 528, 2003.
4. A. Tripathy, A. Anand, and S.K. Rath, "Document-level Sentiment Classification using Hybrid Machine Learning Approach," Knowledge Information Systems, vol. 53, no. 3, pp. 805- 831, 2017.
5. B. Liu, "Sentiment Analysis and Subjectivity," Handbook of Natural Language Processing, pp. 627-666, 2010.
6. K. Schouten and F. Frasincar, "Survey on Aspect-Level Sentiment Analysis," IEEE Trans. Knowledge and Data Eng., vol. 28, no. 3, pp. 813-830, 2016.
7. J. Wang et al., "Mining Multi-Aspect Reflection of News Events in Twitter: Discovery, Linking and Presentation," Proc. 2015 IEEE Int'l Conf. Data Mining, pp. 429-438, 2015.
8. W. Wang, S.J. Pan, D. Dahlmeier, and X. Xiao, "Recursive Neural Conditional Random Fields for Aspect-Based Sentiment Analysis," Proc. 2016 Conf. Empirical Methods in Natural Language Processing, pp. 616–626, 2016.
9. Y. Ma, H. Peng, T. Khan, E. Cambria, and A. Hussain, "Sentic LSTM: A Hybrid Network for Targeted Aspect-Based Sentiment Analysis," Cognitive Computation, vol. 10, no. 4, pp. 639-650, 2018.
10. H. Luo, T. Li, B. Liu, B. Wang, and H. Unger, "Improving Aspect Term Extraction with Bidirectional Dependency Tree Representation," arXiv Preprint arXiv:1805.07889, p. 12, 2019.
11. H. Li and W. Lu, "Learning Latent Sentiment Scopes for EntityLevel Sentiment Analysis," Proc. 31st AAAI Conf. Artificial Intelligence, pp. 3482-3489, 2017.
12. M. Pontiki, D. Galanis, H. Papageorgiou, S. Manandhar, and I. Androutsopoulos, "Semeval-2015 Task 12: Aspect Based Sentiment Analysis," Proc. 9th Int'l Workshop on Semantic Evaluation, pp. 486-495, 2015.
13. J. Yang, R. Yang, C. Wang, and J. Xie, "Multi-Entity Aspect-Based Sentiment Analysis With Context, Entity and Aspect Memory," Proc. 32nd AAAI Conf. Artificial Intelligence, pp. 6029-6036, 2018.

DOI: 10.5281/zenodo.3943130

14. Y. Tay, A.T. Luu, and S.C. Hui, "Learning to Attend via WordAspect Associative Fusion for Aspect-based Sentiment Analysis," Proc. 32nd AAAI Conf. Artificial Intelligence, pp. 5956-5963, 2017.
15. J. Wang et al., "Aspect Sentiment Classification with both Wordlevel and Clause-level Attention Networks," Proc. 27th Int'l Joint Conf. Artificial Intelligence, pp. 4439-4445, 2018.
16. S. Angelidis and M. Lapata, "Multiple Instance Learning Networks for Fine-Grained Sentiment Analysis," Trans. Association for Computational Linguistics, vol. 6, pp. 17-31, 2018.
17. C.G.Q. Chi, Z. Ouyang, and X. Xu, "Changing Perceptions and Reasoning Process: Comparison of Residents' Pre-And PostEvent Attitudes," Annals of Tourism Research, vol. 70, pp. 39-53, 2018.
18. T. Young, D. Hazarika, S. Poria, and E. Cambria, "Recent Trends in Deep Learning based Natural Language Processing," IEEE Computational intelligence Magazine, vol. 13, no. 3, pp. 55-75, 2018
19. L. Lee and B. Pang, "Opinion Mining and Sentiment Analysis," Foundations and Trends® in Information Retrieval, vol. 2, no. 1–2, pp. 1-135, 2008.
20. H. F. Tang, S.B. Tan, and X.Q. Cheng, "A Survey on Sentiment Detection of Reviews," Expert Systems with Applications, vol. 36, no. 7, pp. 10760-10773, 2009.
21. M. Wiegand, A. Balahur, B. Roth, D. Klakow, and A. Montoyo, "A Survey on The Role of Negation in Sentiment Analysis," Proc. Workshop on Negation and Speculation in Natural Language Processing, pp. 60-68, 2010.
22. 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.
23. K. Ravi and V. Ravi, "A Survey on Opinion Mining and Sentiment Analysis: Tasks, Approaches and Applications," Knowledge-Based Systems, vol. 89, pp. 14-46, 2015.
24. A. Giachanou and F. Crestani, "Like It or Not: A Survey of Twitter Sentiment Analysis Methods," ACM Computing Surveys, vol. 49, no. 2, pp. 1-41, 2016.
25. L. Zhang, S. Wang, and B. Liu, "Deep Learning for Sentiment Analysis: A Survey," Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, vol. 8, no. 4, p. e1253, 2018.
26. D. Zimbra, A. Abbasi, D. Zeng, and H. Chen, "The State-of-theArt in Twitter Sentiment Analysis: A Review and Benchmark Evaluation," ACM Trans. Management Information Systems, vol. 9, no. 2, pp. 1-29, 2018.
27. L. Yue, W. Chen, X. Li, W. Zuo, and M. Yin, "A Survey of Sentiment Analysis in Social Media," Knowledge and Information Systems, pp. 1-47, 2018.
28. E. Kim and R. Klinger, "A Survey on Sentiment and Emotion Analysis for Computational Literary Studies," arXiv preprint arXiv:1808.03137, p. 45, 2018.
29. A. Vaswani et al., "Attention is All You Need," Proc. 31st Conf. Neural Information Processing Systems, pp. 5998-6008, 2017.
30. D. Hu, "An Introductory Survey on Attention Mechanisms in NLP Problems," Proc. SAI Intelligent Systems Conf., pp. 432-448, 2019.
31. Y. Ma and J.C. Principe, "A Taxonomy for Neural Memory Networks," IEEE Trans. Neural Networks and Learning Systems, 2019.
32. A. S. Manek, P.D. Shenoy, M.C. Mohan, and K. Venugopal, "Aspect Term Extraction for Sentiment Analysis in Large Movie Reviews using Gini Index Feature Selection Method and SVM Classifier," World Wide Web, vol. 20, no. 2, pp. 135-154, 2017.
33. T. Mikolov, K. Chen, G. Corrado, and J. Dean, "Efficient Estimation of Word Representations in Vector Space," Proc. 2013 Int'l Conf. Learning Representations pp. 1-12, 2013.
34. J. Pennington, R. Socher, and C. Manning, "Glove: Global Vectors for Word Representation," Proc. 2014 Conf. Empirical Methods in Natural Language Processing, pp. 1532-1543, 2014.
35. M. Peters et al., "Deep Contextualized Word Representations," Proc. 2018 Conf. North American Chapter of the ACL: Human Language Technologies, pp. 2227-2237, 2018.
36. J. Devlin, M.W. Chang, K. Lee, and K. Toutanova, "Bert: PreTraining of Deep Bidirectional Transformers for Language Understanding," arXiv preprint arXiv:1810.04805, p. 13, 2018.
37. Y. Yin, F. Wei, L. Dong, K. Xu, M. Zhang, and M. Zhou, "Unsupervised Word and Dependency Path Embeddings for Aspect Term Extraction," Proc. 25th Int'l Joint Conf. Artificial Intelligence, pp. 2979-2985, 2016.
38. P. Liu, S. Joty, and H. Meng, "Fine-Grained Opinion Mining with Recurrent Neural Networks and Word Embeddings," Proc. 2015 Conf. Empirical Methods in Natural Language Processing, pp. 1433- 1443, 2015.
39. A. Giannakopoulos, C. Musat, A. Hossmann, and M. Baeriswyl, "Unsupervised Aspect Term Extraction with B-LSTM & CRF using Automatically Labelled Datasets," Proc. 8th Workshop on Computational Approaches to Subjectivity, Sentiment and Social Media Analysis, pp. 180–188, 2017.
40. T.A. Rana and Y.N. Cheah, "Sequential Patterns Rule-Based Approach for Opinion Target Extraction from Customer Reviews," J. Information Science, p. 0165551518808195, 2018.
41. S. Jebbara and P. Cimiano, "Improving Opinion-Target Extraction with Character-Level Word Embeddings," Proc. 1st Workshop on Subword and Character Level Models in NLP, Association for Computational Linguistics, pp. 159–167, 2017.
42. Z. Toh and J. Su, "Nlangp at Semeval-2016 Task 5: Improving Aspect Based Sentiment Analysis using Neural Network Features," Proc. 10th Int'l Workshop on Semantic Evaluation, pp. 282-288, 2016.
43. R. Agerri and G. Rigau, "Language Independent Sequence Labelling for Opinion Target Extraction," Artificial Intelligence Review, vol. 268, pp. 85-95, 2019.
44. D. Ma, S. Li, and H. Wang, "Joint Learning for Targeted Sentiment Analysis," Proc. 2018 Conf. Empirical Methods in Natural Language Processing, pp. 4737-4742, 2018.
45. X. Li and W. Lam, "Deep Multi-Task Learning for Aspect Term Extraction with Memory Interaction," Proc. 2017 Conf. Empirical Methods in Natural Language Processing, pp. 2886-2892, 2017.
46. R. He, W.S. Lee, H.T.Ng, and D. Dahlmeier, "An Unsupervised Neural Attention Model for Aspect Extraction," Proc. 55th Ann. Meeting of the ACL, pp. 388-397, 2017.
47. W. Wang, S.J. Pan, and D. Dahlmeier, "Multi-Task Coupled Attentions for Category Specific Aspect and Opinion Terms CoExtraction," Proc. 31st Conf. Neural Information Processing Systems, pp. 1-11, 2017.
48. W. Wang, S.J. Pan, D. Dahlmeier, and X. Xiao, "Coupled MultiLayer Attentions for Co-Extraction of Aspect and Opinion Terms," Proc. 31st AAAI Conf. Artificial Intelligence p. 7, 2017.

DOI: 10.5281/zenodo.3943130

49. X. Li, L. Bing, P. Li, W. Lam, and Z. Yang, "Aspect Term Extraction with History Attention and Selective Transformation," Proc. 27th Int'l Joint Conf. Artificial Intelligence, pp. 4194-4200, 2018.
50. S. Angelidis and M. Lapata, "Summarizing Opinions: Aspect Extraction Meets Sentiment Prediction and They Are Both Weakly Supervised," Proc. 2018 Conf. Empirical Methods in Natural Language Processing, pp. 3675–3686, 2018.
51. M. Schmitt, S. Steinheber, K. Schreiber, and B. Roth, "Joint Aspect and Polarity Classification for Aspect-based Sentiment Analysis with End-to-End Neural Networks," Proc. 2018 Conf. Empirical Methods in Natural Language Processing, pp. 1109–1114 2018.
52. X. Wang, W. Jiang, and Z. Luo, "Combination of Convolutional and Recurrent Neural Network for Sentiment Analysis of Short Texts," Proc. COLING 2016, the 26th Int'l Conf. Computational Linguistics, pp. 2428-2437, 2016.
53. X. Gu, Y. Gu, and H. Wu, "Cascaded Convolutional Neural Networks for Aspect-Based Opinion Summary," Neural Processing Letters, vol. 46, no. 2, pp. 581-594, 2017.
54. H. Xu, B. Liu, L. Shu, and P. S. Yu, "Double Embeddings and CNN-based Sequence Labeling for Aspect Extraction," Proc. 56th Ann. Meeting of the Association for Computational Linguistics, pp. 592–598, 2018.
55. D.H. Pham and A.C. Le, "Exploiting Multiple Word Embeddings and One-Hot Character Vectors for Aspect-Based Sentiment Analysis," Int'l J. Approximate Reasoning, vol. 103, pp. 1-10, 2018.
56. Y. Liu, S. Li, X. Zhang, and Z. Sui, "Implicit Discourse Relation Classification via Multi-Task Neural Networks," Proc. 13th AAAI Conf. Artificial Intelligence, pp. 2750-2756, 2016.
57. M. Tubishat, N. Idris, and M. Abushariah, "Implicit Aspect Extraction in Sentiment Analysis: Review, Taxonomy, Oppportunities, and Open Challenges," Information Processing and Management, vol. 54, no. 4, pp. 545-563, 2018.
58. I. Cruz, A. F. Gelbukh, and G. Sidorov, "Implicit Aspect Indicator Extraction for Aspect based Opinion Mining," Int'l J. Computing Linguistics and Applications, vol. 5, no. 2, pp. 135-152, 2014.
59. S. Chatterji, N. Varshney, and R. K. Rahul, "AspectFrameNet: A Framenet Extension for Analysis of Sentiments Around Product Aspects," The J. Supercomputing, vol. 73, no. 3, pp. 961-972, 2017.
60. T. Gaillat, B. Stearns, R. McDermott, G. Sridhar, M. Zarrouk, and B. Davis, "Implicit and Explicit Aspect Extraction in Financial Microblogs," Ann. Meeting of the Association for Computational Linguistics, 2018.
61. S. Poria, I. Chaturvedi, E. Cambria, and F. Bisio, "Sentic LDA: Improving on LDA with Semantic Similarity for Aspect-Based Sentiment Analysis," Int'l Joint Conf. Neural Networks, pp. 4465- 4473, 2016.
62. T.A. Rana and Y.N. Cheah, "Aspect Extraction in Sentiment Analysis: A Comparative Analysis and Survey," Artificial Intelligence Review, vol. 46, no. 4, pp. 459-483, 2016.
63. Y. Wang, M. Huang, and L. Zhao, "Attention-Based LSTM for Aspect-Level Sentiment Classification," Proc. 2016 Conf. Empirical Methods in Natural Language Processing, pp. 606-615, 2016.
64. M. S. Akhtar, D. Gupta, A. Ekbal, and P. Bhattacharyya, "Feature Selection and Ensemble Construction: A Two-Step Method for Aspect Based Sentiment Analysis," Knowledge-Based Systems, vol. 125, pp. 116-135, 2017.
65. C. Wu, F. Wu, S. Wu, Z. Yuan, and Y. Huang, "A Hybrid Unsupervised Method for Aspect Term and Opinion Target Extraction," Knowledge-Based Systems, vol. 148, pp. 66-73, 2018.
66. S. Deng, A.P. Sinha, and H. Zhao, "Adapting Sentiment Lexicons to Domain-Specific Social Media Texts," Decision Support Systems, vol. 94, no. C, pp. 65-76, 2017.
67. T.A. Rana and Y.N. Cheah, "A Two-Fold Rule-Based Model for Aspect Extraction," Expert Systems with Applications, vol. 89, pp. 273-285, 2017.
68. P. Zhang, J. Wang, Y. Wang, and Y. Wang, "A Statistical Approach to Opinion Target Extraction using Domain Relevance," 2016 2nd IEEE Int'l Conf. Computer and Communications, pp. 273-277, 2016.
69. L. Luo, X. Wang, S. Hu, C. Wang, Y. Tang, and L. Chen, "Close yet Distinctive Domain Adaptation," arXiv preprint arXiv:.04235, p. 11, 2017.
70. X.N. Kong, M.K. Ng, and Z.H. Zhou, "Transductive Multilabel Learning via Label Set Propagation," IEEE Trans. Knowledge and Data Eng., vol. 25, no. 3, pp. 704-719, 2013.
71. R.M. Marcacini, R.G. Rossi, I.P. Matsuno, and S.O. Rezende, "Cross-Domain Aspect Extraction for Sentiment Analysis: A Transductive Learning Approach," Decision Support Systems, vol. 114, pp. 70-80, 2018.
72. L. Shu, H. Xu, and B. Liu, "Lifelong Learning CRF for Supervised Aspect Extraction," Proc. 55th Ann. Meeting of the Association for Computational Linguistics, pp. 148–154, 2017. Authorized licensed use limited to: University of Melbourne. Downloaded on May 04,2020 at 02:41:25 UTC from IEEE Xplore.
73. S. Poria, E. Cambria, and A. Gelbukh, "Aspect Extraction for Opinion Mining with A Deep Convolutional Neural Network," Knowledge-Based Systems, vol. 108, pp. 42-49, 2016.
74. M. Al-Smadi, B. Talafha, M. Al-Ayyoub, and Y. Jararweh, "Using Long Short-Term Memory Deep Neural Networks for AspectBased Sentiment Analysis of Arabic Reviews," Int'l J. Machine Learning Cybernetics, pp. 1-13, 2018.
75. X. Zhou, X. Wan, and J. Xiao, "CLOpinionMiner: Opinion Target Extraction in a Cross-Language Scenario," IEEE/ACM Trans. Audio, Speech and Language Processing, vol. 23, no. 4, pp. 619-630, 2015.
76. T. Alvarez-López, J. Juncal-Martinez, M. Fernández-Gavilanes, E. Costa-Montenegro, and F. J. González-Castano, "Gti At Semeval2016 Task 5: SVM and CRF for Aspect Detection ad Unsupervised Aspect-Based Sentiment Analysis," Proc. 10th Int'l Workshop on Semantic Evaluation, pp. 306-311, 2016.
77. S. Jebbara and P. Cimiano, "Zero-Shot Cross-Lingual Opinion Target Extraction," arXiv preprint arXiv:1904.09122, p. 10, 2019.
78. K. Schouten, F. Frasincar, and F. De Jong, "Commit-p1wp3: A Cooccurrence Based Approach to Aspect-Level Sentiment Analysis," Proc. 8th Int'l Workshop on Semantic Evaluation, pp. 203- 207, 2014.
79. Y. Liu, F. Wei, S. Li, H. Ji, M. Zhou, and H. Wang, "A Dependency-Based Neural Network for Relation Classification," Proc. 53rd Ann. Meeting of the Association for Computational Linguistics and the 7th Int'l Joint Conf. Natural Language Processing (Volume 2: Short Papers), pp. 285-290, 2015.
80. W. Maharani, D.H. Widyantoro, and M.L. Khodra, "Aspect Extraction in Customer Reviews using Syntactic Pattern," Procedia Computer Science, vol. 59, pp. 244-253, 2015.
81. L. Qiu, "An Opinion Analysis Model for Implicit Aspect Expressions Based on Semantic Ontology," Int'l J. Grid Distributed Computing, vol. 8, no. 5, pp. 165-172, 2015.

DOI: 10.5281/zenodo.3943130

82. Q. Liu, B. Liu, Y. Zhang, D. S. Kim, and Z. Gao, "Improving Opinion Aspect Extraction using Semantic Similarity and Aspect Associations," Proc. 13th AAAI Conf. Artificial Intelligence, pp. 2986-2992, 2016.
83. S. Kim, J. Zhang, Z. Chen, A. Oh, and S. Liu, "A Hierarchical Aspect-Sentiment Model for Online Reviews," Proc. 27th AAAI Conf. Artificial Intelligence, 2013.
84. S. Ruder, P. Ghaffari, and J. G. Breslin, "A Hierarchical Model of Reviews for Aspect-Based Sentiment Analysis," Proc. 2016 Conf. Empirical Methods in Natural Language Processing, pp. 999–1005, 2016.
85. S. Poria, E. Cambria, L.W. Ku, C. Gui, and A. Gelbukh, "A RuleBased Approach to Aspect Extraction from Product Reviews," Proc. 2nd Workshop on Natural Language Processing for Social Media, pp. 28-37, 2014.
86. A. Mukherjee and B. Liu, "Aspect Extraction through SemiSupervised Modeling," Proc. 50th Ann. Meeting of the Association for Computational Linguistics, pp. 339-348, 2012.
87. K. Schouten, O. Van Der Weijde, F. Frasincar, and R. Dekker, "Supervised and Unsupervised Aspect Category Detection for Sentiment Analysis with Co-Occurrence Data," IEEE Transactions on Cybernetics, vol. 48, no. 4, pp. 1263-1275, 2017.
88. S. Manandhar, "SemEval-2014 Task 4: Aspect Based Sentiment Analysis," Proc. 8th Int'l Workshop on Semantic Evaluation, pp. 19- 30, 2014.
89. M. Pontiki et al., "SemEval-2016 Task 5: Aspect Based Sentiment Analysis," Proc. 10th Int'l Workshop on Semantic Evaluation, pp. 19- 30, 2016.
90. M. Hu and B. Liu, "Mining and summarizing customer reviews," Proc. 10th ACM SIGKDD Int'l Conf. Knowledge Discovery and Data Mining, pp. 168-177, 2004.
91. M. Mitchell, J. Aguilar, T. Wilson, and B. Van Durme, "Open Domain Targeted Sentiment," Proc. 2013 Conf. Empirical Methods in Natural Language Processing, pp. 1643-1654, 2013.
92. H. Jangid, S. Singhal, R. R. Shah, and R. Zimmermann, "AspectBased Financial Sentiment Analysis using Deep Learning," Companion Proc. The Web Conference 2018, pp. 1961-1966, 2018.
93. M. Saeidi, G. Bouchard, M. Liakata, and S. Riedel, "SentiHood: Targeted Aspect Based Sentiment Analysis Dataset for Urban Neighbourhoods," Proc. COLING 2016, The 26th Int'l Conf. Computational Linguistics, pp. 1546-1556, 2016.
94. M. Wojatzki, E. Ruppert, S. Holschneider, T. Zesch, and C. Biemann, "Germeval 2017: Shared Task on Aspect-Based Sentiment in Social Media Customer Feedback," Proc. GermEval 2017 – Shared Task on Aspect-based Sentiment in Social Media Customer Feedback, pp. 22-29, 2017.
95. G. Ganu, N. Elhadad, and A. Marian, "Beyond the Stars: Improving Rating Predictions using Review Text Content," Proc. 12th Int'l Workshop on the Web and Databases, pp. 1-6, 2009.
96. J. McAuley, J. Leskovec, and D. Jurafsky, "Learning Attitudes and Attributes from Multi-Aspect Reviews," Proc. 2012 IEEE 12th Int'l Conf. Data Mining, pp. 1020-1025, 2012.
97. Z. Luo, S. Huang, F.F. Xu, B.Y. Lin, H. Shi, and K. Zhu, "ExtRA: Extracting Prominent Review Aspects from Customer Feedback," Proc. 2018 Conf. Empirical Methods in Natural Language Processing, pp. 3477-3486, 2018.
98. H. Wang, C. Wang, C. Zhai, and J. Han, "Learning Online Discussion Structures by Conditional Random Fields," Proc. 34th Int'l ACM SIGIR Conf. Research and Development in Information Retrieval, pp. 435-444, 2011.
99. H. Ye, Z. Yan, Z. Luo, and W. Chao, "Dependency-Tree Based Convolutional Neural Networks for Aspect Term Extraction," Pacific-Asia Conf. Knowledge Discovery and Data Mining, Springer, pp. 350-362, 2017.
100. K. Schouten, F. Frasincar, and F. de Jong, "Ontology-enhanced aspect-based sentiment analysis," Int'l Conf. Web Eng., Springer, pp. 302-320, 2017.
101. M. Z. Asghar, A. Khan, S.R. Zahra, S. Ahmad, and F.M. Kundi, "Aspect-Based Opinion Mining Framework using Heuristic Patterns," Cluster Computing, pp. 1-19, 2017.
102. A. Konjengbam, N. Dewangan, N. Kumar, and M. Singh, "Aspect Ontology based Review Exploration," Electronic Commerce Research and Applications, vol. 30, pp. 62-71, 2018.
103. Z. Hai, K. Chang, and J.j. Kim, "Implicit Feature Identification via Co-Occurrence Association Rule Mining," Int'l Conf. Intelligent Text Processing and Computational Linguistics, pp. 393-404, 2011.
104. W. Wang, H. Xu, and W. Wan, "Implicit Feature Identification via Hybrid Association Rule Mining," Expert Systems with Applications, vol. 40, no. 9, pp. 3518-3531, 2013.
105. Y. Zhang and W. Zhu, "Extracting Implicit Features in Online Customer Reviews for Opinion Mining," Proc. 22nd Int'l Conf. World Wide Web, pp. 103-104, 2013.
106. H. Sayyadi and L. Raschid, "A Graph Analytical Approach for Topic Detection," ACM Trans. Internet Technology, vol. 13, no. 2, p. 4, 2013.
107. S. de Kok, L. Punt, R. van den Puttelaar, K. Ranta, K. Schouten, and F. Frasincar, "Aggregated Aspect-Based Sentiment Analysis with Ontology Features," Progress in Artificial Intelligence, vol. 7, no. 4, pp. 295-306, 2018.
108.]K. Khan, B. Baharudin, and A. Khan, "Identifying Product Features from Customer Reviews Using Hybrid Patterns," Int'l Arab J. Information Technology, vol. 11, no. 3, pp. 281-286, 2014.
109. S. Kiritchenko, X. Zhu, C. Cherry, and S. Mohammad, "NRCCanada-2014: Detecting Aspects and Sentiment in Customer Reviews," Proc. 8th Int'l Workshop on Semantic Evaluation, pp. 437- 442, 2014.
110. J. Wagner et al., "Dcu: Aspect-Based Polarity Classification for Semeval Task 4," Proc. 8th Int'l Workshop on Semantic Evaluation, pp. 223-229, 2014.
111. L. Dong, F. Wei, C. Tan, D. Tang, M. Zhou, and K. Xu, "Adaptive Recursive Neural Network for Target-Dependent Twitter Sentiment Classification," Proc. 52nd Ann. Meeting of the Association for Computational Linguistics, pp. 49-54, 2014.
112. D.-T. Vo and Y. Zhang, "Target-Dependent Twitter Sentiment Classification with Rich Automatic Features," Twenty-Fourth Int'l Joint Conf. Artificial Intelligence, 2015.
113. D. Tang, B. Qin, X. Feng, and T. Liu, "Effective LSTMs for TargetDependent Sentiment Classification," Proc. COLING 2016, the 26th Int'l Conf. Computational Linguistics, pp. 3298-3307, 2015.
114. D.H. Pham and A.C. Le, "Learning Multiple Layers of Knowledge Representation for Aspect Based Sentiment Analysis," Data and Knowledge Eng., vol. 114, pp. 26-39, 2018.
115. K. Ghasedi and H. Huang, "Sentiment Analysis via Deep Hybrid Textual-Crowd Learning Model," Proc. 32nd AAAI Conf. Artificial Intelligence, pp. 1563-1570, 2018.

DOI: 10.5281/zenodo.3943130

116. B. Huang and K. Carley, "Parameterized Convolutional Neural Networks for Aspect Level Sentiment Classification," Proc. 2018 Conf. Empirical Methods in Natural Language Processing, pp. 1091- 1096, 2018
117. M. D. Smucker, D. Kulp, and J. Allan, "Dirichlet Mixtures for Query Estimation in Information Retrieval," Center for Intelligent Information Retrieval, Univ. Massachusetts, pp. 1-13, 2005.
118.]W. Xue and T. Li, "Aspect Based Sentiment Analysis with Gated Convolutional Networks," Proc. 56th Ann. Meeting of the Association for Computational Linguistics, pp. 2514–2523, 2018.
119. W. Zhao et al., "Weakly-Supervised Deep Embedding for Product Review Sentiment Analysis," IEEE Trans. Knowledge and Data Eng., vol. 30, no. 1, pp. 185-197, 2018.
120. X. Li, L. Bing, W. Lam, and B. Shi, "Transformation Networks for Target-Oriented Sentiment Classification," Proc. 56th Ann. Meeting of the Association for Computational Linguistics, pp. 946–956, 2018.
121. R. He, W.S. Lee, H.T. Ng, and D. Dahlmeier, "Exploiting Document Knowledge for Aspect-level Sentiment Classification," Proc. 56th Ann. Meeting of the Association for Computational Linguistics, pp. 579–585, 2018.
122. P. Ramachandran, P.J. Liu, and Q.V. Le, "Unsupervised Pretraining for Sequence to Sequence Learning," Proc. 2017 Conf. Empirical Methods in Natural Language Processing, pp. 383-391, 2017.
123. M.-T. Luong, Q.V. Le, I. Sutskever, O. Vinyals, and L. Kaiser, "Multi-Task Sequence to Sequence Learning," Proc. 2015 Int'l Conf. Learning Representations pp. 1-10, 2015.
124. Q. Hu, J. Zhou, Q. Chen, and L. He, "SNNN: Promoting Word Sentiment and Negation in Neural Sentiment Classification," Proc. 32nd AAAI Conf. Artificial Intelligence, pp. 3255-3262, 2018.
125. R. He, W. S. Lee, H. T. Ng, and D. Dahlmeier, "Effective Attention Modeling for Aspect-Level Sentiment Classification," Proc. 27th Int'l Conf. Computational Linguistics, pp. 1121-1131, 2018.
126. S. Sukhbaatar, J. Weston, and R. Fergus, "End-to-End Memory Networks," Proc. 28th Int'l Conf. Neural Information Processing Systems, pp. 2440-2448, 2015.
127. Y. Tay, L. A. Tuan, and S. C. Hui, "Dyadic Memory Networks for Aspect-Based Sentiment Analysis," Proc. 2017 ACM on Conf. Information and Knowledge Management, pp. 107-116, 2017.
128. D. Tang, B. Qin, and T. Liu, "Aspect Level Sentiment Classification with Deep Memory Network," Proc. 2016 Conf. Empirical Methods in Natural Language Processing, pp. 214-224, 2016.
129. P. Zhu and T. Qian, "Enhanced Aspect Level Sentiment Classification with Auxiliary Memory," Proc. 27th Int'l Conf. Computational Linguistics, pp. 1077-1087, 2018.
130. M. Zhang, Y. Zhang, and D. T. Vo, "Neural Networks for Open Domain Targeted Sentiment," Proc. 2015 Conf. Empirical Methods in Natural Language Processing, pp. 612-621, 2015.
131. J. Liu and Y. Zhang, "Attention Modeling for Targeted Sentiment," Proc. 15th Conf. European Chapter of the Association for Computational Linguistics, pp. 572-577, 2017.
132. S. Zheng and R. Xia, "Left-Center-Right Separated Neural Network for Aspect-based Sentiment Analysis with Rotatory Attention," arXiv preprint arXiv:.00892, p. 7, 2018.
133. P. Chen, Z. Sun, L. Bing, and W. Yang, "Recurrent Attention Network on Memory for Aspect Sentiment Analysis," Proc. 2017 Conf. Empirical Methods in Natural Language Processing, pp. 452- 461, 2017.
134. C. Fan, Q. Gao, J. Du, L. Gui, R. Xu, and K.F. Wong, "ConvolutionBased Memory Network for Aspect-based Sentiment Analysis," Proc. 41st Int'l ACM SIGIR Conf. Research & Development in Information Retrieval, pp. 1161-1164, 2018.
135. P. Liu, X. Qiu, and X. Huang, "Deep Multi-Task Learning with Shared Memory," Proc. 2016 Conf. Empirical Methods in Natural Language Processing, pp. 118-127, 2016.
136. X. Liu, J. Gao, X. He, L. Deng, K. Duh, and Y.-Y. Wang, "Representation Learning using Multi-Task Deep Neural Networks for Semantic Classification and Information Retrieval," Proc. 2015 Conf. North American Chapter of the Association for Computational Linguistics, pp. 912–921, 2015.
137. G. Balikas, S. Moura, and M.-R. Amini, "Multitask Learning for Fine-Grained Twitter Sentiment Analysis," Proc. 40th Int'l ACM SIGIR Conf. Research and Development in Information Retrieval, pp. 1005-1008, 2017.
138. C. Li, X. Guo, and Q. Mei, "Deep Memory Networks for Attitude Identification," Proc. 10th ACM Int'l Conf. Web Search and Data Mining, pp. 671-680, 2017.
139. Y. Yin, Y. Song, and M. Zhang, "Document-level Multi-Aspect Sentiment Classification as Machine Comprehension," Proc. 2017 Conf. Empirical Methods in Natural Language Processing, pp. 2044- 2054, 2017.
140. M. Yang, W. Tu, J. Wang, F. Xu, and X. Chen, "Attention Based LSTM for Target Dependent Sentiment Classification," Proc. 31st AAAI Conf. Artificial Intelligence, pp. 5013-5014, 2017.
141. Q. Liu, H. Zhang, Y. Zeng, Z. Huang, and Z. Wu, "Content Attention Model for Aspect Based Sentiment Analysis," Proc. 2018 Conf. World Wide Web, pp. 1023-1032, 2018.
142. S. Gu, L. Zhang, Y. Hou, and Y. Song, "A Position-aware Bidirectional Attention Network for Aspect-level Sentiment Analysis," Proc. 27th Int'l Conf. Computational Linguistics, pp. 774- 784, 2018.
143. N. Majumder, S. Poria, A. Gelbukh, M. S. Akhtar, E. Cambria, and A. Ekbal, "IARM: Inter-Aspect Relation Modeling with Memory Networks in Aspect-Based Sentiment Analysis," Proc. 2018 Conf. Empirical Methods in Natural Language Processing, pp. 3402-3411, 2018.
144. B. Wang and W. Lu, "Learning Latent Opinions for Aspect-level Sentiment Classification," Proc. 32nd AAAI Conf. Artificial Intelligence, pp. 5537-5544, 2018.
145. D. Hazarika, S. Poria, P. Vij, G. Krishnamurthy, E. Cambria, and R. Zimmermann, "Modeling Inter-Aspect Dependencies for Aspect-Based Sentiment Analysis," Proc. Conf. North American Chapter of the Association for Computational Linguistics, pp. 266- 270, 2018.
146. P. Lin, M. Yang, and J. Lai, "Deep Mask Memory Network with Semantic Dependency and Context Moment for Aspect Level Authorized licensed use limited to Sentiment Classification," Proc. 28th Int'l Joint Conf. Artificial Intelligence, pp. 5088-5094, 2019.
147. J. Li, H. Yang, and C. Zong, "Document-level Multi-aspect Sentiment Classification by Jointly Modeling Users, Aspects, and Overall Ratings," Proc. 27th Int'l Conf. Computational Linguistics, pp. 925-936, 2018.
148. J. McAuley, R. Pandey, and J. Leskovec, "Inferring Networks of Substitutable and Complementary Products," Proc. 21th ACM SIGKDD Int'l Conf. Knowledge Discovery and Data Mining, pp. 785-794, 2015.

DOI: 10.5281/zenodo.3943130

149. B. Pang and L. Lee, "A Sentimental Education: Sentiment Analysis using Subjectivity Summarization based on Minimum Cuts," Proc. 42nd Ann. Meeting on Association for Computational Linguistics, pp. 271-278, 2004.
150. H. Wang, Y. Lu, and C. Zhai, "Latent Aspect Rating Analysis on Review Text Data: A Rating Regression Approach," Proc. 16th ACM SIGKDD Int'l Conf. Knowledge Discovery and Data Mining, pp. 783-792, 2010.
151. M. Yang, Q. Qu, X. Chen, C. Guo, Y. Shen, and K. Lei, "FeatureEnhanced Attention Network for Target-Dependent Sentiment Classification," Neurocomputing, vol. 307, pp. 91-97, 2018.
152. M. Zhang, Y. Zhang, and D.-T. Vo, "Gated Neural Networks for Targeted Sentiment Analysis," Proc. 13th AAAI Conf. Artificial Intelligence, pp. 3087-3093, 2016.
153. S. Wang, S. Mazumder, B. Liu, M. Zhou, and Y. Chang, "TargetSensitive Memory Networks for Aspect Sentiment Classification," Proc. 56th Ann. Meeting of the Association for Computational Linguistics, pp. 957-967, 2018.
154. D. Ma, S. Li, X. Zhang, and H. Wang, "Interactive Attention Networks for Aspect-Level Sentiment Classification," Proc. 76th Int'l Joint Conf. Artificial Intelligence, pp. 4068-4074, 2017.
155. X. Ma, J. Zeng, L. Peng, G. Fortino, and Y. Zhang, "Modeling Multi-Aspects within one Opinionated Sentence Simultaneously for Aspect-Level Sentiment Analysis," Future Generation Computer Systems, vol. 93, pp. 304-311, 2019.
156. F. Fan, Y. Feng, and D. Zhao, "Multi-Grained Attention Network for Aspect-Level Sentiment Classification," Proc. 2018 Conf. Empirical Methods in Natural Language Processing, pp. 3433-3442, 2018.
157. Ma, K. Wang, T. Qiu, A.K. Sangaiah, D. Lin, and H.B. Liaqat, "Feature-based Compositing Memory Networks for AspectBased Sentiment Classification in Social Internet of Things," Future Generation Computer Systems, vol. 92, pp. 879-888, 2017.
158. J. Qiu, C. Liu, Y. Li, and Z. Lin, "Leveraging Sentiment Analysis at the Aspects Level to Predict Ratings of Reviews," Information Sciences, vol. 451, pp. 295-309, 2018.
159. H.X. Yang, Z.X. Wu, C. Zhou, T. Zhou, and B.-H. Wang, "Effects of Social Diversity on the Emergence of Global Consensus in Opinion Dynamics," Physical Review E, vol. 80, no. 4, p. 046108, 2009.
160. J. Aronfreed, "The Socialization of Altruistic and Sympathetic Behavior: Some Theoretical and Experimental Analyses," Altruism and Helping Behavior, pp. 103-126, 1970.
161. K. Kaplanidou et al., "Quality of Life, Event Impacts, and MegaEvent Support among South African Residents Before and After The 2010 Fifa World Cup," J. Travel Research, vol. 52, no. 5, pp. 631-645, 2013.
162. F. Fu and L. Wang, "Co-evolutionary Dynamics of Opinions and Networks: From Diversity to Uniformity," Physical Review E, vol. 78, no. 1, pp. 0161041-1060414, 2008.
163. T. Carletti, S. Righi, and D. Fanelli, "Emerging Structures in Social Networks Guided by Opinions' Exchanges," Advances in Complex Systems, vol. 14, no. 01, pp. 13-30, 2011.
164. K. Sznajd-Weron and J. Sznajd, "Opinion Evolution in Closed Community," Int'l J. Modern Physics C, vol. 11, no. 06, pp. 1157- 1165, 2000.
165. P. Fan, P. Li, H. Wang, Z. Jiang, and W. Li, "Opinion Interaction Network: Opinion Dynamics in Social Networks with Heterogeneous Relationships," Proc. ACM SIGKDD Workshop on Intelligence and Security Informatics, p. 8, 2012.
166. H. Liang, Y. Yang, and X. Wang, "Opinion Dynamics in Networks with Heterogeneous Confidence and Influence," Physica A: Statistical Mechanics and its Applications, vol. 392, no. 9, pp. 2248-2256, 2013.
167. L.T. Nguyen, P. Wu, W. Chan, W. Peng, and Y. Zhang, "Predicting Collective Sentiment Dynamics from Time-Series Social Media," Proc. First Int'l Workshop on Issues of Sentiment Discovery and Opinion Mining, p. 8, 2012.
168. P.C. Guerra, W. Meira Jr, and C. Cardie, "Sentiment Analysis on Evolving Social Streams: How Self-Report Imbalances can Help," Proc. 7th ACM Int'l Conf. Web Search and Data Mining, pp. 443-452, 2014.
169. L. Zhang, Y. Jia, X. Zhu, B. Zhou, and Y. Han, "User-Level Sentiment Evolution Analysis in Microblog," China Communications, vol. 11, no. 12, pp. 152-163, 2014.
170. N. Charlton, C. Singleton, and D.V. Greetham, "In the Mood: The Dynamics of Collective Sentiments on Twitter," Royal Society Open Science, vol. 3, no. 6, pp. 160-162, 2016.
171. S. Poria, N. Majumder, D. Hazarika, E. Cambria, A. Gelbukh, and A. Hussain, "Multimodal Sentiment Analysis: Addressing Key Issues and Setting Up the Baselines," IEEE Intelligent Systems, vol. 33, no. 6, pp. 17-25, 2018.

DOI: 10.5281/zenodo.3943130